

Horrific Suffering, Divine Hiddenness, and Hell: The Place of Freedom in a World Governed by God

Keith DeRose

Part One: A Powerful Problem	1
1. A Brief Look at Where We're Going: The Problem of Horrific Suffering, Two Other Forms of the Problem of Evil, and the Place of Human Free Will in a World Governed by a Wholly Good God	1
2. The Problem of Horrific Suffering	2
3. What's the Problem Here, Anyway?: Just Look at the Horrific Evils of Our World!	8
4. OK, So What's the Argument, Then?: The "Simple" Argument from Horrific Evils.....	13
5. OK, But What's the Case for That Premise, Then?	17
6. But What's the Case for That Premise, Then?: Rowe's Appeal and the "Good Cop, Bad Cop" Routine	19
7. "Making the Case" for the Premise of the Simple Argument from Horrific Evils	21
8. A Very Powerful Appearance, A Very Powerful Argument.....	25
9. Wykstra's Challenge: CORNEA and Suspect Appearances	33
10. The Importance of the Basis of an Appearance Claim	37
11. What Our Atheologian Is Relying on and How They Pass CORNEA: An Appearance Reasonably Relied Upon	39
12. "But We're All Newbies Here!": Sensible Skeptical Theism vs. the More Aggressive Varieties of It	42
13. Plantinga and "Probabilistic" vs. "Deductive" Arguments from Evil	45
14. Plantinga's Attack on Probabilistic Arguments from Evil and What the Attack Gets Right ..	49
15. After Plantinga's Attack: Our Still Extremely Powerful "Probabilistic" Argument from Evil ..	51
16. Distinguishing Probabilistic Cases from Evil from the Trash Talking often Associated with Them	54

17. Plantinga on the Powerful Non-Argument from Horrific Evils: “Isn’t It Just Apparent, Just Evident?”	56
Part Two: Solutions: Theodicy, Anti-Theodicy, and Goalposts	60
18. Theodicy, in the Form of a Credible Account of What God’s Reasons for Allowing the Horrific Evils of Our World Might Be	60
19. Anti-Theodicy and Some Explanation of Our Goalposts.....	62
20. God, Freedom, and Immortality: Theism as Part of a Package	66
Part Three: The Basic Free Will Defense and the Potentially Immense Importance of Libertarian Freedom	69
21. The Free Will Defense: The Theist’s Best Hope, Only Hope?	69
22. The Famous, Intuitive Free Will Defense: The Basic Idea	71
23. The Power of the Free Will Defense: The Potentially Immense Importance of Libertarian Freedom.....	76
24. The Free Will Defense as an Attempt to Get God’s Relation to Evil on the Good Side of Key Ethical Distinctions: Evil as a Merely Allowed Side-Effect	77
Part Four: Freedom without Risk?: God’s Knowledge of What We Would and Will Freely Do and the Two Gods of the Free Will Defense.....	80
25. The Gods of the Free Will Defense: Plantinga’s Molinist God	80
26. Against Molinism	85
27. Middle Knowledge and Simple Foreknowledge: Standing or Falling Together	89
28. The Freewill Theist’s Risk-Taking God, the Problem of Evil, and Providential Control	92
Part Five: The Meaning of “Free” and the Potentially Immense Importance of Libertarian Freedom	100
29. The Meaning of “Free” in Philosophy-Speak: An Elucidation by Means of Examples	100
30. The Meaning of “Free”: Our Example of a Free Action	102
31. Actions that Do Not Count as “Free” in Ordinary Talk, but Are “Free” in Our Philosophical Sense.....	104
32. The Meaning of “Free” in Philosophy-Speak: An Example of an Action (or Is it, Really?) that Is Not Free	106
33. Pessimists and the Meaning of “Free” in Philosophy-Speak: But What if This Distinction Is Bogus?.....	109
34. The Libertarian Hope and the Potentially Immense Importance of Libertarian Freedom..	112

<i>Part Six: The Credibility of the Libertarian Hope</i>	116
35. The Libertarian Hope: Initial Appearances.....	116
36. Intuitions, Arguments, and the Credibility of the Libertarian Hope	118
37. Taking Compatibilism, and Other Positions, Seriously: Thinking through the Issues from Some Starting Points	119
38. “The Best Argument for Incompatibilism”	125
39. “The Best Argument for Compatibilism”	129
40. An Argument for the Unacceptability of Incompatibilism?	132
41. Is the Libertarian’s Allegiance to Our Freedom Too Fragile?	134
42. Buoyant Libertarianism.....	135
43. Buoyant Libertarianism: A Completely Legitimate Libertarian Option.....	140
44. “Is Incompatibilism Intuitive?”: “What if Determinism Turns Out to Be True?” Cases	145
45. “What of Worlds of Which Determinism Is True?”, “What if Determinism Had Been True?”	149
46. The Experimental Philosophy Battle over Intuitions: Attempts to Explain Away the Apparent Intuitions of the Other Side.....	154
47. Asking an Abstract Question	160
48. Surveying Our Freedom	164
49. The Great “Upshot” Argument against Free Will	166
50. Freedom’s Place in the Causal Structure of the World: The Shape of the Libertarian Hope	173

Horrific Suffering, Divine Hiddenness, and Hell: The Place of Freedom in a World Governed by God

Keith DeRose

Part One: A Powerful Problem

1. A Brief Look at Where We're Going: The Problem of Horrific Suffering, Two Other Forms of the Problem of Evil, and the Place of Human Free Will in a World Governed by a Wholly Good God

Asked in a public forum by a fellow philosopher¹ to give examples ("without getting into details") of one argument for and another argument against the existence of God that I thought a reasonable person could find plausible, I didn't even have to think about which argument to use on the negative side:

I'm going to have to be conventional here and go with the usual suspect: the argument from evil. Without getting into any details, you can feel the force of the argument by choosing a suitably horrific example (the Holocaust, children dying of cancer) that leads you to say, "There's no way a perfectly good God would have allowed *that!*" There is a huge, often fascinating, discussion that tries to refute such arguments. But I find this intuitively powerful case does stand up to scrutiny, at the very least to the extent that someone could reasonably accept it at the end of the day. I suspect that even God thinks there is something wrong with you if you are not at least tempted by such an argument from evil.

That argument from horrific suffering, and the problem it gives rise to for standard theism (belief in a God that is wholly good, as well as all-powerful and all-knowing), is the main focus of this book, though our wrestling with this problem will involve us in two further problems for standard theism: the problem of divine hiddenness and the problem of hell. Though I hope to uncover what relief there is to be found from these problems, I certainly

¹ "Why Take a Stance on God?", an interview Gary Gutting did with me in *The Stone*, the *New York Times* philosophy blog, 18 September 2014; reprinted with slight revisions as "Religion and Knowledge," in Gary Gutting, *Talking God: Philosophers on Belief* (New York: W.W. Norton, 2017) pp. 172-186.

won't be finally settling any of them. But at least there will be three, and not just one, problems that we will be failing to solve here.

"Free will defenses" have constituted perhaps the main type of theistic response to problems like these, and not just in the thoughts of philosophers and theologians, but in wide stretches of the broader culture, to the point that if you ask some theist why there is so much evil in the world, despite its being governed by a perfectly good God, or why some people will end up in hell, what you're likely to hear as the "short answer" response is some quick version of the free will defense. It turns out that I accept, in about the strongest way possible, the account of freedom that fuels such responses, and I am convinced that human freedom must play a huge role in God's relation to the horrors of our world—well, as it pretty much must play a huge role in all of God's dealings with us (if God does exist and we are free). Yet I still think that standard free will defenses fail, and quite miserably, at least when directed at the most gripping forms of our problems, and so the "go to" short answer to such questions, as well as longer, more careful answers, should be flipped to something else. So a main concern will be explaining why such accounts fail, but then also what role I think human free will might play in a credible account of God's relation to evils.

"The problem of evil" is the problem of squaring the existence of the (wholly good) God of standard theism with the occurrence of evils, where "evil" is used broadly to designate anything bad, including instances of suffering, whether or not those "evils" are connected with any immoral actions anyone might be performing. "The" problem of evil is best viewed as a family of related problems (or as a problem that comes in several different forms) that vary from each other in terms of which evils are being considered. Our three problems—hiddenness, hell, and horrors—are then forms of the problem of evil. Indeed, three of its toughest forms.

2. The Problem of Horrific Suffering

But the greatest of these is horrors.

Well, I don't mean that the problem of horrific suffering is straightforwardly *tougher to solve* than is the problem of hell. The difficulty of the latter varies greatly with how it is construed. As some readers were no doubt thinking when I threw hell into the mix, it has an importantly different character from the other problems, as it is not based on something bad that we the living have observed in the world, but is rather a theological commitment that some theists take on. And of course, how tough it is to square that commitment with God's goodness depends on just what kind of hell is in question. Some unspeakably brutal (and

sometimes literally inconceivably brutal), yet quite common and traditional, doctrines of hell explicitly specify that hell is chock full of horrific suffering, including physical pain, that somehow manages to be far worse than anything ever suffered in any earthly life, and that befalls not just “Hitler types,” but also what we take to be ordinarily decent people, who just have not done what’s needed to have their run-of-the-mill sins forgiven. If that’s the version of theism that constitutes one’s live theistic options, the problem of hell *includes* an extreme form of the problem of horrific suffering. (Though henceforth, we will construe the “problem of horrific suffering/evil” so that it encompasses only the horrors of this earthly life.) But others, for whom the live theistic options are different, can instead opt for some version of “hell lite” — or easily solve the problem of hell by jettisoning hell altogether. (More on this when we get to hell!)

But I do think the problem of (earthly) horrific evils is the “greatest” problem of evil in terms of being the most influential. Where atheism is based on argument at all, some form of the argument from horrific evils is among the first arguments cited, and many actual atheists join me in having this as *the* first and main argument for atheism that they reach for.

Provocative as my indented statement about the problem quoted at the start of the previous section may seem, especially at the end of that passage, I really was trying to be Mr. Moderate Peacemaker there: What had prompted the request for examples of pro- and anti-God arguments was my claiming that, because there are arguments on both sides of the issue that one could reasonably accept, there was plenty of room for the rationality of both theism and atheism, and I was appealing to the argument from horrific suffering to make that room for reasonable atheism.

But that may badly undersell the argument’s power. It can (and, to many, does) seem to be an intellectual steamroller, destroying the possibility of reasonable belief in God and leaving atheism as the *only* intelligent position available on the question of the existence of the God of standard theism (which we will construe as positing the existence of a *wholly good* God).

I was recently confronted with a stark case of this while reading a draft of a book about religious disagreement by a colleague of mine. Seeking examples of sharp disagreement over religious questions, they had found a terrific example of someone who rejects the wholly good God of standard theism and expresses that judgment with great and blunt assurance. My colleague quoted this “very accomplished philosopher” as taking up the question of the existence of a “benevolent almighty creator,” and then writing this:

Well, really, now, how credible do you find BENEVOLENT ALMIGHTY CREATOR [the claim that there is such a being]? Or, for that matter, how credible do you find any proposition even just remotely like that traditional, or traditional-sounding, doctrine? Well,

you can put me down for a big loud, “Not Very!” And, boy, it’s a veritable landslide here, as the case for a negative answer is almost absurdly overwhelming.

What is this “almost absurdly overwhelming” case? Our case from horrific evils. As my colleague explained, our prominent philosopher then takes a few paragraphs to catalog some of the worst suffering of our world, before concluding:

How horrible is *that*, I ask you, all you who dare to uphold, quite as heartlessly propounded as it’s brainlessly affirmed, any claim that’s even the least bit like the utterly incredible BENEVOLENT ALMIGHTY CREATOR? It’s certainly too horrible, I tell you, to have anyone believing, even the least bit reasonably, or the least bit intelligently, anything remotely like that extremely dubious doctrine.

Heartless *and* brainless, oh my! Upholders of this “utterly incredible” tenet brace ourselves for being painted as other characters from *The Wizard of Oz*.

I was jolted by these passages as I read them quoted—but for different reasons and in a quite different way from how other upholders of BENEVOLENT ALMIGHTY CREATOR would likely be provoked by them. It was for me in large part a jolt of recollection. The philosopher being quoted was one of my oldest friends in the field, and their quoted words seemed eerily familiar. Peter Unger and I had worked closely together for a number of years, reading and critiquing each other’s writings, but (largely due to my moving) I had dropped out of that collaboration as he began working on the book that my current colleague had found.² So I was *reading* Unger’s remarks for the first time. But I had heard similar remarks from him (just once, but memorably) before—in fact, directed at me, the first time we met in person.

I was in my last year in graduate school, and was on the (extremely tight—as it is today, though it’s even worse now) philosophy job market. I had no job yet, but had a few possibilities going. New York University, where Unger worked (and still works) was interested enough that when they heard that my flight between two job talks I was giving made a stop at JFK, they told me to get off the plane there, grab a taxi to NYU, and they’d grill me for a couple of hours, before putting me on a train to Philadelphia, where I would be grilled at Temple University the next day. If it went well, and if NYU’s first choice didn’t take their offer, this could result in my getting NYU’s job.

² The book is Unger’s *All the Power in the World* (Oxford University Press, 2005). Unger takes up God’s existence at pp. 501-508 of that book. The indented quotations above are from p. 506 and p. 507, respectively. I did know the basic topic of the book as Peter was writing it, but did not realize it included a foray into philosophy of religion. I sometimes wonder how things would have gone if I had continued to work with Peter as he wrote that book. I can reasonably suppose at least this much: There would have been some very interesting phone conversations.

NYU didn't want me to read my "job talk" paper to them before taking questions. Instead, I sent the paper to them so they could read it ahead of time, and when I arrived, they would proceed directly to the "grilling", asking questions about that paper, but then also about other things. My paper was not about religious matters, but my CV listed philosophy of religion as one of my "areas of competence." I thought things were going extremely well as we discussed my paper. But then when the questions moved away from that paper to a more general discussion of my work, Peter jumped in and first asked whether I was, as he suspected, an upholder of BENEVOLENT ALMIGHTY CREATOR (to use the allcaps label he later came to use for this doctrine). Having ascertained that I was, he launched into a diatribe that was in tone as well as substance very much like the portions of his later book quoted and described above.

Unger's ramped-up rhetoric was in a way shocking, in that I would not have at all expected to encounter it in such a setting (or in many other settings, for that matter). But why not? What perhaps should have been strange is that I hadn't been subjected to such charges before. Since so far back in childhood that I can't recover its beginning, when I learned of some of the horrific evils this world contains (I was always most moved by accounts of some of the gruesome examples of intentional torture people have been subjected to at the hands of others), I have felt the threat these constitute for belief in a wholly good God. "Wouldn't such a God have stepped in and prevented *that*?"; "If not *that*, what *would* it take to get God to prevent an evil?"; "What possible reason could God have for not preventing *that*?". Like many before, after, and around me, I imagine I often wondered in words that were, other than the distinctive three-capitalized-word phrase used to refer to God, close to these of Unger's: "In all this, where, on God's good green earth, is the Benevolent Almighty Creator? Nowhere, that's where."³ Do the amounts and kinds of suffering this world contains render belief in a wholly good God "utterly incredible"? Well, why not? Are they horrible enough to prevent me from believing in a good God in a way that's at all reasonable? Well, that's the worry. If not, how bad *would* the evils of this world have to be to yield such conclusions?

Providence or Fortune had brought me, six years earlier, to what was likely for me the best place in the world for me to wrestle with such questions. As I fairly blindly stumbled into UCLA's graduate program in philosophy, I had no idea that Marilyn McCord Adams, who taught there, would be working on her paper on "Horrendous Evils and the Goodness of God," and then her book of the same name,⁴ and that I would get to study the problem of evil with her and with Robert Adams, and read much of the material in philosophical theology best suited to help me. This gave me plenty to talk about, and as the occasion called for it (a paper due for a

³ *All the Power in the World*, p. 506.

⁴ "Horrendous Evils and the Goodness of God," *Proceedings of the Aristotelian Society*, Supplementary Volumes, Vol. 63 (1989): 297-310; *Horrendous Evils and the Goodness of God*, Cornell University Press, 1999.

seminar), to write about. But I was never faced with the threat that horrific suffering posed for my theism. To her great credit, Marilyn was facing that challenge squarely.⁵ And that enabled me to engage the challenge at a critical distance, by discussing the pros and cons of Marilyn's approach (and those of others). But I did not face the pointed challenge, directed squarely at me: "Well, so, in light of all that stuff you've been reading, but also in light of the horrors of this world, for real now: isn't it still utterly incredible to suppose there is a wholly good God?" To the extent that anyone approached issuing that challenge, some lighthearted semi-joking (or fully joking) response about how hard that problem is ("I plan on solving that next quarter") was enough to satisfy them—or at least to get them to politely act satisfied. Or at least to let the matter go. And I very politely was no longer pressing myself.⁶

But in the philosophy department conference room at NYU, my luck had run out. I was approaching a moment of truth.

Well, or maybe not. The then-chairperson of NYU's philosophy department was visibly upset, I think by the tone of Unger's line of questioning, and had tried, extremely unsuccessfully, to interrupt and derail it a couple of times. And as soon as the lecture/question was finished, they jumped right in and told me, "You don't have to answer that." Reprieve!

⁵ Well, Adams was at least *in effect* facing the problem. In a move I always found frustrating, she consistently set up her discussion so that the problem she was explicitly addressing was the *logical* problem of horrific suffering, which is focused on the question of whether it is logically possible that there should have been a perfectly good God and also horrific evils (see, e.g., "Horrendous Evils and the Goodness of God," p. 298, where Adams sets the task of exhibiting "compossibility of Divine perfection with evils in the amounts and of the kinds found in the actual world," and *Horrendous Evils and the Goodness of God*, p. 15). I find the problem of whether, in light of the horrific suffering our world contains, it is *credible* or *believable* to suppose that a wholly good God actually does exist to be a far more pressing and important problem. Thankfully, much of what Marilyn writes is germane to this more important problem—which is in large part why her work was and is so exciting.

⁶ The paper I wrote for the seminar that Marilyn Adams and Bob Adams co-taught on the problem of evil came out a few years later as "Plantinga, Presumption, Possibility, and the Problem of Evil," *Canadian Journal of Philosophy* 21 (1991): pp. 497-512. There I was writing about the more aggressive, and more easily answered, form of the problem of evil that we will consider in section 13, and that alleges a logical *inconsistency* between the existence of the God of standard theism and the existence of any evil at all in the world. I argued that Plantinga's attempt to deal with such form of the problem was not, as it was widely held to be, successful. So, I was there pushing for a conclusion that falls on the anti-theism side of things. However, my conclusion was very limited, and, by focusing on a limited question concerning a form of the problem of evil that was not its most gripping form (as it doesn't even get into just how horrible the evils of this world often are), my work on that paper is a good example of how I was able to feel I was "working on the problem [of evil]" without having to face the pointed question that the problem in its most gripping form should have pressed on me.

But then I heard a surprising response come out of my own mouth: “No, I want to answer it.”

“You’d better have something really good up your sleeve!” I thought at that part of me that had just spoken.

As I started talking, trying to explain how one might believe in a wholly good God in the face of the horrors of our world, two things occurred to me. First, and somehow surprisingly, the things I had read in the previous few years that tried to answer the problem really did seem helpful, and did provide material for significant relief, this relief coming almost entirely in the form of accounts of why God might allow horrific evils—so, in a common use of the term, in the form of theodicies! (And further helpful writings have appeared since then.) My not having to squarely face the big question before had kept me from realizing that. But second, there was no piece of writing that I could point to as really providing that much help by itself. (And that would still be the case today.) I was cobbling various things together, throwing in an idea of my own here and there.

No, actually, I didn’t, and still don’t, have anything great up my sleeve. Well, I did and will be here appealing to some truly great things, as I suppose one has to. (“The worst evils demand to be defeated by the best goods,” as Marilyn taught me.⁷) But the result won’t be a response to the problem that is “great” in the sense of being extremely effective in a way to make the problem lose its force. Still, I did in that moment come to think of this problem as something that could and should be squarely faced, and didn’t have to be joked aside. It’s a fearful challenge, but not a clearly destructive steamroller. Or so I think. Well, or at least hope. I’m sure many will vehemently disagree, even after reading what I have to say. But I think the time has come to at least say it.

My audience’s reaction? Well, I did end up getting that job.⁸ On the other hand, Unger did later still write what he wrote, so I think it’s fair to say I didn’t assuage his worries. And my guess is: probably not those of his (then soon to be our) colleagues, either. But, then again, I’ve improved the explanation I give since then.

And you, dear reader? Well, I don’t know where you stand with respect to BENEVOLENT ALMIGHTY CREATOR, so it’s very hard to say how you should respond to any of this. Those of

⁷ And as she also wrote, at “Horrendous Evils and the Goodness of God,” p. 309.

⁸ As I recall, while NYU’s first choice did turn them down to take another job, the philosophy department convinced the NYU administration to offer me a job before he made his decision, and to hire both of us if we both accepted. I suppose this is further reason to think my presentation was at least not a total flop, though it’s quite possible that I failed badly in responding to Unger’s question, but they decided to ignore that part of the interview, or at least not weigh it too heavily.

you who are completely unimpressed by our “argument from horrific suffering” so far, and can’t see why anyone would find it at all persuasive or even troubling: I’ll try to move you to feel its power starting in the next section. Even if I fail to get you personally worried, I can perhaps still help you understand how others might find it so gripping, and reasonably so. But supposing we can all get on board in thinking the case can be a powerful one, where to go from there, as we see how the theist or would-be theist might seek relief from this problem, radically depends on where one is coming from. In the extreme case, if you are somehow infallibly certain that (a wholly good) God exists, you might be perfectly rational (well, at least insofar as your complete certainty of God’s existence is itself rational) in not being worried, at least so far as your own belief goes, in any of this, thinking, “I have absolutely no idea why God allows this awful stuff to happen, and I can see how that could stand in the way of others accepting that God exists, but *I* am infallibly certain that God does exist, and so that God’s existence is *somehow* compatible with even the worst things that have happened.” On the other extreme, if you are somehow certain on independent grounds that God does not exist, then none of this will matter to your own stance on theism, either. Of course, extremists of both kinds may still be interested in this as an interesting problem for others. For some in the middle regions between those two extremes, your ability to embrace or stick with something like BENEVOLENT ALMIGHTY CREATOR may largely hinge on how things turn out here. But even so, where you end up will of course depend on what other grounds you have for and against the claim, in addition to how well, if at all, this problem can be dealt with. But for reasons that will emerge, I think we can all profitably journey together.

3. What’s the Problem Here, Anyway?: Just Look at the Horrific Evils of Our World!

First, we must address whether our problem really is so gripping—to the extent that it should challenge the belief of almost any theist. I face quite a divide: Aside from how upfront he is about it, even in relevantly “mixed company,” my friend Peter is utterly typical of a very wide swath of skeptics in seeing this challenge from horrific evils as such a killer problem that it decisively settles it that there is no wholly good God. And some theists also immediately and viscerally feel the gripping challenge this problem presents for them. This includes some prominent Christian intellectuals. Here, for example, is one of the most prominent theistic philosophers alive, the eminent Richard Swinburne, with a strong statement indeed:

Just reflect on some of the horrors that we read about in our newspapers and history books: the prolonged cruelty of parents to lonely children, the torture of the innocent, the long-

drawn-out acute physical pain of some disease, and so on. If we cannot see all that as a reason for believing that there is no all-good and all-powerful being, when we cannot think of any reason why such a being should allow it all to happen, there really is something deeply wrong with us. We have lost our sensitivity to the good.⁹

And Swinburne is just one prominent example of thoughtful theists who admit this.

But you will also find those who claim to see *no good reason whatsoever* in all the horrific suffering of our world for thinking there is no wholly good God—and this not because (or in some cases, not just because) they have an account of why God might allow the horrors in question that they are happy with. Rather, they just profess bafflement at what the reason for doubting God’s existence here might possibly be. These baffled theists in turn, of course, *really* bewilder those who insist that not only is there a real problem here, but one so bad that it renders belief in God beyond respectable acceptance—and sometimes beyond even respectable discussion.

So, just what *is* the problem here? One way to press it brings in probability theory to structure the discussion, resulting in a “Bayesian Argument from Evil,” where one argues that the evidence of the suffering in our world is antecedently much more probable on some hypothesis that’s an alternative to theism than it is on theism itself, in order to argue (via steps that I’m skipping here) that theism is probably false. (This will be the path not traveled, so those who don’t follow this and the following paragraph need not worry about losing the discussion. Just pick it back up two paragraphs down, at “As our problem is usually pressed...”)¹⁰ For a good example, in his paper, “Pain and Pleasure: An Evidential Problem for Theists,”¹⁰ Paul Draper argues that crucial evidence concerning the pain and pleasure in our world is antecedently far more probable on his “Hypothesis of Indifference” that “neither the nature nor the condition of sentient beings on earth is the result of benevolent or malevolent actions performed by non-human persons” (p. 332) than it is on theism, and argues that this gives us “a *prima facie* good epistemic reason to reject theism—that is, a reason that is sufficient for rejecting theism unless overridden by other reasons for not rejecting theism” (p. 331).

⁹ Swinburne, *Providence and the Problem of Evil* (Oxford University Press, 1998), p. 23. Like the statement of mine that I quote at the very start of section 1, Swinburne says that there is something wrong with those who don’t feel the power of the argument—though Swinburne doesn’t put that swipe into the mind of God. I don’t think I encountered this passage until after doing that interview: I believe I finally read that book of Swinburne’s when doing my literature check for writing this book. It was very interesting (and more than a bit reassuring to me) to see Swinburne express such an opinion.

¹⁰ *Noûs* 23 (1989), pp. 331-350: Here Draper executes what he elsewhere calls a “Bayesian argument from evil”; he calls that type of argument that name in, e.g., his “The Problem of Evil,” in Thomas P. Flint and Michael C. Rea, ed., *The Oxford Handbook of Philosophical Theology* (Oxford University Press, 2009), pp. 332-351; see esp. p. 339.

But I will seek to vindicate what I take to be the intuitively powerful argument from horrific suffering as that intuitive argument is usually made, often by non-philosophers, and often powerfully: with no explicit appeal to probability theory. I view the two ways of running the argument as running parallel to one another, often turning on similar issues.¹¹ On the particular not-explicitly-probabilistic way of running the argument that I find most powerful, and will endorse as such, the reader will be left to judge the plausibility of a key premise that states that there is horrific evil in the world that wouldn't be there if there were a perfectly good God. I will be pointing toward various powerful ways that atheologians have tried to motivate acceptance of this premise, without appealing to the apparatus of probability theory. But it's certainly true that another way to advance the discussion is to bring in probabilities and start to ask the kinds of questions that drive Bayesian arguments: How antecedently likely are these evils on theism (or this or that particular theistic hypothesis), and how likely on this or that non-theistic hypothesis? This is a way I myself am often happy to pursue, and I will note *here and there* where what I'm doing overlaps with Draper's efforts, mainly taking advantage of how some of Draper's later works answer some recent attempts by theists to avoid the problem of horrific suffering, citing Draper's counters to those moves.

As our problem is usually pressed, much of the attention is directed at conveying just how horrific our world is at many points and at making this vivid to the audience. Unger's attack, described in the previous section, is typical (but for its extraordinary vituperativeness—and folksiness of style). Skeptics often catalogue some of the major types of evils that occur in our world, giving their audience a sense of just how immense, widespread, and common they are, but they also often zoom in on a few particular evils, or a few particular kinds of evils, to describe them in some excruciating detail, in order to give the listener a vivid sense of just how horrific they are. As in Unger's telling, the litany of horrors is often peppered with variations on the haunting question: "Wouldn't a wholly good God have prevented *that*?"

The focus horrors vary. In a case that has been important in philosophical discussions for the last several decades, and to which we will return shortly, William Rowe chooses to highlight an example of intense animal suffering:

¹¹ Draper has a very different attitude to the relation between these two types of running an argument from evil. Using "MT arguments" to denote the non-Bayesian variety (see *note 22*, below), he writes: "Problems such as these make me doubt that a convincing MT argument from evil can be constructed. I'm much more optimistic about what I call 'Bayesian' arguments from evil" ("The Problem of Evil," p. 339).

Suppose in some distant forest lightning strikes a dead tree, resulting in a forest fire. In the fire a fawn is trapped, horribly burned, and lies in terrible agony for several days before death relieves its suffering.¹²

Rowe's choice of animal suffering is likely guided by the thought that some of the often-proposed reasons God might have for allowing people to suffer horribly don't apply (at all, or as smoothly) to non-human animals. And Rowe's choice of suffering caused by natural events, rather than harm intentionally inflicted by human actions, is likely motivated to choose a case where "free will defenses" seem inapplicable.¹³

But the *standard* choice of highlighted evils seems to be horrific suffering, often at human hands, ending in death, by young human children—children developed enough to not only suffer greatly, but to appreciate how greatly they are suffering, while remaining young enough to be quite innocent and obviously underserving of any great suffering; to have their misery compounded by their complete inability to make any sense of what's happening to them; and, let's face it, to trigger our natural instinct to protect young children. (Note that Rowe, though he reaches outside our species for the victim of his focus horror, still chooses a (presumably cute) fawn, rather than a full-grown deer.) Focusing on victims of human tormenters does perhaps invite the invocation of the free will defense, but, for reasons we will consider in section xx, the standard free will defense seems quite powerless in the face of such truly horrific suffering, anyway, and the human perpetrator adds an element of interpersonal violation that makes the evil seem worse and for many intensifies the call on God to intervene should God exist.

The massively best-selling novel *The Shack*¹⁴ engaged many readers largely because (as its fans will tell you) it "deals with" or "does not back away from" the "hardest" questions/problems facing believers in a good God—where it's clear that the problem mainly in view is something in the vicinity of our problem of horrific suffering. The novel centers on its narrator's encounter with God after his already rocky relationship with God is destroyed when his young daughter is abducted and killed, in a grisly, disgusting, and horrific manner. Like Rowe's fawn, this fictional case is clearly meant to raise to our consideration the many actual cases it approximates.

The speech by Ivan in Dostoevsky's novel, *The Brothers Karamozov*, is perhaps the most famous prosecution of our problem, and it focuses largely on the sufferings of young humans. Let me here use the theologian David Bentley Hart's telling description of the speech. Hart is

¹² William Rowe, "The Problem of Evil and Some Varieties of Atheism," *American Philosophical Quarterly* 16 (1979): 335-341; p. 336.

¹³ [Note here](#) on the problem of animal suffering.

¹⁴ *The Shack*, at first self-published by the author, William P. Young, in 2007.

comparing it with Voltaire's *Poème sur le désastre de Lisbonne* (which I suppose should also be mentioned as an example of the pressing of our problem, anyway):

For all its power, however, Voltaire's poem is a very feeble thing compared to the case for "rebellion" against "the will of God" in human suffering placed in the mouth of Ivan Karamazov by that fervently Christian novelist Dostoevsky; for, while the evils Ivan recounts to his brother Alexey are acts not of impersonal nature but of men, Dostoevsky's treatment of innocent suffering possesses a profundity of which Voltaire was never even remotely capable. Famously, Dostoevsky supplied Ivan with true accounts of children tortured and murdered: Turks tearing babies from their mothers' wombs, impaling infants on bayonets, firing pistols into their mouths; parents savagely flogging their children; a five-year-old-girl tortured by her mother and father, her mouth filled with excrement, locked at night in an outhouse, weeping her supplications to "dear kind God" in the darkness; an eight-year-old serf child torn to pieces by his master's dogs for a small accidental transgression.¹⁵

Ivan comments explicitly on his choice of the young as his focus here, as he relates one of his several accounts (mentioned above in Hart's summary) of horrific mistreatment of children:

These educated parents subjected the poor five-year-old girl to every possible torture. They beat her, flogged her, kicked her, not knowing why themselves, until her whole body was nothing but bruises; finally they attained to the height of finesse: in the freezing cold, they locked her all night in the outhouse, because she wouldn't ask to get up and go in the middle of the night (as if a five-year-old child sleeping its sound angelic sleep could have learned to ask by that age)—for that they smeared her face with her excrement and made her eat the excrement, and it was her mother, her mother who made her! And this mother could sleep while her poor little child was moaning all night in that vile place! Can you understand that a small creature, who cannot even comprehend what is being done to her, in a vile place, in the Dense and the cold, beats herself on her strained little chest with her tiny fist and weeps with her anguished, gentle, meek tears for 'dear God' to protect her—can you understand such nonsense, my friend and my brother, my godly and humble novice, can you understand why this nonsense is needed and created? Without it, they say, man could not even have lived on earth, for he would not have known good and evil. Who wants to know this damned good and evil at such a price? The whole world of knowledge is not worth the tears of that little child to 'dear God.' I'm not talking about the suffering of grown-ups, they ate

¹⁵ David Bentley Hart, "Tsunami and Theodicy," a post on the blog, *First Things*, March 2005. An expanded (and diluted in power, I think) version of this summarizing passage occurs in Hart's book, *The Doors of the Sea: Where Was God in the Tsunami?* (Eerdmans Publishers, 2011).

the apple and to hell with them, let the devil take them all, but these little ones! I'm tormenting you, Alyoshka, you don't look yourself. I'll stop if you wish.¹⁶

There's much going on there, but I would just like to comment on Ivan's "to hell with them" dismissal of the suffering of adults, and say that one of the gravest of the many grave problems facing young children who find themselves in our horror-strewn world is attitudes like those of Ivan: that if they survive to that point, they will soon grow up, inevitably "eat the apple," and perhaps more importantly lose their cuteness, and then they will find that many of their fellows have lost their empathy toward what suffering they are subject to.

In addition to the examples already mentioned, I should also cite the at-least-philosophically-famous example of the presentation (that does not focus on children) that David Hume puts in mouth of his character, Philo (with some help from Demea), in Part 10 of the *Dialogues concerning Natural Religion*.¹⁷

My goal is not to add in any meaningful way to the depictions of the horrors of our world—which is why I instead here lean on the descriptions of others and point you to some noteworthy examples of that. What I am seeking to address is: Once we have looked in an effective way at the horrors of our world, where does the case go from there?

4. OK, So What's the Argument, Then?: The "Simple" Argument from Horrific Evils¹⁸

Our baffled theists typically don't deny that the horrors of our world present *some* kind of challenge to their worldview. But they see it as a challenge to the completeness of their view, not as pointing toward a potential error in it. They see our world's evils as pointing to a hole in

¹⁶ *The Brothers Karamazov* (tr. Richard Pevear and Larissa Volokhonsky; New York: Farrar, Straus and Giroux, 1990), Part II, Book Five, Chapter 4, "Rebellion," p. 242. I do love Alyosha's response, which immediately follows the quoted passage: "'Nevermind, I want to suffer, too,' Alyosha murmured."

¹⁷ First published in 1779, after Hume's death.

¹⁸ Though we do all this in significantly different ways, in seeking to construe the argument from horrific evils in a simple way, and then defending the power of the argument so construed, I take myself to be working along lines somewhat similar to those being explored by Trent Dougherty in his championing of what he calls the "The Common Sense Problem of Evil" (see esp. Dougherty's ms of that title, though some of it is prefigured in Dougherty's paper, "Epistemological Considerations Concerning Skeptical Theism," *Faith and Philosophy* 25 (2008): pp. 172-176.) One important difference between us (well, at least this is important to epistemologists like Dougherty and myself) is that I do not, like Dougherty, construe the key premise(s) of the argument as being immediately justified: See esp. [note 29](#), below.

our understanding of God's ways, not as pointing to God's non-existence: "Oh, it's all truly horrific, I agree, of course! And I don't mind admitting I have no idea why God allows it all. So I see how all this raises a very vexed question that I can't answer. And, yes, I can see how that can deeply trouble someone. It does me. But what I'm not seeing is how, beyond raising an unanswerable question of why God operates as God does, any of this produces any *good argument* against the existence of a wholly good God, or good *reason* for thinking that such a God *doesn't exist*. Just what is the argument supposed to be?"

"Umm...Well....Hmm." Just what *is* the argument? One might stumble around a bit in response.

In what has proven to be an important paper on our problem (that I will wrestle with later), in which he goes on to argue that the horrors of our world don't provide even a weak reason against theism, Stephen Wykstra first acknowledges that it can be hard to escape the sense, the "feeling," that they give us strong reason indeed for rejecting it:

Many of us—believers as well as nonbelievers, car mechanics as well as philosophers—have at some times in our lives felt instances of suffering in this world to be evidence against theism, according to which the universe is the creation of a wholly good Being who loves his creatures, and who lacks nothing in wisdom and power. If it has proven hard to turn this feeling into a good argument, it has, perhaps, proven just as hard to get rid of it.¹⁹

In the absence of a carefully formulated, "good argument" (which we might have to hire a highly-trained—and hopefully reasonably paid—philosopher to produce), are we left with just a *feeling* here? Perhaps one to be explained away as stemming from our natural frustration over not understanding an aspect of how God operates with great relevance to what may lie ahead for us and our fellow creatures? A feeling that's difficult to shake, but, at least insofar as it's directed against belief in the very existence of a wholly good God, just a feeling, all the same?

Our problem (as Wykstra might agree²⁰) is that the challenge to produce the argument is suckering us into seeking for the wrong kind of thing before being willing to call it a good argument. There is a powerful (perhaps all too powerful) argument against standard theism

¹⁹ Stephen Wykstra, "The Humean Obstacle to Evidential Arguments from Suffering: On Avoiding the Evils of 'Appearance'," *International Journal for the Philosophy of Religion* 116 (1984): pp. 73-93; p. 73.

²⁰ I *think* the way to read Wykstra here is as saying: Though it has proven difficult to turn this feeling into a good argument, Rowe has at least come close. Wykstra's strategy is then to critically examine this best way of trying to turn the feeling into a good argument, and show that, even then, the argument suffers from the killer problem that Wykstra raises. And Wykstra's often extremely perceptive treatment of Rowe at key points indicates that he sees the power in moves that Rowe makes that in various ways keep the argument simple. I'm thinking primarily of Wykstra's defense of Rowe's appeal to what "appears" to be the case, as opposed to mounting explicit arguments, in motivating his key premise.

here, but (as I at least think often tends to be the case for powerful philosophical arguments) it is simple and obvious. So simple and obvious that it is understandable why those presenting the problem of horrific suffering would not bother to formulate it, but just focus (as they tend to) on how horrific our world can be; and also so simple that it might fly under the radar screens of those looking for something more sophisticated.

We can start with this stunningly simple “formulation” of our argument by Peter van Inwagen:

We find vast amounts of truly horrendous evil in the world; if there were a God, we should not find vast amounts of horrendous evil in the world; there is, therefore, no God.²¹

To facilitate discussion, we can engage in some mild philosophy-write, and divide van Inwagen’s argument into numbered/lettered steps (adding ‘v’s on our labels, to indicate that these are steps from van Inwagen’s formulation, numbering our argument’s premises, and using “C” for its conclusion):

van Inwagen’s Argument from Horrendous Evil

1v. We find vast amounts of truly horrendous evil in the world

2v. If there were a God, we should not find vast amounts of horrendous evil in the world

So, C. There is no God

Now you have the argument in premise-and-conclusion form, suitable for a “Critical Reasoning” class!

But before presenting baffled theists with our formulation of the argument for discussion, I think we can tweak it to better capture how the intuitively compelling, but more informal, presentations of the problem operate. (Some readers may think van Inwagen’s argument itself should have been charitably understood as containing the upcoming tweak. But as this argument is our focus, it is worth making sure to capture it as best we can.) Van Inwagen’s formulation has the argument proceeding just from the fact that there are “vast” amounts of “truly horrendous” evils in the world. Though those terms are in a way quite vague (though harmlessly so, in this context, I think), we can still see this limitation in the formulation of the problem: Once an evil passes whatever the threshold is for this, it counts as “truly horrendous,” and once enough such evils pass whatever the threshold is for the world to count as containing “vast amounts” of such evils, 1v is satisfied. If even more truly horrendous evils occur, or if evils not only pass the threshold for counting as “truly horrendous,” but *far* exceed

²¹ Peter van Inwagen, *The Problem of Evil*, p. 56. I put “formulation” in quotation marks, not because van Inwagen says he’s “formulating” anything here, but to try to capture the sense of the challenger to the existence of any good argument here, who might declare themselves to be waiting to see something properly “formulated.”

it, that does not affect this problem for theism, at least so far as van Inwagen's formulation of it goes.

We can fix this, while retaining much of van Inwagen's wording, by means of what we can call the following "Simple" argument from horrendous evils—and we will tack an 's' on to our label for the argument's premise, to indicate it's the premise from our Simple formulation. Like van Inwagen, we take it that to count as "God," a being must be perfectly good, as well as omnipotent and omniscient. And like van Inwagen, our argument is driven by the vast amounts of truly horrendous evils in the world. But our reformulation makes facts about just *how* truly horrendous these evils are, and just *how* vast is the amount of such evils, relevant to the strength of the problem that the evil of the world poses for theism. And I think it's best to actually pare the argument down to one that has just a single premise²²:

²² Here I break from van Inwagen, and with Rowe (whose formulation we're about to encounter), and I suppose with standard practice, in not employing two premises, one stating the facts about horrible evils that are being appealed to, and then another stating that those facts would not obtain if there were a (wholly good) God. (I suppose I thus strictly depart from the form of argument that in "The Problem of Evil" Draper labels as "MT arguments" (for "modus tollens") in order to contrast with the "Bayesian Arguments" that he prefers (see Draper, "The Problem of Evil," in Thomas P. Flint and Michael C. Rea, ed., *The Oxford Handbook of Philosophical Theology* (Oxford UP, 2009)—though I think that the argument as I formulate it should still be thought of as being of the same basic type as these two-premise formulations, as similar issues will arise with respect to it. And I suspect Draper would think his reasons for preferring the Bayesian over the MT arguments would also be reasons for preferring the Bayesian argument over my formulation.) We could have followed that tradition, I suppose, by using some term, like italicized occurrences of *that*, to point, as it were, at degrees of horrendousness and of vastness of our world's evils, and arguing *something* like the following (as you consider different ways I might have taken this, you might start to get a sense for what's lost by taking any of the particular possibilities):

1s-alt. There is *that* vast amount of truly horrendous evils in our world, and many of them are *that* horrendous

2s-alt. If there were a God, there would not be *that* vast amount of truly horrendous evils in our world, with many of them getting *that* horrendous

So, C. There is no God.

But I don't think our intuitive problem is best represented as following from two judgments, one about the nature of the world, and one about whether it would be like that if God existed. I think the problem is best represented as issuing from a single (though I suppose, a bit more complex) *comparative* judgment, concerning a glaring apparent mismatch between how the world is, and how it would be if there were a God. This I think is best captured by a single premise, like our 1s.

Our formulation also makes room for the relevance of features of horrendous evils beyond how horrendous they are. Perhaps some truly horrendous evils, though not being more horrendous than other truly horrendous evils (they're not *particularly* horrendous, so far as truly horrendous evils go), do nevertheless have some other special feature which make them particularly hard to square with God's existence.

The Simple Argument from Horrendous Evils

1s. If there were a God, we should not find the evils that we find in the world, especially including the truly horrendous evils we find, and in such vast amounts as we find them.

So, C. There is no God.

“There’s your argument!”²³

“Ah, well, OK, but what’s the argument for that premise, 1s, then?”

“It’s a premise! There’s no argument for it. Then it would be a conclusion.”

5. OK, But What’s the Case for That Premise, Then?

Alright, that last response is too brusque—though there is a good point behind it. Forceful prosecutions of the problem of horrific evils typically spend their energy on confronting us with

It will perhaps help for me to acknowledge that the second half of 1s (starting with the word “especially”) is in a way idle: The first half of 1s already involves the “the evils that we find in the world,” including those that the “especially” clause highlights. We keep the second half, despite that idleness, because it helpfully calls attention to an important aspect of what’s in the first half—the aspect that provides much of our problem’s kick, at least according to those most moved by horrendous evils. Note, then, that our formulation has the special problem of horrendous evils taken up into the more general problem of squaring the evils of our world with the existence of God, and that other forms of the problem of evil could then easily be substituted into this scheme, by replacing the “especially” clause of our premise.

²³ For present purposes, I imagine the argument being put forward aggressively. It would be in keeping with my own manner, at least in the moods I mostly inhabit, and also with my estimation of the power of even intuitively powerful, old arguments, to instead follow the presentation of such an argument with humbler words like: “That’s it, really, wrapped up in old newspaper and string.” I’m borrowing these words from my late teacher, Rogers Albritton, who was writing about another familiar, simple, and powerful old argument. The opening of the relevant talk (posthumously published as a paper):

I have been intermittently obsessed for years with a certain form of skeptical argument “from possibility,” as I will say. The idea of it is ancient and familiar. It’s that “anything’s possible,” as we say, so “you never know,” as we also say. Anything’s possible, so you never know. More expansively: you can always or practically always be wrong; but if you know, you can’t be wrong; so, you never or practically never know. That’s it, really, wrapped up in old newspaper and string. (“On a Form of Skeptical Argument from Possibility,” *Philosophical Issues* 21 (2011): 1-24; p. 1.)

I don’t know if I picked up my habit of spending my philosophical attention obsessing over simple, old, familiar, and intuitively powerful arguments from Albritton, but if I didn’t, our shared habit of doing so (that he already possessed, and I was disposed toward) is then no doubt a large part of why I ended up working with him.

the horrors of this world, and what I'm proposing is that, insofar as they are challenging the existence of the wholly good God of standard theism, that energy is being expended in support of our 1s (or some claim much like 1s). The simplicity of our argument, together with how the matter of just how horrendous our world can get seems relevant to our simple premise, explain why these forceful presentations proceed as they do. But then these presentations do seek to somehow support, or make some kind of case for, that premise, even if they offer no proper argument for it in the form of deeper premises from which the target premise is alleged to follow. In fact, that's often about all that they do. So let's scratch that quick and blunt response, and try a different tack:

"Haven't you been listening? I've just spent the last several [pages/minutes] making the case for 1s!"

"OK, then, but what's your *argument* for 1s? What are your premises for it? Do you mean to be putting forward all these (true, I admit) claims about the horrors of our world—that this happened, that that happened, that this or that kind of thing happens a lot, etc., and that it's all so awful—as premises from which you conclude 1s? Because I don't see how 1s is supposed to *follow* from all that. Beyond such claims, you ask a lot of questions like 'Wouldn't a wholly good God have prevented *that*?'. But I can't see how to convert those questions into implied assertions from which, together with your other claims, 1s might plausibly follow. I mean, I have some ideas about how one might start to try to make a proper argument out of all this, but my attempts on your behalf aren't turning out very well. I'm having a very hard time outlining your argument!"²⁴

²⁴ Here I'm mainly seeking to characterize a type of stance one will hear often enough from a certain type of Christian "apologist"—a type that, among other things, bears the marks of having some philosophical training (perhaps a "Critical Reasoning" course, or various skills picked up in philosophy classes). This is part of a wider phenomenon one can observe on social media (among other places), if, I guess, one has the "right" contacts there: argumentative battlers for various causes, set on "winning" whatever argumentative game they take themselves to be playing, using jargon and tools learned in the study of philosophy, whether in defense of religious (or anti-religious), or political/ideological, positions, or other stances. These are the folks who will label and count the "fallacies" that occur in their opponents' arguments. And, more to our current concerns, these are the folks who ask what their opponents' arguments and premises might possibly be, and who then try to "helpfully" clarify their opponents' arguments by outlining them in premise-and-conclusion form, before attacking them. They find the claim that functions as their opponents' argumentative starting point that they think is the most vulnerable, and then know to ask something like this about it: "But why in the world should we accept *that*?" When they are the ones making or defending an argument, they have labeled their most vulnerable starting claim as among their "premises," and when it comes under scrutiny, know to respond in something like this way (which puts presumption on their side): "Well, like all arguments, mine begins from premises, which others might reject. But my opponent has failed to refute my premise; so that premise stands." (Yes, I realize that I sound here like a grumpy old man, complaining

6. But What's the Case for That Premise, Then?: Rowe's Appeal and the "Good Cop, Bad Cop" Routine

No, these claims about the evils of our world are best not construed as premises from which 1s is supposed to follow. We instead follow the example of Rowe, whom I already mentioned (remember: he's the fawn guy), in making the case more effectively.

Though our formulation of the argument diverges from his, we need to look briefly at how Rowe formulates his argument, so we can see and convert the insightful moves he makes in defense of it. Rowe uses three numbered steps: two premises leading to his atheistic conclusion (we add an 'r' to the steps' labels, to distinguish them from the steps of other formulations of the argument):

Rowe's Argument from Intense Suffering

1r. There exist instances of intense suffering which an omnipotent, omniscient being could have prevented without thereby losing some greater good or permitting some evil equally bad or worse.

2r. An omniscient, wholly good being would prevent the occurrence of any intense suffering it could, unless it could not do so without thereby losing some greater good or permitting some evil equally bad or worse,

So, Cr. There does not exist an omnipotent, omniscient, wholly good being.²⁵

about the "kids these days." But sometimes grumpy old men painting with broad brushes have helpful, if caustic, observations to make. Or so I like to think.)

For the most part, I don't suspect instructors of philosophy classes of teaching their students to carry on in this way. My suspicion (and hope) is rather that these warriors have learned certain things (jargon and tools) from their philosophy training, and that, once they get out onto the "battlefield", they quickly figure out how to use these things to so battle their "enemies"—and/or learn this from argumentative warriors who went before them.

But while I am characterizing a stance one will mainly encounter in non-academic settings, one can find it in professional philosophy, too. For me, the prime example of this (at least of trying to figure out an opponent's argument might be, in decidedly unhelpful ways) is Alvin Plantinga's response to arguments from horrific evil (that I have already mentioned, and will return to soon) in Plantinga's paper, "The Probabilistic Argument from Evil," *Philosophical Studies* 35 (1979) pp. 1-53, and particularly the opening few pages of that paper. However, I should draw attention to the age of that paper (it's from 1979), and report that Plantinga has displayed a refreshingly better attitude toward at least the "case," as he puts it, from horrific suffering in more recent work, as we will see in [Section 17](#)—though, as we'll also see, he still thinks there is no good "argument" from horrific evils.

²⁵ William Rowe, "The Problem of Evil and Some Varieties of Atheism," *American Philosophical Quarterly* 16 (1979): 335-341; p. 336. Rowe's formulation points to "intense," rather than "horrific" or "horrendous" suffering, but when he starts pointing to instances to back up his claims, it becomes clear that the instances of intense suffering he has in mind can be well characterized by our h-words.

1r is Rowe's key premise; it's in defense of it that Rowe tells us about his fawn, and more generally cites the evils of our world. But Rowe does not call the claims he makes in defense of 1r "premises" from which 1r is supposed to follow—and, more importantly, as I'll explain, he doesn't *treat* them as premises, either. He instead issues these claims, along with some related reasoning, and then *appeals* to us to evaluate his premise, 1r, in light of them. How does that work, and how (one might suspiciously ask) does it differ from advancing premises of a proper argument from which 1r is supposed to follow?

Let's look. Here again is Rowe's fawn, but also now what follows it:

Suppose in some distant forest lightning strikes a dead tree, resulting in a forest fire. In the fire a fawn is trapped, horribly burned, and lies in terrible agony for several days before death relieves its suffering. So far as we can see, the fawn's intense suffering is pointless. For there does not appear to be any greater good such that the prevention of the fawn's suffering would require either the loss of that good or the occurrence of an evil equally bad or worse. Nor does there seem to be any equally bad or worse evil so connected to the fawn's suffering that it would have had to occur had the fawn's suffering been prevented. Could an omnipotent, omniscient being have prevented the fawn's apparently pointless suffering? The answer is obvious, as even the theist will insist. An omnipotent, omniscient being could have easily prevented the fawn from being horribly burned, or, given the burning, could have spared the fawn the intense suffering by quickly ending its life, rather than allowing the fawn to lie in terrible agony for several days. Since the fawn's intense suffering was preventable and, so far as we can see, pointless, doesn't it appear that premise [1r] of the argument is true, that there do exist instances of intense suffering which an omnipotent, omniscient being could have prevented without thereby losing some greater good or permitting some evil equally bad or worse[?] (Rowe, p. 337)

Note the nature of the appeal Rowe makes on behalf of 1r. After citing his example, he discusses it a bit, and even does some reasoning (as I think it's fair to call it) concerning evil and its relation to 1r. But instead of claiming that 1r follows from some deeper claims he is using as "premises" in his reasoning, in the last sentence of the above, he instead asks us: In light of his example, and also his discussion of it, including claims that he puts forward in support of his premise, "doesn't it appear that" his premise is right?

Expanding his case, his reasoning, for 1r, Rowe then goes on to broaden out the considerations he's raising, and asks us to consider 1r in light of "*all* the instances of seemingly pointless human and animal suffering that occur daily in our world,"²⁶ before returning to his

²⁶ Rowe, p. 337, emphasis in the original.

evaluation of 1r, which no longer employs just a mild questioning appeal,²⁷ but has become something of a “good cop, bad cop” routine, with some much more assertive (note especially the “extraordinarily absurd” part), “bad cop” parts mixed in. Rowe asks whether it is reasonable to resist 1r in light of all that horrible suffering he has just mentioned, and answers:

And surely the answer to this more general question must be no. It seems quite unlikely that *all* the instances of intense suffering occurring daily in our world are intimately related to the occurrence of greater goods or the prevention of evils at least as bad; and even more unlikely, should they somehow all be so related, tha[t] an omnipotent, omniscient being could not have achieved at least some of those goods (or prevented some of those evils) without permitting the instances of intense suffering that are supposedly related to them. In the light of our experience and knowledge of the variety and scale of human and animal suffering in our world, the idea that none of this suffering could have been prevented by an omnipotent being without thereby losing a greater good or permitting an evil at least as bad seems an extraordinarily absurd idea, quite beyond our belief. It seems then that although we cannot *prove* that [1r] is true, it is, nevertheless, altogether *reasonable* to believe that [1r] is true, that [1r] is a *rational* belief.²⁸ (pp. 337-338)

Mixed in here with his more assertive language, Rowe is still appealing to his readers to note an appearance—here phrased in repeated (well, twice, here) talk about what “seems” to be.

7. “Making the Case” for the Premise of the Simple Argument from Horrific Evils

I propose we understand presentations of the problem of horrendous evils as similarly appealing to the audience to consider 1s, or something much like 1s, in light of the evils of our world that our skeptics are pointing to. They confront us with evils of this world and then ask us, sometimes with some associated reasoning, and often urgently: In light of the facts and reasoning just displayed, doesn’t it seem that if there were a wholly good God, we should not

²⁷ Note that while Rowe does say that *something* is “obvious” in the quotation above, it is not his premise, but rather a claim he makes in the course of defending that premise, namely, that an omnipotent, omniscient being could have prevented the fawn’s suffering, by which he seems to mean only that such a being would have the power to have done so, however bad an idea it might or might not be to exercise that power.

²⁸ Rowe, pp. 337-338. Toward the bottom of this passage, how Rowe’s article actually reads is that the denial of his premise in “an extraordinary absurd idea”. In part because there is no comma between the two adjectives there, my best guess is that this is some kind of misprint, and the phrase should be: “an extraordinarily absurd idea,” as I have rendered it.

find the evils that we find in the world, especially including the truly horrendous evils we find, and in such vast amounts as we find them? In other words: Doesn't it appear to be true?

The distinction I'm seeking to make between more generally "making a case" for a claim and more specifically offering an argument for a conclusion in the form of premises from which that conclusion is supposed to follow may seem subtle—perhaps suspiciously so. But it is important (though the particular terms I have decided to employ are not), and we can make good sense of it by starting with the more general type of case, and asking when such a case becomes a proper argument.

Imagine someone just "making a case" for there not being any (wholly good) God, having not yet presented any "proper argument" for that conclusion, in the form of something that can be well captured in a premise-(and-perhaps-sub-conclusions)-and-conclusion outline. In the course of making their atheistic case, they make a lot of claims, some, but perhaps not all, of them about the evils of our world, that they think militate against the existence of God, hoping that their audience will both agree with these claims and agree that they count against God's existence. They may also explain a bit why they think those facts militate against the existence of God.

Well, if they are making claims in support of God's non-existence, and especially if they are also explaining how they think those claims support their contention that God does not exist, are they not arguing against God's existence? Indeed they are, on a (and perhaps the) perfectly good and ordinary use of "argue." And we can well call the contention that God does not exist the "conclusion" of this "reasoning." But they are not yet giving what we are narrowly calling a "proper argument," in the form of premises from which their conclusion is supposed to follow. Instead, they can be construed as making their supporting claims, and engaging in some reasoning concerning them, and then, appealing to us to consider God's existence in light of their claims and the reasoning they have offered, and asking us, in light of all that: Doesn't it seem that, doesn't it appear that, their conclusion is right?

Such a case could be quite effective. Perhaps the considerations advanced really do create, and quite rationally so, at least for many who consider the matter, a strong appearance that God does not exist. And if it does create such a strong appearance for someone, and they thereby come to believe that God doesn't exist, this belief will likely be at least partially *based on* some of the claims made in the case that moved them.²⁹ But just *how* that conclusion is

²⁹ This point is pivotal to the point of difference between Dougherty and myself over whether such a belief in the conclusion of such a "case" is immediately justified, i.e., justified in a way that does not involve its being based on other beliefs. I think such beliefs are best construed as being based on others (usually including some of those cited in the making of the "case") via an inference that, if it is made at all, is not made consciously, and often cannot be explicitly stated by the believer themselves. Dougherty,

based on those claims may not be fully articulated—and importantly, our arguer may not know how to make it fully articulated. Crucially, the connection may be loose, hard to specify, and may essentially involve other, unidentified beliefs that are lurking in the background, unnoticed. That’s why our contender didn’t claim their conclusion follows just from any of the claims they explicitly made in support of it: because, likely, it doesn’t! (And if you “help them out” by “clarifying” their argument, and present it as if their conclusion is supposed to follow from some of those claims, you are “setting up a straw man,” or the gender-neutral form of that, or however it’s put in “Critical Reasoning” classes these days.) Still, it is utterly common for us to be able to reasonably judge that some considerations, perhaps with the aid of background beliefs that we may well share with others but can’t yet identify as playing a key role, support some conclusion—perhaps strongly, perhaps sometimes even with overwhelming strength—before we can say just how they do.³⁰ (And if we *can’t* do this, then we might as well give up on all arguments for controversial and philosophically substantive conclusions. For all such arguments will rest on at least one premise that is controversial and quite deniable, and for which no proper argument is offered, and so for which the kind of “case” that we are now describing is the most that’s on offer.)

Perhaps for other listeners, the case fails, either because they don’t see the considerations put forward as really militating so strongly, or perhaps at all, against the existence of God, or perhaps because they think other considerations militate even more strongly in favor of God’s existence. In the latter case, if these pro-God considerations are put forward by our skeptic’s opponent, we might soon arrive at an interesting discussion, which in

by contrast, sees such a belief as not being based on other beliefs at all. This seems to be due to Dougherty’s deep general suspicion of “unconscious inferences.” And I might be happy enough to cede the matter of whether there’s an unconscious “inference” here to Dougherty. For me, the crucial question concerns when one belief is *evidentially based on* others. This crucial question is also often exceedingly tricky, especially where there is no process of inference that the believer is aware of. One thing I think it makes sense to look for before ruling in such a case that one belief is based on others is whether, in holding that first belief, the believer is sensitive to the support provided for it by those other beliefs, where such sensitivity consists in such facts as that the believer wouldn’t hold that first belief, or at least wouldn’t hold it to the degree that they do, were it not for the evidential support provided for it by those others. One could be in that way sensitive to the support relations that hold among one’s beliefs even where one gives no conscious thought to those relations, and yet it seems to me that epistemic justification might well be transmitted from some beliefs to others in virtue of such sensitivity to relations of support. For an idea of some of the considerations that I think are involved, see my discussion of whether “I am not a brain-in-a-vat” is typically based on our ordinary perceptual beliefs, or is instead immediately justified/known, in Chapter 7, esp. sects. 18-20, and super-esp. sect. 19, of (DeRose 2017).

³⁰ As Wykstra nicely puts this point, in defense of Rowe, one’s conclusion in such a case “rests upon some tacit inference; but the evidential import of some things is more readily ‘seen’ than ‘shown;’ and the ‘appears’ idiom concerning us is characteristically used for just such things” (Wykstra, p. 80).

perfectly good English would be described as a two-way “argument,” and perhaps a heated one at that, with two sides advocating for opposing claims, each advancing many considerations that they think push toward their claim, appealing to each other, and perhaps to a wider audience, in light of the considerations they have pointed out, “Don’t I seem right?” But since both sides are just asking us to consider their desired conclusion in light of the considerations they have raised, and are not claiming that their conclusion follows from some of their claims, which function as their “premises,” we still have no proper argument, in the restricted sense we are using, and are just about to explain, on either side of this discussion.

But, crucial to our current concern, sometimes, a more focused line of reasoning will emerge from the wider tangle of such a case, and an arguer will see how some of the considerations they are urging lend particularly powerful support to a small group of claims from which their desired conclusion does logically follow. It may be wise to then focus on these particular claims and the support they lend to the desired conclusion by taking them to be premises of a proper argument to that conclusion. Though this argument can still be seen as part of a larger overall cumulative case for its conclusion, insofar as we focus on the argument in question, our arguer is no longer appealing to us to agree that their desired conclusion appears to be true in light of all the considerations they are bringing forward, but is placing special focus on a set of claims—their argument’s “premises”—from which they claim that their desired conclusion follows, and is instead, for now, appealing to us whether *those premises* seem to be right. And then they might make a “case” for those premises, with that case not being a proper argument for them.

Why do that? Well, they may think that this set of claims from which their desired conclusion does follow is such that each of the claims gives, or can be made to give, a particularly strong appearance of being true. If so, the special focus would be wise.

That’s how we should view our Simple argument from horrific evils—though it’s a special case where the set of claims which function as premises consists of just a single claim. From among the tangle of considerations which can seem to militate against the existence of God, what seems to be a particularly powerful line of thought, at least to the backers of the argument, runs through 1s (or something like it), from which the atheistic conclusion follows, and which, according to fans of the argument, in light of various facts about the evils of our world, does give a very powerful (perhaps even overwhelmingly powerful) appearance of being true.

The case for 1s can be given in a variety of ways—some of which we have been looking at, and will continue to consider. (In [sect. 17](#), with some expert help, we’ll see the case made in a compelling way.) But to the question: “Well then, what’s your *argument* for 1s, what are your premises from which 1s is supposed to follow?,” our skeptics’ answer is: We’re not giving that

kind of “proper argument” for 1s. We are making the case for 1s primarily by means of a heavily guided tour of some of our world’s horrors, and then appealing to you: Doesn’t 1s seem to be, appear to be, true? Doesn’t it seem that if there were a (perfectly good) God, we should not find the evils that we find in the world, especially including the truly horrendous evils we find, and in such vast amounts as we find them?³¹

8. A Very Powerful Appearance, A Very Powerful Argument

Appearances can seem (!) feeble things: “mere appearances,” as we might say. We usually say something “seems” or “appears” to be so when we don’t feel in a position to flat-out declare that it is so. And perhaps the most proverbial thing we know about appearances is that they can be deceptive. How solid, and how compelling, can our premise, and then our argument, be, then, if it is based on an appearance?

³¹ It is vital that the appearance claim is *not* construed as the kind of step in the argument that should make it on to a premise-and-conclusion outline of the argument. (That seems no more appropriate here than for any other philosophical argument. One might as well take every current starting premise, P, in every outlined argument, and reconstrue it as a sub-conclusion, supported by the “true premise” “It seems that/appears that P.” This would not only be inelegant, but would misrepresent the function of those appearance claims.) The outline of the argument, as I think it’s best done, is given above, at the end of Sect. 4—and the numbered steps Rowe actually provided for his argument are given toward the start of Sect. 6. In neither case does any appearance claim make the outline—and rightly so. In their “Evil and Evidence” (in *Oxford Studies in Philosophy of Religion*, Vol. 7 (2016): 1-31), citing the papers of Rowe and Wykstra we are discussing, Matthew Benton, John Hawthorne, and Yoaav Isaacs write that “Early work on the evidential problem centered around epistemic seemings or appearances (Rowe 1979, Wykstra 1984; more recently Matheson 2014 and Tucker 2014). The evidence brought to bear against the existence of God was not evils, but rather what we think about how the evils appear to us. This is, to our minds, a mistake.” And this indeed *would be* a mistake, and would rightly subject the argument to mocking that immediately follows: “We are moved by someone crying out, ‘How could a good God allow such suffering as mine?’ We are less moved by someone crying out, ‘How could a good God allow such epistemic states as mine?’” (p. 9). But I think the real mistake here is the mischaracterization of the skeptic’s case. The appearances cited aren’t being put forward as *evidence* for the premises they are applied to (and then for the atheistic conclusion reached). The “evidence brought to bear” for the premises in question and then against the existence of God are the claims about the evils of our world that are cited in support of those key starting premises (premise 1s, in the case of the argument we’re most focused on). The appearance claim functions as an appeal to the reader/listener to accept the starting premise in light of the claims and reasoning (short of a proper argument) made on its behalf. It’s the kind of appeal that any substantial argument must in effect make on behalf of its substantial starting premises, insofar as it seeks to support them at all. This doesn’t mean that evidence cited by all substantial philosophical arguments concerns the arguers’ epistemic states.

Some of the apparent (!) general shakiness of appearances is due to a misleading trick of language: As I suggested above, when appearances are strong and reliable, we're usually in a position to say something stronger than just that things appear to be the way we take them to be. For instance, we'll often then be in a position to say that things simply are that way—and sometimes even that we know them to be that way. And when we're in a position to say something stronger like that, we typically *will* say, and generally should say, that stronger thing, rather than making the needlessly weak “It appears/seems that...” claim. So, though appearances can be strong and steady, our use of “appears”/“seems” *talk* is generally called for when shaky appearances are in play, and this can make us think of appearances as being generally shaky things, when in fact they run the gamut.³² Some appearances certainly can be deceptive, as well as shaky. That's quite consistent with some others of them, as well as some of what's based upon them, being very solid.

Indeed, some of us, myself included, think that *all* we know about the world, including the things we know most solidly, are ultimately based on appearances. According to this “Phenomenal Conservatism,” as it's known in philosophy, in developing our view of the world and our place in it, we all start with some appearances, play them off of each other, getting rid of some as deceptive, but, thankfully, find many of them to fit together well to form a good, coherent picture of what the world is like. Because much of the resulting picture fits together so well, and incorporates incoming experience so well, we end up knowing many things, some of them being things we know very solidly indeed, and being very justified in many of our beliefs—despite it all being ultimately based on appearances.³³

³² I here adapt H.P. Grice's explanation, in “The Causal Theory of Perception,” *Proceedings of the Aristotelian Society*, Supplementary Volume 35 (1961), pp. 121-152, for why the likes of “It looks red to me” generates an “implication” (Grice later came to use the label “implicature” here) to the effect that there is some “doubt or controversy” about whether the object is red, though the sentence doesn't actually say that there is such doubt or controversy.

³³ Michael Huemer has been the most important champion of Phenomenal Conservatism, perhaps most influentially in his paper “Compassionate Phenomenal Conservatism,” *Philosophy and Phenomenological Research* 74 (2007): pp. 30-55, but also then in other work. I'm going beyond PC as Huemer tends to formulate it—though in ways I've always suspected that (and have now confirmed that, via p.c.) Huemer would agree with. Huemer construes Phenomenal Conservatism as the position that (to take his gloss in the just-cited paper) “appearances of all kinds generate at least some justification for belief” (p. 30), which falls short of saying that appearances are where all our justification ultimately comes from (and indeed, falls short of even saying that appearances generate very much justification). But I think one could always find in Huemer's work pushes toward accepting a more thorough-going “Phenomenal Conservatism” than what he himself officially endorses. This occurs mainly where Huemer gives his central argument against non-conservatives: that their position is self-defeating in that it cannot be both true and justified.

Admittedly, the question of what we so “start with,” and what our knowledge of the world is “ultimately based on,” is an extremely tricky one, and this “Phenomenal Conservatism” provides only a controversial answer to it. Thankfully, though (since I don’t think anything like 1s could be one of our ultimate “starting points,” anyway), good arguments to substantial and interesting philosophical conclusions don’t have to—and I think never do—reach so far back into the cognitive mists as to begin exclusively from the beliefs or belief-like states from which our knowledge of the world ultimately begins, but instead make use of appearances that emerge some distance down the cognitive road³⁴ (though the early steps down the cognitive road we took to get to the appearances we now appeal to can be notoriously hard to reliably reconstruct).

So, based on experience in philosophy, including a good deal of studying of its history (though my knowledge of its history is admittedly spotty), and not on any controversial grand claim about where all our justification for our beliefs ultimately comes from, I feel pretty comfortable saying: Where its conclusion is indeed substantial and interesting, any philosophical argument will have at least one premise—let’s say its most controversial and shaky starting premise; we can call it the argument’s “key premise”—that is itself interesting and substantial.³⁵ And potentially deniable.³⁶ And what can be said in defense of such a

³⁴ I should perhaps here cancel any suggestion that if we could dig down to where our knowledge of the world begins, we would find at the foundation of our picture of the world something unusually and unshakably solid, because I for one suspect the appearances we begin with are mostly quite uncertain, shaky ones, that do give rise to all our knowledge, including some very certain knowledge, only after a lot of playing these appearances off each other, and so quite a ways down the cognitive road from its fabled beginnings.

³⁵ I’m here assuming that arguments’ premises imply their conclusions. If an argument is not like that, and some of its controversial aspects instead concern whether its conclusions really follow from the premises listed in support of them, rather than just whether the premises are true, it can be converted to the needed form by adding conditional bridge premises (“If [premise[s]], then [conclusion]”), which will then bear the weight of the controversy. Otherwise, one can construe both the premises and the inferences as “steps” in an argument, and then say that where its conclusion is substantial and interesting, any philosophical argument will have at least one “key step” that is itself interesting and substantial and deniable, where an inference is “denied” where one claims that its conclusion doesn’t follow from its premises.

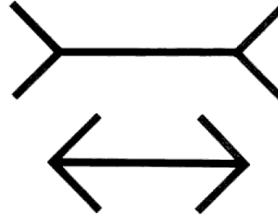
³⁶ “So much the worse for philosophical arguments!”—I can hear the response. And I do agree that philosophical arguments generally aren’t strong enough to produce anything close to knowledge of their conclusions. In fact, though I don’t think this could provide an exception-free definition, I think there is something right about a characterization of philosophers as specialists in addressing some of the questions we find important, but which nobody has yet figured out a knowledge-producing way to get answers to, generating answers to such questions, and good (even if not knowledge-producing) support for those answers—sometimes in the form of “proper arguments” for these conclusions, and sometimes by means of other “cases” for them. None of this is to put philosophy down. Philosophy is wonderful—and in large part precisely because it deals with those important but hard questions. (And I think there’s something to the idea that once we get to the point that philosophy is producing actual knowledge

premise? Well, one can offer a proper argument for it, but then it becomes a conclusion (or a sub-conclusion: a claim that one provides an argument for, but that then one uses to argue for further conclusions), with premises supposedly leading to it, and not an initial premise of one's argument, and what I'm claiming here is that at least one of those *initial* premises of any argument for a substantial conclusion will itself be substantial and deniable. And, short of offering a proper argument for that premise, which we can now suppose our arguer is not in any position to do effectively, what can our arguer do in defense of their shakiest *initial* premise, other than to present it in its best light, perhaps asking their audience to evaluate it in light of certain considerations, and appeal to their audience, to us, in light of the considerations adduced: Doesn't that *seem* right? Doesn't it *appear* to be true? When the pusher of our Simple argument makes their case for their premise, but then ends with a "mere" appeal to appearances or their key premise seeming true, they are doing what all arguers must at least in effect do (though they may use different wording when doing so) when they reach their initial premise(s)—well, insofar as they seek to support those premises at all.³⁷

Since such seemings or appearances are so important to our argument, and indeed to any substantial philosophical argument, it's worth pausing to remark on how the appearances that underwrite good arguments compare with some other appearances. Sometimes we say that something "seems" or "appears" to be the case, even though we are not at all tempted to believe that things really are as they appear to be. For instance, even after being effectively informed (perhaps before even encountering it) that it really is an illusion, and that the two horizontal lines included here in it really are the same length, many will report that the horizontal line on top here "appears to be longer" (or "seems to be longer" or "looks longer") than the other, lower one in this display of the Müller-Lyer illusion:

about a topic, then the area of philosophy that deals with the area is likely to break off and no longer be thought of as philosophy.) Philosophy is sometimes attacked as a waste of time for dealing with such questions, but quite inconveniently for them, those making such an attack appear to be engaging in philosophy in the very making of the attack—and often not very well! (I suppose it's no great surprise when those who don't value an activity can tend not to be so good at it.)

³⁷ Some back-up on this from a leading philosopher: After an evaluation of an argument of his own (in fact, one that will be important to us later), Peter van Inwagen writes: "And this, *mutatis mutandis*, is all that can be asked of any philosophical argument. At any rate, no more can be said for any known philosophical argument than this: it is valid and its premises seem to be true" (van Inwagen, "Free Will Remains a Mystery," *Philosophical Perspectives* 14 (2000): 1-19; p. 10).



Though they are fully convinced that the top line is not longer, I think that when they say “Wow, the top line seems to be longer!”, such a person is still reporting some push toward believing that the top line is longer that they can feel within their soul. This push or inclination to believe may never have had any chance of resulting in a belief that the top line is longer, given their unwavering trust that the lines are in fact the same length, but the push is still there, and can still be felt, and is I think what we are reporting with the relevant “seems” or “appears” claims.³⁸

³⁸ I believe that “seems” and “appears” claims generally report some push toward believing the proposition that one is saying “seems” or “appears” to be true. Here, I go against Huemer, who in his main work, “Compassionate Phenomenal Conservatism” (see [note 33](#)), argues as follows:

Nor should appearances be identified with dispositions or inclinations to form beliefs. One reason is that one might be so convinced that an appearance was illusory that one was not even inclined to believe its content. One could even be convinced in advance that one was going to experience an illusory appearance, so that there would be no time at which one had the relevant inclination to believe. (p. 31)

But as I pointed out in the main text (and I had Huemer’s argument in mind when I did so), there can be a push toward (and the same point would hold for an inclination toward) belief that is preemptively and thoroughly checked by some other force.

In later work (“Phenomenal Conservatism Über Alles,” in Chris Tucker, ed., *Seemings and Justification: New Essays on Dogmatism and Phenomenal Conservatism*, Oxford University Press, 2013); pp. 328-50), Huemer argues as follows:

A natural approach to analyzing seeming is to appeal to dispositions to believe. One might hold that its seeming to one that P is simply a matter of one’s being in a state such that one would believe that P, were there no other factors interfering with one’s forming such a belief.

It is worth briefly reminding ourselves of one of the main reasons for rejecting that approach. This is the fact that one can be disposed to believe P for different sorts of reasons, other than its seeming to one that P. If I am disposed to believe in the afterlife because I *want* there to be an afterlife, this is quite different from my being disposed to believe in the afterlife because that seems true. The lesson is that appearances are only one sort of ground for the disposition to believe. (p. 329)

But I think that to someone inclined to believe something because they want it to be so, the thing does seem to them to be the case. Of course, there are different levels of inclinations to believe things, and not everything that can be called a “disposition” to belief in someone forces us to say that the thing “seems” or “appears” to them to be so. My desire for there to be an afterlife may not yet have generated any felt push in me toward believing that there is an afterlife, though it may have made me particularly open to the thought, should it be urged by those around me, and in that way, may have in

By contrast, the seemings that we hope underwrite the premises of our philosophical arguments are those reported by what I call “all-in” appearance or seeming claims. Here, the claimant is not just reporting a push toward belief at some initial stage of cognitive processing that may have been checked by some other processing occurring elsewhere in their soul, but is reporting what they are inclined to believe at their final (so far) stage of inquiry on the issue, taking into account all the relevant considerations they have access to.

Well, our philosophical arguer is likely not just reporting their own all-in inclination to believe, but is in effect also appealing to their audience, “Hey, doesn’t it seem that way to you, too?” Our arguer has likely just raised various considerations which (they at least think) push toward the judgment that things are as they’re claiming they appear, and they are asking us to consider the matter with those considerations especially in mind, but they are inviting us, appealing to us, to agree, all-in, that, yes, it sure seems as if that’s the case.

And often, of course, to many of us, the arguer’s key premise does *not* seem true: When we bring all the relevant considerations to bear, including those the arguer has just stressed, we are not all-in inclined to think that the key premise is right. Indeed, I suppose this is the usual case—and that the unusual cases are quite unusual indeed. But even when unconvinced, we can often recognize the power of the argument: “Well, that key premise still doesn’t seem to me to be true, but I can certainly see how someone else might reasonably find it plausible—and perhaps even compelling.” At that point, you might be able to explain a bit why things seem otherwise to you, and the conversation may advance. Or maybe you can’t. These things go in different ways.

It’s rare, but the appearances behind the key premises of some substantial philosophical arguments can be (or can be made to be) extremely strong: Sometimes it *really* seems that something is the case. (And sometimes, even when the key premise of an argument doesn’t give the appearance of truth to you yourself, you are able to see how it not only gives such an appearance to others, but does so very powerfully.) Insofar as philosophical arguments for

some sense, I suppose, “disposed” me toward belief. But this is not a case, to use Huemer’s words (in the above quotation), of “one’s being in a state such that one would believe that P, were there no other factors interfering with one’s forming such a belief,” and the fact that some ways of being “disposed” toward belief don’t generate what we would call “appearances” of truth doesn’t show that “appears” and “seems” claims don’t generally report pushes or dispositions toward belief.

One important way that I agree with Huemer about the meaning of appearance claims is that, like him, I reject the multiple senses of “appears” and “seems” that were posited by Roderick Chisholm (whose approach was adopted by Wykstra; see note 43), and instead hold that there is just a single sense: see sect. 1.3, “One Sense of ‘Appear,’” pp. 323–332 of “Phenomenal Conservatism Über Alles.” However, this single sense, on which, on my view, one reports some push toward belief, displays a lot of context-sensitivity, as one can report these pushes (or lacks thereof) at various stages of cognitive processing.

substantial conclusions go, where the shakiest of the starting premises are ones whose credentials are that they give a very strong appearance of being true, so far from being a feeble argument based on “mere appearances,” *that’s the good case!* That’s what the *best* of our arguments are like. Would that more of them had such claims as their shakiest initial premises! And those pushing our problem of horrendous evils often think their case is that powerful. Rowe, for example, starts the last of the passages indented above by asserting that his premise’s answer to the question it’s directed at “must” be right, then softens a bit by saying “it seems quite unlikely” that it’s wrong, but soon gathers enough conviction to declare that its denial is “an extraordinarily absurd idea, quite beyond our belief.” Such an aggressive declaration rhetorically clears the way for the meek (“good cop”) claim that follows, that his premise is “*reasonable* to believe,” hopefully (from the arguer’s perspective) making it seem a wildly generous understatement.

Such a defense can also (and I think even better) be applied to the premise of our “Simple” argument and its key (and only) premise. To appreciate the power our problem has for many, and perhaps to also begin to feel its power in your own thinking, you may find it helpful to imagine encountering in person a Rowe (or an Unger, or an Ivan), moved by the horrors of our world to declare the likes of that “there is just no way” that a wholly good God would have allowed all that. It may even help to imagine them getting more verbally aggressive, and stating forcefully how absurd it strikes them that there might be a God actually allowing *that*, and *that*, and all *that*. Perhaps even *ridiculing* the very idea that there is a wholly good God who has actually decided to let *that* – I mean look at it, if you dare: *that!* — happen.³⁹ We

³⁹ In a passage whose importance Michael Bergmann’s work brought to my attention (see Bergmann’s “Epistemic Circularity: Malignant and Benign,” *Philosophy and Phenomenological Research* 69 (2004): 709–27; p. 723), though I must have read right over it myself before reading Bergmann, the Eighteenth Century Scottish common sense philosopher Thomas Reid had this to say about the role of absurdity and ridicule in thinking about “first principles”:

We may observe, that opinions which contradict first principles are distinguished from other errors by this; that they are not only false, but absurd: And, to discountenance absurdity, Nature hath given us a particular emotion, to wit, that of ridicule, which seems intended for this very purpose of putting out of countenance what is absurd, either in opinion or practice.

This weapon, when properly applied, cuts with as keen an edge as argument. Nature has furnished us with the first to expose absurdity; as with the last to refute error. Both are well fitted for their several offices, and are equally friendly to truth when properly used. (Reid, *Essays on the Intellectual Powers of Man*, D. R. Brookes (ed.), Edinburgh: Edinburgh University Press, 2002; Essay VI, Chapter IV, p. 462)

I am very sympathetic to the importance of recognizing absurdity, in cases where argument fails us, and even in such cases where what’s seen as absurd is only contingently false (and Reid does see this as operative in cases of contingent falsehood): How the suggestion that we are brains vats strikes us as absurd and ridiculous plays a key role in my account (in Chapter 7 of *The Appearance of Ignorance* (Oxford UP, 2017), see esp. pp. 227–28) of how we come to know (at least by ordinary standards for knowledge) that we are not brains in vats (which belief I think is immediately justified). Of course, these

may often respond to such verbal bullying very defensively (to the point of even resorting to the likes of *Wizard of Oz* jokes to deflect the pressure), but sometimes I think it does help to appreciate the power of a claim to see how someone else can find it, not just plausible, but overwhelmingly powerful—and then perhaps ask how horrible things would have to get before *you* would find the case powerful, too. (Or is there really no conceivable limit for you?) And having imagined such a convicted display, one can then evaluate how reasonable our skeptics might seem in then stepping back into “good cop” mode, toning things down a bit, and simply saying that, well, it sure seems to them that if there were a God, we should not find the evils that we find in the world, especially including the truly horrendous evils we find, and in such vast amounts as we find them.

It can also help to get comparative (and a bit personal) and imagine them asking you for what philosophical arguments for substantial and interesting conclusions you *do* like, and then critically comparing your key premise (the shakiest of the starting premises of your darling) with 1s in terms of which gives the stronger and more secure appearance of being true. Though I don’t know what your favorite argument is, and so can only be so confident of this, I have to guess that this wouldn’t be a pleasant exercise for you, and that I would have a lot of sympathy for the claim of our Rowe-ish character were they to exclaim, “Wow, I find 1s a whole lot more

tools can be misapplied; and ridicule—especially when publicly ridiculing the claims of another person—should be used with caution.

But I think I am inclined to apply Reid’s insight to a broader scope of judgments than Reid himself did—though it isn’t always easy to map Reid’s epistemological terminology on to my own. Reid is speaking of “first principles,” but what are those? Bergmann says that for Reid, first principles are believed “noninferentially” (Bergmann, p. 722), and indeed, in the passage Bergmann cites for this, Reid says that in the case of first principles, “There is no searching for evidence, no weighing of arguments; the proposition is not deduced or inferred from another. . .” (Reid, p. 452) And Bergmann seems to speak of what is believed and known noninferentially as being known “directly” (Bergmann, p. 722, n. 28), and, very much in keeping with this characterization, Reid’s description of first principles that I’ve just quoted immediately continues with these words: “. . . it has the light of truth in itself, and has no occasion to borrow it from another.” (And it is not for nothing that Reid is often referred to as a “direct realist.”)

But, as I conveyed in note 29, I think many of our beliefs are *based on* others of our beliefs, and so, in Reid’s terminology, “borrow their light of truth” from other beliefs, though we have performed no conscious inference from the basis to the based, and have not and cannot formulate any arguments. When one or more belief does evidentially support another, and we are sensitive in our holding the supported belief to the support it gets from those other beliefs—in that we would not have held the supported belief, or would not have held it so strongly, were it not for the support it receives from those others of our beliefs—then I think the supported belief is not held “directly,” but on the basis of the beliefs from which it receives support, even though no conscious inference has been performed, and no argument has been formulated, and is perhaps beyond the ability of the believer to formulate. And I think the role Reid assigns to absurdity and ridicule should be applied to such “indirectly” held beliefs: Where a judgment is very strongly supported by others of our beliefs, but in a way we can sense but cannot articulate, our sense of absurdity often kicks in to be a guide to truth, where argument fails us.

powerful than *that!*” Or a bit more snarkily: “Wow, after all your complaints about 1s, *that’s* what you *do* find powerful? Really?!” This guess of course is based on my judgment that 1s is extremely intuitively powerful for the shakiest premise of a philosophical argument to a substantial and interesting conclusion. And since it is our Simple argument’s only premise, and since that argument’s conclusion really does follow from it, that would make our Simple argument a powerful philosophical argument.

One of the most powerful I know.

9. Wykstra’s Challenge: CORNEA and Suspect Appearances

Starting in [section 18](#), I will be urging that the way to get relief from this powerful argument against theism is/would be via the avenue I reported provided me with some relief: by providing credible accounts of why God might allow the horrific suffering of our world—to engage in “theodicy,” in a prominent use of that term. To the extent this effort is successful, it can reduce the degree to which the premise of our simple argument all-in seems to be true. And the rest of this book will then largely be an attempt to obtain such relief.

But is theodicy needed here? We are for now wrestling with swipes at the power of our argument that don’t depend on the swiper producing any theodicy—expressed bafflement at what the argument could possibly be having been the first such complaint.

But we now move on to consider a couple of other important attacks on such arguments as our Simple argument from horrific evils that are in the philosophical literature and that also seek to cut these arguments off before the need for theodicy ever arises—attempts to show that there is no good argument against theistic belief here worth worrying about, even in the absence of any account of what God’s reasons might be for allowing horrific suffering. These attacks, by Stephen Wykstra and by Alvin Plantinga, are not only influential in some philosophical circles, but, because they are the working out of intuitive thoughts that are likely widely shared in not-so-well-worked-out forms, answering them may help address related intuitive worries about our argument that might naturally arise for many readers, or that they might encounter in discussion of these issues.

Stephen J. Wykstra’s wonderful essay, “The Humean Obstacle to Evidential Arguments from Suffering: On Avoiding the Evils of ‘Appearance’” ([note 19](#)), sparked the movement in philosophical theology known as “skeptical theism” (that we will discuss more generally a bit in [section 12](#)), and also, taking Rowe’s argument as its example target, posed an important

challenge to evidentialist arguments from evil that we must answer. Wykstra has followed that paper up with more recent work, some of it co-written with Timothy Perrine, and all of it well-executed, but, for reasons that will emerge, I find the most valuable lessons to be in Wykstra's original.

Wykstra spends much of the beginning of his paper insightfully defending Rowe, and, by extension, evidentialist arguments in general. As Wykstra sees, any argument must start with some premise(s) that is/are not properly argued for (see esp. Wykstra, pp. 80-81), and starting with a premise that gives a strong appearance of being true seems a good place to start. Wykstra thus concludes that there is "nothing inherently suspect about Rowe's appeal-without-further-argument to the claim" that some instances of intense suffering appear to not serve any outweighing good.⁴⁰ And Wykstra shows great openness in general to a principle—a principle expressing a form of epistemic/methodological "conservatism"—that would underwrite Rowe's move from appearances to reasonable beliefs in what appears or seems to be the case. As Wykstra puts it: "And *this* move is, on the face of it anyway, licensed by what many philosophers regard as a proper, indeed primary and indispensable, principle of justification: what Richard Swinburne calls 'the Principle of Credulity,' or what might more descriptively be called 'the "seems so, is so" presumption'" (p. 83).

Yet, after all this able defense of Rowe, and, by extension, evidential arguments generally, Wykstra ends up arguing that the great suffering of our world does not even weakly support the atheistic conclusion of arguments like Rowe's (pp. 77-79), because, although there is nothing *inherently* suspect in Rowe's starting his argument by an appeal to how things seem to be, and while appearances can *often* be presumed to be a guide to how things are, the appearances that ground apparently powerful evidential arguments from horrific evils fall into a problematic class of appearances that cannot be rationally relied upon.

Wykstra attempts to show this by means of examples in which appearance claims seem out of place. Here is one of his cases (slightly modified in the way explained in the attached note):

Searching for a table, you look through a doorway. The room is very large—say, the size of a Concord hangar—and it is filled with bulldozers, dead elephants, Toyotas, and other vision-

⁴⁰ Wykstra, p. 80. The claim Rowe actually evaluates is that these instances of suffering "do not appear to serve any outweighing good." But as Wykstra argues (esp. pp. 82-83), this is best understood as the claim that they "appear to not serve any outweighing good." ("Does not appear to" in its "ordinary sense", as opposed to its "strict sense," generally means "appears to not," Wykstra tells us.) I am here skipping this whole kerfuffle by just putting things in "appears to not" terms in the first place.

obstructing objects. Surveying this clutter from the doorway, and seeing no table, should you say: “It appears that there is no table in the room”?⁴¹

Especially if the table you are seeking would be a small one,⁴² and the objects cluttering your view are very large and many, you would seem not to be in a position to make the appearance claim being considered. Wykstra thinks that this, and similar verdicts about his other cases, are explained by, and support, his proposed “Condition of Reasonable Epistemic Access,” or “CORNEA” principle:

CORNEA: On the basis of cognized situation *s*, human *H* is entitled to claim “It appears that *p*” only if it is reasonable for *H* to believe that, given her cognitive faculties and the use she has made of them, if *p* were not the case, *s* would likely be different than it is in some way discernible by her. (p. 85)

In short, the problem with your claim as you look into the large cluttered room, is that it’s not reasonable for you to think that if there were a table in there, things would look (visually appear) any different to you. Well, at least different in any way likely to impact your judgment about there being a table in the room—an addition that probably should have been reflected in Wykstra’s formulation of his principle. Suppose it’s reasonable for you to think things would look discernibly different if there were a table in the room—if, say, you’re reasonable to think that one of the large objects you see from the doorway would then be in a somewhat different location, maybe because the objects in the room are so tightly packed that any significant addition would have called for a little reorganization. Still, if it’s not reasonable for you to think any of these differences would have any impact on your judgments about whether or not there is a table in the room (or whether or not there may be, or whether or not there appears to be, a table in the room), you are still in no position, just standing there in the doorway and peering in, to declare: “It appears that there is no table in the room.”

Something like Wykstra’s CORNEA principle seems to be right, at least as applied to what I called “all-in” appearance claims to which one wishes to apply the “seems so, is so presumption,”⁴³ as we can see more securely by expanding on his case for it. As Wykstra hoped

⁴¹ Wykstra, p. 84. Wykstra phrases the claim he is asking about at the end of this case as “It does not appear that there is a table in the room.” In keeping with the previous note, I have modified this to instead consider the claim phrased in “It appears that not” terms.

⁴² Wykstra would have done better here to have chosen something smaller than a table to be the object of the search. The key elements of this example seem to be the result of Wykstra basing this case on an example quickly given by Swinburne in a passage quoted at (Wykstra, p. 83).

⁴³ Appearance claims which report some push toward a belief at some stage of cognitive processing that has been thwarted at a later stage, so the push does not constitute an “all-in” inclination to belief, are not governed by anything like CORNEA. Wykstra intends CORNEA to apply only to some appearance claims, though he has a different approach than me to the varying meaning of those claims, on which he

we could sense, the character in his case is not entitled to the appearance claim being considered, and she also fails his CORNEA test, so the CORNEA test gives the correct result here. And the same happens with Wykstra's other two cases. But just giving the right result on three cases isn't much of a reason for thinking a general principle like this is correct—especially when the three cases are structurally as similar to one another as Wykstra's are. Yet, when considering his examples, it's easy to get the sense that something like CORNEA doesn't just spit out the intuitively correct results here, but is really getting at the correct explanation of *why* the characters can't issue the relevant appearance claims. Perhaps the best way of see why this is so is by considering series of cases made by making alterations to the originals, as we'll now do with the table case we are focused on.⁴⁴ Imagine changing Wykstra's example in various ways, slowly (by a series of small alterations), in order to approach and then move beyond a point where our judgment on whether the speaker is entitled to their appearance claim flips. For instance, you can try imagining slowly reducing the number and size of vision-obstructing objects in the room into which you are peering. Once you get down to where there are only a few scattered objects, all too small to effectively hide a table, then, as we can sense, you *are* in a position to declare: "It appears that there is no table in the room."⁴⁵ And you will also then pass CORNEA, so chalk up another success for that principle. But by structuring our cases, we see that the very changes that make our judgment flip on whether the appearance claim is in order make us flip on whether CORNEA is satisfied. Also, as you move through the "grey zone" of cases—close-call cases in which it's hard to say whether you're entitled to the appearance claim—you should also find yourself in a CORNEA grey zone, where it's a tough call whether you pass the CORNEA test ("Is it really reasonable for me to think that...?"). And

intends his principle to apply to those made in the "epistemic sense." But he steps lightly over the various meanings: "Though the term 'appears' is ubiquitous in his case, Rowe provides no explication of its meaning. I shall provide one, suggesting that he uses the term in something close to what Swinburne, following Chisholm, calls its 'epistemic' sense (as distinct from its 'comparative,' 'phenomenal,' and 'hedging' senses). But the 'appears idiom' is a philosophical swamp with an enormous literature, most of it of dubious relevance to Rowe's use. To avoid getting bogged down in it, I propose the following as a paradigm of Rowe's use that will give us our initial bearings. . . ." (p. 80). And I trust that, despite our different ways of navigating the swamp, Wykstra and I end up focusing in on roughly the same range of claims as those to which to apply CORNEA: it's claims to the kind of appearance of truth that one hopes attaches to the premises of one's arguments.

⁴⁴ Here I pursue the argumentative strategy I employ in defense of an "insensitivity" explanation for when we will seem to not know that something is the case at (DeRose 1995: 23-26 = DeRose 2017: 18-20). Wykstra's CORNEA principle is an insensitivity principle for when we are entitled to appearance claims.

⁴⁵ Indeed, in keeping with the points made in the second paragraph of Sect. 8, at a certain point down this path, you should no longer say "It appears that there is no table in the room," because you should say something stronger instead: that "There is no table in the room," for instance. But this is a case where you are entitled to the appearance claim, but still shouldn't issue it, because you are entitled to, and ought to say, something stronger, instead.

throughout the cases we consider, whether you pass the CORNEA test intuitively seems tightly connected to whether you can make the appearance claim: it seems to be pointing to quite relevant considerations on which the matter of whether you can make the claim quite plausibly turns. Similar results obtain for Wykstra's other cases. It's considerations like these that show us that CORNEA is on to something—again, when we take it as a principle about our entitlement to the kind of “all in” appearance claims we are interested in here.

10. The Importance of the Basis of an Appearance Claim

Nailing down all the needed details of how to formulate and apply CORNEA would be a very difficult and long task, most of which we'll skip here. But before getting to why Wykstra thinks his principle undermines the key appearance claims of arguments like ours, and why I think he's wrong about that, we do have to take a look at one very important feature of how to wield CORNEA properly.

Consider the example of a company that buys old houses, updates them and fixes them up, and then sells them. As the construction and remodeling team nears the end of a very big job a bit ahead of schedule, they are looking for a new smaller project that they can squeeze in to the couple of weeks open on their schedule before they are committed to starting work on another big job. They are considering a few houses that are on the market, and that look from their listings like plausible candidates. From the age and size of the houses, which can be ascertained from their listings, the company can start to get an idea of whether fixing them up would be too big a job for them to take on in their current circumstances, and they can start to weed out some potential candidates and zero in on others. But they need a better basis than what they can gather from the listing before actually taking on a job. As it happens, they have just sent one of their workers to the vicinity of one of the candidate houses, anyway (to deliver a contract, say), so they decide to have that worker take a look at the house. Unfortunately, the look has to be a very quick walk-through, and not a careful inspection, because the worker has another important appointment to get to.

At this point, our story divides into two versions. On one variant, the worker is an Old Pro, who has looked through many old houses, and can often ascertain quite reliably the likely size of a job from even a quick look-through. On our second variant, the worker is a Newbie, who has little idea of what to look for. As they approach the house, both take a fairly quick look at the roof, since if the roof needs to be replaced, that adds considerably to the size of the job. We can suppose that the need for a new roof would not by itself mean the project was

unfeasible, but would take things a significant distance in that direction. Neither of our inspectors sees any indication that the roof needs to be replaced, but Newbie is in no position to be very confident that it isn't needed. If the roof were in horrible enough shape, even Newbie would likely see signs of trouble, but it could easily happen that more subtle signs of big problems elude Newbie. But Old Pro, knowing just what to look for, is in a position to be at least fairly confident that the roof is in good enough shape. This is emblematic of what happens throughout the look-overs of the house, with Old Pro being confident enough that they won't be plagued with each of the various sources of big trouble that she ends up being quite confident indeed that there won't be enough total major problems to make the project unworkable. Imagine the cases so that Old Pro, but not Newbie, ends up in a position to declare that, given the company's time and financial constraints, "It **seems** that we can do this job," with Newbie left in a position to only say something weaker like, "Well, there weren't many indications of any really major problems that I was able to catch."

To check if and how it can explain our contrast, let's look at CORNEA again:

CORNEA: On the basis of cognized situation *s*, human *H* is entitled to claim "It **appears** that *p*" only if it is reasonable for *H* to believe that, given her cognitive faculties and the use she has made of them, if *p* were not the case, *s* would likely be different than it is in some way discernible by her. (p. 85)

And it seems that CORNEA can account for the difference here, because it seems that it *is* reasonable for Old Pro, but not for Newbie, to think that "given her cognitive faculties and the use she has made of them," if the job had been unfeasible, the "cognized situation" likely would have been different in a way that would have been not only discernible to her, but (to take account of the little tweak I suggested to the principle) would have affected how she was inclined to judge the feasibility of the project. Note that to handle the cases correctly, we had to pay attention to what our inspectors' judgments were based on, what they were "going by": this was an aspect of the use Old Pro made of her cognitive faculties that was vital to the correct application of CORNEA to her case. It's because Old Pro was taking into account various signs of no-big-trouble and various lacks of signs-of-big-trouble that Newbie was in no position to appreciate that Old Pro, but not Newbie, ended up in a position to issue the appearance claim. This is the key lesson of how to use CORNEA that we have to apply as we consider our argument from evil.

11. What Our Atheologian Is Relying on and How They Pass CORNEA: An Appearance Reasonably Relied Upon

Wykstra thinks CORNEA undermines Rowe's crucial appearance claim, and more importantly, since Rowe is just the exemplar Wykstra is using to address them, evidentialist arguments generally. But I think Wykstra attains this result by underestimating the basis for these claims. In this important passage, I add emphasis to the key part where Wykstra characterizes what our atheologian is relying on as Wykstra opens his crucial section, "CORNEA Applied":

Return, then, to Rowe's fawn, suffering in a distant forest. Rowe's claim is that the suffering appears not to serve any God-justifying purpose. It is clear, I think, that *the feature of the cognized situation crucial to Rowe's claim is that there is no outweighing good within our ken served by the fawn's suffering*. CORNEA thus forces us to ask the following question: if there were an outweighing good of the sort at issue, connected in the requisite way to instances of suffering like this, how likely is it that this should be apparent to us? (pp. 87-88)

And Wykstra's answer (p. 88), given how much greater God's mind would be than ours, is that it isn't very likely at all that the outweighing good God was acting toward is one that would be apparent to us, and thus, in brief, we, and Rowe, cannot reasonably judge that we would see what that good was, if it existed, and thus, we, and Rowe, are not entitled to Rowe's appearance claim. This because there being "no outweighing good within our ken" is what we're going by in making our key judgment; it's the basis of our appearance claim. Well, at least about this one case involving the fawn. As we saw, Rowe then steps back and considers how things appear given all the horrible suffering of our world. But Wykstra thinks this will not help:

And it is evident that the case cannot be repaired merely by pointing out the number of instances of suffering which have no evident purpose—unless one has justification for believing that these instances belong to a class such that if Divinely purposed goods exist in connection with all known instances of suffering in this class, these goods would always or usually be within our ken. (p. 88)

It is puzzling why Wykstra thinks we'd have to be reasonable in thinking that the goods in question would always or usually be within our ken, rather than just that they would be so more often than they in fact are.

But the larger problem here is that Wykstra isn't taking into account all that our atheologians are going by as the basis of their key appearance claim: Wykstra is in that way treating Old Pro as if she were Newbie. Those who press our argument are not just going by their or our inability (or even "complete inability," said with whatever intonation one wishes) to discern (or even begin to discern) the goods for the sake of which God allows the horrors of our

world (for Rowe), or (more generally) the good reason God has for allowing them. If that were all they were going by, their claim would be in jeopardy of failing a fair application of CORNEA, for the reasons Wykstra gives.⁴⁶ But our skeptics are also, and in a way I'm about to explain, independently, going by just how vast and horrible so many of these evils are.

Of course, anyone who has encountered the arguments in question, and certainly Wykstra, realizes that they *somehow* make use of just how awful and extensive the evils of our world are. But the way the arguments often go can make it seem that those features of our world's evils are just used as theodicy killers: to knock out as insufficient particular proposed reasons God might have for allowing evil, and thus to reinstate the claim that we can't discern (or more aggressively: can't even begin to discern) God's reasons for allowing the evils of our world: "Oh, yes, that would [or "might," to taste] be a good reason for allowing *some* evils, but it can't account for why God might allow all the truly horrific evils of our world, so we're right back where we started: We have no idea why God allows those evils." For instance, in Ivan's famous speech in *The Brothers Karamozov* that we looked at in Sect. 3, Ivan targets a particular account of why God might allow evils, claiming it is not good enough to account for the horrific evil he has been conveying:

[C]an you understand why this nonsense is needed and created? Without it, they say, man could not even have lived on earth, for he would not have known good and evil. Who wants to know this damned good and evil at such a price? The whole world of knowledge is not worth the tears of that little child to 'dear God.'

So, yes, the nature and extent of our world's evils are used to counter particular proposed reasons God might have had. But, you know, it's not only those theists invested in the particular theodicy Ivan targets who are made to squirm by Ivan's attacks. Fans of other theodicies, or combinations of theodicies, can also have their faith (in the very existence of a perfectly good God) shaken—and rationally so—by a real look at the horrors of our world, as they wonder whether the accounts they've accepted can really handle all that. And important to our current concerns: The same can happen to those who haven't latched on to any particular theodicies, but have just supposed that God allows evils for reasons they can't see. They can be quite reasonable in taking that stance with respect to some (perhaps minor) evils, but in then changing their stance and thinking it is very unlikely indeed that God (exists and) has some good unknown reason for allowing all *that*, as they point to, or have their attention pointed to, unspeakably horrific suffering. There's nothing irrational in finding it more unlikely

⁴⁶ Here, I echo Draper's evaluation of how a limited kind of evidential argument from evil (in Draper's terms, "arguments from the failure of theodicy" (see Draper, "Confirmation Theory and the Core of CORNEA," esp. pp. 138-9)) are challenged by Wykstra's arguments, but Draper is issuing his evaluation with respect to more recent efforts of Wykstra's.

that there is a God-justifying reason for the world's evils as one finds those evils to be vastly more horrifying and finds horrifying evils to be vastly more extensive.

"Well, but isn't it still possible that God has some good reason, even for all *that*, that you can't discern?"

"I suppose anything is *possible*. I just judge it [incredibly] unlikely. After looking at all that unspeakably horrific evil our world contains, it's really [—and I mean *really*—] seeming to me that if there were a God, we should not find the evils that we find in the world, especially including the truly horrendous evils we find, and in such vast amounts as we find them. I mean, have you really looked at those things?"

That seems a perfectly reasonable stance. And in keeping with it, and now turning more explicitly toward Wykstra, there seems to be a perfectly reasonable, and closely related, response to a CORNEA-based challenge to it, when we keep in mind that our atheologian is in large part going by just how horrible and extensive the evils of our world are:

"Well, do you think that, given your cognitive faculties and the use you have made of them, if there *were* a God-justifying reason for our world's evils, the situation would likely be different in a way discernible by you? Is it *reasonable* for you to think that?"

"Well, yes, in the relevant way, I do think that, and quite reasonably so. I am basing my appearance claim in large part on just how unspeakably horrible so many of the world's evils get, and I do think it likely that if there were a God-justifying reason for the world's evils, there wouldn't be so many of them that were so unspeakably horrible. In fact, for many particular ones, I'm finding it quite likely that if there were a good God watching over things, those particular ones would not have been allowed to get so horrible. But especially when taking them all together, do I think it's likely that things would be discernibly different if there were a God-justifying reason for all the world's worst evils?: Yes!"

"But isn't it possible that if there were a God-justifying reason for all the world's evils, the relevant aspects of our situation—for you, how many of the world's evils get so unspeakably horrible—would be just as they actually are, and so not discernibly different at all?"

"I suppose I can't really rule out that being *possible*, but you didn't ask whether it's possible, but whether it's likely. And on the matter of whether the 'cognized situation' would be discernibly different in the way I spoke of, I'm finding that extremely likely indeed. And that's putting it very mildly."

And reasonably, I'd add.

12. “But We’re All Newbies Here!”: Sensible Skeptical Theism vs. the More Aggressive Varieties of It

When, in the above section, I wrote that Wykstra is “treating Old Pro as if she were Newbie,” I was just making the point that Wykstra is neglecting some of the basis of the atheologians’ claims. (Note my careful “in that way” qualification.) But some readers were probably yelling at me: “We’re all Newbies here!”, having in mind something in the vicinity of the good point that none of us have had the opportunity to observe many worlds, some governed by a perfectly good God, and others not, so that we could compare the two types in terms of the worst evils they tend to contain.⁴⁷ In that way, we are much like the Newbie of my house inspection story. Let me quickly assure the provoked that I agree it’s wise to keep in mind this lack of experience of ours, and that it can sensibly be taken to provide those moved by our argument from horrific evils with some reason for displaying some modesty in how certain they take their judgment to be here: We would indeed be better positioned to judge the likes of premise 1s if we had the experience to make us in the relevant way like Old Pro. Some caution here seems like a very sensible bit of “skeptical theism,” where (to take a good encyclopedia statement of it) we take that in general to designate strategies “for bringing human cognitive limitations to bear in reply to arguments from evil against the existence of God.”⁴⁸

But Wykstra is not just sounding that sensible, “don’t be so absolutely certain” note of caution, but is making the much more aggressive skeptical theist claim that the horrific evils of our world don’t even count as weak evidence against the existence of a perfectly good God. And here, our atheist can remind us that even Newbie would be in a position to see signs of great trouble, if they were bad enough, and would then be entitled to the pessimistic appearance claim. I should point out that in later work, Wykstra has toned down his skepticism a bit, arguing instead that arguments like Rowe’s do not provide “leveraging” evidence—evidence strong enough to rationally move someone from belief in God to disbelief.⁴⁹ However, the story here gets complicated, because Rowe is there being construed (alas, it seems, with his consent) as relying on (what Wykstra dubbed) a “noseeum” inference,⁵⁰ from our inability to

⁴⁷ Hume passage?

⁴⁸ Opening sentence of Trent Dougherty’s entry on “Skeptical Theism,” in the *Stanford Encyclopedia of Philosophy*; entry first published Jan. 25, 2014; <https://plato.stanford.edu/entries/skeptical-theism/>

⁴⁹ Wykstra, “Rowe’s Noseeum Arguments from Evil,” in Daniel Howard-Snyder, ed., *The Evidential Argument from Evil* (Indiana University Press, 1996), pp. 126-150; see esp. pp. 130-131 for an explanation of “leveraging evidence.”

⁵⁰ See the very opening of the Wykstra paper referenced in the previous note. The term seems fairly entrenched in the literature, but I use it with some hesitation. It is introduced as a Midwestern colloquialism for “tiny flies...which are so small that you ‘no see ‘um’,” but I fear that its origin may lie in Midwestern attempts to mock how Native Americans speak English.

see a good reason for God to allow the evils of our world to there (probably) being no such reasons—a move which very considerably weakens the argument, but which, sadly, much of the philosophical literature on skeptical theism has come to emphasize.

My interest here has been to defend our “Simple” argument from the aggressive form of skeptical theism embodied in Wykstra’s CORNEA principle. This involved skipping the many twists and turns in the subsequent literature on skeptical theism, in order to see how our argument escapes an aggressive, though in its way quite intuitive, attempt to undermine it—and thereby arrive at a better idea of *how* our argument’s premise, 1s, is based on the horrific evils contained in our world. In particular, the lessons we’ve drawn from this look concern how, in its best form, our argument does not utilize the types and amounts of horrific evils our world contains just to wipe out particular proposed theodicies and thereby clear the way for a mere “noseeum inference,” but utilizes those features of our world to generally motivate the thought that the evils of our world are far too vast and far too awful for it to be likely that God (exists and) has a sufficient reason for allowing them.

But to update you a bit on the state of the discussion of skeptical theism: It is not at all unusual for skeptical theists to adopt a stance as aggressively skeptical as Wykstra’s was in his original paper, and that may indeed deserve to be thought of as the “standard” position of skeptical theists.⁵¹

As I noted above, Wykstra himself has toned down the skepticism in his skeptical theism a bit. However, as we saw, it’s still quite aggressive—and certainly goes well beyond the “don’t be so certain” stance that I labelled “sensible skeptical theism.” And what’s more, by my lights, at least, his arguments still rely on an overly narrow view of the atheologian’s basis for their argument that I pointed out in [sections 11-12](#), above. In “Confirmation Theory and the Core of CORNEA,”⁵² a critique of Wykstra’s more recent work, Draper notes how Wykstra toggles (my term for it) between claiming that the vast horrific evils of our world fail to provide strong reason against God’s existence and claiming that our inability to fathom God’s reason for allowing that fails to provide such strong reason (see esp. section 6, pp. 139-141 of Draper). My best explanation for this toggling is based on how, as I noted in [Sect. 11](#), above, the vastness and horrifying nature of the world’s evils is often seen as functioning: as just knocking down

⁵¹ See section 1.2 of his *Stanford Encyclopedia of Philosophy* article on “Skeptical Theism” (article copyright 2014; accessed at SEP on 26 July 2020) where Trent Dougherty presents the places at which several of the most prominent skeptical theists make strong claims of the likes that, for the kinds of reasons skeptical theism gives, arguments from the evils of our world provide no good reason whatsoever for thinking that God does not exist. Dougherty labels the strong claim of this sort the “No Weight Thesis.”

⁵² In Trent Dougherty and Justin P. McBrayer, ed., *Skeptical Theism: New Essays* (Oxford University Press, 2014), pp. 132-141.

particular theodicies, so that we are returned to a state of not knowing God's reasons and are thereby ready to perform a "noseeum" inference. Given this (mistaken) view of how those features of the world's evils function in the best forms of the argument, those features can be seen as playing some role (which they obviously are), while the atheist's inference can still be viewed as wholly going through our inability to discern God's reasons. As I claimed with respect to Wykstra's original effort, his argument does not go through on a proper construal of the basis of the athologian's case. Similarly, but with respect to Wykstra's more recent work, Draper argues that while Wykstra's case does threaten attempts to wrest God's non-existence from *our inability to see* God's reasons for allowing evils, it "fails to show that the evils themselves are not strong evidence against theism" (p. 132).

And construing the atheistic argument as relying on a "noseeum inference" has not just illicitly made Wykstra's job easier, but has aided skeptical theism generally. Indeed, Michael Bergmann, a leading skeptical theist himself, sees the resistance to such inferences as the skeptical theist's signature move:

[T]he fact that humans can't think of any God-justifying reason for permitting an evil, doesn't make it unlikely that there are no such reasons; this is because if God existed, God's mind would be far greater than our minds, so it wouldn't be surprising if God had reasons we weren't able to think of. This very natural sort of response is precisely the move the skeptical theist is known for.⁵³

Bergmann realizes that not all arguments from evil rely on "noseeum inferences," and so says that "skeptical theists don't claim that their skeptical theses undermine all arguments from evil."⁵⁴ But what we have seen is that the intuitive argument from horrific suffering, understood *in its best form*, does not rely on such an inference, and so is not neutralized by the signature moves of the skeptical theist, which target only needlessly weak forms of the argument.⁵⁵

⁵³ Bergmann, "Skeptical Theism and the Problem of Evil," in Thomas P. Flint, Michael C. Rea, ed., *The Oxford Handbook of Philosophical Theology* (Oxford UP, 2009), pp. 374-375.

⁵⁴ Bergmann, "Skeptical Theism, Atheism, and Total Evidence Skepticism," in Trent Dougherty and Justin P. McBrayer, ed., *Skeptical Theism: New Essays* (Oxford UP, 2014), p. 210.

⁵⁵ Well, I suppose what we've actually seen is just that Wykstra's original effort, but also, I suppose, any instance of what Bergmann sees as the signature move of skeptical theism, doesn't neutralize the atheist's argument in its best, intuitive form. Beyond that, my claim that the efforts of skeptical theists have generally failed to disarm our argument is trash talking that is subject to correction if some unexpected (to me) piece of successful skeptical theism turns up somewhere.

13. Plantinga and “Probabilistic” vs. “Deductive” Arguments from Evil

I certainly cannot be making claims about how powerful our Simple argument is without grappling with the attack on arguments like it made by the ace defender of theism, both in general, and in particular from problems of evil, Alvin Plantinga, perhaps most vociferously in his paper, “The Probabilistic Argument from Evil” (henceforth “PAE”), and updated in some later works.⁵⁶ Especially in Sections 3-6 above, I considered challenges to “produce the argument!” issued by those who might claim to be having a hard time seeing what the argument from evil is even supposed to be. I was construing the professed bewilderment as one is likely to encounter it from various “apologists” for theistic belief (on this, see note 24, above), but insofar as such a response to the problem is directly or indirectly inspired by material in the professional philosophical literature, I imagine it is Plantinga’s attack that is often responsible. Plantinga certainly professes an inability to see what his opponents’ argument is supposed to be at crucial points, as well as a willingness to attempt to help those hapless opponents out by exploring what their arguments might be—only of course, to find all attempts to help them out to come to naught. But, puzzlingly, Plantinga is in part responding to Rowe’s argument, laid out in numbered steps, so his problem really can’t be one of not knowing how the argument for the non-existence of God might be formulated. Rather, as I’m about to urge, Plantinga manages to be puzzled by construing Rowe and others who might use similar arguments for atheism as aiming to establish further conclusions related to but going beyond the claim that God does not exist, and what Plantinga turns out to be at a loss for is seeing how those further conclusions might be established. In executing this strategy, Plantinga is saddling his opponents’ argument with completely unreasonable burdens, and if someone is inclined to find an argument like our Simple argument from horrific suffering to be a powerful one (as I certainly am), there is nothing in Plantinga’s attack that should lead them to retract such a judgment. Or so I will argue.

But first: I keep talking about arguments “like” our Simple argument—and Rowe’s and van Inwagen’s argument. What is the type in question? Answer: These are what are called in the philosophical literature “probabilistic” (or sometimes “evidential”), as opposed to “deductive” arguments from evil. Plantinga has prominently attacked both kinds of argument, but we are here considering his attack on the former. To evaluate the success of Plantinga’s defense, we need a decent idea of the scope and nature of the arguments he intends to be defending against.

⁵⁶ Plantinga, “The Probabilistic Argument from Evil,” *Philosophical Studies* 35 (1979): 1-53.

Plantinga opens PAE by introducing the label “atheologian” for those who offer arguments against God’s existence (p. 1), and he then gives a brief history of atheology, writing that “The vast majority of those who offer an atheological argument from evil have held that the existence of evil (or of the amount and kind we find) is *inconsistent* with the existence of a wholly good, omniscient and omnipotent God,” but then relates that atheologians had recently retreated from such bold charges of inconsistency to what he calls “the probabilistic argument from evil”:

More recently, however, those who make an atheological argument from evil have conceded that the existence of evil may be logically consistent with that of God; nevertheless, the former renders the latter unlikely, or improbable, or provides evidence against it, or gives a good reason for accepting its denial. (PAE, p. 1)

Though the distinction here might not be as clear as one might like (see [notes 58-59](#)), our “Simple” argument, along with van Inwagen’s and Rowe’s arguments would all seem to be such “probabilistic” arguments, that do not trade in charges of inconsistency, as opposed to the old-style “deductive” arguments, which do, and so all three fall in the scope of Plantinga’s attack in PAE—though Rowe’s is the only one that was around when Plantinga wrote PAE, and so the only one he explicitly addressed there. (Plantinga had previously, and very prominently, attacked deductive arguments in other works. Indeed, his attack is usually credited with being largely responsible for the retreat of at least some atheologians to the less aggressive arguments.) And an atheologian who endorsed any one of our probabilistic arguments would certainly think that it “gives a good reason for accepting” the non-existence of God. We are wondering if there is anything in Plantinga’s attack to show that they’re wrong to think so.

Note that none of our “probabilistic” arguments actually have anything about probability in their formulations, so one might wonder how they earn the label. To better see the difference between deductive and probabilistic arguments, consider:

The Classic Deductive Argument from Evil

- 1d. Evils exist
- 2d. An omniscient and perfectly good being eliminates every evil it can eliminate
- 3d. An omniscient and omnipotent being can eliminate every evil
- So, C. There is no God

“God” in the conclusion here denotes the God of standard theism, who is by definition perfectly good, omnipotent, and omniscient, which is why the non-existence of such a being follows from the premises of the argument: If such a being existed, it would have both the ability and the will to eliminate all evils (or so the argument’s second and third premises claim), so there would be no evil whatsoever; and since there is evil, we can conclude that there is no such God. We can

now contrast this “deductive” argument with the three “probabilistic” or “evidential” arguments we have considered (reproduced here for convenience):

van Inwagen’s Argument from Horrendous Evil

- 1v. We find vast amounts of truly horrendous evil in the world
- 2v. If there were a God, we should not find vast amounts of horrendous evil in the world
- So, C. There is no God

Rowe’s Argument from Intense Suffering

- 1r. There exist instances of intense suffering which an omnipotent, omniscient being could have prevented without thereby losing some greater good or permitting some evil equally bad or worse.
- 2r. An omniscient, wholly good being would prevent the occurrence of any intense suffering it could, unless it could not do so without thereby losing some greater good or permitting some evil equally bad or worse,
- So, Cr. There does not exist an omnipotent, omniscient, wholly good being.

The Simple Argument from Horrendous Evils

- 1s. If there were a God, we should not find the evils that we find in the world, especially including the truly horrendous evils we find, and in such vast amounts as we find them.
- So, C. There is no God.

None of our probabilistic arguments explicitly conclude just that God *probably* doesn’t exist. Instead, like the deductive argument, their conclusion is simply that God does not exist. (Rowe explicitly spells out the properties of God that are in question, while the other arguments build that into how they are using the term “God.”) What’s more, like their deductive cousin, the premises of our probabilistic arguments are supposed to logically (deductively) imply this simple atheistic conclusion. What’s the difference, then?

The difference is in the nature of the premises. Atheologians who peddle the classic deductive argument think that their premise 1d, which simply reports that there is evil, is *inconsistent* (as Plantinga puts it) with the claim that such a God exists. They may similarly claim that there is a *contradiction* in holding that both evils and God exist. In keeping with such claims, they think that the only other premises they need to get from that evil-reporting premise to their atheistic conclusion—those other premises being 2d and 3d—simply unpack the meaning of the key terms in their argument (“perfectly good,” “omnipotent,” and “omniscient”), so that the very meaning of the terms involved rule out there being both God and evil. (Whether these arguments succeed is another matter: In what I’m calling the “Classic” formulation of the argument, 2d, in particular, seems wrong, and is certainly quite problematic

as a mere unpacking of the meaning of its key terms—especially of “wholly good.”⁵⁷ But we are currently interested in what proponents of such arguments were at least *trying* to do.) In its evil-reporting premise, the Classic argument appeals only to the fact that there are some evils or other, no matter how few, how small, or of what kind. One could beef up that premise to one that appeals to something about the amounts and kinds of evils that exist, and still be offering a “deductive” argument, if the premises of one’s argument are still only a) reports of the evils our world contains⁵⁸ and b) premises that simply unpack the meanings of the key terms in one’s argument: that way, one will hold that the existence of the evils mentioned in one’s evil-reporting premise are *inconsistent* with the existence of God.

Our probabilistic arguments are not like that. Van Inwagen’s first premise I think should count as just a report about the evil of our world. Unlike the classic deductive argument, he relies not just on there being some evils or other, but on the fact that there are “vast amounts of truly horrendous evil in the world.” But if he could somehow get from that premise to the non-existence of God by using only other premises that simply unpacked the meanings of key terms, he would be advancing a deductive argument to show that it’s inconsistent to accept both that there are vast amounts of truly horrendous evils and that God exists. But that’s not his goal.⁵⁹ His second premise offers a judgment that atheologians pushing his argument will

⁵⁷ We will later see reasons for thinking this argument does not work, but those who find the matter urgent can right now think about the kinds of considerations that might make one retreat from 2d to something like Rowe’s more limited premise 2r, and then consider whether it really follows just from something’s being perfectly good that it eliminates every evil it knows about and can eliminate.

⁵⁸ There is an issue of what are the kinds of evils whose existence we can just “report,” in the sense in question. Clearly, one cannot just so report that there are “evils which would not exist if a perfectly good God existed.” The existence of evils of *that* kind, so designated, certainly is inconsistent with the existence of a perfectly good God, but that’s not a kind that we can in this context just report to be exemplified in our world. That there are “horrific” evils, or “natural” evils (to cite a kind much discussed in connection with the problem of evil), by contrast, can I think be in the relevant sense just reported. Likewise for certain ways of relating the *amounts* of evils that exist: that there are, say, “vast” amounts of horrific evils can be reported (however vaguely); but not that the amount of horrific evils is “more than a perfectly good God would allow.” One idea here is: for a type of evil whose existence can just be reported, characterize it independently from the likes of whether God could allow it, or could have prevented it without jeopardizing some greater good, etc.

⁵⁹ van Inwagen seems unhappy with the distinction between the two types of arguments, but classifies his own as “evidential” nonetheless: “Since I find the distinction artificial and unhelpful, I am, of course, not going to allow it to dictate the form that my discussion of the argument from evil will take. I am, as it were, jumping right into the evidential problem (so-called; I won’t use the term) without any consideration of the logical problem” (*The Problem of Evil*, p. 68). His skipping to the evidential problem seems due mainly to his not being very impressed by the logical problem. Having found the importance of the discussion on the evidential side, and finding the distinction a mess, anyway, he didn’t expend effort on the logical problem or the distinction between the two forms of the problem. I don’t imagine my discussion here would satisfy van Inwagen that the division has been made clear: the distinction I rely on between those premises that do and those that don’t simply “unpack the meanings” of the key

find plausible (or perhaps even compelling), but there is no pretension that it simply unpacks the meanings of some of the terms involved in the argument. Rowe's argument and our Simple argument don't even offer a premise that is just a report of the evils in our world,⁶⁰ much less do they propose to get from such a report to the non-existence of God just by the meanings of any terms used in the arguments. So none of our probabilistic arguments purport to show any *contradiction* or *inconsistency* between God and evils (or of evils of certain reportable amounts or kinds).

But they do still aim to use evils of our world to argue that God does not exist. And those who endorse these arguments do of course think they provide good reason for accepting that atheistic conclusion.

14. Plantinga's Attack on Probabilistic Arguments from Evil and What the Attack Gets Right

So, again, does Plantinga's attack show that such probabilistic arguments fail to so "give a good reason" for atheism? Or even that they fail to give *an exceedingly powerful* reason for atheism? Things very quickly fly off the rails. Though the atheologian I'm defending (one who endorses our "Simple" argument) falls squarely in Plantinga's announced intended range, Plantinga's punches seem only to strike things that they aren't saying. And it's not just a matter of whether our atheologian uses forms of the word "probable"; there's a more serious disconnect here.

Plantinga cheekily construes the atheologian as utilizing the claim:

E There are 10^{13} turps of evil,

continuing: "where the *turp* is the basic unit of evil, so that ' 10^{13} turps' is a name of the evil, past, present, and future, the actual world (call it ' α ') contains" (PAE, p. 2). So Plantinga's atheologian is appealing to the amount of evil the world will ever contain. *Our* atheologian is appealing not to just how much evil there is, but also to what a lot of it is like, but let's let that

terms is quite vague—as is the distinction between those premises that do and those that do not simply "report" on the evils of our world (see the immediately previous note). But I suppose I would be joining van Inwagen as classifying the four arguments in front of us as I do, placing three of them on the "evidential" side, and one on the "logical" side, of this vague distinction.

⁶⁰ See two notes above for an explanation of why Rowe's first premise is not the relevant sense just a report of a kind of evil we find in the world. Rowe certainly doesn't treat it as just such a report, as we have seen when we took a good look at Rowe's defense of that premise.

pass. The real problem here is still coming. Not being one of those old-style ones, the “atheologian” Plantinga is dealing with here does not claim that E is logically *inconsistent* with

G God exists and is omniscient, omnipotent, and wholly good.

Instead:

[T]he probabilistic argument contends that E is evidence against G, or that given E, it is unlikely that G is true. (p. 2)

This *may* be a fair characterization, so far. I mean, as we’ve noted, our atheologian’s argument is simply for the conclusion that God does not exist, with nothing explicitly in that conclusion about either evidence or unlikelihood. But in advancing that argument, our atheologian certainly intends to convey that the evil of our world is “evidence against G,” and also that (as I’ve already said) it provides us with good reason to reject G, whether or not they care to pronounce on the overall likelihood of G. So, this may be characterizing them well so far, if this is an easy-going “or” (which can be satisfied by satisfying either of the items the “or” connects; the other reading would be the “or” of reformulation—the “or in other words” “or”). But then Plantinga’s punches begin to miss our atheologian wildly, striking only at strong claims they need not be making; immediately following the above, we read:

And we shall have to ask why the atheologian thinks this is so; what is his reason for thinking G is improbable with respect to E? But before we ask that question, let’s suppose, for the moment, that the probabilistic atheologian is right and that the fact is, G is improbable on E. What is supposed to follow from that? How is that to be construed as an objection to theistic belief? How does the atheologian’s argument go from there? It doesn’t follow, of course, that theism is false. Nor does it follow that one who accepts both G and E (and, let’s add, recognizes that G is improbable with respect to E) has an irrational system of beliefs or is in any way guilty of noetic impropriety. For it could be, of course, that G is improbable with respect to E but probable with respect to something else we know. (p. 2)

Wait a minute! How did our atheologian get into the business of accusing theists of having “an irrational system of beliefs” or of being “guilty of noetic impropriety”? Of course, some may make such charges. But is an atheologian who endorses a probabilistic argument from evil as providing a strong reason (or, say, an overwhelmingly strong reason) for denying God’s existence, but who does not get into the business of calling all who disagree with them irrational, off the hook, so far as Plantinga’s attack goes?

Things get worse. After an illustration to show that one claim can indeed be improbable on another, yet still be probable on something else that one might know (which we should all grant), Plantinga continues:

What the atheologian must show, if he wants to show that there is a viable objection to theism here, is that on some relevant body of *total* evidence—his own, perhaps, or the theist's, or perhaps a body of total evidence shared by all those who are party to the dispute—G is improbable. If he could show that G is improbable on his *own* total evidence, the atheologian could show that *he* has a good reason for rejecting theism; if he could show that G is improbable on the *theist's* total evidence then he could show, perhaps, that the *theist* is irrational or guilty of noetic impropriety in accepting G. . . .

So the atheologian must show that G is improbable with respect to the relevant body of total evidence, whatever exactly that is. To do this, he would be obliged to consider all the sorts of reasons natural theologians have invoked *in favor of* theistic belief—the traditional cosmological, teleological and ontological arguments, for example. He would also have to consider more recent versions of the moral argument as developed, for example, by A. E. Taylor, and still more recently by Robert Adams, along with the sorts of broadly inductive arguments. . . (p. 3)

Alright, let's cut that off (mid-sentence) right there. How did we get to the point that our atheologian cannot be offering a decent argument against God's existence unless they have (as part of that argument?) considered and countered every argument offered for the opposing view?

Plantinga has somehow saddled his atheological opponents with the burdens of showing that their conclusion is likely on some body of total evidence, in a way which seems to require them to counter every argument against them before their own argument can be any good; and of showing that those who disagree with them are irrational.

And here is what Plantinga certainly gets right about probabilistic arguments from evil: Though I haven't construed the atheologian as trying to do so, I readily agree with Plantinga that probabilistic-argument-wielding-atheologians have not met those burdens that Plantinga has assigned to them: Atheologians have not by means of such arguments shown that their own arguments are better than all the arguments that are on the other side of the issue, nor have they shown that theists who disagree with them are irrational. And I also readily grant that there are good considerations in Plantinga's work for saying that atheologians have failed to show such things.

15. After Plantinga's Attack: Our Still Extremely Powerful "Probabilistic" Argument from Evil

But boy howdy, those would be tough things to show! And not just on the matter at hand: If you somehow get maneuvered into the position that you must meet such burdens in order to

be credited with having a good argument (or a powerful argument, or an extremely powerful argument) for your side on *any* controversial matter, I'm just going to give you up for dead. (So while, as I said, I find that Plantinga has helpful things to say in arguing that atheologians have not met these burdens, that is all in service of a conclusion that I would have been inclined to just spot him from the get-go anyway, on completely general grounds.) These are unreasonable burdens. And, yes, that holds in the special case of the atheologian's opponent: If in presenting arguments *in favor of* God's existence (one of Plantinga's favorite activities, it seems) theists have to counter all arguments for atheism before their argument is a decent one, or are somehow necessarily dragged into the project of charging atheists (perhaps all of them?) with irrationality and showing that charge to be right, I'm going to find their cause hopeless, too.⁶¹

⁶¹ It's interesting that the way Plantinga sets up the positions at the start of his paper discourages turning around some unreasonable burdens to apply to theists like him. Note the opening two sentences of PAE:

Ever since the days of Epicurus there have been philosophers who believed that the existence of evil constitutes a formidable objection to theistic belief and a powerful argument for atheism. We might call those who urge this argument 'natural atheologians'; just as the natural theologian offers arguments for the existence of God, or for the rational propriety of theistic belief, so the natural atheologian offers arguments for the non-existence of God, or for the rational impropriety of theistic belief. (p. 1)

"Natural atheologians" are here portrayed as offering arguments for God's non-existence, just as natural theologians offer arguments for God's existence. So far, so fair. But we then for some reason move right into what these folks are trying to show about the rational status of some beliefs, and there, while the natural theologians are construed as trying to show the "rational propriety" of their own belief (that God does exist), the "natural atheologians" are portrayed, not as defending the rationality of their own belief (that God does not exist), but as contending for the "rational impropriety" of their opponent's belief.

I don't think this difference in construals can be justified by a claim that it's only those on the atheological side of this debate who have taken swipes at their opponents' rationality. One thinks immediately here of what is perhaps the best known bit of natural theology, where St. Anselm, kicking off the presentation of his famous ontological argument, approvingly quotes the Psalmist, saying, "The fool says in his heart that there is no God."

Plantinga thinks that the old-style, "deductive" argument from evil, if it worked at all, would work in a way that would relieve its backers from having to consider arguments for God's existence: "Herein lies the beauty of the deductive argument from evil (from the atheological point of view); if that argument could only be made to work the putative favorable evidence could be gracefully ignored" (PAE, pp. 3-4). Perhaps some arguments in favor of God's existence are of that type as well, then? If so, that might at least explain why their backers aren't being asked to address all the arguments against their conclusion before their arguments can be credited as good. However, if I have a good sense of the type in question, this would apply to only a few theistic arguments, and not really the best ones. At any rate, I don't buy these different treatments of the two types of arguments, anyway. If the "deductive" argument from evil did work, it would derive God's non-existence from reports about the evils in our world together with analyses of the meaning of the attributes being assigned to God (omnipotence, omniscience, perfect goodness). But even if we did in fact happen to hit upon correct analyses of those

Here's the problem: That there is a good (or successful, or cogent, or powerful, or very powerful, or even exceedingly powerful) argument for some conclusion, by itself, doesn't *show* anything about the rationality of believing that conclusion, or the irrationality of failing to believe that conclusion, or of believing the opposite of that conclusion. For all that, there could possibly be other, even more powerful arguments, or other powerful rational considerations, pointing against your conclusion. And how could you rule that out? I suppose the most natural way to get from the existence of even a very powerful argument for some claim to the conclusion that it's rational to believe that claim, and/or irrational not to believe it, and/or to believe that the claim is false, would appeal to there being no arguments close to as powerful that point away from your claim (and good luck with actually *showing* that, no matter how right you are about it), together with *some* principle connecting rational belief with arguments (and good luck with formulating that, too). But what a tough argumentative row to hoe!

But then, since all of the above is so general, and holds even for extremely powerful arguments, merely making the case that some opponent hasn't by means of their argument shown the contrary view to be irrational, nor met the other unreasonable burdens in question, doesn't show that their argument isn't a good one—nor even that it isn't extremely powerful.

Nor even that it isn't in fact (despite your opponent's failure to *show* this) so powerful as to render belief in its conclusion rational and disbelief irrational for anyone who encounters their argument.

Of course, I don't find our Simple argument *that* powerful. But I do find it very powerful indeed, and am finding little in Plantinga's attack to dissuade me from that evaluation. (And those who find it even more powerful than I do should also find nothing in Plantinga's attack to dissuade them from their evaluation.) I'm just finding reasons for thinking that atheologians haven't achieved by means of their argument something that I think is never achieved by *any* philosophical arguments for controversial conclusions.

attributes, and demonstrated God's non-existence from our world's evils along with those correct analyses, I would still think that, being human, and so fallible on matters of analysis of terms, as well as on other matters, we'd still have to harbor some worry that we got those analyses wrong, and so would be open to claims that we should consider arguments for conclusions opposite of our own to weigh them against each other, even if our own argument was of the "deductive" variety, especially if we were throwing around charges of irrationality.

16. Distinguishing Probabilistic Cases from Evil from the Trash Talking often Associated with Them

“But wait! Why wonder how atheologians ever got into the business of charging theists with irrationality? Haven’t they—or at least many of them—put themselves in that business? Aren’t they an aggressive bunch? Haven’t many of them in fact made such charges?⁶² Didn’t we ourselves see Unger totally go there, with great gusto? And in his ‘bad cop’ moments, didn’t Rowe at least strongly insinuate such a charge? So why then say that Plantinga has saddled probabilistic-argument-wielding atheologians with unreasonable burdens? Haven’t they taken those burdens on themselves? And, really, DeRose, haven’t you been insinuating such things yourself, though you hide behind the skeptics/atheologians you are defending as you do so?”

Well, yes, some actual atheologians have been aggressive in this way. Plantinga certainly thinks of them as being so—and I presume this is based on plenty of first-hand experience.

But this surely does not mean we should require that atheologians succeed in such things as showing that their opponents (all of them?) are irrational, or that all the arguments on the other side are lacking—in meeting requirements that no proponents of arguments for any

⁶² Of the places where atheologians have in print explicitly targeted the rationality of theism on the basis of the problem of evil, a passage from J.L. Mackie is worth singling out, not just because it is well-known and clear on the point in question, but also because it is often quoted by Plantinga in setting up his discussions of the problem of evil. It’s from the opening paragraph of Mackie’s paper, “Evil and Omnipotence” (*Mind* 64 (1955): pp. 200-212). Mackie opens by remarking on (what he takes to be) the failure of arguments for God’s existence, but also how the theologian can handle this failure:

The traditional arguments for the existence of God have been fairly thoroughly criticised by philosophers. But the theologian can, if he wishes, accept this criticism. He can admit that no rational proof of God's existence is possible. And he can still retain all that is essential to his position, by holding that God's existence is known in some other, non-rational way.

Then, in the passage Plantinga likes to quote, Mackie turns toward the problem of evil, and shows that he is indeed targeting the rationality of theists:

I think, however, that a more telling criticism can be made by way of the traditional problem of evil. Here it can be shown, not that religious beliefs lack rational support, but that they are positively irrational, that the several parts of the essential theological doctrine are inconsistent with one another. . . (p. 200)

Plantinga began quoting this passage in his famous paper, “The Free Will Defense” (in Max Black, ed., *Philosophy in America*, Cornell UP, 1965; pp. 204-220), p. 204. Mackie is presenting the *logical* problem of evil as the challenge to the rationality of theism, which logical problem is the focus of Plantinga’s famous 1965 paper. But in later treatments, Plantinga worked the quotation from Mackie into Plantinga’s overall narrative on which atheologians later retreated from such a logical form of the problem, but then started pressing the probabilistic or evidential argument from evil, as a new way to pursue (among other goals) the old goal of attacking the rationality of theists: See *Warranted Christian Belief*, pp. 459-464, where the Mackie passage is quoted at p. 460.

controversial philosophy can ever meet—before we can recognize them as having a good or powerful argument for their atheistic conclusion.

And as our atheologians will be quick to point out, when on offense, theists, at times, can be pretty triumphalist themselves.

To straighten this all out and sensibly evaluate the strength of the case for atheism before us, we should start by distinguishing that case itself from any associated trash-talking by the atheologian that may accompany it. (A parallel procedure is in order in evaluating pro-God arguments.)

What makes for successful or strong or powerful arguments, beyond the “proper argument,” as I’ve been calling it (the thing that can be outlined in premise-and-conclusion form), is the arguer’s ability to make a “case” for their key initial premise(s): to present the issue of whether that premise is true in a way that makes it rationally seem that it is. This is a matter of getting the audience to view the matter from the proper angle (or, more suspiciously: from an angle favorable to the arguer’s cause) in order to see that apparent truth. And it will often involve quite a bit of reasoning for the truth of that key initial premise—though not in the form of presenting deeper premises from which the premise in question follows, in which case that premise wouldn’t be an initial premise. The arguer presents their case, including what reasoning they have for that premise, and in light of that case, appeals to their audience: “Doesn’t that premise seem true?”

As we’ve seen, our arguer might also register, and perhaps quite aggressively, how the matter strikes them (as I have repeatedly described our argument from evil as “powerful”)—and perhaps even what they think of those to whom things might seem different (a particularly aggressive move I haven’t myself made, but did quote Unger as making). Or in stating just how strong they find the appearance of truth attaching to their key premise, and then how powerful they find the argument itself, they might venture opinions—often quite bold—about how that argument compares with what could or has been said on the other side of the issue. All this can be viewed as trash-talking: attempts to play up the strength of one’s own side of the debate and/or denigrate the other. It’s an interesting aspect of intellectual battles, and one’s ability to engage in it successfully often turns on how well one can subsequently back up that trash-talking—or how well one can appear to do so. It’s almost inevitable in philosophical arguments, though the degree of aggressiveness varies greatly from case to case.

And of course, it’s wise not to be cowed by an arguer’s trash-talking into an evaluation of their argument matching theirs. But the line between what I’m calling “trash-talking” and the more rational aspects of an arguer’s case is subtle. As I indicated toward the end of section 8, I

think it can be enlightening to encounter an aggressive trash-talker, and through them to get a feel for how matters look to one who finds a key appearance to an argument very compelling.

But however you feel about the potential for any positive value to trash-talking, you should not let the fact that the arguer hasn't shown various of the trash-talking claims they made in association with their argument to be true lead you to think the argument itself is weak. That would lead you to disparage all substantial arguments. It's in the nature of trash-talking that its claims haven't been duly demonstrated—and often that they'd be very hard to demonstrate.

Ultimately, dear reader, the call on how powerful is premise 1s, and then our Simple argument that it fuels, as always, is yours to make. But I urge you in making that call not to think that Plantinga has demolished such arguments. Or even that he's really done anything that should make you lower your estimation of how strong the argument is. Well, unless you actually were thinking that atheologians had shown their trash-talking claims to be true, in which case I suppose Plantinga can set you right about *that*. Otherwise, if you found the argument convincing, or strong, or powerful, you should never retract that judgment based on an attack like Plantinga's.

17. Plantinga on the Powerful Non-Argument from Horrific Evils: “Isn't It Just Apparent, Just Evident?”

Ratcheting things up a bit, I want to close this part of the book by actually enlisting (or, I suppose, conscripting) Plantinga's aid *in defense of* the argument from horrendous evils. This move may be surprising to those who know Plantinga's work (including perhaps Plantinga himself!), since he has never sounded very impressed at all by anything he's called an “argument” from evil, whether deductive or probabilistic, and has spent much energy fighting against such arguments. Nevertheless, though he would not himself categorize what he's doing in these terms, I think that in his monumental book *Warranted Christian Belief* (henceforth “WCB”),⁶³ Plantinga himself in essence makes a presentation of our argument from horrendous evils—and very forcefully. Updating my response to Plantinga's attacks on the argument to include those in his book will thus allow me to end this Part with a rousingly effective presentation of the argument itself.

⁶³ Plantinga, *Warranted Christian Belief*, Oxford University Press, 2000. Note that we have moved forward in time 21 years from PAE.

After countering some attempts at “probabilistic” arguments from evil,⁶⁴ Plantinga suggests that perhaps no *argument* is needed for the facts about the evils of our world to work against belief in God:

Perhaps the defeating power of these facts in no way depends on the existence of a good antitheistic argument (deductive, inductive, abductive, probabilistic, whatever) from the facts of evil. (WCB, p. 482)

In showing how this “defeat” might work, Plantinga starts to convey the horrors of our world that such a non-argument (as we might for now call it) might use, and in doing so, he sounds a lot like the “atheologians” we have considered, pushing the problem of horrific evils, even quoting Dostoevski’s Ivan. Plantinga sums up his tour of our often hellish world, and then enlisting the help of Eleonore Stump (whom Plantinga is quoting after the colon, below), adds to the horrific picture, to yield the most graphic snapshot of one of world’s horrors that the reader will encounter in this book:

The list of atrocities human beings commit against others is horrifying and hideous; it is also so long, so repetitious, that it is finally wearying. Occasionally, though, new depths are reached:

A young Muslim mother in Bosnia was repeatedly raped in front of her husband and father, with her baby screaming on the floor beside her. When her tormentors seemed finally tired of her, she begged permission to nurse the child. In response, one of the rapists swiftly decapitated the baby and threw the head in the mother’s lap.⁶⁵

Plantinga then very insightfully reflects on how such facts can militate against the existence of a wholly good God:

These things are absolutely horrifying; it is painful even to consider them, to bring them squarely before the mind. To introduce them into cool philosophical discussion like this is distressing and can seem inappropriate, even callous. And now the question: wouldn’t a

⁶⁴ Here is Plantinga summing up his take on arguments in the vicinity by Rowe and by Draper: He finds some nice things to say about them in terms of the issues they cause us to think about, but his estimation of the arguments’ power seems quite negative indeed: “These new arguments by Rowe and Draper are subtle and sophisticated; many deep and interesting topics come up in considering them. Upon close examination, however, they fail, and fail resoundingly. They fail to provide a defeater for theistic belief and, indeed, give the person on the fence little if any reason to prefer atheism to theism. . . . If the facts of evil really do provide a substantial challenge to Christian or theistic belief, it must be by a wholly different route,” Plantinga, *Warranted Christian Belief*, p. 481.

⁶⁵ WCB, pp. 483-484, quoting Eleonore Stump, “The Mirror of Evil,” in *God and the Philosophers: The Reconciliation of Faith and Reasons*, ed. Thomas Morris (New York: Oxford University Press, 1994), p. 239. Looking back, I find it interesting that I provided this most graphic look at a horrific evil at two removes.

rational person think, in the face of this kind of appalling evil, that there just couldn't be an omnipotent, omniscient, and wholly good person superintending our world? Perhaps he can't give a demonstration that no perfect person could permit these things; perhaps there isn't a good probabilistic or evidential atheological argument either: but so what? Isn't it just apparent, just evident that a being living up to God's reputation couldn't permit things like that? (WCB, p. 484)

Well, yes, *isn't* it "just apparent, just evident" that a wholly good God would not permit such things? That's just the kind of appeal I have construed our argument from horrific evils as being based on. (I take "evident" here to mean something in the vicinity of "very apparent.") Plantinga does studiously avoid calling this an "argument" against God's existence, but he does call it a "case" (which seems for him a broader category, which would include "arguments" or what I'm calling "proper arguments" as a sub-category) and, apparently a good one—or at least the "best" one:

Something like this, I think, is the best version of the atheological case from evil. (WCB, p. 484)

Well, that's a comparative evaluation, and the "best" needn't be very good, I suppose. But, though Plantinga has an answer to this "case" that he thinks is effective (and that we will soon examine), one definitely gets the impression in this material (and I trust that readers have gathered that impression from just the bits I have quoted) that Plantinga thinks one needn't be stupid to find it exceedingly powerful indeed.

And then, I don't see why this shouldn't mean there is a powerful *argument* here (even in the narrow sense we've been using for "proper arguments"). Can't this very "case" that Plantinga is conveying be helpfully seen as running through something like our premise, 1s, a premise that would seem intuitively powerful, given Plantinga's case, and which implies our argument's atheistic conclusion? What more would one want from a powerful argument, other than perhaps complications, to make things seem more sophisticated?

Indeed, I imagine some readers by now find the argument so compelling that they are wondering how Plantinga, ace defender of theism, manages to escape the problem he has so powerfully set up. And it turns out that the nature of the escape Plantinga resorts to⁶⁶ in WCB betrays the intuitive power of the argument. As I have agreed (see the last paragraph of section

⁶⁶ I don't mean to be insinuating that Plantinga's response is silly. It is provocative—and I think unsuccessful, given my view of what a response ought to achieve. But I say he "resorts" to it because I think that if he had a response that could be more helpful to those who don't find themselves already so completely certain of God's existence, and who also don't buy his contentious anthropological and methodological machinery, according to which, in an important way, we ought to be completely certain of God's existence, he would have (also) presented that response.

2), Plantinga notes that those who are somehow rationally *absolutely certain* of the existence of a wholly good God can reasonably react to the evils of this world by retaining their confident belief in such a God, and supposing that God has a reason for allowing those evils that completely escapes them (WCB, p. 485). But, moving beyond this correct point, Plantinga thinks that the just described reaction is in an important way *the* rational response to our problem, since he thinks that a human whose cognitive faculties were functioning properly *would* be completely certain of the existence of such a God—indeed, they would be “as convinced of God’s existence as of her own” (WCB, p. 485)!—and it is only because of a (very common indeed) cognitive disorder that any of us are ever less than absolutely certain of God’s existence.

Plantinga’s response rests on some fascinating theological anthropology, on which our cognitive faculties were designed to have a sense of God (“what [John] Calvin calls a *sensus divinitatis* or sense of divinity” (WCB, p. 172)) so lively that we would be completely certain of God’s existence, and on which cognitive malfunctions that are the result of human sin are what stand in the way of our enjoying such certainty. Also essential is some highly provocative epistemology and (more to the point) philosophical methodology on which, for Plantinga’s purposes, the success of the “case” from horrific evils can be measured by how one would rationally respond to it if one’s cognitive faculties were functioning in their non-fallen, proper way.

Without contending with that elaborate theological and methodological framework Plantinga has constructed, and from within which he answers the problem of horrific evil in WCB, we will instead seek what rational relief we might find that does not so evade the problem by appealing to any normative absolute certainty of the existence of a wholly good God. This is largely because engaging that fortress would be a huge undertaking.⁶⁷ But it is worth noting that if Plantinga’s maneuver does work against our argument, it would similarly work against *any* case against God, no matter how powerful, since if you were absolutely certain that a good God exists, you could reasonably judge that any case against God must have some mistake somewhere. If then cases against God are to be judged by how we could rationally react to them if we were so absolutely certain of God’s existence, they’re all doomed

⁶⁷ I will note this, however (in addition to what I note in the text above), remarkable only given a certain aspect of the climate of current philosophy. One hears lots of complaints these days about the use of “ideal theory” in philosophy, these complaints somewhat descended from Charles W. Mills’s “‘Ideal Theory’ as Ideology” (*Hypatia* 20 (2005): 165-184). I must say that I tend to be skeptical of such complaints. I suppose I tend to think there are problems with the targets complained about that are not a thousand miles away from their employment of ideal theory, but I tend to think the real problem is not the use of ideal theory itself. But here in Plantinga’s use of the just described “highly provocative” bit of philosophical methodology we may have a real example of ideal theory truly run amok.

from the get-go, no matter how powerful they might seem—or, as I would say, how powerful they might be.⁶⁸

Part Two: Solutions: Theodicy, Anti-Theodicy, and Goalposts

18. Theodicy, in the Form of a Credible Account of What God's Reasons for Allowing the Horrific Evils of Our World Might Be

Well, then, how *will* we here seek relief from our problem? By means of a very natural, and very common, if also much maligned, type of response to it: by attempting to give a credible account of what God's reasons for allowing the evils of our world might be.

This is to attempt to give a “theodicy,” in a philosophically common use of that term that I have been utilizing already in this book, and will continue to use. But this sense is a bit narrower than what seems to be the term's most common use outside of philosophical theology. “Theodicy” ordinarily means something quite general, along the lines of “a defense of God's goodness and omnipotence in view of the existence of evil,” as many dictionaries have it, more or less. Our more narrow use differs from this general one in that ours covers only such

⁶⁸ Indeed, more generally, for any claim you might seek to defend, whether that claim involves God's existence or not, if you're willing to assert that humans are designed to be absolutely certain that your claim is right, and that it's only through some cognitive malfunction that we fail to have such certainty, and if you are ready to evaluate the success of cases against your claim by reference to how we would rationally respond to them if we were functioning in the way you contend we were designed to, you are set to so defend that claim, whatever it is, against any case against it, no matter how powerful. (Well, I suppose that's just so long as the case against you isn't strong enough to produce absolute certainty that you are wrong—and even then, it isn't clear what should result from the resulting hypothetical clash of opposing absolute certainties. Perhaps you should claim that no case against you could be that strong?) But it should be acknowledged that a key move in this Plantinga-style defense—namely, the assertion that humans are designed to be completely certain of the truth of the claim one is defending—might be considerably less plausible when employed in defense of many claims than it is in defense of the existence of God. So, though this type of defense can in a way be applied very broadly, that doesn't mean it can be applied *just as well* in defense of any claim whatsoever as it is in Plantinga's own actual use of it. Evaluating the plausibility of this aspect of Plantinga's defensive maneuver would be “contending with that elaborate theological and methodological framework Plantinga has constructed,” and so is just what I'm avoiding here. But interested readers should consult *Warranted Christian Belief*, where Plantinga most fully develops all this.

defenses that proceed in a particular way: by proposing potential *reasons* God might have for causing or allowing evils. Thus, those who defend belief in God in the face of the evils of our world, and so are offering a “theodicy” in the normal, quite general sense of the word, but who eschew any speculation about God’s reasons or possible reasons, and instead defend theism from the problem of evil in some other way, are rejecting the project of “theodicy” in the more narrow sense I will be using, even as they engage in the project in the more general sense of that word. These rejecters of our more narrow project of theodicy may hold that God *has* reasons for allowing the evils of our world (though some may not, instead holding God is somehow beyond reasons⁶⁹); they then just reject the project of trying to discern what those reasons are or even might be. “Skeptical theism,” which we have briefly looked at, is one way—or I suppose, a family of ways—of engaging in the project of “theodicy” in the broad sense without attempting it in our narrower sense.⁷⁰ By arguing that our inability, or even total inability, to come up with good reasons God might have for not preventing evils does not damage the credentials of our belief in God, the skeptical theist is very much defending that belief in the face of the problem of evil, while very decidedly refusing to provide any potential God-justifying reasons for allowing our world’s horror. In fact, a key claim of much of what flies under the banner of “skeptical theism,” put in our terminology, is: You don’t have to engage in theodicy in the narrow sense to produce a successful theodicy in the broader, ordinary sense.

Moving beyond terminology: *Why* seek relief from our problem in the form of a theodicy, in our more narrow sense? Why propose reasons God might have for allowing the evils of our world? Well, the challenge that horrific evils pose for theistic belief presents itself to some worriers in the first place in a way that explicitly involves the notion of God’s possible reasons, in the form of thoughts like: “There couldn’t possibly be a good reason for God to allow *that!*” Possible reasons are of course just what such worries cry out for. But even if one’s worry is instead guided by a more general thought like “There’s no way that a world governed by a wholly good God would contain *that!*” (which doesn’t explicitly mention God’s reasons, or lack thereof), a cogent account of why God might allow the horrors in question, if such an account could be given, would still seem a good way to show how those evils might well occur despite the existence of a perfectly good God. Imagine, abstractly (and for a brief moment,

⁶⁹ Some may hold that God is also beyond goodness, but they are then moving beyond how I am here using the term “God.”

⁷⁰ The line between those approaches that do and those that do not propose potential reasons is not a sharp one. Some skeptical theists will allow that there are some goods or others that are tied up with God’s allowing of even the most horrible evils, and that God allows the evils for the sake of those goods, whatever they are. This is, at least very vaguely, to propose a reason for God’s allowing of the evils – and it will indeed thereby draw the ire of *some* anti-theodacists. But as I employ our narrow use of “theodicy,” one has to be saying something a bit more about just which goods are in question, to count as proposing a potential reason God might have for allowing evils and to thereby count as offering a “theodicy” in our narrow use of the term.

don't worry about how abstract this imagining is), learning of a reason God might have for allowing even the most challenging evils you know of, and coming to see that, yes, that really would be a good reason for allowing them. Would this not put a damper on the thought that there couldn't possibly be such evils in a world containing a perfectly good God? Of course (to now allow the worry about how abstract our imagining exercise was), one who finds our problem so powerful that they think it's generally hopeless to try to answer it will also likely think it's particularly hopeless to try to come up with a good reason God might have for allowing our world's evils. But pessimists might still agree that a good account of God's reasons *would* help, if only (*per impossibile*, they might want to add) it could be given. So, for those who have not succumbed to hopelessness, it is natural for a defense of standard theism to concern itself with what God's reasons might be—for a “theodicy” in the more general sense to take the form of a “theodicy” in our more narrow use.

19. Anti-Theodicy and Some Explanation of Our Goalposts

But while natural, attempts at such narrow theodicy are often disparaged—with the attacks coming from wildly different directions. The attempt to specify reasons for allowing truly horrific suffering can seem appalling and oblivious to how bad our world's worst evils are. Some who find it so are atheists we've just been considering who judge the argument from horrific evils to be overwhelmingly powerful, and find theodacists' attempted responses to it to be both doomed and disgusting. But others are themselves theists who suppose God's reasons are beyond us and who think that for us to speculate about those reasons is both presumptuous and inevitably insensitive to the horrific nature of the evils being treated, in addition to being unpromising in the extreme. And still others are theists who suppose God is beyond having or needing reasons, and think that attempts to speculate as to what God's reasons might be are for that reason destined to end in inglorious failure. But whatever is the direction from which it comes, these questions can be very pointed: How can the reasons we can possibly come up with even come close to justifying the allowance of the horrors of our world, and how can we think we know the mind of God?

You will be in a better position to assess the challenges of and around “anti-theodicy”—as the movement of those opposed to the project of giving and discussing theodicies is aptly called—after we have considered, and even put forward, proposed potential reasons God might have for allowing horrific evils. But as we now turn to briefly comment on the goalposts for our inquiry, which are already set as providing a credible account of such divine reasons, a

few things can be said here at the outset, to perhaps reduce the sense of hopelessness, at least a bit.

First, rest assured that I won't be claiming to be articulating what I think I know to be God's actual reasons for allowing horrific evils. On the other hand, we will steer a middle course, and not merely seek a *logically possible* account of what the reasons of a wholly good God *might have been*. What I'm aiming for is a *credible* account of God's reasons *might be*.

On these matters, I align fairly closely with the wise moderate approach laid out by David Lewis, one of the great philosophers of the second half of the twentieth century, who made his name as a metaphysician, but who happily turned some of his attention to the problem of evil toward the end of his life, and especially in his paper "Evil for Freedom's Sake?".⁷¹ Lewis was into evaluating "theodicies," but in light of influential work by Plantinga, Lewis had to clarify that he was not interested in "theodicies" in the technical, special use of that term that Plantinga introduced and made prominent (at least in some philosophical circles), and is for us yet a third, most narrow of all, use of that term (but one which will only appear briefly now, as an example of what we won't be attempting), on which those hawking "theodicies" really do claim to know God's actual reasons for allowing evils:

Alvin Plantinga, our foremost modern authority on free-will theodicy, would recoil from that name for his subject. He has taught us to distinguish 'theodicy' from 'defence'. 'Theodicy', for Plantinga, means an audacious claim to know the truth about why God permits evil. And not just a trivial bit of the truth—God permits evil for the sake of some good or other—but something fairly substantive and detailed. One who claims to know God's mind so well (especially if he claims to know without benefit of revelation) will seem both foolhardy and impudent.⁷²

After rejecting the project of attempting such an "audacious" (in Lewis's term) "theodicy," Plantinga retreated to seeking a mere "defense", which in Plantinga's special terminology, is just a logically possible account of what a perfectly good God's reasons for allowing evils might have been.⁷³ Lewis sees this as an over-retreat, and, in a passage that I think provides a

⁷¹ David Lewis, "Evil for Freedom's Sake?", *Philosophical Papers* 22 (1993): 149-172. See also Lewis's paper specifically on the problem of hell: "Divine Evil," in Louise M. Antony, ed., *Philosophers without Gods: Meditations on Atheism and the Secular Life* (Oxford UP, 2007). The latter paper was actually put together by Philip Kitcher from an outline of it that Lewis had sent him, based on Kitcher's memory of discussions of the outline with Lewis, and published after Lewis's death. I think anyone who had discussed these issues with Lewis, as I was privileged to on a couple of occasions, can hear Lewis's distinctive voice coming through Kitcher's fine work on that project.

⁷² David Lewis, "Evil for Freedom's Sake?", *Philosophical Papers* 22 (1993): 149-172; p. 151.

⁷³ Here I present the force of logical "defenses" a bit differently than the defenders themselves tend to. They will tend to speak of the reasons God *might have*, or that God *possibly does have*, and to say that

wonderful roadmap for how to approach our problem (and a roadmap I will largely follow here), proposes a better, middle course for the theist (and specifically, here, the Christian):

Defence is too easy; knowing God's mind is too hard. I think the topic worth pursuing falls in between, and has no place in Plantinga's scheme of theodicy versus defence. *Pace* Plantinga, I'll call that topic 'theodicy', but I don't mean the know-it-all theodicy that he wisely disowns. Rather I mean tentative theodicy, even speculative theodicy. The Christian needn't hope to end by knowing for sure why God permits evil. But he can hope to advance from a predicament of not having a clue to a predicament of indecision between several not-too-unbelievable hypotheses (maybe still including the hypothesis: 'none of the above'). The job is to devise hypotheses that are at least somewhat plausible, at least to the Christian, and to find considerations that make them more plausible or less.⁷⁴

Though I take myself to be roughly following the path Lewis pointed to, I should add that Plantinga-style "defense" is only "too easy" if you are dealing with the (or some) "evidential" problem of evil. For "logical" forms of the problem (which Plantinga was engaging when he was being "defensive"; recall [section 13](#) for the distinction between the two forms of the problem), a Plantinga-style "defense" is just what is called for: If you are championing the logical possibility that a wholly good God might have existed despite the existence of evil (or of evil in a certain amount and/or of a certain kind), coming up with a good reason some God logically might have had for allowing evil (or for allowing evils in the amount and/or of the kind in question) is the right strategy. But for our much more gripping "evidential" form of the problem, more is needed—though attempting *too much* more still seems unwise.

In seeking a theodicy in the form of *a credible account of what God's reasons for allowing evils, and in particular, horrific evils, might be*, I am following Lewis's favored middle option, roughly as he describes it. My aspiration for a "credible" account may be setting the bar a bit higher than Lewis's proposed goal of hypotheses that are "not-too-unbelievable," though perhaps it's quite closely aligned to his other description of the goal as hypotheses that are "at least somewhat plausible."

their aim is to show in the face of evil that God *might* exist, or that it is possible that God does exist. I believe that they have to misuse language a bit to put their goals in these terms, and that what they really should say is that they seek what reasons God, or a god, *might have had*, or *possibly could have had*, or what God's reasons *might have been*, and that if their defenses work, they'd be showing something like that, despite evils, God *might have existed*, or that it's possible that, despite evils, God *should have existed*. Yes, it's subtle, and even their goals, properly expressed, if attained, would answer logical problems of evils. I fear, though, that their way of putting things make their defenses sound more germane to answering the problem of evil generally and in its intuitively most powerful forms (and particularly evidentialist forms of it) than those defenses actually are.

⁷⁴ "Evil for Freedom's Sake?", p. 152.

But *why* set that as the goal? We've already seen in general terms why locating credible hypotheses as to God's reasons for allowing horrific evils would help here. If a credible reason could be given, and especially if it were very credible indeed, well, then, one supposes, it would rationally become *less apparent* that the horrific evils of our world wouldn't exist if there were a wholly good God. But why set the bar at "credible"—or at "somewhat plausible," to the extent that that's different? And just how credible, how plausible, and just how detailed, would it have to be to help? And credible to whom? And how much relief would it actually provide? All good questions—that for the most part I can't give definite answers to—but for what I hope is a good reason.

Critical readers will be especially alert, and perhaps even suspicious, as I set the goalposts for this inquiry. The worry here would be that I started with an idea of what I could achieve, and then, in a bit of reverse-engineering, am now setting my goalposts to that point, with no reason for thinking that just that is what is needed for success other than that I thought that's the standard I could meet, and I will then declare victory as I make the case that I have met it. And the vagueness one may sense in "credible" (or "plausible") may add to the suspicions of the wise, as I could try to get that squishy term construed in a stronger, more robust way to motivate the thought my goalposts are plausible as ones that it would be worthy to reach, but then angle for a weaker reading as I argued that I have reached them.

But don't worry: that's not how this is going to go down. I suppose there is something to the charge of reverse engineering, as my aspirations are no doubt as low as they are because they indeed have been disciplined by my sense of what can be attained. But there will be no declaring of victory. Readers will have gathered by now that, though I find our problem a gripping one, I must have found enough relief to continue in my theistic ways. But while I have in some sense thereby signaled that I think the relief can be sufficient for that response, that's not something I'm in any good sense declaring. I'm sure some readers will think this response is a mistake brought on by my failure to really face up to the power of our problem, and while I will disagree with them, their judgment won't be in conflict with any real declarations I'm making. But what I will be about is, in Lewis's words, attempting to "devise hypotheses that are at least somewhat plausible." How plausible, how credible? As plausible as they can be made. The more credible, the better. I'm looking for as much relief as I can find. I'm really not setting goalposts (or at least not very definite ones) so much as trying to get as far down the field as I can. While I may say things that betray my sense of how far down the field we've come, this will all be in the spirit of realizing that those who journey with me will make their own judgments about how far we've gotten, and whether that's far enough for whatever purpose. When I reject some alternative as not good enough, you can perhaps best understand that as a judgment that we can and will do better, at least in my judgment.

Picking up again on some of the material in that second Lewis quotation indented above (our “roadmap”), I should add that if such a credible hypothesis could be found, the way that would help, in my view, is largely by making the “none of the above” possibility that Lewis mentions more reasonable: If you’re in a state where you’re inclined to declare “There’s just *no way* a wholly good God would allow *that!*”, it seems nothing would reasonably loosen you up to the possibility that God might have *some* good reason for allowing *that* that you cannot discern as well as encountering what seems a credible account (that you are able to discern) of what such a reason might be. This is an important way in which one does not have to think one is discerning the actual thought of God in order to find some relief in a proposal as to what God’s reasons *might* be or might include.

And, to pick up on another issue raised the above quotations from Lewis: In various ways, the account won’t be very detailed. Even if it were right, it won’t be complete enough to be able to explain why, for instance, one individual, rather than another, suffered horrifically. Though I will seek to explain why our world might be such as to subject some of us or other to horrific suffering, and why we might not be able to explain or predict just on whom the worst horrors will fall. And indeed I will propose how it could be that we not only can’t see the reason why this person rather than that suffered horrifically, but why, in an important way, there might not be a reason, known even to God, for that.

On the matter Lewis raises of just *to whom* our account should seek to be credible, I can quickly say that I do hope that it will be credible to you, the reader, and, failing that, to your idea of a “reasonable person”: Even if our proposed reasons cannot seem credible to you, I still hope (realizing this hope won’t be realized in all cases) to allow you to see how a reasonable person could find them so.

20. God, Freedom, and Immortality: Theism as Part of a Package

Lewis, recall (second of the two quotations indented in the previous section), describes the task of worthwhile theodicy as formulating hypotheses that are “at least somewhat plausible, at least *to the Christian*, and to find considerations that make them more plausible or less.” This focus on how the Christian would find things may make the task sound like an evaluative and a defensive one, executed by or on behalf of the Christian (or theist, given our broader target): When we are evaluating X’s belief for rationality, or something like that, what seems most directly at issue is how it fits with X’s general picture of the world. (Perhaps it’s a bonus if the

hypotheses also make sense to others; hence the “at least”?) But Lewis makes it clear that he has his eyes on potential persuasion, and not just evaluation.⁷⁵

This focus on what is plausible to the Christian, or theist, in our case, when aimed at persuasion, may result from the recognition that theism often comes as part of a package of philosophical views that tend to go together. Most importantly for a discussion of the problem of evil, theists tend to accept life after death and Libertarianism—the view that we are free, and that free will is incompatible with determinism. We will discuss the connections of both of these with theism and the problem of evil as our discussion unfolds, but it is the theist’s Libertarianism that draws Lewis’s attention. And while most philosophers, and also most of them who specialize in the most relevant areas of philosophy, tend to be compatibilists, like

⁷⁵ The following immediately follows the quotation just given and shows that Lewis sees this as part of the project of attempted persuasion: “Robert M. Adams has written that ‘the atheological program . . . need not be one of rational coercion. It might be a more modest project of rational persuasion, intended not to coerce but to attract the minds of theists and agnostics, or perhaps to shore up the unbelief of atheists.’ Right; and the same, *mutatis mutandis*, goes for theodicy.” —Lewis, “Evil for Freedom’s Sake?”, p. 152; quoting Robert M. Adams, “Plantinga on the Problem of Evil,” J.E. Tombererlin and P. van Inwagen, eds., Alvin Plantinga (D. Reidel Publishing Company, 1985), pp. 225-255; at p. 240. The ellipses in the quotation from Adams are as Lewis has them. As one might guess from the title of his paper, and from the book it appears in, Adams is writing about Plantinga, and how Plantinga is construing what the atheologian is attempting, and so is covering the territory we discussed in Sects. 14-16, above.

Lewis,⁷⁶ Libertarianism, which combines incompatibilism with the claim that we sometimes do act freely, is correlated among philosophers with theism.⁷⁷

The Libertarian's incompatibilism makes a key list of points that, in his set-up, Lewis announces he will "suspend disbelief" on (p. 150) and grant for purposes of discussion: "I shall not make heavy weather—well, not for long—of assuming incompatibilism. . ."⁷⁸ As we'll see, Lewis relates what he takes to be the best arguments for both compatibilism and incompatibilism, and admits the power of the argument for the latter—even as he clearly himself sides with the former. But Lewis's compatibilism doesn't mean that a compatibilist variety of theism is the type that would be most attractive to him. I suspect that granting Libertarianism was actually for Lewis a way to explore the form of Christianity that had the best shot at winning him over. I had the chance to discuss these matters with Lewis a couple of times, and if he were still with us, and I got another chance to talk with him about these matters (updating the discussion with things I've learned, or at least take myself to have

⁷⁶Results from an updated PhilPapers Survey are now in a note in **Part Six**. This will note and the next one will have to be changed, probably to just point ahead to the new note about the new results. The 2009 PhilPapers Surveys <<http://philpapers.org/surveys/>> results for "Free will: compatibilism, Libertarianism, or no free will?" strongly favored compatibilism over Libertarianism, the variety of incompatibilism needed to fuel the Free Will Defense, both among philosophers in the "target faculty," generally, and also among those whose specialization were in the relevant areas:

Among "All respondents" in the "target faculty":

Accept or lean toward: compatibilism 550/931 (59.1%)

Other 139/931 (14.9%)

Accept or lean toward: Libertarianism 128/931 (13.7%)

Accept or lean toward: no free will 114/931 (12.2%)

Among specialists in "metaphysics" in the "target faculty":

Accept or lean toward: compatibilism 130/234 (55.6%)

Accept or lean toward: Libertarianism 50/234 (21.4%)

Accept or lean toward: no free will 31/234 (13.2%)

Other 23/234 (9.8%)

Among specialists in "Philosophy of Action" in the "target faculty":

Accept or lean toward: compatibilism 23/43 (53.5%)

Accept or lean toward: Libertarianism 8/43 (18.6%)

Other 7/43 (16.3%)

Accept or lean toward: no free will 5/43 (11.6%)

⁷⁷ Indeed, the correlation between "God:theism" and "Free will:Libertarianism" is one of the strongest correlations between pairs of answers in the PhilPapers Surveys mentioned in the previous note; see: <http://philpapers.org/surveys/linear_most.pl>.

⁷⁸ Lewis, "Evil for Freedom's Sake?", p. 151. Lewis continues: ". . . or even of assuming the Molinist doctrine of middle knowledge." I turns out that I will make heavier weather over middle knowledge, as I think the free will defense is best, and most intuitively, executed without that doctrine, and I will explain in Part Five why I think that—though I will also, as I move forward, explain how the free will defense can proceed if Molinism is accepted.

learned, since the previous ones), I would certainly take promoting Libertarianism as part of my best hope of making a theist out of him. But I should quickly point out the comparative nature of “best hope,” “best shot,” and “most attractive”: Theism could not in any form be very attractive to Lewis, and the chances of his adopting it in any form were close to zero. And my imagined discussion with him, like the actual ones I had years ago, would not involve any real hope of either winning the other over. Rather, finding the version of my view that would have the best chance with him, even if this best chance were extremely slim, seemed to function as a way to have some kind of meeting of minds, even in the face of pretty-much-inevitably continued disagreement.

A similar approach may make sense to many readers, who may be, like Lewis, compatibilists about human free will, or who may believe that we don’t have free will—a position not very popular among philosophers (see note 76), but more so among various groups of non-philosophers. I will in Part Six give reasons for finding Libertarianism credible, and this may promote a positive evaluation of the rationality of the free will theist’s stance. But it may also render a Libertarian version of theism the variety that is most attractive to you, even if you are not yourself a Libertarian.

Part Three: The Basic Free Will Defense and the Potentially Immense Importance of Libertarian Freedom

21. The Free Will Defense: The Theist’s Best Hope, Only Hope?

We were largely following the no-nonsense lead of Peter van Inwagen when we formulated our argument from horrific suffering in section 4. Van Inwagen is equally blunt about where the theist might profitably turn for any hope of relief from this problem:

There seems to me to be only one defense that has any hope of succeeding, and that is the so-called free-will defense.⁷⁹

⁷⁹ *The Problem of Evil*, p. 70. “Seems to me,” of course, softens things (as compared with, say, replacing that phrase with “is”); it’s the “any hope” that I’m finding so blunt. And van Inwagen makes this strong statement despite having a low bar for the “success” of a defense. If you would like to compare his goalposts with ours, here is his explanation of the type of “defense” he seeks, which immediately

The “free will defense” is indeed usually so-called, and we will follow the crowd here. And you have just seen “defense” used in schemes where defenses are contrasted with “theodicies,” as humbler alternatives to their bolder, more know-it-all big siblings. But the free will defense is a “theodicy” in our more modest use of that term: it gives a reason God might have for allowing evils. In brief slogan form: Evil is the price God pays for the good of human free action.

I don’t know how many who have studied our problem would agree with van Inwagen about this theodicy being the theist’s only hope. Some hold out hope for other escapes; others find no hope at all, anywhere. But I do think many indeed would count it the theist’s *best* hope, and among ordinary theists, or at least many ordinary Christians (the kind of ordinary theist I am most familiar with), as well as professional worriers, it is the go-to defense that immediately springs to mind, and then to mouth or keyboard, when trying to explain how our world might contain its evils despite being governed by the completely good God of standard theism—or, for those convinced that our world couldn’t then contain its actual evils, the Free Will Defense is taken as the most serious threat to that conviction.

And as I indicated in this book’s opening section, I will contend that the theist’s go-to defense should be flipped to something else. But the free will defense ain’t the champ for nothing. The power of this theodicy must be reckoned with, even as we examine its limitations and seek to replace it—especially since the best replacement will have to incorporate key aspect of the Free Will Defense. Van Inwagen himself judges the “simple” free will defense a failure, and decisively so, but the hope put in it to yet be rewarded by the “expanded” free will defense that he raises out of the ashes and that he thinks proves more successful. We will again follow van Inwagen, at least in basic direction here, finding the “basic” (which word functions for me much as “simple” does for van Inwagen) free will defense to fail, and for reasons somewhat similar to those cited by van Inwagen (which are widely recognized by others, too), before looking for a related account that might fare better. However, we will find that we have to move beyond the boundaries of what can be counted as even an expanded “free will defense,” as we scramble to regain hope for theism. This all means we need to account for both the power and limits of the popular free will defense—and the limits of it in two senses: both in terms of where and how it fails, and also in terms of what are the boundaries to what can be counted as a “free will defense” as we seek to replace the free will defense with a better

follows the sentence I have quoted: “In saying this, I place myself in a long tradition that goes back at least to St Augustine, although I do not propose, like many in that tradition, to offer a theodicy. I do not claim to know that free will plays any central part in God’s reasons for allowing the existence of evil. I employ the free-will defense as just precisely a defense, a story that includes both God and evil and, given that there is a God, is true for all anyone knows” (p. 70).

account—though one on which human free will plays a crucial role—of why God allows the evils of our world.

22. The Famous, Intuitive Free Will Defense: The Basic Idea

So, just what is this famous free will defense? And where can we find the giants of theistic philosophy making this defense through the centuries? Our champion proves elusive, as neither of those questions turns out to be entirely easy to answer. As a young teacher, when I would mention in class, or when I was asked by students about, say, the cosmological argument for God's existence, or the problem of evil, I would know just where to send the curious for some standard presentations of these things, classic and current. But I was embarrassed not to know where to send them for a standard presentation of the free will defense. (This was before van Inwagen's book, *The Problem of Evil*, came out, and Plantinga's free will defense was not suitable, at least for me, for a reason about to emerge in this section.) Now I don't feel so embarrassed about not having known. I was relieved to find that Plantinga, the philosopher now most closely associated with the free will defense, admitted similar problems. In a lovely "Self-Profile," Plantinga conveys his state as he began work on his own famous free will defense:

I knew in a vague sort of way that theistic thinkers had appealed to free will in responding to the problem of evil, but I had never come across an explicit free will defense. In fact they aren't easy to find.⁸⁰

"Aren't easy" holds out some hope: Perhaps Plantinga is about to tell us about some places to go that were tough to find? But the suggestion we get, while it does concern a giant of Western philosophical theology, starts out in a way that—to put it mildly—doesn't promise much by way of a definitive treatment: "Augustine sometimes seems to give something like a free will defense. . .".⁸¹ I had myself somehow associated the free will defense with Augustine, and had

⁸⁰ Plantinga, "Self-Profile," in J.E. Tomberlin and P. van Inwagen, eds., *Alvin Plantinga* (D. Reidel Publishing Company, 1985), pp. 3-100; at p. 41.

⁸¹ Plantinga immediately continues (after a semi-colon): "there are passages in which it is reasonably plausible to interpret him as holding that it wasn't within God's power, despite his omnipotence, to create free creatures and also cause them to exercise their freedom in such a way that they do only what is right. Augustine isn't entirely clear on this matter, however. He also displays a considerable list towards theological determinism, and towards theological compatibilism: the view that human freedom and divine determinism are compatible" (p. 41). Here Plantinga indicates that a key feature of anything he'd be inclined to call a "Free Will Defense," is that it's fueled by an incompatibilist account of free will.

hoped to find it in his writings (I remember especially hoping to find it in *On Free Will*), but, like Plantinga, could find only vague hints of it that sat in Augustine's thought along with some substantial counter-hints, that seemed to indicate, at least to my admittedly non-expert eyes, that he couldn't have been pushing, at least consistently, anything like what I think of as the free will defense.⁸² Perhaps Plantinga's attempt to find a predecessor was doomed to failure, at least if he was looking for someone who had presented a free will defense that was at all worked out: Michael Murray and Michael Rea now tell us that the Free Will Defense "was first developed in detail by Alvin Plantinga."⁸³

But not to worry! The basic idea is simple and obvious, Plantinga assures us:

The central idea of the Free Will Defense is simplicity itself and has occurred to nearly every thoughtful theist (by the time I was seventeen I had devoted a fair amount of thought to it). (p. 41)

But here's the problem: Plantinga gets it wrong! Well, at least his presentation seems a bit off-target to me in a key respect, if he's trying to capture the idea that I so naturally mulled over as a teenager (and since), and found so intuitive. So I will here give the basic idea of the FWD as I think this intuitively powerful defense, that seems elusively somehow in the air, is best understood, using key elements of Plantinga's presentation of the basic idea where we are

⁸² Like Plantinga (see previous note), I was disturbed (for the purpose of finding a free will defense) by the suggestions of theological determinism I found in Augustine. But I recall thinking I faced an even more fundamental problem: Augustine seemed to be primarily addressing a different problem of evil than the one that most challenged me (and that seems to have most challenged Plantinga), and against which I thought of the free will defense as being directed. Augustine seemed concerned with something like what I came to think of as "the medieval problem of evil," as this problem prominently came to occupy later medievals as well. This is (roughly) the problem of accounting for where evil could have come from, given that (as those who suffer from this problem are *prima facie* committed), in some good sense, (a) God is the cause of everything, yet (b) God cannot be the cause of evil. To this problem, accounts of the metaphysical nature of evil (e.g., that it is a privation of good, or that it is a limitation or absence of being) that can be used with an understanding of, or perhaps a limited reformulation of, (a), to get evil outside the scope of what God is the cause of, seem just what is needed—and thus Augustine seemed very occupied with views on the nature of evil that could help in this way. His thoughts about human free will seem designed to play a key role in the kind of account of the origin of evil that could solve this medieval problem. By contrast, I thought of the free will defense as potentially providing help with the (no doubt deeply related) question (whose answer would no doubt be somehow deeply related to—but distinct from—the answer to the Augustine's question): Why wouldn't a good God have prevented our world's evils?

⁸³ Michael J. Murray and Michael C. Rea, *An Introduction to the Philosophy of Religion* (Cambridge University Press, 2008), p. 164. More expansively: "This argument, known as the Free Will Defense, was first developed in detail by Alvin Plantinga. The argument is meant to show that there are, for all we know, some goods (like the good of free choice) which even God cannot bring about without also allowing certain evils (specifically, morally evil choices) to occur."

seeing similarly enough to allow me to do so. In Part Four, I will explain Plantinga's way of doing things. As we then move forward, I'll use notes to explain how some issues may work out differently on Plantinga's way of running the Defense.

The central idea of the FWD really is simple, and in its way (a way that coexists with its containing extremely controversial bits), it is indeed quite intuitive. Moving a bit beyond mere slogan, we start with a basic statement of what we can think of as the Defense's six key elements (the quotations mark spots at which I am able to use Plantinga's exact words, from his statement of the defense's "central idea," in my exposition):

The Basic Idea of the Free Will Defense: First, the good God seeks, for the sake of which the possibility of evil is allowed, is creatures who are "free with respect to morally significant actions" (actions where they are free to do what's right and also free to do what's wrong) and/or creatures freely doing what is right in those situations in which they are significantly free. But, second, not even omnipotent God can (directly or indirectly) causally determine creatures to always freely act rightly because, "if [God] causes [creatures] always to do only what is right, then they don't do what is right freely." Because God cannot determine creatures to freely do what is right, third, even omnipotent God can't assure that if creatures are allowed to be free in morally significant situations, they won't act wrongly. So, fourth, to get the goods They seek here, even God must allow creatures to act freely in morally significant situations, and take a real risk of their acting wrongly. Fifth, evil results on those occasions where that risk doesn't pay off. But sixth, God was right to take the risk in order to get the important goods They seek.

Initial comments: On the first element, I went on to complicate Plantinga's statement (within quotation marks) of the good God seeks. Plantinga here construes God as seeking creatures acting freely in morally significant situations as that good in itself, but Free Will Defenders in good standing can also, or instead, have God specifically shooting for creatures freely doing what is right when they make significant choices. Indeed that seems to me the more natural good to appeal to—and a few pages after the passage I'm quoting, Plantinga himself toggles over to that being the good God is seeking on the FWD.⁸⁴ It's a subtle difference, to be sure (and I don't get the feeling that Plantinga was cognizant of his shifting on it), and I suppose the best particular path to take might be to have God seeking both as goods. But for maximal flexibility, accommodating different views (among those geeky enough to notice this difference), I went with the (geeky) "and/or" formulation.

⁸⁴ Plantinga, "Self-Profile," p. 45: "A crucial part of the free will defense is the contention that it is possible that a wholly good God consider it valuable that there be moral good—good, that is, that results from the moral activity of free creatures freely doing what is right—and is willing to put up with evil in order to achieve this end."

Our freely acting in the right way need not be the end of the story, of course: many theists who emphasize the importance of human freedom think a reason it is so important is that it's an ingredient in other valuable states. Perhaps most importantly, some, for instance, think there is an immensely great value (perhaps the greatest we can ever attain) realized by our being in certain types of relationship with God which can only be, or else can most valuably be, entered into freely.

The second element of the Defense—intuitive as I'm convinced it is—is based on what's known as an “incompatibilist” account of freedom, which turns out to be extremely controversial. We will have to address this controversy, and the question of the credibility of the Free Will Defense given its reliance on this controversial account, and will do so especially in **Part Six**. This will be tied up with the crucial issue of whether creatures acting and/or acting rightly in situations where they are free is really such a great good, and so will be tangled up with the first element of the Defense.

Plantinga simply says that creatures don't act freely “if God causes” them to act as they do, but it's worth adding that the clear freedom killer, according to incompatibilists, would be God's, directly or indirectly, causally *determining* creatures' actions: One's freedom can presumably survive a gentle, and even a substantial, causal nudge toward an option. Now, if one is only pushed toward the path one then in fact chooses (or to use the philosophically famous terminology of Leibniz (well, in English translation): if one is only *inclined without being necessitated* by outside forces to act as one does), it may not be correct to say simply that one “was caused to” act as one did by those forces. But perhaps when the push gets strong enough and/or works in particular ways, we can reach the point that, if the pushed for action is performed, it's true to simply say that the push “caused” the action, even where it didn't strictly determine one to so act. (Perhaps the push left open some possibility that you wouldn't act in the way pushed for, but still worked in such a way that if you did act in the way pushed for, it would be true to simply say that the push “caused” you so to act.) And maybe reaching that point is enough to make the push a freedom killer, as opposed to a mere incliner. Still, it seems God would have to *determine* us to act as They want to avoid all risk of our acting otherwise, and that this determining is a graver threat to our freedom than is non-determining causing, so it's worth at least often using the terminology of being “causally determined” to make clear that, for our current purposes (which involve evaluating what would be God's role in a world of creaturely freedom) we're not talking about any mere non-determining causal nudges.

The third and fourth elements are presented as following from what precedes them, but you may sense a bit of a gap over which that jumps (from can't *determine* to can't *assure*). If so, and, in particular, if you are wondering whether God might use Their knowledge of what we

would freely do in various possible circumstances to get the goods sought without determining our actions but also without taking any risks of our going wrong, then a) you are thinking very much along lines Plantinga himself does, and b) your worries require getting into some extremely complicated territory, and will be addressed in **Part Four**. (For those who know the relevant label, it may help to say it is there where we will address Plantinga's *Molinism*.)

The main difference my presentation of the free will defense has with Plantinga's is that mine has God taking *risks*—and in fact, puts that front-and-center, so that it takes a prominent place in even my basic statement of the defense. Plantinga not only keeps such a notion out of his basic statement of the defense, but as he more fully develops his version of the defense, he retains a God who takes no risks as They deal with free creatures. Still, to give a cryptic hint of what's to come in the difficult explanation of Plantinga's defense (and perhaps also to hint toward why it's so difficult), risk is so inescapable to any intuitive version of the free will defense, or so I will suggest, that Plantinga's God does turn out to be in a key way *subject to risk*, even though They *take* no risks. In the meantime, to call in some perhaps needed prominent back-up here, I close this section by quoting a nice passage by Swinburne, which is in a book that came out just a bit too late to help me as a young teacher, but that features a very strong statement of the centrality of the free will defense to at least Christian theology, as well a basic statement of the free will defense which shows the importance to the defense of the notion of God taking risks, as that aspect plays quite a starring role in even this very brief statement of it:

Many Christian theologians have claimed not merely that God has some good purposes for the sake of which he allows the bad to occur, but that such-and-such were God's actual purposes; or at any rate that such-and-such were possible purposes to realize which God (omnipotent and perfectly good) could have allowed the bad to occur. In other words, they have provided a theodicy for many of the world's bad states. And central to the view of almost all of them has been the 'free-will' defence in respect of moral evil—that the bad states caused (or negligently allowed to occur) by human beings result from their own free choice; and that it is such a good thing that humans should have free will that it is worth the risk that they will abuse it in various ways.⁸⁵

⁸⁵ Swinburne, *Providence and the Problem of Evil* (Oxford University Press, 1998), p. 33

23. The Power of the Free Will Defense: The Potentially Immense Importance of Libertarian Freedom

At the heart of any theodicy is the judgment that there are goods so valuable that they can plausibly be thought to justify God in seeking to secure them – even at the cost of the world’s horrors.

.....

How could anyone possibly believe that the evils of this world are outweighed by the good inherent in our having free will? Perhaps free will is a good and would outweigh, in Theist’s words “a certain amount of evil”, but it seems impossible to believe that it can outweigh the amount of physical suffering (to say nothing of other sorts of evil) that actually exists.

Though arrestingly bold, the top quotation above, by Daniel Speak,⁸⁶ is inescapably, if troublingly, right, at least about any theodicy that hopes to address our problem of horrific evils. And indeed, the great power of the Free Will Defense lies in the good that it appeals to, human “Libertarian free will,” as it is sometimes called, seeming to have the twin features of being, first, something not even an omnipotent being can control, and so being something even God might have to risk evil in order to secure; while being, second, potentially an immensely great good.

Great enough even to cast doubt on the second, skeptical passage above, which was penned by van Inwagen⁸⁷—or so I will argue in Part Five.

Now, this doubt turns out to only slightly complicate how our world’s horrors refute the Free Will Defense, so, as I’ve indicated, I will not be promoting the Free Will Defense. There’s no need to be mysterious about what is behind the Defense’s failure, so to briefly preview the salient cause: Though freedom may be so important that on a simple comparative evaluation, it might appear to be possibly worth even the world’s horrors to get it, just slightly more careful thoughts about how God *might get this great good without all the horrors* sink the Free Will Defense, as we’ll finally see in **section xx**. Still, the ability of this good to stand up against even the horrors of our world, at least in advance of such more careful thought, will be a testimony to its potential immensity—and to the wisdom of keeping human free will in our account of why God allows the horrors of our world, even as we discard the Free Will Defense itself.

⁸⁶ Daniel Speak, “Free Will and Soul-Making Theodicies,” in Justin P. McBrayer and Daniel Howard-Snyder, eds., *The Blackwell Companion to the Problem of Evil* (2013), pp. 221; p. 205.

⁸⁷ Peter van Inwagen, *The Problem of Evil*, p. 73.

24. The Free Will Defense as an Attempt to Get God's Relation to Evil on the Good Side of Key Ethical Distinctions: Evil as a Merely Allowed Side-Effect

So in **Part Five**, we will take a closer look at the potentially immense importance of Libertarian free will, which is the good that the Free Will Defense appeals to. But as we here wrap up our look at the Basic Free Will Defense, we will very quickly look at another source of that defense's intuitive power: that it seems to get God's relation to evil on the good side of two important moral distinctions, a) that between what an agent does vs. what they merely allow to happen, and b) that between what an agent intends as their goal or as a means toward their goal vs. the merely foreseen but not directly intended side-effects of their plans of action. I first studied these distinctions from Warren Quinn, and his great work on the "Doctrine of Double Effect" (DDE), the name already common for the claim that the second of the distinctions mentioned above is a morally relevant one, and the "Doctrine of Doing and Allowing" (DDA), which name Quinn (so far as I can tell) coined to express the analogous claim about the first of those distinctions.⁸⁸ Though we won't here look at examples, both doctrines can be illustrated, and also intuitively supported, by pairs of cases, which, in Quinn's words, "have the *same* consequential profile"⁸⁹—where the agents face options that are the same as each other's in terms of their consequences (e.g., in terms of how many people die, or are saved, or are hurt, etc.), and where the agents choose options with the same consequences as each other's. However, when, in these cases, the different agents' courses of action put them on opposite sides of one, or the other, or both, of our doctrines, this can result, at least intuitively, in sharp

⁸⁸ As I recall, I encountered Quinn's work as a long typescript paper that Quinn had written about both doctrines/distinctions, and that he would have the UCLA bookstore make copies of to sell to students in his relevant ethics classes. For purposes of publication, Quinn later broke this down into two papers, which appeared as "Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing," *The Philosophical Review* 98 (1989): 287-312 and "Actions, Intentions, and Consequences: The Doctrine of Double Effect," *Philosophy & Public Affairs* 18 (1989): 334-351. (I was thrilled that my first philosophical publication appeared in the same (July 1989) issue of *The Philosophical Review* as the first of those papers by friend and teacher.) Quinn certainly writes as if he's coining the phrase we're following him in abbreviating as "DDA": "In the case of the distinction between action and inaction, the view has no common name, so for convenience we may call it the Doctrine of Doing and Allowing (DDA)" ("The Doctrine of Doing and Allowing," p. 287). These topics were very much in the air at UCLA philosophy when I was a graduate student there, due to the presence of Philippa Foot, as well as Quinn, on the faculty, and play a role in the work of Marilyn Adams, who was Quinn's and Foot's colleague: See esp. her section on "Double Effect, Doing/Allowing, New Intervening Agent," at pp. 34-35 of *Horrendous Evils and the Goodness of God*.

⁸⁹ "The Doctrine of Double Effect," p. 335.

differences in the moral status of their behavior.⁹⁰ The doctrines can appear in absolutist forms: You may never under any circumstances actively bring about evil (though in some circumstances you may allow evil); You may never under any circumstances directly intend evil (though in some circumstances you can bring it about as a foreseen but not directly intended side-effect of your action). But they are more defensible in the weaker form in which they just assert a morally relevant difference, which difference may sometimes make the difference between a course of action being morally justified or not, or may show up sometimes as a difference in just how much good it takes to justify a certain amount of evil, or sometimes as a difference in just how bad an agent's course of action was. The non-absolutist forms of the doctrines can be important to those who embrace the Free Will Defense as a key part of a mixed strategy: they may see God's relation to evil sometimes falling on the dark sides of these distinctions (which is allowed because the doctrines are not accepted in their absolutist forms, or at least not as they are applied to God), where another defense from the mix is needed, while being relieved that, at least where the Free Will Defense does apply, God is on the sunny side of these distinctions.

I suspect that many who are strongly drawn to the Free Will Defense, are so drawn precisely because that Defense gets God's relation to evil on the good sides of these intuitively important distinctions.

And I am not alone in so suspecting. Here are Murray and Rea, with a nice statement of the importance to "most theists" of the "buffer" provided by getting God on the good side of the Doing/Allowing distinction (though they use "permit" rather than "allow"):

Of course, all varieties of theism must confront the problem of evil. But most theists are insistent . . . that God does not cause evil but rather merely (justifiably) permits it. Seeing

⁹⁰ In well-known earlier work, James Rachels had used a such a pair of cases with similar consequentialist profiles, but one where there seemed (at least to him) no difference in the moral status of the behavior, to argue *against* (what we are now calling) DDA: Rachels argued that where Smith drowns his young cousin in the bathtub in order to inherit a family fortune, he behaves no worse than does Jones, who instead finds his young cousin already unconscious under the water, and, in order to inherit the fortune, does not save him (Rachels, "Active and Passive Euthanasia," *New England Journal of Medicine*, 292 (1975): pp. 78–86). While some who have responded to Rachels have denied that Smith's behavior is no worse, the more interesting criticisms have argued that the moral significance between killing and letting die (or, more generally, between doing and allowing), need not manifest itself as a difference in the degree of wrongness in all such consequentially similar cases. For discussion, see Quinn's treatment at pp. 288–290 of "The Doctrine of Double Effect," but also, and more decisively, Shelly Kagan ("The Additive Fallacy" (*Ethics* 99 (1988): 5–31), who argues against the assumption that "if a factor has genuine moral relevance, then for *any* pair of cases, where the given factor varies while others are held constant, the cases in that pair will differ in moral status" (p. 12).

God as merely permitting evil provides a moral buffer that would be unavailable if God were the full and direct cause of evil as is the case with occasionalism.⁹¹

Now, formulating either of our Doctrines in detail so that they apply to all cases in the intuitively correct way is a very complicated matter (that we won't be attempting here), and, as we now consider applying DDA to God, we face the fact that the God of standard theism, as the Creator and sustainer of all things, is no mere totally passive bystander in the occurrence of any evil. Perhaps rather than simply saying that the Free Will Defense gets God on the sunny side of this distinction, it might be better to say that it does an unusually good job of doing so—as opposed to occasionalism, the view Murray and Rea mention, on which God directly causes everything that happens to take place, and which view seems an extreme failure here. And it certainly does seem that in some potentially morally relevant sense, the God of the Free Will Defense *allows* the evils done by God's free creatures. Avoiding messy details, I here instead just appeal to your sense of whether the terminology of allowing is in order on the picture of God's relation to evil painted by the Free Will Defense. True, God's creatures couldn't have done evil without God. In fact, like everything else, we wouldn't have existed, and so couldn't have done anything, without God's creative activity. And on some theistic views, God must not only create us, but also constantly and actively uphold our existence at every moment for us to continue to exist and so for us to be able to do anything. But in some good—and, one feels, possibly important—sense, when the God of the Free Will Defense upholds us in existence, yet does not control how we act, and so *allows* free creatures to do our thing (as one quite naturally describes it), and we mess up, the God of the FWD is allowing the evil that results.

And the God of the Free Will Defense allows this evil that we do as a side-effect of Their plans, at least on an intuitive initial division of side-effects from ends and means. That resulting evil is no part of God's ends, nor the means by which God achieves those ends. What the God of the Free Will Defense wants is our freely acting rightly, especially in morally significant situations. To get this, God must put us in, or allow us to be put in, morally significant situations in which we are free, and let us act freely. When we then do evil, this does not help God in achieving any of Their goals. God needs to allow the real possibility of evil, but the evil itself does not help God. This allows God to be whole-heartedly against the evils that so happen, in a way that seemingly could not be the case if those evils were chosen means to any of God's ends, or, of course, if they themselves were among God's ends. And that can be intuitively important, especially when the evils are great.

⁹¹ Michael J. Murray and Michael C. Rea, *An Introduction to the Philosophy of Religion* (Cambridge University Press, 2008), pp. 25-26.

Part Four: Freedom without Risk?: God's Knowledge of What We Would and Will Freely Do and the Two Gods of the Free Will Defense

25. The Gods of the Free Will Defense: Plantinga's Molinist God

While I used the work of Alvin Plantinga, the philosopher most closely associated with the Free Will Defense, where I could, in explaining the basic idea of that defense in [sect. 22](#), I also indicated that my construal of the Free Will Defense differs from his in a key way that involves the notion of risk. Here in [Part Four](#), we will look at these two different ways of running the Free Will Defense, which amount to two different ways of understanding how God exercises providential control over a world containing free creatures, even as both views operate with a Libertarian view of human freedom.

This Part ([sects. 25-28](#)) will be quite complicated material, especially in the footnotes, and some readers may want to skip it. Others, I imagine, may really get into it—including, perhaps, even the notes. You'll probably be able to tell which camp you fall in pretty quickly. If you do skip these difficult sections, you should be able to pick up the story at the start of Part Five. There, and moving forward from there, I will evaluate the Defense as I think it is best, and most intuitively, construed, but will explain how various considerations fare on Plantinga's way of running the defense in notes.

My part and section titles speak, perhaps rather dramatically, of there being different Gods on the two pictures we will be considering, but that is a bit misleading. God's attributes are the same on each picture. The very different ways God interacts with the world (different enough to explain my talk of "different Gods") on the two views stem from differences in the relation between the future and present in a world containing free creatures, not from differences in what God is like. These are perhaps more accurately viewed as two sketches of how the same omnipotent and omniscient God operates given different views of the relations between future freedom and current truth, as I hope will become clear as this Part unfolds.

Here is Plantinga's own statement of the basic idea of the Free Will Defense (some phrases of which I have already quoted and used), which comes immediately after Plantinga assures us that "the central idea of the Free Will Defense is simplicity itself":

Plantinga on the Basic Idea of the Free Will Defense: This idea is as follows: It is possible that God (who is omnipotent, omniscient and wholly good) thought it good that there be significantly free creatures—creatures free with respect to morally significant actions—but wasn't able to create such creatures in such a way that they always exercise their

freedom to do good; for if he causes them always to do only what is right, then they don't do what is right freely. If so, then it is possible that there be evil even though God is omniscient, omnipotent and wholly good.⁹²

I suppose that can sound simple and intuitive enough. But I find it extremely complicated and counter-intuitive (though here, your mileage very well may vary) as soon as I try to scratch below the surface at all in order to explain why Plantinga thinks that evil and God can go together. (Truth be told, I think this version of the Free Will Defense only sounds intuitive insofar as it can be mistaken for the version I opted for in [sect. 22](#), which version I think really is quite intuitive.)

The glaring difference between Plantinga's statement of the Defense and mine is that Plantinga does not say anything about God taking any risks—a notion that plays a pivotal role in my statement of even the basic idea of the defense. And it's not that Plantinga just judged that the bit about God's taking risks isn't important enough to get into his short statement of the basic idea. It's instead that Plantinga's God takes no risks. Plantinga thinks God has what is known (for reasons we need not discuss) as "middle knowledge": knowledge of what every free creature would freely do in every possible situation in which it would act freely.⁹³ And this is

⁹² Plantinga, "Self-Profile," in J.E. Tombererlin and P. van Inwagen, eds., *Alvin Plantinga* (D. Reidel Publishing Company, 1985), pp. 3-100; at p. 41.

⁹³ There are actually two different construals of "middle knowledge" in the philosophical discussion, which the participants on both sides of the debate must think amount to the same thing, but which really come apart. (Well, I speak here just about the 20th Century (and now continuing) version of the discussion, without addressing any 16th Century disputes.) Both are knowledge of conditionals—"if ____ then ____" statements—about what creatures or possible creatures would freely do in various possible situations. But they differ over just which conditionals are the objects of this "middle knowledge." On the first, and most important, construal, middle knowledge would be of conditionals the knowledge of which would allow God to exercise no-risk-taking providential control over the world despite its containing creatures acting freely. This would presumably be knowledge of such future directed conditionals as "If Eve is in situation S_{498} , she will sin," or its "'were'-d up" future-directed counterpart, "If Eve were to be in situation S_{498} , she would sin." Strangely, though, when the writers of the founding documents of the (20th Century episode of the) debate gave their explicit flagship examples of what middle knowledge would be knowledge of, they used past-directed, counterfactual conditionals like "If Eve had been in situation S_{498} , she would have sinned," which would seem to be of help only in after-the-fact "Monday morning quarterbacking" (as the saying goes, based on the fact that professional American football is usually played on Sundays; the saying derives from earlier times, back when quarterbacks would decide which plays to run, and so might well wonder on Mondays what would have happened if they had tried a different play). Those giving the examples, but also thinking that "middle knowledge" would be what God would need to exercise the kind of providential control in question, presumably thought that the past-directed counterfactuals are true and known to God after it's too late to do anything about the situation if and only if the future-directed, potentially useful conditionals were true and known to God before the relevant time. But as it is extremely doubtful that any such relationship between the various conditionals really holds, the middle knowledge debate splits into two

absolutely certain knowledge (as is befitting God). And Plantinga's God uses this middle knowledge to exercise no-risk-taking, no-dice-throwing providential control over the world, even where that world contains free actions by creatures (and perhaps other indeterministic happenings). Plantinga's God doesn't riskily put us in a situation in which we are free and hope for the best, for They know full well exactly what we'll do if put in that situation. This system, on which God has middle knowledge and uses it to exercise no-risk-taking providential control of the world is called "Molinism," after Luis de Molina, the 16th Century Jesuit who devised it.⁹⁴

But then how does the Free Will Defense work for Plantinga? Why couldn't, and why doesn't, God use this seemingly very handy middle knowledge to see to it that creatures are put only in situations where we will freely do right, and thereby get the goods God seeks (of us acting freely, and freely doing good, in morally significant situations) while assuring that evil does not result?

Well, the answer, at least as I would put things, is that while They *take* no risks, Plantinga's Molinist God was in an important way *subject to risk*⁹⁵—in a way that gives rise to the possibility of God having to tolerate evil in order to get moral good. Suppose that what God really would have liked is Eve doing the right thing in situation S₄₉₈, a situation in which she would be free to do right and free to do what's wrong.⁹⁶ Well, then, since Eve would be free, there are truly

non-equivalent debates. For (much) more on all this, see my paper, "The Conditionals of Deliberation," *Mind* 119 (2010): 1-42; the explicit application to the "middle knowledge" debate is at pp. 3-6 (importantly including the notes at those pages).

The distinction is especially important to me, as I think omniscient God can't have middle knowledge in the first, more important way of construing it, but will often know the providentially useless conditionals of the second construal. (This gets a bit complicated, because the meaning of the counterfactuals is context-dependent, and they *can* be read in such a way that they too are almost never true of potential past free actions. However, they can also mean something that is often enough true, and where they are indeed true, omniscient God certainly knows them to be so.)

I will here use "middle knowledge" in the first, more important, way, unless I specify otherwise.

⁹⁴ As we will see, it is also definitional of Molinism that the conditionals are true pre-volitionally—that is, that God does not cause them to be true (see [note 101](#)).

⁹⁵ This way of putting things is apparently due to me. I was going to ascribe it to Dean Zimmerman, but I had forgotten events, and I now see that where he uses the formulation, he writes: "I thank Keith DeRose for this nice turn of phrase" (Zimmerman, "Yet Another Anti-Molinist Argument," in S. Newlands, L.M. Jorgensen, eds., *Metaphysics and the Good: Themes from the Philosophy of Robert Merrihew Adams* (Oxford UP, 2009), pp. 33–94; p. 46, n. 17). I had hoped to have to just endorse Zimmerman's phrasing, citing him as an expert on the relevant bit of metaphysics, but now seem to see that I am really laundering my own phrasing through him!

⁹⁶ "Situation S₄₉₈": I take this to be a completely specified situation which Eve might find herself in, down to the smallest detail, of the world up to the point in time where Eve performs her free act. I take it that the answer to the question of how Eve would act in situation S_a (a not-completely-specified situation, that could be realized in somewhat different ways) is typically: "Well, that depends on just how S_a is realized," and that the middle knowledge debate most importantly concerns whether God can know

possible worlds (ways things might have been) in which Eve does the right thing in S_{498} and also truly possible worlds in which she there goes wrong. But Plantinga supposes that there is a fact of the matter as to which she would do, and God eternally knows what that fact is. Let's suppose that, sadly, the fact is that she would do wrong. Well, then, Plantinga's God can't get what They would really want here—the world in which Eve freely acts rightly in S_{498} —because, as God knows full well, Eve won't cooperate if she is put in S_{498} . Alright, then, why not put Eve in a nearby possible situation in which she would do better, say, S_{499} ? Or try some other possible free creature in some nearby situation? Well, it's *possible* that God was *deplorably* unlucky, and these facts about what creatures would freely do in various situations were very much lined up against God. So perhaps there just are no nearby situations in which Eve, nor some other possible creature much like her, would cooperate with God's hopes. More: Perhaps the facts about what creatures would freely do in various situations are generally so hostile to God's hopes and purposes that the only possible situations in which any possible creatures would ever act rightly include that they have elsewhere acted wrongly (in which case every possible free creature suffers from what Plantinga dubs "transworld depravity").⁹⁷ It's

how we will freely act in completely specified possible situations, like S_{498} (which, for omniscient God, turns on whether there are facts to be known about how we will act in such situations).

"Eve": We are discussing a point in the story at which Eve does not yet exist (and indeed, it is being decided, among other things, whether she will exist). So at this point, Eve is a non-existent individual—which is a bit of an embarrassment for "actualists" like both Plantinga and me, who don't believe there really are any non-existent individuals! Plantinga supposes that before Eve existed, there was the property of being Eve, which property existed eternally (and necessarily, as Plantinga thinks is the wont of properties generally), and what God is really considering when God considers what would happen if (as we might say) Eve existed and were put in S_{498} , is what would happen if the property of being Eve (perhaps this property could be called "Evity"?) were instantiated, and the individual who instantiated that property were put in S_{498} . I don't believe in the "uninstantiated individual essences" of Plantinga's account (no Evity without Eve for me), and instead suppose that God is thinking in terms of what would be freely done by creatures having certain qualitative properties (properties which don't include those of being identical with this or that individual). There are "singular propositions," as they are called, about you as an individual, and there is the property of being you, when, but only when and after, you come into existence. But I assume that after you come into existence, the singular proposition that you will do something if put into a particular situation has the same truth value, if any (and also has the same probability) as does the completely specified non-singular, qualitative proposition that now turns out to be about you that God has been thinking in terms of all along: that if there is a creature with such-and-such properties that is in such-and-such a situation (all completely specified, at least qualitatively), it will freely do such-and-such. All that said, you will be relieved to learn that, for ease of exposition, I will just be writing as if there were non-existent individuals, some of whom later come into existence, trusting that those who are bothered by such talk can now convert it into metaphysically more scrupulous terms.

⁹⁷ In line with the metaphysical scruples mentioned in the previous note, Plantinga writes of every free "creaturely essence," rather than every possible free creature, suffering from transworld depravity. Note that it would be enough for Plantinga's purposes that the relevant conditionals might have had it that no possible free creature would ever do good in any situation unless they *or others* had elsewhere

possible.⁹⁸ Well, then, if *that's* the kind of array of facts (about what creatures would do in various possible situations in which they are free) that faced God, then despite God's omnipotence, and despite Their middle knowledge by which They can exercise no-risking-taking providential control, it was not only certain particular possibilities, like Eve acting rightly in S₄₉₈, but even the general possibility of there being creaturely moral good but no moral evil, that were beyond God's reach. That would seem to me *immensely* bad luck, but it seems metaphysically possible that God was so forlornly unfortunate. So it's possible that God, despite Their omnipotence, couldn't get moral good without moral evil. That is the somewhat more filled-in basic idea of Plantinga's solution to the logical abstract form of the problem of evil—the problem of logically squaring God's existence with the existence of any evil at all.

Some may find it helpful to think of the control God exercises over the world on this Molinist scheme in terms of Plantinga's God ordering what They want quite definitively, but from a limited menu, where the "menu" is the list of possible worlds that God can make actual (the list of "realizable," or sometimes "feasible," worlds, as they're sometimes technically called in the field of Plantinga studies). There are some really possible situations—like in our abstractly imagined case, Eve acting rightly in S₄₉₈—that just aren't on the menu, because, as God knows full well, free creatures won't do their part in making those scenarios happen. How then are they really possible? Because the facts that put those situations beyond God's reach really could have been otherwise. In some good sense, while there never was a time when they could be made actual, they still *could have been* actual, if only those pesky conditionals of freedom had been different—as they really could have been. The risks God was subject to were all "front-loaded," as we might say, and are the risks of which possibilities do and which do not show up on the menu of options God can make real. But God *takes* no risks: They know with certainty and from eternity exactly what is and what is not on the menu, and know that and how They can assuredly get any item that is on the menu, with no risk of missing. As we may well say, it is because the risks of dealing with free creatures were already front-loaded on this scheme that God did not have to *take* any (further) risks in getting good actions. And Plantinga's solution to the logical abstract problem of evil on this picture is that it is possible that the menu of realizable worlds that confronted God contained no dishes on which there

gone wrong. In that case, while possible free creatures would not individually suffer from transworld depravity, it would still be beyond God's reach to get moral good without moral evil.

⁹⁸ Well, that this is possible is a premise that drives Plantinga's argument that God and evil are compossible. Whether he is really entitled to that premise, or advances the case for the compossibility of God and evil by using such a premise, are questions I wrestled with way back in "Plantinga, Presumption, Possibility, and the Problem of Evil," *Canadian Journal of Philosophy* 21 (1991): pp. 497-512. Briefly and perhaps cryptically: I argue that Plantinga doesn't really advance the case in a worthwhile way that goes beyond where the theist would be by just directly presuming that God and evil are compossible, absent good reason to think them not compossible. Here, where I'm looking to get on to the questions of credibility rather than mere possibility, I'll just grant the compossibility.

was moral good but no moral evil. If so, then to get moral good, God had to put up with moral evil. (Of course, how there could be as much evil as there is in the world, and how that's not only metaphysically possible, but credible, despite the existence of a wholly good God is a more difficult Free Will Defense story to tell. (I believe impossibly difficult to tell plausibly.))

26. Against Molinism

Now, I am not only suspecting, but also hoping, that if you are following this account, some of you are finding some of it rather fishy, with perhaps the most pressing concern being the question: Well, if there are really possible worlds in which Eve acts rightly in S_{498} , what made it instead the case that she would do wrong in that situation? Or more generally: What makes these facts about what creatures would do in various situations true?⁹⁹

⁹⁹ Here we wonder what, if anything, on the Molinist scheme is causally responsible for the facts about what creatures would freely do being as they are. This is to be distinguished from the matter of what, if anything, *grounds* the truth of these facts—the matter of what their truth *consists in*. The thought that there is not a good Molinist answer to the latter question is the basis of the (philosophically) famous “grounding objection” that Robert Adams raised against Molinism in his paper, “Middle Knowledge and the Problem of Evil,” *American Philosophical Quarterly* 14 (1977): pp. 109–17; esp. at pp. 110–111.

I endorse Adams's objection, where it is aimed against middle knowledge as that is construed in the more important of the two ways described in note 93, but recognize its limited effectiveness against some Molinists, who can respond that they don't see why the facts about what creatures would freely do in various situations have to be grounded in supposedly more fundamental facts. As Plantinga nicely counters, one could similarly ask, about some free action one actually performed in the past, what grounds the fact that one did perform that action, and one might find that one has little answer here, other than the fact that one did indeed do the act. Why can't the Molinist similarly simply respond that what grounds the fact that Eve would sin if she were put in S_{498} (to use our example here) is just that she would indeed sin if she were in S_{498} ? (Plantinga, “Replies,” in J.E. Tombererlin and P. van Inwagen, eds., *Alvin Plantinga* (D. Reidel Publishing Company, 1985), pp. 313–398; at p. 374). I think the difference here is that facts about what would happen in various possible future situations have to be grounded in more fundamental facts in the way demanded by Adams's grounding objection, and in a way that facts about the actual past do not have to be, since those facts about the simple past are fundamental enough themselves. So, I accept Adams's objection, but realize that its success is based on some very deep, and also controversial, judgments about the relative fundamentality of different kinds of facts.

But there is also a sense in which I accept middle knowledge (see note 93 again). How do the counterfactual conditionals that are the objects of that form of middle knowledge escape Adams's grounding objection? I accept that facts about what would have happened if certain counterfactual situations had occurred have to be grounded in the more fundamental facts that Adams supposes. However, I do not accept Adams's requirement that the grounds of these counterfactuals have to be necessitating in way Adams claims (Adams, “Middle Knowledge and the Problem of Evil,” p. 111). Adams himself thinks that there are facts about what creatures *probably* would have done in various situations

in which they would have been free, and gives an account of what kinds of more fundamental facts can ground these probabilistic counterfactual truths (Adams, p. 111). For me, those same more fundamental facts are often sufficient to ground the truth of the counterfactuals that have no “probably” inserted into them. These not-explicitly-probabilistic counterfactuals are context-sensitive in their meaning, but the likes of “If Eve had been in S_{498} , she would have sinned” can, and often enough do, mean something fairly easy to make true, and indeed mean roughly what Adams would take “If Eve had been in S_{498} , she *probably* would have sinned” to mean (which is also quite context-sensitive in its meaning), and when it does mean that, the former’s truth can be grounded by the same kinds of facts that Adams supposes are enough to ground the truth of the latter.

In close connection with his claim that the grounds of the relevant counterfactuals must be necessitating, Adams suggests (at esp. p. 110) an argument against the truth of those counterfactuals based on “might” counterfactuals: If Eve is free in S_{498} , then “If Eve had been put in S_{498} , she might have sinned” and “If Eve had been put in S_{498} , she might not have sinned” must both be true; but the former of those is incompatible with “If Eve had been put in S_{498} , she would not have sinned”, and the latter is incompatible with “If Eve had been put in S_{498} , she would have sinned”; thus, neither of those “would” counterfactuals is true. Here, I deny, and argue against, the “duality” account of the relation between “would” and “might” counterfactuals that drives this argument, while accounting for and thereby undermining the appearance of truth that can attach to that account: See my “Can It Be That It Would Have Been Even Though It Might Not Have Been?” *Philosophical Perspectives* 13 (1999): 385-413.

But what then of an analogous “might” argument aimed at the future-directed conditionals that I think would be the object of the more important form of middle knowledge? That argument would also fail, and for the analogous reason. But this would be a failed argument for a true conclusion. Though it is directed at simple (non-conditional) statements about the future, rather than future-directed conditionals, my paper, “Simple *Might*’s, Indicative Possibilities, and the Open Future,” *The Philosophical Quarterly* 48 (1998): 67-82, may give a good indication of how I think this all shakes out.

Why think the grounds of the counterfactuals need not be necessitating, while taking a more demanding approach to the meaning of the future-directed conditionals? To be quick here, I will just say that when propositions are construed in the way I like (importantly, with times built into them), I don’t believe propositions can “flip”: go from being true to being false, or from false to being true—though they can go from having no truth value to have one truth value or the other (in fact, this happens all the time—I suppose quite literally, as infinite multitudes of propositions gain truth values every moment as time moves forward). Because reality can no longer bite us and make the antecedent of one of the counterfactuals true while its consequent is false, such a counterfactual can, and often does, express a proposition with what we can think of as “lax” truth conditions. Not so for the future-directed conditionals that would be the objects of the more important form of middle knowledge—nor for the non-conditional future-directed would-be objects of simple foreknowledge. I think the more consistently demanding accuracy conditions of the objects of the more important kind of middle knowledge are deeply tied to how they would be providentially useful if they were known to God: It is no accident that the useful ones are so hard to know (impossibly hard, I think, in the relevant cases).

Finally, and perhaps most cryptically of all, I will take the chance to here say that I think a main upshot of seeing just which conditionals would be the objects of middle knowledge on the more important construal of middle knowledge is that the issue of whether God can have that important kind of middle knowledge can be seen more clearly to stand or fall together with the matter of whether God knows (with certainty) the simple (non-conditional) open future. (And I for one say they fall together.) This goes against a hopeful suggestion of Adams’s, early in his paper, where he can be read as suggesting that we can give up on God’s middle knowledge while holding on to their simple foreknowledge: Adams, at the end of p. 110.

Plantinga likes to talk about deciding how things go in a world of free creatures (or deciding which possible world becomes actual, to put it in those terms) being a “co-operative venture,” that is in an important way only partly up to God. In fact, he will say that this is essential to the FWD, as in this nice passage from *The Nature of Necessity* (NN):

The essential point of the Free Will Defence is that the creation of a world containing moral good is a co-operative venture; it requires the uncoerced concurrence of significantly free creatures. But then the actualization of a world W containing moral good is not up to God alone; it also depends upon what the significantly free creatures of W would do if God created them and placed them in the situations W contains. Of course it is up to God whether to create free creatures at all; but if he aims to produce moral good, then he must create significantly free creatures upon whose co-operation he must depend. Thus is the power of an omnipotent God limited by the freedom he confers upon his creatures.¹⁰⁰

But if what happens in situations of creaturely freedom is only partly up to God, on what or whom is it also partly up to? As the passage above would lead you to expect, Plantinga likes to answer that it is partly up to us, God’s free creatures. Here he is a few pages earlier in NN, writing of the corrupt but free character Curley (an old mayor of Boston, not the member of The Three Stooges!), who functions for Plantinga much like Eve is functioning for us, and a situation in which Curley is free to take a bribe or not:

Accordingly, there are possible worlds such that it is partly up to Curley whether or not God can actualize them. It is of course up to God whether or not to create Curley, and also up to God whether or not to make him free with respect to the action of taking the bribe at t . But if he creates him, and creates him free with respect to this action, then whether or not he takes it is up to Curley—not God. (NN, p. 184)

But going back to our question of what makes these facts about what creatures would freely do in various situations true, and our example of that, it seems that it cannot be Eve who made it true that she would go wrong if put in S_{498} : She doesn’t exist yet. Indeed, she may never exist: God may decide not to create her after all, and perhaps precisely because of how uncooperative she would have been. Still, the facts about what she freely would do if she were to exist and were in various situations contributed to the limiting of God’s options, on Plantinga’s scheme.

And God doesn’t make it true¹⁰¹: The whole idea here is that this is a point where God doesn’t alone decide how the world turns out, and free creatures (or at least facts about them,

¹⁰⁰ Plantinga, *The Nature of Necessity* (Oxford University Press,), p. 190.

¹⁰¹ In the closing two paragraphs of this section, we will see that conditionals of freedom are “pre-volitionally” true on Molinism: they are “known by God prior to his creative act of will,” as Thomas P. Flint puts it in *Divine Providence: The Molinist Account* (Cornell University Press, 1998), p. 46.

or about their essences, I suppose Plantinga must say) also have some say. But then we're left with no good account of how it came to be that, though the world could have been such that Eve would have done right if she were put in S_{498} , it instead turned out that she would have sinned if put in that situation. As my fellow anti-Molinist, Dean Zimmerman, nicely puts it:

According to Molinism, then, it is as though God 'wakes up' to find certain contingent things true—there is an independent source of contingent fact at work 'before' God has a chance to do anything about it. Although Molinists may reject such talk as tendentiously impious, there is an important (and potentially troubling) truth behind it. The Molinist conditionals really are supposed to be contingent truths discovered by God, not determined by Him; and discovered 'before' He creates—at least, 'before' in the order of explanatory priority.¹⁰²

What's potentially objectionable there is the talk of God "waking up" (though, presumably, that is why it's put into scare quotes). The Molinists' God was never asleep: There was no earlier (pre-waking) time at which They didn't know what those facts were. But the extended sense in which it is *as if* God wakes up to these facts, to avoid Zimmerman's perhaps also contentious use of "discovers" (which also may be taken to imply an earlier pre-discovery time at which God does not know, since one can perhaps only discover what was once covered to one), is that God sees these contingent facts to be true, without having made them true and without having had any say in their being made true. Of course, that's fine, at least to those of us with free-will-theist sensibilities, if it is the free creatures themselves who instead make them true. But as we have seen, at the relevant time, or relevant point in the order of explanation, those creatures, being non-(yet)-existent, cannot have done anything yet. These facts are true and limiting God's options before we even exist (and even where "*we*" *never* exist), so it isn't us who are making them true. There is, and can be, no good account on the Molinist scheme of how these God-constraining facts, which really could have been otherwise, came to be as they are.

Molinists may be unbothered by this. Why think there has to be something that accounts for how these facts came to be as they are? Can't some facts, even some contingent ones (ones that could have been otherwise), just be true? To anti-Molinist eyes, like mine, Molinists' tolerance for brute truths of different sorts (unexplained truths, both in the sense of what is responsible for their being true, and what their truth consists in (see n. 99)) is one of their most remarkable, and puzzling, features.

Even beyond their having to tolerate brute facts that look like odd candidates for such (at least to the eyes of others), Molinists really ought to be bothered here, for the problem with

¹⁰² Dean Zimmerman, "Yet Another Anti-Molinist Argument," in S. Newlands, L.M. Jorgensen, eds., *Metaphysics and the Good: Themes from the Philosophy of Robert Merrihew Adams* (Oxford UP, 2009), pp. 33–94; p. 46.

the conditionals being true without the agents involved having made them true goes beyond leaving a gap in our account of how they come to be true: It also conflicts with the freedom of the agents.

Molinists themselves can be alert to this problem in the special case where it is God who is seen as the cause of the truth of the relevant conditionals. As is perhaps most clearly explained by the great Molinist, Thomas Flint, conditionals of freedom are “pre-volitionally” true on Molinism: they are “known by God prior to his creative act of will.”¹⁰³ On a Thomist alternative to Molinism (that’s “Thomist” or “Thomism” as named after St. Thomas Aquinas, of course—not after Thomas Flint!), by contrast, conditionals of freedom are true because they are made to be true by God (Flint, p. 86). Flint sees well that this Thomist view is no good for Libertarians, for “to be a genuine Libertarian,” one “can hardly view such counterfactuals as under divine control, for if God were to determine the truth or falsity of such conditionals, they could hardly rank as conditionals of freedom” (p. 75). This seems Flint’s main reason for choosing Molinism over the “Thomistic” picture on which God does cause the relevant conditionals of freedom to be true.

But then Flint should be able to sense that the problem here isn’t just with *God* causing the relevant conditional of freedom to be true. Your freedom is no more secure if it is, say, Baal, who has made the relevant conditional of “freedom” true of you, or if something else other than you yourself has made it true – or indeed, if it is true while nothing has made it true. The real problem in its general form here is with supposing agents can be free in situations where the relevant conditional of freedom about them is already true, but is not caused to be true by the agents themselves. God causing them to be true would be just a special case of the real problem here.

27. Middle Knowledge and Simple Foreknowledge: Standing or Falling Together

The libertarian alternative to the “Molinist God” is God as portrayed in the version of the Free Will Defense we saw in sect. 22: the God who does take risks (and is not just subject to risks) in order to get the goods of our significant freedom. This is the God of what is known as “Freewill Theism” or “Open Theism,” whom you will meet more properly in the following section. First, though, here, we have to get the Molinist God and the God of Freewill Theism properly squared

¹⁰³ Thomas P. Flint, *Divine Providence: The Molinist Account* (Cornell University Press, 1998), p. 46.

off against one another, as the two views seem, at least officially and definitionally, aimed at answering different questions about God's knowledge.

I take it that the defining feature of Freewill (or Open) Theism is the denial that God has simple (non-conditional) foreknowledge of all that will happen in the future—and, in particular, that God knows what we will freely do. But while Molinism does affirm some future-directed knowledge that God has of our free actions, it is focused, as we have seen, on affirming that God has “middle knowledge” of providentially useful *conditionals* (if-then statements), which affirmation seems to be the view's defining feature. So what the Molinist officially affirms is not quite what the Freewill Theist officially denies. But though the two views are directed at somewhat different questions, at least so far as their defining features go, they do end up being opposed to one another, for the difference over whether God can know such conditionals as *If Eve is in S₄₉₈, she will do the right thing* naturally results in a correlated difference in whether, having determined to put Eve in S₄₉₈, God has simple (i.e., non-conditional) foreknowledge of the likes of *Eve will do the right thing (in S₄₉₈)*. That is, God's having middle knowledge (at least of the providentially useful kind that we are concerned with: see [note 99](#)) of what we will freely do if put in various situations stands or falls together with God's having simple (non-conditional) foreknowledge of what we will freely do. Or so I claim—though this is not just me, as here I'm agreeing with wise Molinists who also see the two as standing or falling together, but who judge that they together stand, while I see them as together falling.¹⁰⁴

I think this connection can be seen intuitively. So let's start by slightly misconstruing the Open Theist, imagining them as focused on denying the very middle knowledge that the Molinist affirms (and so as being properly opposed to Molinism). We will see this difference results in the Molinist affirming simple foreknowledge in cases of freedom and the Open Theist denying it—in other words, we will reach the denial that is in fact definitional of the Open Theist Stance, with the Molinist affirming what the Open Theist officially denies.

In our complicated world, God must manage an immensely complex system, where what situations creatures face depend on, among other things, what free choices they and others have made before on a multitude of earlier occasions. But for the purpose of seeing the connection between God's having middle knowledge and Their having simple foreknowledge, we can cut all that clutter. So, for the moment, suppose you are God, and are considering creating a simple world in which there is only one free creature, performing only one free

¹⁰⁴ See Flint, *Divine Providence*, pp. 128-130, following Alfred Fredosso's "Introduction" to Molina, *On Divine Foreknowledge* (Cornell UP, 1988), pp. 70-71. Here we go against what I've always taken to Robert Adams's suggestion (in the last full paragraph on p. 110 of "Middle Knowledge and the Problem of Evil") that we can hang on to God's having simple foreknowledge of what we will freely do even if we give up on God having middle knowledge.

action in an otherwise deterministic world. You would love to have Eve freely do the right thing in, say, S_{32} , a (completely specified) situation in which she is free to do right and free to do wrong, as the only free act in the world. That would be the perfect completion of the really great, if somewhat small, world you are planning. However, if she went wrong in S_{32} , that would totally ruin things: You would rather not create Eve at all, and just make do with a nice, but completely deterministic world that has no free creatures, than tolerate her messing up. What's a God to do?

As omniscient God, can you use your future-directed knowledge to avoid having to take any risk in order to get what you really want here—Eve doing the right thing in S_{32} ? On Molinism, as we have seen, you do know what Eve will do if put in S_{32} , and if she would in fact do the right thing, you are good to go: You can put her in S_{32} , knowing full well that everything will turn out great. However, you were subject to the front-loaded risk that the fact about what Eve would do would be hostile to your hopes, and if in fact she would sin if she were in S_{32} , you know that fact, and thereby know full well that you can't get what you want here. By contrast, as we are now (slightly mis-) construing Open Theism, it denies that God has this middle knowledge. Then, since you, as God, do not have the conditional knowledge of what Eve will do if put in S_{32} , to get what you want, you must take a risk, putting her in S_{32} without the assurance of knowing what she would do if put in that situation (though you will know whatever probabilities are involved).

Now, on both pictures, assume that you do decide to put Eve in S_{32} , and you are divinely resolute in this decision in a way that renders you certain that you will divinely do this. On the Molinist picture, we were supposing you know with certainty what Eve will do if put in S_{32} . Presumably, since you have decided to put her in S_{32} , it has turned out (as you knew all along) that she would do the right thing if put in that situation. We add that you now know that she will be in S_{32} , having now decided to put her in that situation. Quite clearly, you should now know for certain the simple (non-conditional) fact that Eve will do the right thing (in S_{32}), for you are certain both that she will be in that situation and also that she will do the right thing if she is in that situation. You have simple foreknowledge. By contrast, when we suppose with the non-Molinist (our Open Theist, as we are now construing them) that you could not know what Eve would do if she were to be in S_{32} , but we then suppose you decide to, and thereby know that you will, put her in S_{32} (you have presumably decided it is worth the risk), it seems pretty clear (though maybe not as clear as the link was on the positive side of things) that you now will *not* know with certainty either that Eve will, or that she won't, do the right thing in S_{32} . You lack simple foreknowledge. After all, you didn't know what she would do if she were to be in S_{32} , and all that's changed as we moved forward in your deliberations (whether or not they take place over an extended period of time, and so whether or not we have moved forward in time)

is that you have now decided to put her in that situation, about which you didn't know what would happen in it.

So, on this non-Molinist, Open Theist view on which God lacks middle knowledge, God also ends up lacking simple (non-conditional) knowledge of everything that will happen in the future (supposing the future contains free actions)—thereby getting our presentation of the Open Theist position up to what is in fact the view's defining feature.

28. The Freewill Theist's Risk-Taking God, the Problem of Evil, and Providential Control

This God who *takes* (and is not just subject to) risks to get the goods of our significant freedom is the God of "Free Will Theism," as the view is known, though sometimes the first two words of that label are jammed together into one, in a way we will follow, as in David Basinger's *The Case for Freewill Theism*.¹⁰⁵ The view also goes by the name of "Open Theism," which I have been using, and has been championed, on both theological and philosophical grounds, in different mixtures, by the likes of Clark H. Pinnock, Richard Rice, John Sanders, and William Hasker, as well as Basinger, and by Gregory A. Boyd, Dale Tuggy, Peter van Inwagen, and Dean Zimmerman.¹⁰⁶ It claims that we have Libertarian freedom (freedom that cannot coexist with predetermination), and typically denies that God can have certain "middle knowledge" of what

¹⁰⁵ Basinger: *The Case for Freewill Theism: A Philosophical Assessment*, Intervarsity Press, 1996.

¹⁰⁶ The first five writers mentioned in the text are the authors of the five essays that together comprise *The Openness of God: A Biblical Challenge to the Traditional Understanding of God* (InterVarsity Press, 1994), an important work in Freewill Theism. To philosophical concerns, the most important works of these five authors are probably Hasker's books, *God, Time and Knowledge* (Cornell University Press, 1998) and *Providence, Evil, and the Openness of God* (New York: Routledge, 2004). Moving to our other three authors: For Boyd, see esp. his *God of the Possible: A Biblical Introduction to the Open View of God* (Baker Books, 2000). Moving to philosophers (beyond Hasker): See Tuggy's "Three Roads to Open Theism," *Faith and Philosophy* 24 (2007): pp. 28-52. For van Inwagen, see *The Problem of Evil*, esp. pp. 79-83. The denial of middle knowledge is at p. 80. There is a discussion of simple foreknowledge starting at the bottom of that page, running through to p. 83, in which van Inwagen denies that God has comprehensive foreknowledge. For Zimmerman, see his "The A-Theory of Time, Presentism, and Open Theism," in M. Stewart, ed., *Science and Religion in Dialogue, Vol. 2* (Malden, Mass.: Wiley-Blackwell, 2010): pp. 791-809; and his "Open Theism and the Metaphysics of the Space-Time Manifold," in W. Hasker, T. Oord, and D. Zimmerman, ed., *God in an Open Universe: Science, Metaphysics, and Open Theism* (Eugene, Oregon: Pickwick Publications, 2011): pp. 125-57.

free creatures would freely do if put in situations in which they are free, leaving God to have to take risks in order to get the goods of our freedom.

This denial of middle knowledge sounds like it should help with the problem of evil. Robert Adams's famous attack on middle knowledge (discussed in n. 99), which predates the Freewill Theism movement, notes the potential connection here in its very title: "Middle Knowledge and the Problem of Evil," and in his introductory section of the paper, Adams explains:

The idea of middle knowledge emerges in recent philosophical discussion chiefly because of its relevance to the second question that I shall discuss, which is whether God could have made free creatures who would always have freely done right. More precisely: Could God have brought it about that He had creatures who made free choices, but none of whom ever made wrong choices? The relevance of this question to the problem of evil is obvious and well known. If He could have, why didn't He? If He couldn't have, that's a good enough reason why He didn't. (p. 109)

And in his closing section, where he develops the connection, Adams makes clear how the lack of middle knowledge is connected with God taking risks, and thereby with the likely occurrence of evil:

Without middle knowledge God must take real risks if He makes free creatures (thousands, millions, or trillions of risks, if each free creature makes thousands of morally significant free choices). No matter how shrewdly God acted in running so many risks, His winning on every risk would not be antecedently probable. (p. 117)

And while Plantinga, on the other side of the middle knowledge debate, makes it clear that he really does think that God has middle knowledge, he also likes to point out that ascribing this knowledge to God makes it more difficult to explain how evil exists despite the existence of God. In accounting for evil, Plantinga is taking on an added burden in supposing that God has middle knowledge. Taking the comparison in the other direction, it looks like denying that God can have middle knowledge, as I do, should make it easier to account for how there could be evil in a world governed by a perfectly good God: If actions that are free in the way the Libertarian supposes are a very great good, and if even God must take risks to get that great good, that looks like it helps in the task of squaring God's existence with evil. Since I will be arguing that the Free Will Defense fails, I too will be taking on an added burden, though like Plantinga on the other side, this burden is one my views commit me to. I suppose the unconvinced can see me as making the case that *even if* we suppose God can't have Middle Knowledge, the Free Will Defense *still* fails.

But note first, that it is one thing to explain why there is some evil or other in the world, but quite another to account for why the world is as evil as it is—and our worries here are focused on the second, more pressing, problem. While it may be unsurprising that the risk-taking of the God of Freewill Theism results in some evil, it can still be quite surprising indeed how bad things are, as Adams recognizes.¹⁰⁷

For, second, Freewill Theism is still theism: This is still omniscient, omnipotent, and perfectly good God we are talking about.

How can God still be omnipotent on this view, if They can't causally determine us to freely act as they want? Here the Freewill Theist joins the Molinist, and also other theists who believe in Libertarian free will, in answering, in brief, that omnipotence is having the power to anything that is possible, and it isn't possible to causally determine someone else's free actions. If They want to determine, and thereby control, your action, God can do that, They have that power. Or They at least have the power to cause your quasi-action, if really acting is incompatible with determinism. But where God causes your action, you don't act freely, if you really act at all. Why can't God determine your action while it is yet free? In short, because that is not possible. The Freewill Theist can thus join with Plantinga and others in this nice thought of Plantinga's that we have already seen (though I have questioned whether Plantinga is as entitled to it as are those who believe in a God who takes risks), but is worth repeating here:

Of course it is up to God whether to create free creatures at all; but if he aims to produce moral good, then he must create significantly free creatures upon whose co-operation he must depend. Thus is the power of an omnipotent God limited by the freedom he confers upon his creatures. (NN, p. 190)

But what of God's knowledge? It's here that Freewill Theism breaks with Molinism, denying that God knows what we would do if put in various situations in which we would be free, and then, denying that God, having decided to take the risks and have free creatures, anyway, knows everything that will happen in future. Using a handy, and established, term to express this lack, the Freewill Theist supposes God lacks what is known as "comprehensive foreknowledge."

But God is still omniscient. As the picture is most often developed, and as I think it is best developed, it contains the claim that there just aren't truths about what people will freely

¹⁰⁷ At the end of his paper, Adams admits that, even with the denial of middle knowledge, it still seems "very plausible" that "If God had acted differently in certain ways, He would probably have had better behaved free creatures, on the whole, than He actually has," and that this suggests that "the necessity of permitting some evil in order to have free will in creatures may play a part in a theodicy but cannot bear the whole weight of it, even if the possibility of middle knowledge is rejected" (p. 117).

do that are there to be known before the actions are taken, and omniscience is a matter of knowing all truths.¹⁰⁸ Likewise, God's lack of middle knowledge is due to there being no true conditionals of the relevant kind to be known. This is not just omniscience "in an attenuated sense," as Molinists and other critics of Open Theism might claim,¹⁰⁹ but full-blown omniscience (or "absolute omniscience," as Zimmerman phrases it¹¹⁰): knowledge of all truths. The Molinist ascribes to God knowledge of things that the Open Theist denies God, but since this difference over what God knows is based on an underlying difference over whether there are truths of the relevant kind to be known, it does not signify a difference over whether God is omniscient. For comparison (one that is quite inexact in some ways, but can still serve to make the relevant point): Suppose that I somehow got it into my head that several runners had run official sub-two-hour marathons toward the end of the Twentieth Century, and that you, by contrast, realize that nobody did that. Suppose we argue about the matter, and suppose also that we are both theists. Now, it would be true to say that I hold that, while you deny that, God knows who the first person to run a sub-two-hour marathon in the 20th Century was. I might even declare, truthfully, if perhaps misleadingly: "Wow, you don't think that God knows who the first person to run a sub-two-hour marathon in the 20th Century was!" But it would be silly,

¹⁰⁸ Here I take what Tuggy calls the "wide road" to Open Theism (in his "Three Roads": see note 106). This "wide road" is so-called because it is the one most often taken—and taken for good reasons, as Tuggy argues well. The other road is (of course) the "narrow road," on which one holds that there are truths about the future that God does not know to be true, but one offers an account of omniscience on which God counts as such despite that fact, and the "shortcut," on which one claims that all statements about future contingents are false. Tuggy points out that, among the Open Theists who preceded him, it was Hasker who did "much of the philosophical heavy lifting" (p. 28) in clarifying and defending Open Theism, but, unfortunately, Hasker took the narrow road. Tuggy reports that "The only other narrow roader I'm aware of is Peter van Inwagen" (n. 17, p. 48), citing an unpublished talk of van Inwagen's, likely because Tuggy wrote his paper before seeing the published remarks of van Inwagen's that I cite in note 106. Tuggy did some heavy philosophical lifting himself, arguing against the other roads, and ably defending the "wide road," not just comparatively as the best version of Open Theism, but as a very plausible and well-motivated view. I note that in a recent paper, "Future Truth and Freedom," *International Journal for Philosophy of Religion* 90 (2021): 109-119, Hasker finally abandons his narrow road view for the wide road.

¹⁰⁹ The quoted phrases is from the opening paragraph of Edward Wierenga's review of *The Openness of God*. Wierenga writes that the authors of that book (see n. 106, above) "hold that God is omniscient, but only in an attenuated sense. In particular, God's knowledge of the future extends only to what is causally determined by present conditions and to what God knows that he himself will do: God does not have foreknowledge of what free agents other than he will do. And divine omniscience does not include middle knowledge or knowledge of what free creatures would do in various alternative circumstances" (review in *Faith and Philosophy* 14 (1997): pp. 248-252).

¹¹⁰ Zimmerman, after noting that van Inwagen does not subscribe to the doctrine, writes: "Most of us who accept the label 'Open Theist', however, subscribe to the following doctrine: *Absolute Omniscience*: If a proposition is true, then God knows it" ("The A-Theory of Time, Presentism, and Open Theism," in *Science and Religion in Dialogue*, Vol. 2 (Blackwell, 2010).

or at least wrong, for me to say that because of this, you don't hold that God is omniscient, or that on your view God is omniscient "only in an attenuated sense." (And this would be wrong of me to say even if, contrary to fact, I was the one who was right, and some runners had run sub-two-hour marathons in the 20th Century.) We do disagree over whether God knows who the first person to run a sub-two-hour marathon in the 20th Century was, but this disagreement over what God knows is based on an underlying disagreement over whether there is a fact there to be known.¹¹¹ Likewise, the Freewill Theist will disagree with the believer in divine Comprehensive Foreknowledge over whether God knows what we will freely do before we do it, but this is based on an underlying disagreement about whether there are facts about what we will freely do that are there to be known before we perform the relevant actions. If I were not a theist, I would still likely use God as a helpful fiction to express my view about the open future, saying: "Not even omniscient God could know for certain what we will freely do before we do it!"

But being omniscient, God does still foresee every evil that occurs, including those we freely bring about, *at least as a possibility*, and has the power to prevent any of it—and all of it. So when the evil does occur, God did at least knowingly allow it.¹¹² And being perfectly good, when God does allow the evil to occur, They presumably have a sufficient reason for allowing it. And more: Beyond just seeing the evil as a *possibility*, God would also know what *probabilities* are involved, and these can sometimes get very high, as even we mortals can often see. And as I have already noted, if God really wants to make sure you do something, or to at least appear (including to yourself) to do it, They can always just make you (at least appear to) do it, and if God wants to know for certain ahead of time that you will (at least appear to) do something, They can always bindingly resolve to make you (at least appear to) do it—though in either scenario, if God ends up making you (at least appear to) do the action, then you don't act freely—and arguably, you don't act at all (but only appear to do so).

All that gives the God of Freewill Theism a great deal of control over the world.¹¹³ Indeed, given that God sees *everything* coming, at least as a possibility, and can prevent it,

¹¹¹ I use the phrasing of there not being "a fact there to be known" to make this line of defense available to some Open Theists who go the other way on the matter described in [note 108](#), taking the "narrow road." If they think the truths here are in a strong way in principle unknowable, they can perhaps still say that those facts are "not there *to be known*," even while they do think there *are* such truths about what we will freely do. Connected with my strong preference for the wide over the narrow road, I admit that I would find the use that those other Open Theists would make over this line of defense strained. And while I'm at it, I'll also admit that I find Wierenga's description of Open Theists' account of omniscience ([note 109](#)) more telling when aimed at Open Theists of the narrow type.

¹¹² "At least": on some occasions God may actively cause it, perhaps.

¹¹³ To see in some detail how God's control can play out where human freedom is involved, consider a case in which what God would most prefer is for me to freely mow the lawn on Saturday, though what's

there can be a point in saying even the likes of that God is still, on this view, “totally in control” of what happens.

Different Freewill Theists may have different attitudes toward how much control it would be good to see God as having. Some may think it is very important to preserve God’s sovereignty, in some good sense of that word which makes God’s great control over events to be very important, and may be relieved to find that God is very much in control of events on their view. Others may be mainly driven by a fear of the problem of evil, and worry that ascribing too much control to God will result in giving God too much responsibility for the evil of the world, and may therefore be dismayed to see their view result in God still being very much in control of events. (As we will see, this worry is well-founded!) But how individual Freewill Theists feel about all this doesn’t matter. What’s calling the shots here is the logic of the position. The “Theism” part of “Freewill Theism” dictates that God remains very much in control of things. There is no getting away from that. Even if a particular Freewill Theist is one who holds that God leaves the course of the world extremely open to our free influence, seldom intervening to prevent evil, their God, for better or worse, was still totally in control of the situations in which evil resulted in that They saw the evils coming, at least as possibilities, and *could have* prevented them (perhaps at the cost of our freedom), but in Their divine control of events, decided to allow them. That They often, or even habitually, so allow evil to occur when They see it (possibly) coming does not relieve God of the responsibility of having allowed the occurrence of evils that were within Their power to prevent and so were very much under Their control.

most important to Them is just that I do it at all (or that I at least appear to do it), freely or not, and suppose that as God can see, it’s extremely probable that I will freely mow the lawn on my own, and that I will do so by noon (though God doesn’t much care just when on Saturday I do it). Here, on Freewill Theism, God can bindingly resolve to determine me to mow (or, again, to at least appear to mow) the lawn on Saturday afternoon if I haven’t freely done it in the morning. In the case where I do what was extremely likely, and mow the lawn on my own on Saturday morning, we at least arguably have an action, at least on some ways of describing it—most naturally here, as my “mowing the lawn on Saturday,” supposing I can be said to be “mowing the lawn,” and not just to be apparently doing so, in the case that God coerces this “action”: and if I can’t be so described, then use the description “at least appear to mow the lawn”—that is something both that God knew with certainty that I would at least apparently do (having determined to make me at least appear to do it if I didn’t do it on my own), but that I did in fact do freely (for as things turned out I did it on my own). God knew with certainty that I would mow the lawn on Saturday (or at least would appear to), and They knew it was *extremely probable* that I would *freely* mow the lawn on Saturday. What God couldn’t foreknow with certainty is just that I would freely do the act—and that not because of a lack of omniscience, but because there was no fact there to foreknow. What facts there were, including any concerning the probability of how I would freely act, were known to God.

One way to visualize key aspects of Freewill Theism is to think of God as setting boundaries within which the world must stay as it moves forward in time, but as allowing our free actions, and perhaps other indeterministic events, to decide where, within those boundaries, the world goes. Where God sets those boundaries with divine resoluteness (and we can resolve to use “boundaries” only for ones that are so set),¹¹⁴ They foreknow with certainty that the world will stay within them, but do not know where within those boundaries things will go—though They again do know whatever probabilities there are. When we push the world to one place rather than another within the boundaries that God has set, we are always pushing it to a “place” that God decided to allow within the boundaries. This picture is incomplete in some ways, including that for some purposes the “boundaries” may be fruitfully construed as fuzzy (there may be some “places” God will *almost* certainly not allow the world to go, but will allow in the special case where an especially great value can be realized by risking going there¹¹⁵), and that, in any case, even within the boundaries, God can be active in coaxing things (and perhaps mostly us) to go in certain directions, without forcing them there. But the picture accurately portrays the vitally important feature of Freewill Theism that God can set the limits within which the world must stay. And as we will see, it is this feature that subjects the view to our problem of horrific evil.

As I understand how the terminology marking the major theological positions here works, the boundaries need not be set very widely for a position to count as Freewill or Open Theism. In fact, any degree of wiggle room at all does the trick. To get into the opposing camp of Divine Determinism, one must hold that God predetermines and foreknows *everything* that happens. On our picture, this would be the boundaries collapsing to the fully-controlling point of allowing only one path forward, with no room for any deviation. (On Molinism, God *foreknows* everything, but arguably only forecontrols where the world goes within the space of possibilities left open by the facts about what free creatures would freely do if put in various situations.) The Open Theist might well lobby for a change in the name of their position to “Not Fully Closed Theism”—mainly because of how that makes us sound better than how our competition would then be described. But I suppose most of us will hold the world to be much more open to our free influence (and also to the influence of other indeterministic happenings)

¹¹⁴ On some views of God’s relation to time, all God’s decisions about how God will act, or how God will act if certain conditions are met, can be viewed as all being automatically made with maximum resoluteness. On other views, all God’s decisions are made with maximum resoluteness, but it is not God’s relation to time so much as Their super-steadiness of will that accounts for this. On still other views, there may be a real distinction, with examples on both sides, between those of God’s decisions that are made with maximum resoluteness and those that are not.

¹¹⁵ One could insist on using “boundaries” to mark where God will definitely now allow the world to go under any circumstances. I’m just suggesting there may be situations in which it may make good sense, for certain purposes, to use “boundaries” to mark out areas where God would risk allowing the world to go only where especially great goods can be realized by doing so.

than is strictly required to make the Open view correct, and so our hearts couldn't legitimately be in these lobbying efforts.

The Freewill Theist may (and I, for one, do) like to (somewhat tendentiously, no doubt, but I think in the end quite fairly) phrase the matter of divine control like this: While God *is in control* of events on Freewill Theism, and in some very good sense (that of seeing everything coming, at least as a possibility, and being able to prevent any of it—though sometimes, and perhaps a lot of times, deciding not to prevent it), *totally* in control of them (and, I suppose, in a good sense, “totally sovereign”),¹¹⁶ God is *not* totally *controlling*.¹¹⁷ And why is God not totally controlling? At least in part, in order to realize the values of freedom that can be obtained by not being totally controlling. And are those values really so great? As we are about to see in **Part Five**: Potentially, yes, very much so. They may be immensely great.

¹¹⁶ I think of this as the way that an omnipotent and omniscient being remains “in control,” and even “totally in control,” of events, even where they forego being totally “controlling.” But there is also a sense in which the God of Open Theism is *not* totally “in control” of events—and on which They would have been more “in control” of events if They had decided to “skip that whole free creature idea” and made do with a more dreary, completely deterministic world. (And even now, They can come to be more “in control” by now taking over (“taking control of,” I suppose) what happens and controlling everything.) They would have then been not only totally controlling, but also, in a good sense, “in control” of events in a way They have chosen not to be when They went ahead with idea of having free creatures. But the way in which the God of Open Theism *is* “in control,” and even “totally in control,” of events is also important.

¹¹⁷ The theme of God not being controlling is an important emphasis in the work of Thomas Jay Oord, especially, of course, in his aptly titled book, *The Uncontrolling Love of God* (InterVarsity Press, 2015; subtitle: *An Open and Relational Account of Providence*). Oord's God is not only not totally controlling, but is also not totally in control in the way I am seeking to hang on to. My efforts in this book can be seen as an attempt to get as much relief from the problem of evil as can be had by appealing (largely) to the values God can realize by not being controlling, even as They, as omnipotent and omniscient, remain in an important way totally in control of events. Oord's work is an exploration of another approach—the path less traveled (in my experience)—by viewing God as more limited in power than They are on “standard theism”, as in (to stick with what is contained in the very titles of his books) his follow-up book, *God Can't: How to Believe in God and Love after Tragedy, Abuse, and Other Evils* (SacraSage Press, 2019).

Part Five: The Meaning of “Free” and the Potentially Immense Importance of Libertarian Freedom

29. The Meaning of “Free” in Philosophy-Speak: An Elucidation by Means of Examples

Crucial to seeing both how immensely important and also uncontrollable Libertarian freedom might be is an accurate understanding of what “free” means in discussions like this. Such an understanding will also aid us in avoiding mistakes that can arise from confounding this use of “free” with distinct, but related and common, uses of the word. I will be saying that we are using “free” in its “philosophical sense,” but, importantly, I don’t mean some strange, institutional sense that members of some exclusive guild of philosophers happen to have given it. Rather, this is a sense which ordinary speakers of English, with no special training, can, and often do, find themselves quite naturally using, especially when their conversation takes certain theological or philosophical turns. And it is the sense in which I have been using the term, trusting even those of you with no philosophical training to pick up on. The time has come to explain that sense. So, for those who have been yelling impatiently, “But what do you mean by ‘free’?!” here we go. But note that I don’t take myself just to be stipulating how I have been and will be using the term, but to be explaining how the term is generally best understood in theological and philosophical discussions like the one we are having.

The best way to locate and clarify our use of “free” is by means of cases, utilizing an account like the following, where *these* and *those* refer to examples we will go on to give:

An agent’s action is free in the philosophical sense if and only if it has that (perhaps complex) property, (a) the possession of which sometimes (in one natural use of the term) leads us to call an action “free” (and the lack of which sometimes, in the same natural use of the term, leads us to deny that an action is “free”), (b) that is had by *these* actions, and not by *those* (at least apparent) actions, and (c) the having of which means that the agent has the kind of control over their action needed for them to be at all morally responsible for it.

As clause (a) in our account makes clear, I’m not trying to give a *definition* that would allow someone who started with no understanding of what “free” means, nor how it is used, to come to know that, but rather to give an *explanation* of meaning that directs someone who does have a normal understanding of “free,” but for whom the meaning of the term slides around on them, perhaps in ways they likely don’t fully understand, to zero in on a particular meaning that they likely already use, or at least naturally would use, especially in discussions about matters

like those we are currently considering, but that they likely don't clearly differentiate from other uses.¹¹⁸ I am relying on my audience's prior competence in using "free." And we are zeroing in on this use in a natural way that is likely closely related to how we come to understand the use of the term in question in the first place: by means of examples of actions that are clearly "free" in the intended sense, and those actions (or at least apparent actions: the need for this wrinkle will emerge in section 32) that clearly are not, along with some commentary on those examples. Clause (b) of our account would seem to promise several

¹¹⁸ For those familiar with van Inwagen's "How to Think about the Problem of Free Will" (*Journal of Ethics* 12 (2008): 327–341), and are wondering how my procedure here relates to the guidance he gives there, here is my take. I realize I am in danger of being in violation of the bit of stern guidance that, despite its occurring in a footnote, that paper is best known for, at least in certain circles in which I move: Here is that note in its entirety (and no, van Inwagen doesn't give any support for his guidance elsewhere in the paper):

Whatever you do, do not define "free will" this way: "Free will is whatever sort of freedom is required for moral responsibility" (or "Free will consists in having whatever sort of access to alternative possibilities is required for moral responsibility"). (p. 329, n. 2)

Now, I am explaining the meaning of "free", and not "free will", but that choice of a different starting term does not absolve. A bit more seriously, my account, indented above in the text, being a bit more involved, is noticeably different in content from either of the examples of bad definitions van Inwagen lists in the indented above in this note. Still, I am taking the opposite turn from van Inwagen in a key fork in the road: In my explicating of the meaning of my basic freedom term, I direct readers to the relevant sub-type of a key type of thing by telling them to focus on the sub-type needed for moral responsibility. Van Inwagen, by contrast, insists on the other path, where one uses some clause in the ballpark of "can do otherwise" or (as he likes it) "is able to do otherwise" to define his key term, making no appeal to what's needed for moral responsibility in his definition (but leaving it as a substantive option to take freedom so defined to be needed for responsibility). Still, I believe I escape his ire, at least here, because I am explicitly not giving a definition. Of course, then, I escape censure on this point only at the cost of putting myself in jeopardy of running afoul of what is literally (since he numbers his bits of advice) van Inwagen's number 1 piece of advice which is:

(1) Define every term you use that is neither a word or phrase of ordinary language nor a technical term of some discipline other than philosophy (in which discipline, one supposes, it will have been given an adequate definition). Do not waive this requirement in the case of some term simply because philosophers use that term a lot. (p. 328)

But this is tricky. I suspect I am an ordinary language philosopher of roughly the variety van Inwagen is revealing himself to be here and share his suspicions about unexplained philosophical terminology, but disagree about how that applies to the case at hand. One of the key differences I have with van Inwagen here is that I think that the "philosophical" use of "free," despite my upcoming contrasting of it with "ordinary" uses, is itself in the relevant way a bit of ordinary language. Since it is a use that ordinary speakers do or can naturally find themselves using, what they need is a way of zeroing in on the relevant use and distinguishing it from related uses of the term with which it can be confounded, rather than a definition. (Other differences with van Inwagen would include my skepticism about the fate of his own attempts at definition here, both in "How to Think about..." and in his later "Ability" (in his book, *Thinking about Free Will*, Cambridge University Press, 2017). I think a good reason for appealing to what's needed for responsibility in our account of the meaning of "free" is how this can make sense of what seem to be the pessimists' options for how to use the term, as explained below in Sect. 33.)

examples, and hopefully the reader will be able to supply their own, but we will make do with one positive and one negative example (so, it will be more of a *this* and *that* show, as opposed to the suggested display of *these* and *those*), though we will also discuss other examples that are in between them. (Our discussion in sect. 31 of cases that that are not paradigmatically or clearly “free” in our philosophical sense, though they seem to not-so-paradigmatically count, should also help in locating the philosophical meaning of the term.)

30. The Meaning of “Free”: Our Example of a Free Action

So we start with our example of an action that is “free” in the intended sense—and, really, as it turns out, in just about any good sense of the word I can think of. I will actually perform the action as I write this. As our clause (c) would indicate, this will be an action over which I have the kind of control needed for me to be morally responsible for it. But partly because I don’t feel like doing anything morally fraught right now, and also to make the point that an action need not actually have much by way of moral impact to feature the kind of control in question, I’ll perform a quite neutral action, for which I deserve no blame nor credit (well, other than the blame for not coming up with a better example, and perhaps for making a bit of spectacle of myself). So, as I sit here in this coffee shop, in a perfectly ordinary situation (well, as ordinary as a situation can be where someone is writing about philosophy and also performing some simple action), I clearly decide to, and then do, raise my right hand off of the table in front of me and hold it up about a foot in front of my face. To follow the great Descartes in my description of this boring action (though he was discussing other philosophical matters as he wrote), “as I stretch out and feel my hand I do so deliberately, and I know what I am doing.”¹¹⁹

How should you understand the example, based on such a short description of it? As you should generally take philosophical examples: in as ordinary and boring way as you can, consistent with the description. I’m just some ordinary guy (or, well, close enough) very ordinarily (again, as ordinarily as you can imagine it) raising his hand in a deliberate way.¹²⁰

¹¹⁹ This is from the fifth paragraph of Descartes’s First Meditation, in the John Cottingham translation: René Descartes, *Meditations on First Philosophy with Selections from the Objections and Replies*, tr. Cottingham, Cambridge UP, 1986, p. 13 (AT VII, p. 19).

¹²⁰ “Boring”: Well, against the background that it’s in relevant ways no more remarkable than countless events happening all around us that we have very much grown used to. Of course, from various big picture points of view, it may be seen as positively awe-inspiring that our world has come to include such intentional actions, and my raising of my hand (which I feel a bit embarrassed over in this crowded coffee shop, and am hoping no one notices, because it seems too boring a thing to be bothering to do)

But, crucially, in trying to understand how to construe the example (“Exactly what’s happening in your example isn’t entirely clear to me in some ways very important to our current concerns...”), you might inquire whether I was causally determined to perform the action by events over which I have no control. And here it is vital that we not specify the example by answering that question with a “yes” or a “no.” In this example of an actual, non-fictional, event, we likely don’t know that answer, anyway. (I certainly don’t here.) But in any case, and also if we’re using a fictional example (inhabiting a fictional world that could be stipulated to be deterministic, or not), our answer is that the situation is ordinary. The listener should suppose that if we are ordinarily causally determined to act as we do, then I was; and if not, then not. More generally, my action fits into the causal scheme of our world in the way deliberate human actions ordinarily do, whatever exactly that way is. So specified, most everyone will agree that our example was one of a clearly free action.

But just *most* everyone: There are dissenters. What do we say to deniers, who think, for a variety of reasons, that the contrast in cases that we will be relying on doesn’t exist, since we really don’t have the needed control over what we do even in what we take to be ordinary cases of “free” actions, like the one just described?

We ignore these dissenters—for now. We will first explain our use of “free” with reference to examples (both positive and negative) that most everyone will agree on, before then considering positions on which the distinction we have drawn is bogus. For now, you can imagine our account of the meaning of “free” being negotiated between two “happy-face” philosophers, one a compatibilist (who thinks that acting freely is compatible with being causally determined to act as you do, so long as the determination happens in a suitable way) and the other an incompatibilist (who thinks not). So, your imagined incompatibilist should be a Libertarian: though they think that our being morally responsible for our actions in incompatible with our being causally determined to perform them, they think that, thanks be to God (perhaps quite literally, as many Libertarians are theists; but, otherwise, in whatever the relevant figurative sense is here), determinism isn’t true, and we are often free. The compatibilist is perhaps best imagined (though it wouldn’t hurt to vary this a bit) as an agnostic over whether determinism is true, but one who thinks that, either way (whether our actions are all causally determined, or not), we often have the kind of control over our actions needed for us to be responsible for them, which control is perfectly consistent with our being determined to act as we do. What’s important is that both negotiators hold “happy-face” positions, in that they think that those who use “free” in the philosophical way in question are generally right (i.e., we correctly divide the cases in terms of whether the type of control and possibility of

can be thought of in tandem with the dramatic raising and swinging of an animal bone by that hominid in the “Dawn of Man” scene of Kubrik’s *2001: A Space Odyssey*, worthy of being accompanied by the same dramatic music.

responsibility are present) when we divide examples of actions that are “free” in that sense from those that are not, and that, in particular, the action we just considered does indeed feature the kind of control in question, and the one we will look at in Sect. 32 indeed does not.¹²¹ This agreement over examples will allow us to forge a common understanding of the meaning of “free” that is neutral between compatibilism and incompatibilism, and on which their difference over whether “freedom” and determinism are compatible is a substantive disagreement, and not just a verbal matter of their using “free” with different meanings. We will then consider how not-so-happy-face thinkers can enter the discussion in Sect. 33.

31. Actions that Do Not Count as “Free” in Ordinary Talk, but Are “Free” in Our Philosophical Sense

Filling out our account with examples requires us now to look at examples where “freedom” in the philosophical sense is lacking, and we will do that in the following section. But first, to clearly differentiate our philosophical use of “free” from more ordinary uses, thereby avoiding confusion, we do well to look at examples of actions that *do* count as “free” in the philosophical sense but do not so count in more common talk.

When using “free” as it is typically used in non-philosophical settings, a clear example of an action that is not done “freely” (nor “of one’s own free will”), and in fact, the first kind of example of that that would likely leap to mind would be something like this: A thief points a gun at you, orders you to carry the loot from their heist out to their getaway car, credibly threatening to shoot you if you don’t, and you comply. Here, your carrying the loot to the car is a fairly clear case of an action that is not “freely” performed in the ordinary sense. Or to get a contrasting companion case that more closely parallels our positive example, we can suppose that our thief is robbing the coffee shop I’m sitting in, and orders me to raise my right hand, threatening to shoot me if I don’t. Again, the raising of my right hand here seems a case—and actually, a fairly clear case—of an action that is not “free,” at least where that term is used as it usually is outside of philosophical concerns. However, when “free” is used in the philosophical way, these actions, while *perhaps* not quite *clearly and paradigmatically* free, do seem to count, as “free” actions. (I have crowd-sourced this judgment with a few philosophical crowds.)

¹²¹ Note, though, that our negotiators achieve this agreement over examples by means of a particular way of specifying the examples—namely, by answering the question of whether the agents in the examples of “ordinary” actions are determined to act as they do with the answer: “Just suppose this happens as it ordinarily does: If we’re ordinarily determined to act as we do, then the character in the example is determined; if not, then not.”

Philosophically speaking (if I may put it so, though readers not steeped in the discipline will find they can still follow), I am free to comply with the thief's order, and also free to refuse or ignore it, and then likely get shot (and likely be badly hurt, and possibly killed). I have very strong reason to act as I do, and it would have been extremely unwise not to, but the choice to so act was made freely. If someone were to think that these acts do clearly fail to be "free," I would then take them to be misunderstanding our current "philosophical" use of the term, likely because they were confounding it with more ordinary uses.

The ordinary use of "free," like the philosophical use, seems tied to moral responsibility, though in a somewhat different way that would require some tweaking of our clause (c) if we were to give a full explication of the ordinary meaning of the term that parallels our account of the philosophical use. (Not that I really see just how this tweaking should go.) Supposing it would have been wrong for me to carry the loot out to the thief's car if I had not been acting under threat, the threat that renders my action one that does not count as "free" in the ordinary sense also seems to get me off the moral hook, at least in most ways of imagining this case. It's not fair to blame me: I wasn't acting freely. But the ties between this "freedom", ordinarily so-called, and moral responsibility seem kind of loose, and a bit nebulous, as comes into view when we play around with the severity of the threat and the strength of the call to act against that threat.

Sometimes, rather than our judgment about moral responsibility following our call on whether or not the act in question was "free," our decision as to whether to call it "free" seems more to follow, or at least be influenced by, whether we think the threat the agent was acting under was severe enough to relieve them of moral responsibility. (I focus on threats, but other types of circumstances that provide strong reasons for acting in certain ways seem capable of functioning as ordinary "freedom" killers as well.) I have promised to meet you at a certain location, and I arrive a bit before the designated time. I would have stayed and waited for you, but a bully, yelling at me from some distance through their open window, told me to leave, credibly threatening to walk over and punch me, not with full force in the face, but only moderately hard, on the shoulder, if I refused. While operating with some misconception of the circumstances under which I left, you later confront me (we suppose you know bullies do sometimes chase people off of that location, and that you suspect the account of my leaving that you heard may be incomplete, to explain why you even bring up the issue of my freedom): "Hey, you promised to meet me at our meeting place, and I heard that you were there, but then left [freely / of your own free will] before our meeting time. What gives?" It seems to me the ordinary use of "free" is flexible enough that I can either agree that I did leave "freely," while informing you of and appealing to the threat I was acting under to nonetheless relieve me of at least some of my moral responsibility, or, probably better (since it aligns my lack of responsibility with my lack of "freedom"), I can appeal to the threat I acted under to deny that I

had really left “freely”: I don’t have to answer one way or the other here to reply accurately. As often happens with squishy terms, what matters to the properly informative use of them is not so much whether one affirms or denies, but whether one makes it tolerably clear somehow just where in the range of the term’s possible meanings one is using it so as to accurately present the facts. Well, or so it seems to me.

If this is so, then what I have been calling the “ordinary sense of ‘free’” is really an array of different senses that our ordinary uses of the term squish around among (often remaining rather vague, I think, being sharpened only as the need arises). The way in which our actions performed under extreme threat by a gun-wielder were then particularly clear examples of ordinary non-“freedom” is that they failed to be “free” in just about any ordinary sense. More moderate examples of the type we have just considered, by contrast, occur in an ordinary freedom grey zone, where some precisification of meaning is needed to get them securely on one side or the other of the line dividing “free” actions from those that are not.

We resist diving further into the ins and outs of ordinary “freedom.” (Or I suppose some of us resist, while others are relieved that we now move on.) Our goal was to further our understanding of the not-so-ordinary, philosophical use of “free” that’s in play in discussions like our current one, and we now have a good enough idea of what we are distinguishing that use from to proceed.

32. The Meaning of “Free” in Philosophy-Speak: An Example of an Action (or Is it, Really?) that Is Not Free

A negative example—a clear case of something that is not a “free” action in our intended philosophical sense—would be something like this. Suppose that far away and unbeknownst to me, mad scientists are controlling some of my actions. They don’t do this by directly affecting my wants or desires, or messing with my thought processes as I deliberate over what to do, but just directly zap¹²² the motor center of my brain in such a way that the actions they want are

¹²² The use of “zap” here shouldn’t be taken to imply that the scientists exert only a very momentary force on me, and then they and their gadgetry are passive as that very quick intervention plays out in the raising of my hand. I suppose they make my hand to raise in a way that will appear to be an ordinary, voluntary raising of my hand, and this may take a continuing intervention in the motor center of my brain, as they also receive some form of feedback (either from sensory regions of my brain, or from some kind of direct observation by the scientists of my body) on the progress my hand is making. This is a temporally extended (about a couple of seconds’ worth, perhaps) zapping, which intervention ends, we may suppose, maybe a half-second or so after my hand reaches its raised position, at which

performed. And then they cackle. Suppose they cause me to raise my right hand in this way as I sit here in the coffee shop. Our happy-face Libertarian and happy-face compatibilist agree: Though we often do act freely, *this* raising of my hand is not a free action in our intended sense.

In fact, there is a question of whether it is an action at all, or at least an action of mine (so that perhaps some of the descriptions of it in the previous paragraph should be altered): Maybe we should speak of the scientists as the raisers of my hand? I think this is the fate of any potential clear negative example when it comes to our intended use of “free.” Well, not that there are other potential actors; that won’t always be so. But that it isn’t so clear that the agent in focus is really acting at all. We seek a clear example of an action that isn’t freely performed, and we end up with something which is not clearly an action of the agent’s at all.¹²³

I take the question whether apparent actions like the one we are considering really are actions at all to be a very serious one indeed. Especially in her book, *A Metaphysics for Freedom*, Helen Steward defends “Agency Incompatibilism,” according to which agency itself, which we humans share with many other animals, and not just free action, is incompatible with determinism.¹²⁴ I won’t be supposing that Steward is right—but nor that she is wrong. That agency itself is in question in cases of determinism will be important at several points when we evaluate the cases for the compatibilism vs. incompatibilism with regard to free action.¹²⁵

time I am left to my own devices. (I imagine I would keep my hand up a second or so longer, as I continued to look at it, trying to figure out why I just raised it. I suppose I would have immediately latched on to the vague reason about to be mentioned in the text above: that I must have just felt like doing that, i.e., I must have felt like doing that for no very good reason. But I might still continue wondering after a less vague explanation, that made some dent in the matter of why I came to feel like doing that, despite the lack of a very good reason for doing so.)

¹²³ Here, I seem to be thinking right along in agreement with Ted Warfield, who, when he reaches for cases where we don’t act with “metaphysical freedom” (which seems close to, if not exactly coincident with, what I’m calling freedom “in the philosophical sense”), ends up saying this about his character: “Perhaps his activity did not even constitute an action.” Indeed, Warfield’s initial characterization of an act performed with metaphysical freedom is an act which is “genuinely one’s own.” See Warfield, “Compatibilism and Incompatibilism: Some Arguments,” in Michael J. Loux and Dean W. Zimmerman, eds., *The Oxford Handbook of Metaphysics* (Oxford University Press, 2003), pp. 613–630; at p. 614.

¹²⁴ Helen Steward, *A Metaphysics for Freedom* (Oxford University Press, 2012). Steward had defended the view in earlier papers.

¹²⁵ The degree to which our apparent action in the case before us can seem not to be a genuine action of our agent partly depends on the details of the case I have chosen. If our cackling scientists controlled this “action” by intervening in the agent’s deliberations, rather than just directly zapping the motor center of their brain (and only intervening in the agent’s deliberations in that rather indirect way: “I guess I’m raising my hand....”), as I supposed, this would not to the same extent seem a case of non-agency. Steward would deny agency even in the case where the cacklers intervene earlier in the process—and I think plausibly so. But she would then be defending a verdict about a case where initial appearances were not as favorable to her claim as in the case as I have presented it.

But for our current purpose of understanding the philosophical use of “free,” we can press ahead, only pausing to note briefly that agency itself is in question in cases of determinism. For now, I can just tell you that if you are having trouble coming up with clear negative examples of genuine-actions-that-are-unfree, that’s actually a sign that you are getting the philosophical use of “free” right. We only need examples of things that are clearly not free actions (as opposed to the difficult and perhaps impossible demand for things that are clearly the agents’ actions, and also clearly performed without freedom), and our current example will serve. We can word our account as I in fact did at the start of [sect. 29](#), adding the parenthetical phrase to our clause (b), to yield “that is had by these actions, and not by those (at least apparent) actions.” That raises the question: “Apparent to whom?” But don’t get hung up on that.

Actually, let’s get hung up on that—but only briefly. In the case of me sitting here, it would appear to anyone in the coffee shop who happened to be looking in my direction that I was raising my hand, when, as we are imagining, the scientists zapped me to make it go up. (We imagine that they cause it to be raised with a motion and at a speed that would be typical of someone intentionally raising their hand.) And if nobody else were here and looking, it would still be an apparent action of mine in that it *would have* appeared to be such if someone had been observing. But wait, wouldn’t I be freaking out, or at least showing signs of being deeply dismayed, at my hand going up, seemingly on its own? And what would observers then think? And don’t we want it to be apparent *to me* that I am raising my hand? Well, if we do want this, we can suppose our scientists do what is necessary to produce the illusory sense in their victims that the action (or “action”) was something they somehow decided to do on their own. Perhaps all that would be needed, other than a slow enough and gradual enough start to the motion, is something like releasing a tiny squirt of a chemical (some kind of chilling-out substance?) into the brain to prevent cognitive freak-out long enough (and the need may perhaps be only very momentary—and perhaps might not be needed at all) for the agent to come up with some quick rationalization for why they think that they decided to perform the action (even if, in some cases, it will only be likes of “Well, I guess I just felt like doing that”)?

So now, in our imagined case in which I’m the victim of meddlesome scientists, we have our negative example: a case of an at-least-apparent action (the at-least-apparent raising of my hand), which (whether or not it is an action of mine at all) is a clear instance of something that is not a free action of mine.

33. Pessimists and the Meaning of “Free” in Philosophy-Speak: But What if This Distinction Is Bogus?

So that completes our clarification of the “philosophical” meaning of “free”—at least as negotiated between the happy-face Libertarian and the happy-face compatibilist we were imagining. These characters can now have a substantive, and not merely a verbal disagreement about whether an action’s being “free” in the sense we’ve articulated is compatible with the agent being determined to act as they did. Having agreed that *these*, but not *those*, (at least apparent) actions are free, they can now meaningfully and substantively disagree over whether the freedom displayed by the former, but lacked by the latter, and which allowed for responsibility for actions like the former, could be had by agents who are causally determined to act as they do.

But what of holders of not-so-happy-face positions? Shouldn’t they be able to use “free” in the philosophical sense and join in the conversation?

Before considering possible accommodations, let’s very briefly meet a few of these characters. (At some key points, the introduction will be incomplete, but you should be able to get at least some feel for these positions.) Perhaps the simplest to describe is the well-named “hard determinist.” They are a “determinist”: What we imagined our happy-face compatibilist accepting as a possibility, this character accepts as fact: We are in every case causally determined to decide to act, and to act, as we do. But our hard determinist, like our Libertarian, is an incompatibilist, and thinks that such determination cannot coexist with freedom or the kind of control and responsibility tied to it. Our hard determinist then thinks that even the actions most people would *take to be* clearly free, like my raising of my hand in the ordinary coffee shop scenario, because they are in fact determined, are really ones for which the agents can be no more responsible than are the agents of the actions that most everyone agrees they bear no responsibility for. But one need not be a determinist to worry about whether we are so free as we naturally think in our “ordinary” cases. Consider the *indeterminist* who thinks that, though it doesn’t happen deterministically, our decisions to act, and our actions, including my raising of my hand, are still all just the upshot of mindless, tiny indeterministic physical events, mostly occurring in our brains, over which we have no control, in some way (in some good sense of “upshot,” I suppose) that renders us really no more responsible for any of them than for any others, nor more responsible than victims of our mad scientists are—perhaps because they are all “random” or “just a matter of chance” in some responsibility-killing way. Or consider the most interesting of these dissenters, at least to me, as this is the one whose position I find most tempting, who is simply undecided between the two we have already considered: one who has no opinion on whether determinism is true, but thinks that, either

way, we are really never free in any of our actions, however we may feel about them, because they are all just the (deterministic or indeterministic) upshot, in some responsibility-killing good sense of “upshot,” of stupid little physical events over which we have no control.

There are many pessimists about free action these days, many of them psychologists and those influenced by psychologists. I think their uses of “free” are generally best understood against the backdrop of the account we have been exploring here, as accommodations made to our account, though there are three quite different ways to go here, and clarity can be gained by finding or specifying how any particular pessimist is using the term.

So, suppose you are such a pessimist, and you survey our official account, as it now stands, which, for convenience, I repeat here:

An agent’s action is free in the philosophical sense if and only if it has that (perhaps complex) property, (a) the possession of which sometimes (in one natural use of the term) leads us to call an action “free” (and the lack of which sometimes, in the same natural use of the term, leads us to deny that an action is “free”), (b) that is had by *these* actions, and not by *those* (at least apparent) actions, and (c) the having of which means that the agent has the kind of control over their action needed for them to be at all morally responsible for it.

You don’t share the agreement of the happy-face philosophers who negotiated this account: You don’t think that agents involved can actually be any more responsible for *these* actions than for *those*. But you face a choice, which you should be able to sense is a choice between different ways of using “free” here, as opposed to a choice between substantively different positions you might take: Do you want to divide the actions that are “freely” performed from those that are not roughly as our happy-face philosophers do, but then claim that we lack the kind of control and responsibility in question even in the cases where we act “freely”? Or do you want to say that none of our actions, even *these* ones, are actually “free”?

To take the first route, you change “means that” in clause (c) of our account to “is thought to mean that.” Thought to by whom? Well, by the likes of our happy-face philosophers, and others who think that there’s an important kind of control and responsibility that attaches to agents of *these*, but not of *those*, (at least apparent) actions. The original negotiators of the account should be happy to make the modification to clause (c) and accommodate you. Now, there can be a dialogue, with all three parties using “free” in the same way, and also dividing “free” cases from those that are not in the same way, but substantively disagreeing over the status of what all three call the “free” actions, our Libertarian saying that their agents have the kind of responsibility and control over them in question, but only because they aren’t

determined to perform them;¹²⁶ the compatibilist agreeing that the control and responsibility is had by the agents of ordinary free actions, but holding that that's compatible with them being determined to act as they do (so long as they aren't determined in the wrong way, as may¹²⁷ happen in some of *those* sad cases); and our pessimist claiming that the kind of responsibility and control that the other two *think* attach only to the free actions in fact are not present even there.

If you (and remember, you, at least for now, are a pessimist) instead want to take the second option, and claim that the actions we ordinarily perform, and would ordinarily count as free, are in fact not "free," then, while you might perhaps keep the wording of the account as it is in its original form, to make clear what you are doing, and how it differs from the previous option, you can change "means that" in clause (c) of the original account to "is thought to mean that and actually does mean that." And now, since you think we don't have the kind of control and responsibility in question even with respect to our ordinary, *supposedly* free, actions, and so the property responsible for our calling actions philosophically "free" does not actually mean anything about having the needed control and responsibility, you now deny that even our ordinary actions are "free."

Which of the two options to take is, so far as I can see, just a matter of taste: Do you want to use "free" so as to join our original disputants in which cases you count as "free," or in the significance of "freedom"? Though I won't name names here, I in fact observe some pessimists switching back and forth in which option they take. I suppose that's fine, so long as we're all clear about what's going on there.

There seems to me a third option worth mentioning—one that, though I do not take it here, appeals to the brooding pessimist within me. And that is to say that, if the pessimist is right, and we're really no more responsible for *these* actions than for *those*, the whole distinction here between actions that are philosophically "free" and those that are not is just bogus, built as it is around a mistaken agreement about which (at least apparent) actions, in key cases, are ones for which their agents are responsible. And then one can say what generally

¹²⁶ Our Libertarian should probably also hold that one's freedom is trampled in various cases of one's being caused to act as one does, when this happens in the wrong way, even in cases where the causing isn't fully determining—and that this is likely what's happening in our case of the cackling scientists in Sect. 32.

¹²⁷ Our compatibilist, too (see the immediately previous note), should probably think that if our world isn't generally deterministic, then our cackling scientists from Sect. 32 probably don't determine me to raise my hand, but still do cause my (apparent) action in a way that renders me unfree. It is if our world is deterministic that what divides the free from the unfree is whether the determinism works in the right way. In either case (deterministic or indeterministic), the compatibilist will think that our freedom depends on the causes of one's action being of the right kind.

should be said about the truth-values of claims built on a terms based on such bogus distinctions: either that all claims that an action is “free,” where “free” is being used in the philosophical way, are automatically false (and that all claims that actions are not “free” are automatically true), or, what I instead lean toward, that all such claims, both affirmations and denials of philosophical “freedom,” are automatically neither true nor false. (“Not even so much as false,” as some sometimes like to put it.)

For our part, largely for reasons of ease of exposition, our use of “free” will follow the second option, which use sides with maintaining the supposed significance of calling actions “free” over dividing the cases as is normally done, should those considerations come apart, and on which, therefore, our pessimists (who do think those considerations come apart) will deny that even our ordinary actions are “free.” To make our use clear, you can change “means that” in clause (c) of the original account to “is thought to mean that and actually does mean that.” Or you can keep the official account’s wording as it is, and remember how our chosen, second approach to the term’s meaning differs from the alternative construals.

34. The Libertarian Hope and the Potentially Immense Importance of Libertarian Freedom

We can now see the potentially immense importance of Libertarian freedom. Suppose that the Libertarian about human freedom is right. (Yes, just suppose it, for now. We are just seeing the *potentially* immense importance of Libertarian freedom, and will very soon be worrying about whether it is credible to suppose that Libertarian is right.) That is, suppose that, first, we do often act freely, *as we are now understanding that term*, but that, second, this freedom is incompatible with our being determined to act as we do. Well then, under this supposition, what would have been the cost if God, in order to avoid any risk of evil, had causally determined our actions, rendering us unfree? It looks like the cost would have been absolutely staggering: We would have had no more meaningful control over or responsibility for anything we did (if indeed these could even be well called things that we did) than does the victim of our controlling meddlesome scientists of [sect. 32](#)! This would surely have drained our lives of much of their meaning and value.

A couple of points of clarification are crucial here. First concerns how to understand the phrase “Libertarian freedom.” It sounds like it names one variety of freedom among others, and this may encourage such thinking as: “Well, if we didn’t have Libertarian freedom, we could still have some other variety of freedom (perhaps ‘compatibilist freedom’?), and everything could

be fine—or at least fairly fine.” But “Libertarian freedom” doesn’t designate a variety of freedom, but rather freedom itself (as we are using “free”) according to one view (Libertarianism) of what’s needed for us to have that freedom. If the Libertarian is right about freedom, then Libertarian freedom is that immensely important freedom, and to lose one would be to lose the other.

And, again, to lose our freedom, as we are using “free,” would be to not have the kind of control over our actions that is needed to be at all morally responsible for them—which brings us to our second crucial point: The phrase “to be at all morally responsible for them” is our way of identifying the kind of control in question, but is not meant to exhaust the significance of that control. (It is in fact one of a few different ways of identifying the kind of control in question, and is the one I picked largely just to better connect with some others’ accounts of free will.) We are not here just worried about being morally responsible for any of our actions, but, since the kind of control in question also seems needed also for this, for our actions being in a very important way our own—which is something that can be valuable to us even in cases where the actions in question are morally neutral, and so not subject to praise or blame, anyway. As some put it, our having this kind of control seems needed for our actions to be in an important way something we do, rather than something that just happens to us—as the raising of my hand was just something that happened to me in our case from Sect. 32 where I’m manipulated by cackling scientists (though that is an unusual case of something just happening to me in that what happens to me in that case may seem to me like something I did). To lose this freedom would be to lose a life in which anything one does is in that way one’s own action or is in that way something that one really does, rather than something that just happens to one. And this loss would seem to empty our lives of so much of their value.

You may think: “Empty as that sounds, we would be oblivious to that emptiness.” But first, don’t be so sure that we would be unaware of our grim situation. Our natural sense of ourselves is as free beings. On our current supposition of Libertarianism, that sense is accurate, but the freedom we correctly take ourselves to have is incompatible with determinism. Keeping all of that Libertarian supposition in place about what our actual situation is, if we then ask what would have been our counterfactual situation if God had acted differently (or what would be the counterfactual situation of beings as much like us as is possible, if we could not even exist in such a situation), and determined our (or our counterparts’) actions, it isn’t clear that the closest or most likely such counterfactual situation would be one that preserves our optimistic sense of ourselves, rendering it a (happy) illusion. Why not rather think we (or they: but I’ll stop adding this possibility) would retain the feature that our sense of ourselves on this matter is accurate, so if God had instead determined all our actions, we would be grimly aware of our bleak circumstances? It’s perhaps a bit too much to suppose we’d be specifically cognizant of our lack of freedom in a world where there is no freedom to be found anywhere,

as we would likely have not had a use for the concept, but why not suppose we would at least not positively take ourselves to have this thing that we, and everything around us, lacks, and that our assessment of how meaningful our lives were would be accurate, and therefore bleak? And second, even if it somehow worked out that what would have happened in the counterfactual scenario in which we were determined is that we would have operated under an illusion of freedom and meaningfulness, this would have rendered us pathetically delusional—a very sorry state of affairs God may well seek mightily to avoid.

On the proper philosophical understanding of “free,” the issue of whether we are in that way free is no ho-hum matter. It is the question of whether we ever have the control over our actions to be at all responsible for them, or whether in the way relevant to the meaning of what we do, we have no more control over, nor ownership of, any of our actions (or of our so-called “actions”) than do the victims of our evil scientists. Hardly any threat to the value of our lives can be more menacing than that we might not in this sense ever be free.

So much so that John Martin Fischer, with whom I’ll soon be disagreeing over the compatibility of freedom with determinism, seems to me right about the importance of freedom, (even) when he writes:

Sometimes, when thinking and writing about free will, one feels, as it were, up against the wall. The stakes are so high—people’s lives depend on issues of freedom and responsibility. The very meaning of life is at stake.¹²⁸

And so much so that our freedom in this sense, and the valuable states it makes possible, might even be, to revisit Speak’s words (quoted at the head of [sect. 23](#)), goods “so valuable that they can plausibly be thought to justify God in seeking to secure them – even at the cost of the world’s horrors.” Might be. I’m just reporting uncertainty at first glance on the matter—which by itself speaks strongly to the potential value we are considering. And again, it seems great enough to cast doubt on van Inwagen’s judgment, also from the head of [sect. 23](#), and repeated here for convenience:

How could anyone possibly believe that the evils of this world are outweighed by the good inherent in our having free will? Perhaps free will is a good and would outweigh, in Theist’s words “a certain amount of evil”, but it seems impossible to believe that it can outweigh the amount of physical suffering (to say nothing of other sorts of evil) that actually exists.

Van Inwagen’s is not such an easy judgment to securely reach—certainly not one that, just upon immediate consideration, merits a “How could anyone possibly believe?”-like response.

¹²⁸ John Martin Fischer, “Introduction: A Framework for Moral Responsibility,” in *My Way: Essays on Moral Responsibility* (Oxford University Press, 2006), p. 29.

To reach a verdict on this matter responsibly, we'd have to carefully consider just how widespread and how awful horrific suffering is, and also, to somehow balance that, just how immensely important our free will is, in the sense under consideration. (From another angle: How much significance and value *would be left* if we were made to have no more significant control over nor responsibility for anything we ever do than do the victims of our meddling scientists?) We'd have to consider the lives of all the people who have been subjected to more than their fair share of horrific suffering, but then also consider whether it might even be worth it for many of them (though one naturally thinks: surely, not all of them?) to have had their lives emptied of meaning in the way that completely forgoing all freedom would involve in exchange for avoiding that suffering. And it seems this would involve a lot of careful thought to reach a reasonable answer—careful thought that is not worth our undergoing for current purposes.

For this would all be to reach a judgment on what it would be best for God to do if God faced a once-and-for-all decision of whether to causally determine all human actions, and whether Their doing so, and thereby completely forgoing all human freedom, would be worth it to avoid the vast horrific suffering and evil our world has come to contain and/or that God could have reasonably predicted it would likely come to contain. And what kills the Free Will Defense, at least so far as the suffering brought about by human free actions goes, is that God need not have faced this as a global, all-one-way-or-all-the-other, any-free-will-at-all-or-prevent-horrors decision.

But the death of the Free Will Defense is still ahead of us: We will see it below, in Sect. xx. In the meantime, the possibility that the Libertarian is right, and that there is this good that is both immense and in principle uncontrollable even to God, would seem to promise much hope for theodicy. Even if I will turn out to be right that what can qualify as the “Free Will Defense” fails, another account of how a perfectly good, omnipotent God might promote this important but uncontrollable good might be able to make room for the possibility of vastly much evil in a world governed by a perfect God.

Part Six: The Credibility of the Libertarian Hope

35. The Libertarian Hope: Initial Appearances

But is this a credible hope for theodicy? Is it a serious possibility that Libertarianism is correct?

Well, to begin with, I declare (and we will consider reasons for doubting such declarations later): Libertarianism intuitively seems right, at least to most people. First, though some come to doubt this intuition, it seems to most people that we are at least often free in the philosophical sense we're using the term: It does seem that we often have the kind of control over our actions needed to be at least somewhat responsible for them, as we do in the paradigm positive cases we briefly looked at, but wouldn't have when controlled by meddling scientists in the way we've considered. In fact, it seems to almost all of us that we are also at least often free in the more demanding ordinary senses of the term. And indeed, "seems" is extremely weak here: These are for most very strong appearances. As pessimists about human freedom often admit (though they care about what they are admitting to quite varying degrees), they have strong initial appearances lined up against them.

And second, it does seem to most, and in my experience, quite strongly, that our acting freely is not compatible with our being causally determined to act as we do. Though most philosophers seem to be compatibilists, we incompatibilists among us have initial intuition, and "the people" on our side. Indeed, as many philosophers report (we will look at a couple of these reports in the next section), and has been my own experience, their beginning philosophy students, and other non-philosophers they encounter, often have a hard time even taking compatibilism seriously. I should note that some experimental philosophers have argued that when you survey non-philosophers correctly, they turn out not to be the incompatibilists that philosophers have made them out to be. We will examine their claims in sections 44-47, where I will argue that they don't ask the right questions for testing the intuitive standing of compatibilism vs. incompatibilism.

For now, we continue with my thoroughly unscientific claims about what most people seem to think, and my until-now-implicit appeal to you: "Does it not seem that way to you, too, dear reader?" But what I have so far registered is just to say the two component claims of Libertarianism – that freedom is incompatible with determinism, and that we do act freely – each seem correct when considered individually. But does that appearance survive considering them together? The worry here is not just that the risk of being wrong grows when you consider two claims rather than just one, but that these two claims seem to have some

potential to work against one another in a way that could make it particularly problematic to consider them together.

But I think Libertarianism's appeal does remain strong when its claims are considered together. Think especially here of how things seem to you as you carefully consider, and then perform, a very deliberate action, where the appearance of Libertarianism is strongest. You can revisit here the very deliberate raising of my hand in the coffee shop, described in Sect. 30, and imaginatively project yourself into the role of the very deliberate actor, or, I suppose, you can right now actually perform just such a deliberate action yourself: Decide very deliberately to raise your hand—I suppose for the reason of seeing how things seem to you when you do so. Here is a nice and insightful description of the phenomenology of intentional action, by Timothy O'Connor, one of today's leading Libertarians. O'Connor is here at the point of questioning at which his reasons are being fingered as the potential causes of his decision to act and then of his action: "Well, why did you do it, then?"; "Well, I acted for these reasons...."; "Ah, so *they* are the cause of your action?" O'Connor's answer is that that at least is not how things seem to go:

It does not seem to me (at least ordinarily) that I am caused to act by the reasons which favor doing so; it seems to be the case, rather, that *I* produce my decision *in view of* those reasons, and could have, in an unconditional sense, decided differently.¹²⁹

We will later consider more carefully how this process positively seems to go—and how, according to the Libertarians, it actually does go—including the fascinating suggestion that one's decision and then action are caused, but not by some event or state, but by the actor themselves (Sect. xx). For now, as we stay fixed on how things seem *not* to go, note how in these cases of deliberate actions, one does not seem to be caused to decide, and then to act, as one does—by one's reasons, or by anything else. As O'Connor immediately notes, though they think it is deceptive, even determinists agree that there is this appearance of indeterminism (we elide over his lists of examples of those who agree): "This depiction of the phenomenology of action finds endorsement not only, as might be expected, in agency theorists. . .but also in determinists. . ." And, going beyond what O'Connor explicitly notes right here, there is also an appearance of incompatibilism: Not only does one seem not to be caused to decide as one does, as O'Connor does note, but this seems needed for one's freedom: It seems that, it feels like, if one were to be causally determined to act as one does, by one's reasons (or desires, or whatever), one would not be free. (Though it does seem one's freedom could survive one's receiving a non-determining push from them.) There is an appearance here of freedom, of indeterminism, and of incompatibilism, all together. For whatever this is, or is not, worth, that

¹²⁹ Timothy O'Connor, "Agent causation," in T. O'Connor, ed., *Agents, Causes, and Events* (New York: Oxford University Press, 1995): 173-200, pp. 196.

is how things seem. Or so I declare—and, again, we will later look at attempts to call into question whether things really seem this way to most people.

36. **Intuitions, Arguments, and the Credibility of the Libertarian Hope**

“Forget how things seem! We must follow the strongest *arguments* in deciding what to accept about how things really are!”

Given my thoughts about how philosophical arguments work, and what they’re good for, as found in Part One and especially in Sect. 8, you can guess what I think about this general sentiment: It hardly makes sense to malign the importance of the intuitive plausibility of a philosophical claim as compared with how it fares in argument, for what philosophical arguments do, in the best cases, is enlist intuitions, or how certain claims seem to us upon consideration, to support philosophical conclusions. Philosophical argumentation is just a more indirect way of relying on intuition—which is why it’s well worth it to discuss Libertarianism’s intuitive plausibility.

But there *are* also important arguments to consider here. Perhaps the most influential line of thought against Libertarianism is that it is difficult to see how exercises of Libertarian freedom can fit into a good general picture of the causal structure of the world, whether the world is deterministic or not. This objection can be leveled by either compatibilists or by free will skeptics, and we will consider it in Sections 49-50. But before that, we will wrestle with an argument for compatibilism (and so against Libertarianism’s incompatibilism) that is driven by the commitment to our sometimes acting freely that compatibilists tend to share with Libertarians. But still before that, we will start on the offense, with the most important argument in favor of incompatibilism, starting in the Section 38.

It turns out that if you were to choose between compatibilism and incompatibilism based on the strength of the arguments for and against each, incompatibilism would be the winner of that evaluation. Add a premise that we sometimes act freely, and you have an argument for Libertarianism that is, while far from decisive, stronger than any argument against it. Or at least so I think, will claim, and will argue. But even if you don’t agree with me on that, the hope for Libertarianism should still survive an evaluation of the arguments as a credible one. Or so I hope.

It also turns out that how well Libertarianism fares in our evaluation, both with respect to how intuitively plausible it is, and how well it stands up to the arguments against it, depends

on how the Libertarian position is developed in certain ways. As I've already just suggested (toward the end of the previous section), I will be opting for a form of the view that involves a commitment to what is called "agent causation." In choosing that explanation of how Libertarian free will fits into the causal structure of our world, I will be very much following the path of, and relying on the work of, other Libertarians, especially Timothy O'Connor, already mentioned in the previous section. But I will also be advocating a form of Libertarianism that is "Buoyant," as I will call it, in being especially resilient with respect to its commitment to our freedom. So engaging with the arguments will result in filling out the Libertarian view a bit, as well as leading to an evaluation of the view's credibility.

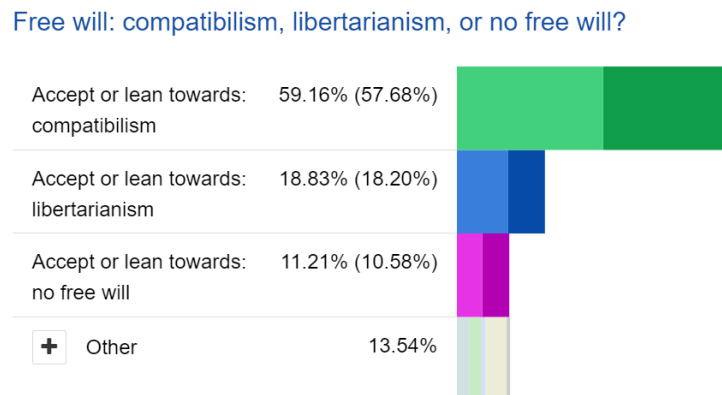
Especially since the rest of this Part of the book will be rather difficult, some readers may wish to grant the credibility of Libertarianism, at least provisionally, and skip ahead to the start of **Part Seven**, picking up the story there, to see how Libertarianism plays out in dealing with the problem of evil. (And some of those readers may find that for them, much does turn on the credibility of Libertarianism, and so they may decide to then return here, and brave the remainder of this Part.)

37. Taking Compatibilism, and Other Positions, Seriously: Thinking through the Issues from Some Starting Points

Before diving into the arguments, I should explain a feature of how my evaluation will proceed that would otherwise be surprising to some readers. Much of the discussion will read as if the main threat to Libertarianism is Compatibilism, even though, if they are at all like my beginning philosophy students, many readers are finding Compatibilism hard to even take seriously, and will think of Pessimism (no free will) as the real main threat to Libertarianism. We will look at, and then answer, the main grounds for Pessimism in the final two sections of this Part, but I should first say what I can here to help motivate those who start out very disinclined to worry that Compatibilism might be right to follow the arguments before we get to those crucial last sections. Indeed, while we're at it, I will in this section say what I can to help those whose thought begins from various starting points to see the attractions of various views they might start out with little sympathy for.

The reason I focus so much attention on the threat of Compatibilism is no doubt that Compatibilism is the majority position of professional philosophers, at least in the English-speaking world, and so would be seen as the main threat to Libertarianism—and an extremely menacing one at that—in the eyes of my discipline (and the discipline in whose domain this

matter falls). It's perhaps best here just to show the bar graph that I find so sad. One of the questions on the 2020 PhilPapers Survey of philosophers of the English-speaking world was "Free will: compatibilism, libertarianism, or no free will?", and among philosophers who met the standards for being in the survey's "Target group," the results were (the lighter portion of the top three bars represents acceptors, and the darker portion leaners):



So, nearly 60% of philosophers surveyed accept or lean toward compatibilism, while fewer than 20% accept or even lean toward Libertarianism, and considerably fewer than that go for No Free Will.¹³⁰ It is thus phrasing things rather kindly to Libertarianism to say that philosophers would see Compatibilism as the main threat to it: It's rather that philosophers as a group tend to be Compatibilists, with Libertarianism being one of the positions that, truth be told, fails to be all that much of a threat to Compatibilism's status as the leader here, even though it is the closest contender.

By extremely sharp contrast, as I have noted, many philosophers, including myself, report that in their experience, beginning students and other non-philosophers seem very much to be incompatibilists. Since many readers may find that the paths that their thoughts take as they think through our question may track those of my students, I will quickly relate my experience here, citing what seems to help many students, and so might also help readers, to take compatibilism seriously.

¹³⁰ The Survey is on-line at: <https://survey2020.philpeople.org/>. Here is the description of the Target group: "The Survey's target population includes 7685 philosophers drawn from two groups: (1) From Australia, Canada, Ireland, New Zealand, the UK, and the US (6112 philosophers): all regular faculty members (tenure-track or permanent) in BA-granting philosophy departments with four or more members (according to the PhilPeople database). (2). From all other countries (1573 philosophers): English-publishing philosophers in BA-granting philosophy departments with four or more English-publishing faculty members. An English-publishing philosopher is defined as someone with one or more publications in the PhilPapers database in a wide range of English-language venues, including English-language journals and book publishers."

And I will start by being a bit fastidious, and saying that, though this is probably due to differences in how compatibilism first arises in our teaching (though perhaps also partly due to subtle differences in our understandings of the key verb here), I would not describe my students' reactions to compatibilism as Derk Pereboom does in this often cited remark in the Introduction to his book, *Living without Free Will*:

Beginning students typically recoil at the compatibilist response to the problem of moral responsibility.¹³¹

Rather, while there may be some scattered "recoiling," I would describe the more common response more along the lines of this—also often cited—remark of van Inwagen's (though I would put things a bit more mildly):

It has seemed obvious to most people who have not been exposed (perhaps 'subjected' would be a better word) to philosophy that free will and determinism are incompatible. It is almost impossible to get beginning students of philosophy to take seriously the idea that there could be such a thing as free will in a deterministic universe.¹³²

That is, while beginning philosophy students do seem to me to mostly find compatibilism "obviously" wrong (or at least *fairly* obviously so), their rejection of it seems in most cases less a "recoiling" matter of seeming to see that those who would go the compatibilist route would be badly misweighing the relevant considerations as an inability to even see why anyone would go that route. Students are genuinely very surprised when I report to them that most philosophers, including most of them who specialize in the relevant areas of philosophy (the PhilPapers survey breaks down responses by areas of specialization), are compatibilists, and they wonder how that could be.

If readers here are like the beginning philosophy students I have encountered, many of you will not only be thinking of Pessimism as the main rival to Libertarianism, but will be harboring at least the suspicion that Pessimism is in the end the more reasonable position. That is, I am guessing that a starting point for many of you is one that seems very common among beginning students, and that can be characterized like this: 1. You find incompatibilism to be fairly obvious (to the point that you need help in seeing why philosophers tend to take compatibilism so seriously). 2. You agree, or can be fairly easily made to agree, that it is quite intuitive to suppose we often act freely: that that is how things at least initially seem. But, 3. You at least suspect that this intuition is, or at least may well be, ultimately wrong.

¹³¹ Pereboom, *Living without Free Will* (Cambridge UP, 2001), p. xvi.

¹³² Van Inwagen, *Metaphysics* (1st ed., Westview Press, 1993), p. 187.

Now, I do not, like van Inwagen, find it “almost impossible” to get students to take Compatibilism seriously, and indeed, I do not even find it all that difficult to do so—in a way we are about to look at. But I do find that getting them to take it seriously is what is needed, as they generally start out at a loss as to why they should do so.

But let’s consider why you might take Compatibilism seriously along with briefer, but hopefully helpful, accounts of the reasons for taking other positions seriously. . .

Taking our freedom seriously: As I’ve reported, I don’t have much trouble getting students to at least agree that it seems to them that we at least often act freely, in the sense of “free” we are working with (explained in sections 29-33). Just talking through something like the second paragraph of section 35 is enough to draw that acknowledgement from most. It can sometimes help, where help is needed, to appeal to a morally charged case where a character acts in what seems a horribly wrong, and horribly harmful, way. About such cases the judgment that the agent is morally responsible for their action, and in a way that victims of cackling meddling scientists are not, and so is acting “freely” in our sense, can be especially forceful. I suppose some may worry that our emotions can be clouding our judgment about such cases, but perhaps, as others will suggest, the forceful nature of the intuition instead reflects that it is here, where we judge cases of potentially great evil, where our verdict is most secure and to be trusted? At any rate, as I’ve noted, the problem here is only very rarely an inability to see how it can intuitively seem that we are free, but is rather the suspicion, which occurs with very different strengths in different inquirers, that this intuition is (or in the cases of some of the suspicious, even must be) wrong. Because this suspicion is usually driven by something like the argument for Pessimism we will look in [Section 49](#), the way to answer it is to address that argument, as we will do in [Section 50](#). My hope there is that even those who end up endorsing this reason for Pessimism will find that the reason for it isn’t the conclusive, knock-down killer that at least some of the natural Pessimists suspect.

Taking Pessimism seriously: Pessimistic students of course often encounter classmates who do not initially agree with them, and some who can’t even take their Pessimism seriously. Those in the latter camp are often those who did not need me to suggest considering agents in morally charged cases who acted horribly wrongly, as they jumped to such cases on their own, which cases to their thinking put it beyond any reasonable doubt that we often act in ways we are morally responsible for, and so which actions, on our use of the term, are free. The Pessimists typically react, at least roughly, by agreeing that, yes, it certainly can or does *seem* that we are free (and sometimes quite forcefully so), but then by trying to articulate reasons for Pessimism at least roughly along the lines of those we will look at in [Section 49](#), thinking their peers just have not yet appreciated those reasons for why our appearance of freedom is illusory. This all

suggest that for many, the case both for and against our freedom will turn on how that argument fares.

Taking Compatibilism seriously: By contrast, in my experience, what most helps students to take Compatibilism seriously is not any of the arguments for Compatibilism (and indeed, what is needed for them to give those arguments a fair hearing), but rather telling them how Compatibilists can distinguish between free and unfree actions. It is perhaps in large part because professional philosophers have a good idea of how that can be done that they are so well able to take Compatibilism seriously.

When you look at clear examples of both free and unfree actions, like our examples in [Section 30](#) and [Section 32](#), it's easy to think that what makes the difference is that in the latter, but not in the former, case, the agent is caused to act as they do by some outside force. And then it's easy to think that there can be no way on Compatibilism for free actions to be distinguished from those that are not free, since the Compatibilist thinks that even being fully determined to act as one does is compatible with one's acting freely.

But this worry can be addressed by realizing that the Compatibilist can sensibly hold that free actions are ones that are caused *in the right way*, giving some idea of what that right way is by looking at example cases. It isn't easy to fully specify what the right way is, and it can be fun—well, at least for a while: the fun does seem to fade rather quickly—to try to develop a Compatibilist account of what that right way is by testing various increasingly advanced proposals against increasingly clever examples. One can start by saying our victim of the cackling scientists does not freely raise their hand because the process by which they came to do this did not even go through their deciding to and willing to do the act. But then you should wonder about different versions of the case where the scientists do cause the subject to raise their hand by causing them to effectively decide to and will to do the act—though you must then suppose such decisions and willings are the kind of things that can be so directly caused by outside forces. The victim here still seems not to raise their hand freely, even though the process by which the action happened went through their decision and willing. A few rounds into this game, if, say, the discussion takes the kind of philosophically famous path it takes in later editions of John Locke's *Essay concerning Human Understanding*, you may find yourself proposing, say, that for an action really to be free, at least in a certain range of cases, the agent must have had the power, even where they did not exercise it, and even if, as perhaps in all things, they were causally determined not to exercise it (but causally determined in the right way!), to “suspend the desire” they are acting for in order to step back and form a judgment about whether they really want to do the act, or something to that rough effect. But then you can very sensibly wonder whether such a complicated line as the one you are drawing can really be in exactly the right place. At any rate, after a few rounds of playing the “theory and

example game,” in advance of arriving at an account that’s exactly precise and clearly right, and even after one despairs of coming up with such an exact and clearly right account, or even a clearly best Compatibilist account, or one that is clearly correct if Compatibilism is true, it can nonetheless become fairly clear that a *reasonable* division can be drawn on Compatibilism.

Those for whom the incompatibilist intuition is strong may think all this wrangling over the details of “compatibilist freedom” is doomed, anyway, because all this “deliberation” and whatnot that we’re trying to assure is or at least could be happening in “the right way” would all just be a sham if one were always determined to act as one does—since the specification of “the right way” for this to happen if the act is to be free should include some provision to the effect that one is not determined to decide to act and to act as one does. Indeed, Libertarians may suspect that real “deliberation” may not have been possible if determinism had been true (that whatever might look like deliberation in a deterministic world would not, could not, really be such), and that if deliberation could occur at all, it would be very radically different for determined beings than it is for us (and they may think similar things of other elements of the process by which we come to perform free actions). I think they are wise to harbor such suspicions (and a key aspect of what’s to come later in this Part will be my in effect advising incompatibilists to remember and to stand by these suspicions). In asking such incompatibilists to take compatibilism seriously, I am then not asking them to agree that it is *really possible* for the process by which we produce free actions, and the various elements of that process, and then ultimately for free action itself, to happen in fully determined beings. To grant that would be to give away the game. I’m rather just asking them to note that this is all *epistemically possible*—that it is possible for all we know that free action and elements of the process by which it occurs can happen in a deterministic universe—and then that some of these epistemic possibilities give compatibilists the tools needed to distinguish free from unfree actions in a way that is reasonable, even if you suspect it is ultimately, and perhaps even profoundly, wrong.

So while the Compatibilist can’t say that what makes the unfree actions unfree is that their agents are caused to perform them, they can quite sensibly hold that we may always be caused to do what we do, and what makes the unfree actions unfree is that their agents are caused to perform them *in the wrong way*—and *clear* cases of not-free-actions are caused in *clearly* wrong ways.

I find that after this exercise, many students are much better able to take Compatibilism, and the arguments for it, seriously. To the extent that this exercise works by showing the incredulous that there is a reasonable way by which free actions can be distinguished from unfree ones on Compatibilism, it is a completely legitimate way to get some of those who are reluctant to do so to take Compatibilism seriously. I have come to worry, however, that it also works in another, quite illicit way, by luring natural Incompatibilists into

accepting that the various ingredients of the process by which we come to perform free actions, and then free action itself, are really possible in a deterministic universe by getting them to play a game—the what-might-distinguish-free-actions-in-a-deterministic-universe game—that can seem to presuppose that these various ingredients of free action appealed to in the accounts unproblematically can occur in deterministic scenarios. At key points in the discussion that follows in this Part, I will be reminding Incompatibilists of the suspicions they should harbor about what is really possible in deterministic universes. These suspicions should also be kept in the back of the mind as Incompatibilists suspend them as it were from operating in the front of their mind to join in the discussion of how free actions might be marked out on a Compatibilist view. Still, insofar as it helps readers to take Compatibilism seriously to see that Compatibilism can reasonably—at least by its own lights—distinguish free from unfree actions, I encourage you to avail yourself of that help.

Taking Incompatibilism seriously: Well, as I’ve reported, students have not in my experience needed help in taking Incompatibilism seriously, as they seem naturally to start out not only taking it seriously, but having a hard time seriously doubting it. And unlike the intuition they have for the other element of Libertarianism (the claim that we sometimes act freely), students’ intuitive attraction to Incompatibilism doesn’t seem to be naturally accompanied by a suspicion that this intuition may well turn out to be wrong. I suppose doubt about Incompatibilism can start creeping in for some, to the extent they are inclined to defer to expert opinion, when I tell them about how popular Compatibilism is among professional philosophers, at least when this is combined with giving them the reason articulated above for taking Compatibilism seriously. At any rate, the best way to help them to continue to take Incompatibilism seriously (or to go back to so taking it, in the cases where that’s needed) is to see how some of the intuitions behind Incompatibilism can be articulated and turned into a good argument—indeed, in my opinion at least, an argument stronger than any on the other side. Which nicely leads us into our examination of the key arguments. . .

38. “The Best Argument for Incompatibilism”

In my judgment that Incompatibilism gets the better of the arguments, I am relying on what the great Compatibilist, David Lewis, in “Evil for Freedom’s Sake?” (a paper we have already been looking at), called “the best argument for incompatibilism.” It is a basic type of argument that

goes way back (as powerful types of philosophical arguments often do),¹³³ but that these days is often called the “Consequence Argument,” the name given to it by van Inwagen, who also gave it its most important treatment, and along the way, provided this handy, reader-friendly presentation of the argument’s basic idea:

If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born, and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (including our present acts) are not up to us.¹³⁴

Following Ted Warfield’s nice formulation, we can spell out this basic idea in a stepwise fashion (to facilitate discussion) like this:

The Consequence Argument

P1. No one has any choice about the past and laws of nature.

P2. Given determinism, the conjunction of the past and laws of nature strictly implies every truth about the future.

P3. No one has any choice about those propositions strictly implied by true propositions that no one has a choice about.

C. So, given determinism no one has any choice about any future truth (including future truths concerning the allegedly free actions and decisions of human agents).¹³⁵

The argument is driven by the principle that the logical consequences of what is not at all up to us (or “what we never had any choice about,” in Warfield’s terms) are also not at all up to us: That’s what stands behind van Inwagen’s “Therefore,” and gets spelled out in Warfield’s presentation at P3. Van Inwagen applies that principle to the state of the world long before one existed, together with the (as we are now supposing, deterministic) laws of nature. Of course, these things are not at all up to us. But if determinism is true, everything that happens, including one’s decisions and actions, are the logical consequences of earlier states of the world together with the laws of nature. So, on the principle in question,¹³⁶ if determinism is true,

¹³³ In his review of van Inwagen’s *An Essay on Free Will* (review in *Philosophy and Phenomenological Research* 46 (1986): 507-522), Gary Watson provides a helpful footnote (n. 6 on p. 510) listing at least many of the 20th Century discussions of the Consequence Argument.

¹³⁴ Van Inwagen, *An Essay on Free Will*, p. 16.

¹³⁵ Ted A. Warfield, “Causal Determinism and Human Freedom Are Incompatible: A New Argument for Incompatibilism,” *Philosophical Perspectives* 14 (2000): 167-180; p. 168.

¹³⁶ It turns out that for the principle to work, we should construe it as saying that if a *single* proposition is true and its truth was something one had no choice about, then one also has no choice about the truth of the propositions strictly implied by that single proposition: The principle is subject to (at least arguable) counterexample if we instead construe it as also saying that if several truths, the truth of each of which one had no control over, together strictly imply some other proposition, then one has no choice about that other proposition. Fortunately for the Consequence argument, the past state of the

one's own actions are also not up to one. Supposing that an action that is not at all up to you can't be a free action of yours, determinism and free action are not compatible. While van Inwagen is here supposing general natural determinism (in order to show what trouble that supposition leads to), a similar argument will apply wherever one's actions are causally determined by some outside cause over which one has no control—like, for instance, divine decrees issued long before the foundation of the world, or God perhaps intervening here and there in ways that causally determine the actions of Their creatures. Like the laws of nature and long-ago state of the world, these things would also not be up to us, and so neither would be their strict consequences, our actions (in cases where we are genuinely determined to act as we do), according to the Consequence Argument. And since freedom requires that our actions are up to us, our actions (or perhaps “actions”) are not free—given determinism. Our acting freely is not compatible with our being determined to act as we do by forces over which we don't have any control.

The principle that drives the argument—P3 in Warfield's formulation—is generally viewed as the argument's weakest link, and it is also difficult to get right, as its correct formulation is tricky. But it certainly seems like an intuitively powerful principle (and its intuitive power, insofar as it is needed to function in our argument, can survive the problems of the trickiness of the principle's formulation, in a way explained in note 137), especially as the weakest link in an argument to such a substantial and important conclusion. And Lewis, who would seem to agree that the principle, which he calls a “closure principle,” is the argument's weakest link, forthrightly admits its plausibility.¹³⁷ Eliding over his presentation of the guts of

universe and the laws of nature are not just things that individually beyond our control, but are such that they can be lumped together into a big conjunction over whose truth we never had any control. That's no doubt why Warfield formulates P2 as he does, utilizing the conjunction of the past and the laws. His P3 should be read as applying only to single premises.

¹³⁷ As I noted, the “transfer” principle (as it's often called, though Lewis calls in a “closure” principle) is tricky—and by this I mean it is difficult to formulate correctly. Indeed, Lewis's formulation of it (see the note immediately following this one for that formulation) is faulty, as was van Inwagen's different formulation of it in *An Essay on Free Will*. How then can I judge the principle to be intuitively powerful? I am judging, and accepting, the principle as ideally formulated: as formulated in the most intuitive way that still has as instances of it the applications needed by the Consequence Argument. I take Lewis's similar judgment (that the principle is plausible) and opposite verdict (he rejects it, despite its plausibility) in the same spirit. Lewis may have thought the formulation he used was free of problems and was already about the best form an advocate of the Consequence Argument could come up with. But if he were shown difficulties, he would not have taken refuge in those technical problems. I am confident that he would instead have rejected the principle in just the way that I accept it, denying it even on its most plausible, useful (to the argument-type) formulation.

The situation here is reminiscent of that of another “closure” principle: the philosophically famous closure principle for knowledge, which says *roughly* that one knows the logical implications of what one knows, and which underwrites some of the most important arguments for philosophical skepticism. There were spirited battles over the closure of knowledge, even as those on both the pro-

the argument, here is Lewis's introduction to the Consequence Argument, followed by how he thinks compatibilists (like him) should—indeed, “must”—respond:

The best argument for incompatibilism rests on a plausible principle that unfreedom is closed under implication. . . . Compatibilists must reject the closure principle.¹³⁸

and anti-closure side recognized that there were problems about the precise formulation of the principle. On the closure of knowledge, Lewis and I were on the same team: the pro-closure team, at least with application to the problem of skepticism. (See my *The Appearance of Ignorance* (Oxford UP, 2017), pp. 163-174, 279-284, for discussion of “intuitive closure,” including explanation of a kind of closure of knowledge (“Oxford closure,” I call it), that is important to a puzzle about knowledge and lotteries, that I do reject.) On the closure of “Unfree” (as Lewis puts it), we are on opposite teams, he anti-, and I pro-, though we both think the principle we differ over, properly understood, is intuitively plausible.

For those uneasy about my endorsing a principle but without having a proper formulation of what I'm endorsing: As it happens, there is a formulation of the closure principle, and a version of the Consequence argument that uses it, that work, so far as I can see, and I'll cite them in the paragraph below. But in case there are still problems that I can't see, and to show how one can proceed when one doesn't yet know a proper formulation: In both epistemology and the metaphysics of freedom, one could run the relevant arguments just with conditionals that represent the applications of the principles in question to the arguments (and in epistemology, that's how I tend to proceed.) So, for instance, in Warfield's formulation above in the text, replace P3 with:

P3c: If no one has any choice about the past and laws of nature, and if, given determinism, the conjunction of the past and laws of nature strictly implies every truth about the future, then, given determinism, no one has any choice about any future truth (including future truths concerning the allegedly free actions and decisions of human agents).

You can accept and use (as plausible, and also as correct) the conditional here. But many philosophers won't feel right about using such particular conditionals, nor the arguments driven by them, without some idea of what the general principles behind them might be.

A Consequence argument that is driven by such a general principle and that “works,” so far as I can see, was presented by David Widerker in “On an Argument for Incompatibilism,” *Analysis* 47 (1987): 37-41. After presenting counterexamples to it, Widerker weakens

β : $Np, N(p \supset q) \vdash Nq$

where “Np” abbreviates “p and no one has or ever had it within his power to make it the case that ~p” to

β' : $Np, \Box(p \supset q) \vdash Nq$

and rejiggers the Consequence argument so that it runs on β' , rather than β . Alicia Finch and Ted Warfield discuss and defend such a reconfigured argument in “The *Mind* Argument and Libertarianism,” *Mind* 107 (1998): pp. 515-528.

¹³⁸ “Evil for Freedom's Sake?”, pp. 155-156. Here's the complete passage, including the material I skip in the main text above: “The best argument for incompatibilism rests on a plausible principle that unfreedom is closed under implication. Consider the prefix ‘it is true that, and such-and-such agent never had any choice about whether’, abbreviated ‘Unfree’; suppose we have some premises (zero or more) that imply a conclusion; prefix ‘Unfree’ to each premise and to the conclusion; then the closure principle says that the prefixed premises imply the prefixed conclusion. Given determinism, apply closure to the implication that takes us from preconditions outside character—long ago, perhaps—and

That it's the "plausible" principle of the argument that compatibilists "must" reject would suggest that it is (easily, one supposes) the argument's weakest link—and also that that weakest link is not all that weak. Given a suitably modest view of what philosophical arguments are generally able to achieve, this would all seem to underwrite van Inwagen's own rather positive assessment of the success of the Consequence Argument (at least relative to other philosophical arguments):

It seems, therefore, that I now have. . . a valid argument for the incompatibility of free will and determinism whose premises seem to be true. And this, *mutatis mutandis*, is all that can be asked of any philosophical argument. At any rate, no more can be said for any known philosophical argument than this: it is valid and its premises seem to be true.¹³⁹

39. "The Best Argument for Compatibilism"

On the other side of the issue, I often have a hard time seeing what the good arguments for compatibilism are even supposed to be.

As I've noted, perhaps the most common reason for rejecting Libertarianism concerns the difficulty of fitting exercises of Libertarian freedom into a good general picture of the causal structure of the world, whether the world is deterministic or not. This objection can be leveled by either compatibilists or by free will skeptics, and we will consider it in Sections 49-50. But we will first wrestle with an argument that is more narrowly for compatibilism and that is driven by the commitment to our sometimes acting freely that compatibilists typically share with Libertarians.

In section 36, I mentioned how "the best argument for incompatibilism" could be turned into a strong argument for Libertarianism: "Add a premise that we sometimes act freely, and

deterministic laws of nature to the predetermined choice. Conclude that the choice is unfree. Compatibilists must reject the closure principle. Let's assume that incompatibilists accept it. Else why are they incompatibilists?"

¹³⁹ This is from van Inwagen's "Free Will Remains a Mystery" (*Philosophical Perspectives* 14 (2000): 1-19; p. 10), and the Consequence Argument he is specifically writing about is result of replacing the old version of principle of the argument that he had famously used in *An Essay on Free Will*, which he calls "Rule β ", with a newly redefined version of that principle that he has just given and defended. (The complete first sentence of the passage quoted in the text above, including what is there elided over, is: "It seems, therefore, that I now have what I thought I had when I thought Rule β was valid on the 'no choice' understanding of 'N': a valid argument for the incompatibility of free will and determinism whose premises seem to be true.")

you have an argument for Libertarianism that is, while far from decisive, stronger than any argument against it.” But if we Libertarians can help ourselves to the premise that we sometimes act freely, shouldn’t compatibilists be allowed that premise, too? Certainly. Simply using that premise in an argument against free will skeptics (our “Pessimists”) would be question-begging, by either Libertarians or compatibilists. And we’ll discuss later how we should evaluate the support for all these positions in light of the arguments and intuitions at play later. But we will construe the battle for now as being between happy-faced (i.e., accepting that we sometimes act freely) versions of incompatibilism (so, Libertarianism) and compatibilism, and so give the compatibilist the premise that we sometimes act freely. The obvious route from that premise to compatibilism would be to argue that determinism is compatible with our acting freely because both are true. But the other premise of that case seems unavailable, because we don’t seem to be in a position to claim that determinism is true—and Lewis certainly didn’t think he was in any position to do so.

Still, I think we should all admit that determinism *might* be true, and Lewis seems to think (and I for one agree) that the claim that we sometimes act freely is not only premise-worthy, but somehow more solid than are the other claims in play here (we will soon be identifying the competition), and one might think one can get a good reason for accepting compatibilism out of such thoughts of comparative solidity. That seems to be the driving idea of what Lewis takes to be the strongest way to make the case for compatibilism:

The best argument for compatibilism is that we know better that we are sometimes free than that we ever escape predetermination; wherefore it may be for all we know that we are free but predetermined.¹⁴⁰

Supplying the implicit premise (1) that provides the most obvious bridge that could get us from Lewis’s explicit premise (2) to his stated conclusion, and following Lewis’s own wording closely, this seems to be what he has in mind:

“Best Argument for Compatibilism”

1. If we know that free will is incompatible with determinism, then we do not know better that we are sometimes free than we know that we ever escape predetermination

But, 2. We do know better that we are sometimes free than we know that we ever escape predetermination

So, C. For all we know, free will is compatible with determinism

Sussing out what could be behind that first premise (which I supplied: so, more generally, what we need to figure out is how to get from 2 to C here), and then evaluating its plausibility, is a

¹⁴⁰ David Lewis, “Evil for Freedom’s Sake?” *Philosophical Papers* 22 (1993): 149-172, p. 155. Yes, that really is the end of the argument as presented there.

fascinating exercise, which can quickly drag one deeply into questions about the logic of the (admittedly somewhat squirrely-sounding) “know better that” relation. I certainly would not say (1) is clearly right, but it has considerable initial plausibility, which is perhaps best seen by stating the general thought that lies behind it in a way that may cut through the claim’s complications a bit (but that retains Lewis’s use of the “know better” relation): If you know that two things are incompatible, you cannot know that one of them obtains any better than you know that the other one doesn’t obtain.¹⁴¹ But beyond that quick effort to uncover the sensibility of the bridge which seems to be taking Lewis from his (2) to (C), I won’t engage in an effort to evaluate Lewis’s proposed “best argument for compatibilism” as he formulates it, for however that all plays out, the glaring problem here is that this isn’t even an argument for compatibilism, but for the mere epistemic possibility (a “for all we know” possibility) of compatibilism. And while I cannot speak for all incompatibilists, I, for one, would have been happy to just grant Lewis that epistemic possibility without argument, anyway. I mean, I am an incompatibilist, but do I *know* I’m right about that? I want to answer in the words of the (at least here very wise) philosopher, Bill Lycan (perhaps using those words in the service of a cause Lycan would not approve of): “Are you nuts? This is philosophy we are talking about.”¹⁴²

¹⁴¹ As well as into questions about the “closure” of knowledge—an issue, discussed a bit in [notes 136-137](#), above, that I am prone to obsess about, as displayed at *The Appearance of Ignorance*, pp. 161-174, 279-284. And into the interplay between closure and the “know better that” relation. Stop me now!

Thanks to those who participated in an online discussion (on Facebook) of Lewis’s argument that helped me see better what could be going on, and especially to Kyle Blanchette, Michael Kremer, and Juan S. Piñeros Glasscock.

¹⁴² William G. Lycan, “On Two Main Themes in Gutting’s *What Philosophers Know*,” *The Southern Journal of Philosophy* 51 (2013): 112-120, p. 115:

Also, suppose I have been thus passionately defending one of my own core philosophical doctrines, say functionalism regarding the mental. I have said in print that functionalism is a view that I would kill or die for. But now suppose I encounter an oracle who knows the truth of the matter, or perhaps God parts the clouds and tells me that in sixty seconds He will reveal whether or not functionalism is true. You invite me to bet \$1,000 on functionalism. Would I take that bet? Of course not. Nor \$10 for that matter. Are you nuts? This is philosophy we are talking about. In reality, I have no idea whether functionalism is true. Materialism is a conviction in Gutting’s sense, and in my work I have given reasons why I think functionalism is the least implausible form of materialism, but I could just as easily be wrong.

So, Lycan is not responding to the question of whether he knows his philosophical position is right, but whether he’d bet on it—though I’m confident that (and *might* even know that?) he’d be happy to give the similar answer to whether he knows. Note that the test of whether you’d double-or-nothing bet on your view is quite different where it is one of two options, as in the issue before us now, rather than one of several, as in Lycan’s case. For my thoughts on such matters (featuring “super-advanced aliens” who really do know the philosophical answer at issue, rather than an oracle or God), see “Do I Even Know. Any of This to Be True? Some Thoughts about Belief, Knowledge, and Assertion in Philosophical Settings and Other Knowledge Deserts,” Appendix C, pp. 268-278 of my *The Appearance of Ignorance* (Oxford

40. An Argument for the Unacceptability of Incompatibilism?

But perhaps Lewis's argument can be reconstrued so as to be for a stronger pro-compatibilism—or at least anti-incompatibilism—conclusion? Though I am happy to grant Lewis the for-all-we-know possibility of compatibilism, I could not so happily cede the related claim that compatibilism is true for all we *should accept*,¹⁴³ for that amounts to the conclusion that we should not accept incompatibilism—that incompatibilism is literally unacceptable.¹⁴⁴ That sounds like it would be bad news for those of us who accept it.

I do not think that a good reason for rejecting incompatibilism, or even for not accepting it, will result from our efforts to best understand this argument. But I think that dodging the criticism that emerges will involve developing the Libertarian position in a certain way that cuts against what many seem to expect from the Libertarian. We can learn a lot, then, from seeing how such arguments can go wrong.

Again, Lewis's explicit premise here is

2. We know better that we are sometimes free than we know that we ever escape predetermination,

though we might now also think of it in this form, converting it from a claim about (relative levels of) knowledge to one about (relative) acceptability:

2a. It is more acceptable to us that we are sometimes free than it is that we ever escape predetermination.

Either way, the claim is driven by the thought, with which I have already agreed, that our judgment that “We are sometimes free” is one that is more solid for us than is any denial of determinism we could now issue.

UP, 2017). My metaphilosophical views are roughly in line with Lycan's—though his expression of them here (“no idea”!) makes me think he just might outdo me in terms of metaphilosophical humility.

¹⁴³ On that close relation: I think that in philosophical discussions (like this one), and in other “knowledge deserts,” as I call them, we often shift into a mode of talk (and writing) in which, in various ways, we speak as if we knew to be correct philosophical claims that we do not know but merely accept. While this mode of talk is handy in ways, it does encourage some confounding of knowledge with acceptance. For some details, see my “Knowledge Deserts” appendix that I cite in the immediately previous note.

¹⁴⁴ And supposing that we should accept either incompatibilism or compatibilism, an argument against accepting the former could be extended to one for accepting the latter. But I will here keep my thoughts about that supposition and extension to myself, for I want to resist this argument before we get to the unacceptability of incompatibilism.

But the simplest way to adjust premise 1 so that it has the same kinds of considerations behind it as in its original form, but so that it now interacts with 2a rather than 2, is to change 1 to 1-ka below, in this adjusted argument:

“Best Argument”-ka

1-ka. If we know that free will is incompatible with determinism, then it is not more acceptable to us that we are sometimes free than it is that we ever escape predetermination
But, 2a. It is more acceptable to us that we are sometimes free than it is that we ever escape predetermination.

So, C. For all we know, free will is compatible with determinism

Unsurprisingly, 1-ka is about as fascinating and as plausible as is 1, and for similar reasons. But we will again resist diving in, for the above adjusted argument that features 1-ka is for the same weak conclusion that Lewis’s original argument was for—the mere for-all-we-know possibility of compatibilism—so the remarks about that original at the end of the previous section apply here as well. We were hoping for something stronger.

The way to get a Lewis-like argument for the *unacceptability* of incompatibilism is to go with something like one of the arguments below, one driven by 2, and the other by 2a:

“Best Argument”-ak

1-ak. We should not accept that free will is incompatible with determinism if we know better that we are sometimes free than we know that we ever escape predetermination
But, 2. We do know better that we are sometimes free than we know that we ever escape predetermination

So, C-a. We should not accept that free will is incompatible with determinism

“Best Argument”-aa

1-aa. We should not accept that free will is incompatible with determinism if it is more acceptable to us that we are sometimes free than it is that we ever escape predetermination
But, 2a. It is more acceptable to us that we are sometimes free than it is that we ever escape predetermination

So, C-a. We should not accept that free will is incompatible with determinism

But neither (1-ak) nor (1-aa) have the intuitive basis enjoyed by (1) or (1k). Recall (from the last paragraph of the previous section) the rough and general thought that seems to stand behind (1):

If you know two things are incompatible, you cannot know that one of them obtains any better than you know that the other one doesn’t obtain

The plausibility that general thought enjoys is due how strong a relation knowledge can seem to be: *Knowing* that two things are incompatible can seem to nail down your commitment to not accepting them both so tightly as to—arguably—preclude you from being any better positioned with respect to one of those things than you are with respect to the denial of the other. But, given how comparatively weak a relation accepting seems to be, neither (1-ak) nor (1-aa) have any intuitive basis like that behind them. They’re really just bridges by which might hope to get from (2) or (2a) to (C-a). So these arguments leave us with the question: Why should we buy 1-ak or 1-aa—which really is just the question of why we should think that the comparative judgment behind 2 and 2a, in either of those two forms, militates against the acceptability of incompatibilism?

41. Is the Libertarian’s Allegiance to Our Freedom Too Fragile?

Well, why, indeed?

Again, I certainly agree with the comparative judgment embodied in Lewis’s second premise, and either in its original (2) (supposing that “know better that” is even a coherent way of talking), or in our modified form (2a). In fact, of the Libertarian answers I give to the three yes/no questions in play here—

Libertarian answers

1. Are we sometimes free? Yes
2. Do we ever escape predetermination? Yes
3. Are freedom and determinism compatible? No

—it’s my happy-faced first answer that I share with Lewis that I take to be the most certain.

And I suspect the problem Lewis has with Libertarians, and is pressing here, is that our incompatibilism allows our freedom to be “held hostage” (as it’s often put) to indeterminism: We Libertarians think that if determinism were true we wouldn’t, and indeed couldn’t, be free. This can seem to amount to a commitment to giving up the judgment that we are sometimes free should determinism turn out to be true—which possibility we should be willing to consider, since we can’t be certain that it doesn’t obtain. And Lewis sees that fragility in our commitment to freedom as betraying the level of allegiance that we, as his fellow-happy-faced philosophers, should have to what we should both take (and what he and I both do take) to be the most solid and stable judgment in play here: that we are sometimes free. The worry that the Libertarian’s

commitment to our freedom is in this way too fragile is at any rate a common one. And I think it's the heart of Lewis's worry about incompatibilism.

In response, I think the Libertarian position, as it is best developed, in what we will call its "Buoyant" form, allows for the kind of relatively robust commitment to our freedom that I take Lewis to be seeking here, while remaining a incompatibilist view, capable of underwriting free will theodicies. Check it out:

42. Buoyant Libertarianism

"Buoyant Libertarianism" is best explained by considering general Libertarianism alongside the other possible answers to our three key questions. This chart gives all the possible combinations of "yes" and "no" answers, with the Libertarian answers occurring (bolded, but not italicized) in row 2:

		1. Do we sometimes act freely?	2. Do ever escape predetermination?	3. Are freedom and determinism compatible?
1	Happy-faced compatibilism-a	Yes	Yes	Yes
2	Libertarianism	Yes	Yes	No
3	<i>Happy-faced compatibilism-b</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>
4	Impossible row	Yes	No	No
5		No	Yes	Yes
6		No	Yes	No
7		No	No	Yes
8	Hard determinism	No	No	No

While eight sets of answers are the combinatorial possibilities here, row 4 can be eliminated as an option (and so is written in strikethrough font): If freedom and determinism are not compatible, then it cannot be both that we are sometimes free but that we never escape predetermination. So this leaves seven coherent options.

"Libertarianism" is defined by its answers to our three questions, so it can be represented precisely as a row on our chart (row 2, as I have arranged things). The same is not

true of “Buoyant Libertarianism,” which, as a sub-type of Libertarianism (in a way we will soon see), is too fine-grained to be captured by a row. But Happy-faced Compatibilism is also not defined by a row on our chart—in its case because it is too course-grained to be so captured. Being “happy-faced” means one answers question 1 positively, and being a Compatibilist means saying “yes” to question 3, but that leaves one free to answer question 2 either way—or to remain agnostic on that matter (by which we here mean: accepting neither answer). Happy-faced Compatibilism thus hovers between rows 1 and 3 on our chart.

And this gives a certain stability to the position, both in general and specifically with respect to the resiliency of the happy-facedness of its adherents (their remaining in the top half of our chart) with respect to the question: “What if determinism turns out to be true? What do you accept then?” Row-3 compatibilists just stay put, of course: They are just imagining one aspect of their position being verified. Those agnostics who actually hover undecidedly between rows 1 and 3 conditionally just settle on row 3. And Row-1 compatibilists conditionally switch to row 3, which switch could be made while remaining in the Happy-faced Compatibilist camp.

Things are very different for the Libertarian: the supposition that determinism is true threatens much greater upheaval to our thought. First, of course, to suppose that determinism is true is to suppose that the Libertarian’s answer to question 2 is wrong. The Row-1 compatibilist is also wrong about that question if determinism is true, but Row-1 compatibilists can switch their answer while remaining Happy-faced Compatibilists, and while keeping their other answers constant over this change in view, all by moving to row 3. By contrast, since Libertarianism is identified by its row on our chart, any change in answers means abandoning the view. What’s more, changing our answer to question 2 is particularly unsettling for Libertarians. For switching our answer to question 2 while holding our answers to other questions steady would land us in row 4, our impossible row! Being a Libertarian means that if you’re wrong about question 2, you’re wrong about at least one of the other questions, as well. And this makes sense; we were right to designate row 4 as impossible: If you think we are free, but also that we could only be free if determinism is wrong, then you couldn’t coherently take in the news that determinism is true while remaining an incompatibilist and while continuing to hold that we are free. Some other switching of answers would be called for to retain coherence.¹⁴⁵

¹⁴⁵ Methodological digression, taken from my *The Appearance of Ignorance*, pp. 43-44, where it is applied to a different set of claims: We actually should not be too quick to assume that we should maintain consistent beliefs or acceptances in such a situation. Especially where the members of the set of claims we are dealing with all seem to have about the same, high degree of plausibility, and superespecially when there are quite a few of them (which does not apply to our current puzzle), the option of continuing to believe all of them, while, of course, realizing that they can’t all be true, and so

And many seem to assume that the rational, or at least the natural, response for the Libertarian to have to the news that determinism is true would be to give up their claim that we sometimes act freely: If you think that two claims are incompatible, you should deny the other one upon coming to know that one of them is true. Thus, it is assumed that Libertarians should move to row 8, adopting the dreaded and bleak “Hard Determinism,” on which we never act freely, should they learn that determinism is true. This is what makes the Libertarian’s commitment to our freedom seem so precarious (and too fragile for the taste of many): It’s something they would give up upon learning some empirical fact that for all we know—since we don’t know whether determinism is true—may really be a fact. And doesn’t that make sense? Shouldn’t someone who thinks that free will and determinism are incompatible conclude that we aren’t really free upon learning that determinism is true?

But when you really think about it (here I admit, I must feign something like surprise), the Libertarian would have two salient ways of restoring themselves to coherence upon learning that determinism is true. Finding out we are wrong about question 2 would put our answers to questions 1 and 3 at odds with another. But that tension could be resolved by changing *either* of those two other answers, thereby moving to either the dreaded row 8 *or* to the happy row 3, which rows are bolded and italicized on our chart, to indicate that they are the Libertarian’s most salient options for a “backup” position of what to accept should determinism be true. (Technically, we would also have the option of changing both of our other answers, deciding we were wrong about everything, and moving from row 2 to row 7. I am assuming we wouldn’t want to do that if we didn’t have to: that would seem a Costanza-like

perhaps tempering the degree of one’s belief in each, seems an attractive possibility, and several philosophers have given strong reasons in support of the rationality of sometimes holding sets of beliefs one knows to be inconsistent. [See Henry Kyburg, “Conjunctivitis,” in M. Swain (ed.), *Induction, Acceptance and Rational Belief* (Humanities Press, 1970), pp. 55–82; Peter Klein “The Virtue of Inconsistency,” *Monist* 68 (1985): 105–35; chapter 4 of Richard Foley, *Working without a Net: A Study of Egocentric Epistemology* (Oxford UP, 1993); and David Christensen, *Putting Logic in its Place* (Oxford UP, 2004).] And of course, in such a puzzling conundrum, the strategy of simply remaining noncommittal on one or more of the claims clearly has its attractions, and indeed, relative to some goals, is just good sense. I recognize that holding inconsistent beliefs or remaining agnostic on one of the claims would often be the way to go relative to various rational goals which concern having the best picture of the world one can right now. The pressure to choose a consistent and complete (taking a position on each of the premises and the conclusions of the arguments in question) package of claims to accept (even if not to really believe, which I suspect is usually beyond reach in most philosophically interesting cases) in these situations seems to me to come from some good philosophical methodology whose purpose is hopefully to promote a better picture down the road a bit: The best way for us to proceed, and hopefully to arrive eventually at the best picture available to us, may be for us to try out various consistent and complete packages of views, perhaps accept them, defend them, attack them, evaluate their comparative merits, etc.

overreaction to learning one of our answers is wrong.¹⁴⁶) To recall our discussion from **Section 40**, the incompatibilist is not supposing that they *know* that incompatibilism is right: Our incompatibilism is just something we *accept*, so when it is put in conflict with something else we accept, it is not a foregone conclusion that it's the other acceptance that we will forsake.

If we gave up our claim that we are sometimes free upon accepting determinism, we would become Hard Determinists, moving to row 8, as discussed two paragraphs above. We can call this Libertarian who would do this a “Dense Libertarian,” given how they would tend to sink to the sad lower half of our chart upon learning determinism is true. But what we’ll call the “Buoyant Libertarian” is one who would take the other option, instead giving up their incompatibilism, and moving to row 3, thereby becoming a Happy-faced Compatibilist, upon learning that determinism is true. Buoyant Libertarianism, then, is defined not just by what its adherents’ answers to our questions actually are, but also by how they would revise those answers. The Buoyant Libertarian says: “I think we are free, and that this is incompatible with determinism. And so, I accept that determinism is false. (So far: I am a Libertarian.) But if I’m shown that determinism is true, and my incompatibilism is thereby put at odds with my conviction that we’re free, it’s my incompatibilism that I will forsake, holding on to our freedom.” This “buoyant” stance could be taken for practical reasons (or so I suppose, though I have worries about accepting positions for such “pragmatic” reasons): One could hold that the consequences of disbelieving in our own freedom (as we are using the term) would be negative, and so it is practically important that our belief in our freedom be robust with respect to determinism. But I think the stance is also the Libertarian’s, or at least my, best option for purely theoretical, “epistemic” reasons: Since our being free is the answer to our questions that we “know best,” to use a Lewis-ism, it’s the one we should stick with when that answer is put at odds with another. Well, unless there’s some good reason to instead stand by the answer we “know worst.” And I see no such reason.

¹⁴⁶ Here I allude to the episode of the old television sitcom *Seinfeld*, in which the character George Costanza, convinced that his “every instinct is wrong,” tries choosing the opposite of what he would ordinarily do in all matters, and things turn out so well for him that he adopts “the Opposite” as his life’s philosophy (“The Opposite,” episode 86, first aired on May 19, 1994). Of course, the viewer notices that as this plays out, George wasn’t reversing himself on everything, but, without realizing it, was being rather choosy about implementing his new philosophy. So, for instance, when the unemployed George amazingly gets a job interview with his beloved New York Yankees, he does, when unexpectedly getting the chance to, and very much against his instincts, berate the Yankees’ owner for his mismanagement of the organization, rather than keeping his complaints to himself, and this does somehow result in George landing his dream job. But, of course, his getting this great result also depended on his actually going to the interview when he got that chance, and accepting the job when it was offered—both choices that seem very much in alignment with, and not at all opposite to, his inclinations. I suppose that, as with other life philosophies, this one needs to be applied judiciously. But when implementing this one, just where does one “do the Opposite”—where one is inclined to, or where one is inclined not to?

The Buoyant Libertarian's commitment to our freedom is not fragile in the way feared, then, but is in fact resilient in the corresponding way hoped for: it is something that would survive learning that determinism is true.

And as it happens (though I think it completely unsurprising), the most prominent Libertarian there is, van Inwagen, takes a position on the matter needed to classify him here, giving the "Buoyant" answer: saying that it's his (very hard won) incompatibilism he would give up, hanging on to our freedom, if he learned that determinism is true.¹⁴⁷

Yet, puzzlingly, it seems often assumed that the Libertarian will be or should be what I am calling a "Dense Libertarian." Or at least this is puzzling to me. I mean, as I admit a few paragraphs above, I suppose it can seem natural to assume that someone who, like the Libertarian, thinks that two things are incompatible will conclude that one of them does not obtain when they learn that the other one does. But isn't that only natural until one realizes the Libertarian has this completely different option available to them? Once we see this other option of conditionally forsaking their incompatibilism, rather than their commitment to our freedom, why think the Libertarian won't, or shouldn't, take it?

¹⁴⁷ *An Essay on Free Will*, pp. 219-221. Van Inwagen, however, had defined "free" in a way different from the account of its meaning we are using here. I assume that his answer would remain the same if he adopted our terminology. (Recall my account of the meaning of "free" is not intended as just an explanation of how I will be using the term here, but the best way to understand "philosophical" uses of the term in general (first paragraph of sect. 29). Its scope includes van Inwagen's uses of the term—despite the vociferousness of his rejection of account of the term's meaning like mine (note 118).) Also, van Inwagen's answer was both more specific and (perhaps in keeping with that) more tentatively accepted than how I am presenting it. His answer to what he would say if science showed determinism to be true was: "I am not quite sure what I would say, but I believe I would say that (β) was, after all, invalid" (p. 219). As it happened, and as we saw in discussing his argument back in note 137, van Inwagen was led to give up his principle (β) without the aid of science, and in one of the most philosophical ways imaginable: by means of counterexamples to it appearing. His tentativeness in fingering (β) as what to give up indeed may have been due to worries about whether it was soundly formulated, and so safe from counterexample. I assume that now, with (β) having been replaced in the Consequence argument, van Inwagen's answer to our question, which is similar to the one he answered, but more course-grained (where we ask what he would think if determinism were to prove true, but at the course level of what answers to our three questions he would then accept, without getting into such details as just where the Consequence argument would be taken to have failed if one chooses to conditionally abandon one's incompatibilism) is that he would move to row 3 on our chart.

43. Buoyant Libertarianism: A Completely Legitimate Libertarian Option

I must tread lightly in trying to make sense of the assumption at issue here (in our terminology, that Libertarians should be Dense Libertarians), because, honestly, it doesn't in the end make any sense to me. But looking for help first from the work of the prominent compatibilist, John Martin Fischer, who defends the assumption, and then to some other suggestion as to what may be behind it, I'll combat that assumption in this section, defending Buoyant Libertarianism as a legitimate Libertarian option.¹⁴⁸

To set the stage, note first that Fischer is also very much on-board with what I've taken to be a point of agreement I have with Lewis (and also with van Inwagen, too, as we've now seen): that the judgment that we at least sometimes act freely is the most solid judgment we have in this area (as compared with our answers to the others of our three questions), and it is a judgment that should be stable, in that we should be ready to maintain it if we were to learn that determinism is true. In a very reflective Introductory essay to his book, *My Way: Essays on Moral Responsibility*,¹⁴⁹ Fischer divulges:

I may as well be up front about this: I am motivated in much of my work by the idea that our basic status as distinctively free and morally responsible agents should not depend on the arcane ruminations—and deliverances—of the theoretical physicists and cosmologists. That is, I do not think our status as morally responsible persons should depend on whether or not causal determinism is true. (p. 5)

Fischer's talk of dependence *may* be importantly ambiguous here, in a way we will discuss later. In some important ("metaphysical," we might say) way, the Libertarian clearly does think that our freedom depends on determinism being false. I mean, as incompatibilists, we Libertarians clearly hold that you *can't* be free if you are determined. But at least part, and maybe all, of what Fischer means here is that *our judgment* that we are free shouldn't ("epistemically," we might say) depend on our judgment concerning whether determinism is true. The "*resiliency* of our fundamental conception of ourselves" (p. 6) that Fischer says he is seeking here seems at least in part, if not in whole, a (n epistemic) resiliency in judgment: Fischer thinks that "the discovery" that determinism is true should not "make us abandon our view of ourselves" (pp. 5-6). Fischer makes clear that he thinks his compatibilism well serves this deep motivation for

¹⁴⁸ Thanks to all those who participated in a Facebook discussion on these matters on 23-24 December 2020, and especially to Matthew Benton, Troy Cross, James Kreines, and Neil Levy, for helping me think through what might be behind the assumption I here combat.

¹⁴⁹ "Introduction: A Framework for Moral Responsibility," in Fischer, *My Way: Essays on Moral Responsibility* (Oxford University Press, 2006).

resiliency.¹⁵⁰ And in other work where Fischer describes this deep motivation of his, he uses the holding-our-freedom-“hostage” talk I mentioned earlier (toward the beginning of Sect. 41):

I believe that we—you and I and most adult human beings—are morally responsible (at least much of the time) for our behavior. Further, I do not think that this very important and basic belief should be “held hostage” to esoteric scientific doctrines. For example, if I were to wake up tomorrow and read in the *Los Angeles Times* that scientists have decisively proved that causal determinism is true, I would not have any inclination to stop thinking of myself, my family and friends, and human beings in general as morally responsible.¹⁵¹

But as we’ve just seen in the previous section, the Libertarian can also have this resiliency in judgment by being a Buoyant Libertarian. While the Buoyant Libertarian’s Libertarianism of course would not survive the discovery that determinism is true, their commitment to our freedom would—and it’s this latter commitment (to “this very important and basic belief”) whose resiliency we are interested in.

So what’s wrong with Buoyant, freedom-even-should-determinism-prove-true-resilient, Libertarianism? Fischer charges that the Buoyant Libertarian (to use our term for it) is guilty of “metaphysical flip-flopping,” which he works out here, where he explains what he takes to be an advantage he has over the “flip-flopping” van Inwagen:

I am thus not subject to a weird sort of dialectic of flip-flopping to which Peter van Inwagen is vulnerable. Van Inwagen is very confident that we are morally responsible. And yet he is also very confident that certain metaphysical principles are valid and that they entail the incompatibility of causal determinism and moral responsibility. So he concludes that causal determinism is false. But van Inwagen also says that in the unlikely scenario in which he were convinced that causal determinism were true, he would not give up his view that we are morally responsible; rather, he would jettison at least one of the relevant metaphysical principles. But it seems bizarre that the tenability of such a principle should depend on the empirical thesis of causal determinism. In contrast to a Libertarian such as van Inwagen, I would not need to reconsider my relevant metaphysical principles (such as the “fixity of the past” and the “fixity of the natural laws”) under the supposition that causal determinism were discovered to be true. I am not vulnerable to this sort of metaphysical flip-flopping—or to the equally unappealing possibility of having to give up my view of human beings as morally responsible agents.¹⁵²

¹⁵⁰ Fischer calls his position “semi-compatibilism,” but in the terminology we are using, it is “compatibilism.”

¹⁵¹ Fischer, “Recent Work on Moral Responsibility,” *Ethics* 110 (1999): 93–139; p. 129

¹⁵² “Recent Work on Moral Responsibility,” p. 130.

About the end of that quotation: Fischer realizes that van Inwagen, like Fischer himself, is not vulnerable to the second “unappealing possibility” of abandoning his commitment to our freedom; Fischer’s point is that van Inwagen achieves that important invulnerability by making himself vulnerable to the first problem, the “flip-flopping,” while Fischer is vulnerable to neither.

“Flip-flopping” certainly doesn’t sound like a good thing to do, and Fischer’s use of “weird” and “bizarre” further indicates that he takes a dim view of van Inwagen’s stance here. But that attack by vocabulary gives little idea of what the problem is supposed to be. Just what is this “flip-flopping,” and why is it so bad to be disposed to it? I take it the flip-flopping in question is supposed to be a problem that the Buoyant Libertarian, but not the Dense Libertarian, is subject to: It’s what stops the Libertarian from legitimately taking our Buoyant option in the first place—indeed, it’s supposedly what makes that choice of which version of Libertarianism to adopt “weird” and “bizarre.” (Having avoided such “weirdness,” the Dense Libertarian still does fall to the problem of insufficiently resilient happy-facedness, but not bizarrely so, as I understand Fischer.) But, as we’ve noted, the Buoyant and the Dense Libertarian would each change their answer to another question as well, upon learning that that they are wrong about whether determinism is true. What makes the one change of view, as opposed to the other, such a bizarrely wrong option? I think our hint is toward the end of the above passage, where Fischer calls the maneuvering van Inwagen would have to do “*metaphysical* flip-flopping” (along with the grousing about abandoning “metaphysical principles”). The idea would be that the Libertarian’s incompatibilism is a metaphysical, or, more broadly, philosophical, view, while their belief in our freedom is more intuitive. Van Inwagen had to work hard to establish his incompatibilism, forging an important form of Consequence argument to win that aspect of his view. And other Libertarians, though they may not themselves have formulated any important arguments, are presumably resting on some kind of philosophical reasoning (perhaps due to others; perhaps their own, but implicit) in adopting incompatibilism, so that is for them, too, a rather philosophical view, as opposed to their acceptance of our freedom, which, once its meaning is clarified, is just an intuition, and somehow pre-philosophical. Being ready to abandon one’s philosophy upon learning an empirical fact that one always had to think could well be true is then to be problematically (weirdly, bizarrely?) irresolute *in one’s philosophy*: Where metaphysics is the area of philosophy in play, it is to be a metaphysical flip-flopper.¹⁵³

¹⁵³ Alternatively, I suppose, one might think of metaphysics as an especially august area of philosophy, and so might find the willingness to take back “metaphysical principles” to be especially, bizarrely unseemly, as compared with flightiness about either intuitive judgments or commitments belonging to other, less solemn areas of philosophy.

At least that's my best guess as to what the problem with opting for the Buoyant version of Libertarianism is supposed to be, at least for Fischer.¹⁵⁴ But my guess may not be worth much, because if that really is it, I don't see any real problem with Buoyancy here at all. Though the matter isn't entirely clear to me, I am happy to grant that, for myself, and, I imagine, for most Libertarians, our acceptance of our freedom is more directly intuitive, and so somehow pre-philosophical, while in some good sense, our incompatibilism is more philosophical, in being more based on reasoning. But in that case, that reasoning is ultimately just a more indirect basing of the view on intuitions (or so one hopes, for the alternative would be reasoning whose starting premises that are counter-intuitive, or at least not-so-intuitive, either of which would surely be worse), after all (though we may try to hide that from ourselves by loftily calling some of the things we intuit "principles") and shouldn't be expected to render one's judgment more respectable than it is in cases of more directly intuiting something to be true. It seems perfectly reasonable—just an apt bit of sensible metaphilosophical humility, I suppose—for one to judge one's acceptance of our freedom to be more solidly grounded than one's acceptance of incompatibilism, and therefore to be the judgment one would and should hold fast to upon learning that one of those judgments has to be jettisoned—and to not reverse that evaluation

¹⁵⁴ In a very interesting discussion of this issue, "Default Compatibilism and Narrativity: Comments on John Martin Fischer's *Ways and Stories*" (*Social Theory and Practice* 37 (2011): 35-45), at pp. 36-40, Michael Nelson proposes an understanding the problem Fischer has with Buoyant Libertarianism and a way to construe van Inwagen's stance such that it would be a principled one (as opposed to one subject to bad, unprincipled, flip-flopping). The problem, according to Nelson, is that this would allow van Inwagen to escape the problem of flip-flopping only *if* van Inwagen's Transfer principle worked, but, Nelson contends, it doesn't: Nelson writes: "I discuss van Inwagen not to recommend the response. The response assumes that the Transfer Principle is sound (given the falsity of determinism). However, there are compelling counterexamples to the Transfer Principle" (pp. 39-40). Nelson concludes: "So, in the end, I don't think van Inwagen's position survives critical scrutiny" (p. 40). A problem with this (and one sufficient to get Buoyant Libertarianism out of the trouble Nelson thinks it dies from here) is that the Transfer principle can be fixed, which would leave Nelson to conclude that van Inwagen does, and the Buoyant Libertarian generally can, have a principled stance here, so far as Nelson's case is concerned.

Fischer considers Nelson's case in his "Reply to Michael Nelson" (*Social Theory and Practice* 37 (2011): 150-156), at pp. 150-154, concluding: "But the Nelson/DeRose strategy to defend van Inwagen against dialectical infelicity (in virtue of being subject to objectionable flipflopping), despite its manifest ingenuity, seems to me to be so strong as to entail unacceptable results" (p. 153). Fischer calls this "the Nelson/DeRose strategy" because Nelson and I had discussed this issue back when we were colleagues at Yale, when I defended van Inwagen's stance as principled and legitimate, and the defense of van Inwagen (or partial defense, as Nelson in the end thinks it doesn't work because of problems with the Transfer principle) Nelson proposes is in part his own working out the suggestions I made. (The ingenuity, I suspect, is due more to Nelson's efforts than to my suggestions.) The reader will have to examine Fischer's argument against the strategy (pp. 153-154) on their own. I myself don't execute the Nelson/DeRose strategy. For better or for worse (though I of course think it's for better), my grounds for the legitimacy of the Buoyant Libertarian stance are much simpler, and are given in the text in the last paragraph of this section.

just so as to favor the judgment that is more indirectly based on intuition. So far from bizarre and weird, the choice of Buoyant over Dense Libertarianism just looks like perfectly good sense (and indeed, the choice of Dense Libertarianism is starting to look a bit strange (and not just mistaken) to me).

A related possibility for why some might think Dense Libertarianism is the version of the view that the Libertarian must or should adopt, or at least most naturally would adopt, whether or not these thoughts might be what's behind Fischer's use of "metaphysical flip-flopping,"¹⁵⁵ Perhaps the analytic/synthetic distinction, or something like it, is lurking in the background here, with the thought being: You don't give up analyticities (things that are true just in virtue of the meanings of the terms involved) because you come to know something synthetic!¹⁵⁵ But while this seems promising as an explanation for why at least some folks might suppose that Dense is the right or the natural option for the Libertarian to take, I don't think it justifies the supposition. I think it's a mistake to suppose that that people's view on whether "freedom" is compatible with determinism is part of what they *mean* by "free." This mistake is promoted in some cases by misuses of such phrases as "compatibilist freedom" and "incompatibilist freedom," which (wrongly, I think) promote the thought that compatibilists and incompatibilists mean different things by "free" and by related terms, with either their compatibilism or their incompatibilism built into the very meaning of their use of the terms. This would imply that compatibilists and incompatibilists aren't having a substantive dispute when they argue. But I think accounts like the one at sections 29-33 better capture what "free" means in philosophical and theological discussions—in significant part precisely because it verifies the sense that these are real, substantive disagreements. And on such accounts, the Buoyant Libertarian would not be reversing themselves on an analytic matter if they were to become a compatibilist.

At any rate, finding nothing in Fischer's contentions, nor in the just-considered related suggestion to make me opt otherwise, like van Inwagen so far and so much more prominently before me, Buoyant Libertarianism is the form of Libertarianism I accept and that I recommend. My grounds for this are just that I think it would make perfect sense for the Libertarian, when forced to choose between them, to stick with the part of their view that seems to them more certain, absent some good reason to instead hang on to the aspect that seems to them less certain—together with my complete inability (despite benefitting from some expert help) to locate any such good reason in the case at hand. Well, and also, of course, that I do judge that we sometimes act freely to be more certain than that incompatibilism is true—and indeed, more certain than my answer to either of the other two of our three questions.

¹⁵⁵ Thanks to Troy Cross for suggesting this.

44. “Is Incompatibilism Intuitive?”: “What if Determinism Turns Out to Be True?” Cases

The “Buoyant Libertarian” stance also helps to address recent (well, at least 21st Century, but also persisting to our current day) worries about how intuitive incompatibilism is.

In the spirit of “experimental philosophy,” and as reported in their prominent paper, “Is Incompatibilism Intuitive?,” Eddy Nahmias, Stephen G. Morris, Thomas Nadelhoffer, and Jason Turner (NMNT) set out to test philosophers’ claims that incompatibilism is intuitive or natural by surveying non-philosophers about the matter.¹⁵⁶ They did not do this in what would seem the most straightforward way: by directly asking their subjects whether freedom and determinism are compatible. This may be because our authors thought people don’t understand “determinism,” at least in the way that philosophers use the term.¹⁵⁷ The natural remedy for that would be to first explain how “determinism” should be understood, and then ask whether free will is compatible with determinism, so understood. But perhaps NMNT worried that non-philosophers would not be able to understand the explanation. Or perhaps they think of “intuitions” as paradigmatically consisting of judgments about particular cases.¹⁵⁸ Or perhaps just that our best intuitions tend to be such judgments. At any rate, they instead presented their subjects with scenarios in which determinism is true, without using the word “determinism” in their description, and then asked them about the actions of a character in the scenario—asking whether that character acted of their own free will, and whether they were morally responsible for their action.

¹⁵⁶ Nahmias, Morris, Nadelhoffer, and Turner, “Is Incompatibilism Intuitive?,” *Philosophy and Phenomenological Research* 73 (2006): 28-53. These results had been reported and discussed in the earlier paper by the same authors, “Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility,” *Philosophical Psychology* 18 (2005): 561–584.

¹⁵⁷ We find this out when they tell us about some earlier surveys (which I find myself wanting to hear more about!) that they had taken:

Notice that we did not actually use the term ‘determinism’ in the scenario. This is in part because in prior surveys we found that most people either did not know what ‘determinism’ meant or they thought it meant, basically, the opposite of free will. If people have internalized the philosophical label “the problem of free will and determinism” and come to understand ‘determinism’ to mean the opposite of free will, that would count as support for the claim that incompatibilism is intuitive only at the cost of making incompatibilism an empty tautology. Rather, the claim that incompatibilism is intuitive should amount to the claim that ordinary intuitions about free will and moral responsibility indicate a conflict with the philosophical conception of ‘determinism’. (p. 37)

¹⁵⁸ Though NMNT don’t directly express that view of what intuitions paradigmatically concern, they perhaps hint at it here, where judgments about cases are the single “for instance” example of intuitions they give: “For our purposes, we take intuitions to be propensities to make certain non-deductive, spontaneous judgments about, for instance, whether or not a particular concept applies in a particular situation” (p. 35).

And NMNT's method here can seem a good and natural response to worries about whether people are ready to directly handle the question whether freedom and determinism are compatible. But as I'll explain, I find their procedure of asking about such cases extremely problematic. And the problem isn't just for experimental philosophy surveys, but can infect your just thinking through the matter on your own. For when you consider the matter—"Ah, are free will and determinism compatible?"—it's natural to break the issue down for yourself in the way that NMNT do for their respondents: by imagining a scenario where determinism has turned out to be true, and then considering whether characters in that scenario are nonetheless morally responsible for their actions, or are acting freely, concluding that free will and determinism are compatible, or at least that they seem to be compatible, insofar as you are inclined to answer "yes." I will suggest that's a very misleading way of thinking through the issue.

Here NMNT describe the results of the testing of the first scenario they describe:

We surveyed people who had not studied the free will debate. In our first study, participants read the following scenario, drawn from a Laplacean conception of determinism:

Imagine that in the next century we discover all the laws of nature, and we build a supercomputer which can deduce from these laws of nature and from the current state of everything in the world exactly what will be happening in the world at any future time. It can look at everything about the way the world is and predict everything about how it will be with 100% accuracy. Suppose that such a supercomputer existed, and it looks at the state of the universe at a certain time on March 25th, 2150 A.D., twenty years before Jeremy Hall is born. The computer then deduces from this information and the laws of nature that Jeremy will definitely rob Fidelity Bank at 6:00 PM on January 26th, 2195. As always, the supercomputer's prediction is correct; Jeremy robs Fidelity Bank at 6:00 PM on January 26th, 2195.

Participants were asked to imagine that such a scenario were actual and then were asked: "Do you think that, when Jeremy robs the bank, he acts of his own free will?" A significant majority (76%) of participants judged that Jeremy does act of his own free will. (pp. 36-37)

Our authors ruled these results to be bad news for the claim that incompatibilism is intuitive, concluding that "pending evidence to the contrary, incompatibilism is not intuitive."¹⁵⁹

¹⁵⁹ p. 39. NMNT continue: "Obviously, these results do not thereby falsify incompatibilism. But they certainly raise a significant challenge for the common claim that ordinary people start out with incompatibilist intuitions and that, hence, the burden is on compatibilists to defend theories purported to be significant revisions of ordinary beliefs and practices. Rather, given this preliminary data, we suggest the burden is on incompatibilists to motivate a theory of free will that appears to be more metaphysically demanding than ordinary intuitions demand" (p. 39).

Other surveys have followed, including more by NMNT that they went on to report in the same paper (and we'll look at their second survey soon), and also by others, and there has been wrangling over how best to test incompatibilism against cases to see how intuitive it is. And there are further details about how NMNT asked their question that we will consider below. But to my thinking, the great problem here is inherent in this whole procedure: You should not test the intuitiveness of Incompatibilism by means of judgments about the freedom of characters in deterministic scenarios. And this is especially so for cases like this inaugural one in NMNT's study, for here, what our authors are here counting as their key anti-incompatibilist intuition—the dominant (76%) response to the question, that their character does act freely—is precisely the response that I, as a Buoyant Libertarian (and therefore as an incompatibilist), have to the example! I am asked to imagine that a scenario concerning the future of our world on which determinism is verified. So I imagine that determinism turns out to be true, and then I am asked if the character in this scenario, who certainly seems to be acting freely, is in fact acting freely. And *I say yes*, not because I am a compatibilist, but because I'm an incompatibilist of the most sensible (at least as I see things) kind: a Buoyant Libertarian who would hold on to their belief in our freedom, and give up their incompatibilism, if determinism turned out to be true. So, when I'm asked to suppose that determinism is true, and I hear about this Jeremy guy, who certainly sounds like he's acting freely, I respond in my Buoyant fashion, *conditionally* giving up my incompatibilism and judging that Jeremy is indeed acting freely. But that doesn't make my response a compatibilist one, other than in the purely conditional way that Buoyant Libertarians have a compatibilist view as their back-up position: It just reveals that I (quite sensibly) *would* embrace compatibilism should it turn out that determinism is true. The answer one gives to this question isn't a good test of whether one's intuitions are compatibilist or incompatibilist, since those holding either view can—and I think both should—answer “yes” (the agent acts freely) here. Questions of this sort are rather at most a test of whether one is an incompatibilist **and** one who would (quite unwisely, to my thinking) hold on to that incompatibilism even should it turn out that determinism is true—a “Dense Libertarian.”

Projecting myself into the role of one of the survey takers, when NMNT ask me to imagine determinism turning out to be true, but then also ask me to also imagine Jeremy nonetheless acting away, seemingly like a normal (well, normal given what I think our world is actually like) free agent, I feel a bit queasy, for, as an incompatibilist, it sounds like I'm being asked to imagine a situation that I suspect may in fact be impossible. When we turn to a fuller description of how their case was presented, it turns out I would be far from alone among NMNT's respondents in so thinking. I have so far been using NMNT's report of their results in a (2006) paper in a philosophy journal, but in a footnote there they send us to another paper they had written on these surveys (NMNT 2005) “for a complete description of this and other

studies,”¹⁶⁰ and in that complete description we learn this about how they asked their question, and what led up to it:

For this study (and those discussed below), participants were undergraduates who had not studied the free will problem. In pilot studies we found that some participants seemed to fail to reason conditionally (e.g., given their explanations on the back of the survey, some seemed to assume that the scenario is impossible because Jeremy has free will, rather than making judgments about Jeremy’s freedom on the assumption that the scenario is actual). To correct for this problem, Question 1 asked participants whether they think the scenario is possible (the majority responded “no”, offering various reasons on the back of the survey). Then they were then asked to “suspend disbelief” for the experimental question:

Regardless of how you answered question 1, imagine such a supercomputer actually did exist and actually could predict the future, including Jeremy’s robbing the bank (and assume Jeremy does not know about the prediction):

Do you think that, when Jeremy robs the bank, he acts of his own free will?

The results indicate that a significant majority of participants (76%) judged that Jeremy robs the bank of his own free will. (2005, p. 566)

When the majority answered that the scenario was impossible, I would take this to be a fairly good display (only “fairly good” because this doesn’t seem a very good way to test for that) that they’re intuitive incompatibilists, whose thought resembled my own, when I would feel a bit queasy: “Well, it sure sounds like Jeremy is acting normally and freely, but he’s in a deterministic world. I wonder if that’s even possible. I doubt it.” But NMNT seemed to be taking it as a sign that their respondents were not conditionally reasoning correctly, in a way that they had to push past. So they pushed their respondents to get past their (incompatibilist, as I think of it) tendencies to get the verdict that NMNT would take (wrongly, I think) to show them to be compatibilists. If I were a respondent, I would have been one who didn’t need this push, but would have pushed past my worries about the possibility of the scenario without any help (with just the question as reported in NMNT 2006, without the additional aspects reported in NMNT 2005). I felt only *a bit* queasy when asked about the scenario, because though I worried that it might be in fact impossible, in some good sense, there is another way in which it is clearly possible: it is possible for all I know, or epistemically possible (though I guess what I’m more exactly doing is taking it to be epistemically possible that it is metaphysically possible), for, though I think freedom is incompatible with determinism, I certainly don’t take myself to be

¹⁶⁰ 2005, p. 37, n. 15: “For a complete description of this and other studies, including methodology, statistical significance, and various objections and replies, see Nahmias et al. (2005),” referring to NMNT “Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility,” *Philosophical Psychology* 18 (2005): pp. 561-584.

certain of this. So, for all I know such a scenario could arise. And going along with that way of taking the scenario to be possible, there's also a way to "imagine" (the directive given in the survey) the scenario, despite my worries about its coherence: I can imagine it "as actual," as some philosophers of mind put this mode (and we will look at the alternative way in the next section), and I am conditionally concluding what I should judge if that scenario could and did present itself. I think the way the case is worded by NMNT asks for this kind of conditional reasoning. And as a Buoyant Libertarian, I judge that Jeremy is acting freely and is responsible for his action, because I am imagining a scenario that would make me give up my incompatibilism, if it occurred: "Ah, if that did happen, I guess I should conclude that Jeremy is free and that free will and determinism turn out to be compatible after all." But if that's how I'm judging things, it doesn't show me to be compatibilist, for I'm only supposing the scenario to be really (as opposed to epistemically) possible in a conditional way. And so my judgment that Jeremy is acting freely doesn't show me to be a compatibilist, but, again, only to be either a compatibilist or an incompatibilist of the more sensible type (as I judge things). The way their scenario was presented to nonphilosophers, even just as reported in 2006, was worded in such a way as to call for the same kind of conditional judgment from them, and the aspects of the presentation we learn of in 2005 further secure that kind of judgment. So NMNT got their positive judgment that Jeremy acts freely from a strong majority of those taking the survey, but only by making that judgment one that is not an indication of compatibilism.

45. "What of Worlds of Which Determinism Is True?", "What if Determinism Had Been True?"

If you want to think through the matter of whether compatibilism or incompatibilism is true, or discern what others think, by considering or asking others to consider deterministic scenarios, it would then seem that, you should not ask what is true if that scenario obtains (imagining the scenario as actual), but should instead consider what would have been true if the scenario had obtained, or alternatively, what is true of alternative worlds or universes (which are ways things could have been) that are not presumed to be our own. Which of these two ways the scenario is presented to survey-takers turns out to matter: In an important paper that came out early in the discussion (2008), Adina L. Roskies and Shaun Nichols reported studies of theirs in which participants were significantly less likely to judge that characters in deterministic scenarios were morally responsible for their actions if those scenarios were presented to them as deterministic universes that were not our own, than if they were asked to imagine our own

universe turning out to be deterministic.¹⁶¹ So peoples' reactions are sensitive to these different ways of considering the cases. However, as I'll urge, thinking through the matter of whether freedom is compatible with determinism by considering what you (or survey-taking others) want to say about deterministic scenarios is hopelessly beset by problems, even when those scenarios are presented as other worlds, not presumed to be our own (or when they are even explicitly stipulated not to be our own).

A deterministic universe not our own was actually NMNT's second case, presented in 2006 immediately after that first case we discussed in the previous section. As you'll see, they think the important difference between their first and their second survey concerns, not the difference in mode of presentation that we are discussing, but rather how simple and salient determinism is made, as they've been worrying that their presentation of determinism was too complicated in their first scenario:

With this in mind, we developed a second scenario using a simpler, and perhaps more salient, presentation of determinism:

Imagine there is a universe that is re-created over and over again, starting from the exact same initial conditions and with all the same laws of nature. In this universe the same conditions and the same laws of nature produce the exact same outcomes, so that every single time the universe is re-created, everything must happen the exact same way. For instance, in this universe a person named Jill decides to steal a necklace at a particular time, and every time the universe is re-created, Jill decides to steal the necklace at that time.

The results were similar to those above. In this case the participants were asked both to judge whether Jill decided to steal the necklace of her own free will and whether "it would be fair to hold her morally responsible (that is, blame her) for her decision to steal the necklace." Most participants offered consistent judgments; overall, 66% judged that Jill acted of her own free will, and 77% judged her to be morally responsible. (p. 38)

So, though people have a greater tendency to respond negatively when asked about the freedom of characters in deterministic scenarios if those scenarios are presented as alternative worlds, just which responses one will get to such questions are sensitive to a lot of different factors, and it is possible to ask about even alternative deterministic worlds such that you will get a majority of positive answers, as this survey shows.

Before considering what to say about whether determined-but-necklace-stealing Jill (and other denizens of alien deterministic universes) is acting of her own free will, let's start by considering what Buoyant Libertarians, in more general terms, should think about what

¹⁶¹ Roskies and Nichols, "Bringing Moral Responsibility Down to Earth," *The Journal of Philosophy* 105 (2008): pp. 371-388.

deterministic worlds would be like—what worlds other than our own and that are deterministic are like, or what our world would have been like if determinism had been true of it. In either case, we’re not thinking about how things *are* if determinism *is* true, but the very different matter of how things *would have been* if determinism *had been* true. (This because possible worlds represent different ways the world could have been.) And about this very different matter, the chief element of the Libertarian’s response should be that *things probably would have been mind-blowingly radically different* from how they actually are, including in ways of great interest to us presently: there probably would have been no free actions and no free beings. In fact, there’s plenty of room to worry whether we could have existed in such a world, or if anything much like us could or at least would have, or whether we or anything much like us could have or would have existed in our world, had it been that way. Those “probably”s and “likely”s are in there, not just because we should be open to the possibility that, despite its potentially great difference from our world, maybe we could somehow exist in a deterministic world even if in fact we don’t, but also because the Libertarian should recognize that we could well be wrong in thinking that determinism is actually false. If, against what we think, determinism is actually true, then to ask what things would have been like if determinism had been true is to (somewhat strangely) ask what things would have been like if the matter of whether determinism is true were kept just as it is—in which case, the answer would be: “Well, I suppose, pretty much as things actually are” (though in NMNT’s second case: pretty much as things actually are, but over and over again). But supposing that determinism is actually false (as the Libertarian thinks is the case), then the Libertarian should think things could and likely would have been radically different, especially with respect to matters involved in human action, if determinism had instead been true. After all, the Libertarian thinks that our acting freely, even in the extremely thin sense under discussion, depends on determinism being false.

Radically enough different to render our intuitions about such a situation very suspect, I would think, especially where those intuitions are evaluations of apparent “actions.” Here we should recognize some legitimacy to a common objection to philosophers’ use of intuitions about particular cases in their arguments. As Jonathan Weinberg, whose spirited complaint against philosophical intuitions is one I like to use in teaching,¹⁶² puts this objection: Our

¹⁶² Jonathan M. Weinberg, “How to Challenge Intuitions Empirically without Risking Skepticism,” *Midwest Studies in Philosophy* 31 (2007): 318-343. To give the flavor of the paper, here is Weinberg in his introduction, describing his side of the debate: “Some philosophers, including Stephen Stich (1990) and Jaakko Hintikka (1999), have argued that intuitions can have no normative epistemic force, are ungrounded in any theory of their correct use, are unreliable, and generally speaking ought to be abandoned with the likes of palmistry and entrail reading. Their arguments have focused on how error-prone such intuitions can be, with a lousy track record both historically (e.g., Kant on the necessity of the Euclidean geometry of space; the comprehension axiom in naive set theory) and scientifically (e.g., Evans and Over 1996; Kahneman, Slovic, and Tversky 1982). Intuitions have also been shown to vary

common philosophical practice of appealing to intuitions about particular imagined cases “appears to set no constraints on how esoteric, unusual, far-fetched, or generally outlandish any given case may be” (p. 321). Now, I think that Weinberg is a bit overly harsh on how philosophy is actually practiced, in general and specifically with respect to this complaint.¹⁶³ Nevertheless, there is certainly some validity to this criticism, where it does apply: we should be especially suspicious of “intuitive” applications of concepts to imagined scenarios that are very far removed, in relevant ways, from the kinds of ordinary situations in which we have learned to apply those concepts—and I think this is so even of cases like the one before us now, which has a feature that may allow it to dodge the suspicion that it deserves. The supposition that determinism is true may be able to fly under our radar because it isn’t *clearly* far removed, or even *clearly* removed *at all*, from our ordinary experience. For all we know, determinism actually is true, in which case the supposition that it is true is the exact reverse of far-fetched. But supposing that determinism is actually false, and supposing further with the Libertarian that our very freedom depends on this falsehood, the supposition that determinism is true is the epitome of a bizarre, remote, far-out hypothetical, at least where it comes to judging normal human behavior, or what can look like it. The proper response of the Libertarian to questions about examples that involve determinism would seem to be to divide their answer into two cases: “If determinism is actually true, then you’re asking about a situation that’s just like actuality; but if, as I think, determinism is false, then I think you are asking about a situation

from group to group in a manner inconsistent with philosophers’ reliance on them (e.g., Machery et al. 2004; Weinberg, Nichols, and Stich 2001), and perhaps to be overly sensitive to what other cases have been recently considered (Swain, Alexander, and Weinberg forthcoming). In addition to intuition’s demonstrable fallibility, some opponents of intuition (e.g., Cummins 1998) have cited problems with the psychological sources of intuition: Our best possible accounts of where intuitions can come from do not square well with our hopes that they are any sort of reliable guide to a truth beyond themselves” (pp. 318-19). In class, I take up the other side of the debate, defending philosophers’ use of intuitions—but not any uses of palmistry or and entrail reading! (I should note that part of defense here can now be that some of the empirical results cited by Weinberg have failed to stand up to further scrutiny. The bit about intuitions varying from demographic “group to group,” especially, has badly failed to replicate; on this see “Minsun Kim and Yuan Yuan, 2015, “No Cross-Cultural Differences in the Gettier Car Case Intuition: a Replication Study of Weinberg et al. 2001”, *Episteme*, 12 (2015): 355–361—an episode in the history of this discussion that I take particular interest in, as one of the authors, Kim, had been an undergraduate student in the basic epistemology course in which I defend philosophers against Weinberg! But I was inclined to defend the philosophical use of intuitions even supposing the empirical results initially reported by critics of that use could stand up.)

¹⁶³ This is not to deny that philosophers often utilize intuitions about quite far-fetched examples, and that the value of these intuitions is jeopardized by how far removed the cases are from the ordinary situations in which we have learned to apply the concepts involved in the intuitions. In class, my defense of philosophy here is largely based on the claim that philosophers are typically addressing hard questions that nobody has come up with a way to answer that’s any, or much, more reliable than by the use of intuitions about far-fetched cases. The use of this weak support is justified largely by the lack of any better competition. (And I claim that just ignoring the questions isn’t a clearly better option.)

so far removed from reality that I shouldn't think I know what to think here." In short, while determinism is not *clearly* far removed from our ordinary experience, it deserves great suspicion since it is potentially (and likely, according to the Libertarian) exceedingly far-fetched indeed.

This especially because the cases in question, including that of necklace-stealing Jill, are not only built on the stipulation that determinism is true in the worlds being described, but then go on to tell us about the adventures of agents who somehow seem nevertheless to be behaving much as we do. This should make Libertarians wonder: What's going on in these wild counterfactual situations? I wouldn't have thought there would be anything much like us if determinism had been true, so why are there things in these deterministic situations that seem to be acting like us, and how are they pulling off the stunt of acting like us (or how is something in their world pulling off the stunt of making them seem to act like us)?

"Stop with all your questions, and just answer ours!: Did Jill act of her own free will in stealing the necklace, or not?"

Well, the problem isn't just that our intuition about these cases is not worth much, as it concerns scenarios potentially extremely far removed from our experience, but that we shouldn't even think we know how to imagine the cases that we are forming a judgment about. The question seems to suppose that fully determined beings (who live in a world where everything they have ever done has been determined) could rather straightforwardly be in a scenario in which they are acting much like we do, and doing things like stealing necklaces. "Straightforwardly": so that there is a clear most boring, least disruptive way to imagine how this scene could be coming about. But how is this even happening? Are these beings moving much as we would when we do things like stealing necklaces, but they are being controlled by some force that for some reason wants them to appear to be acting like us, or is this some giant cosmic accident, or what?

Libertarians should worry, as with NMNT's first case, that the vignette they are being asked about may not be really possible. For they are not told that in the deterministic universe described there are beings who at least appear (would appear to us, I suppose) to be doing things like stealing necklaces, but rather that Jill, for instance, just is stealing a necklace. And perhaps that's not even possible in a deterministic scenario—as the Libertarian especially should worry. But if beings can somehow do things like steal necklaces in deterministic universes (maybe our actually (as opposed to just apparently) doing actions like stealing necklaces is somehow compatible with determinism, even if our doing these things *freely* isn't?) "How is that even happening?" seems like a vital question to ask, given that, when you think about it, the default "Pretty much as it happens in our world" isn't a good answer, except on the hope that our world is like the imagined one on the matter of whether determinism is true.

The problems here are severe, but to speculate a bit on what might be going on when some of us respond to such questions, not with annoying requests for clarification (or with despair in settings like that of taking a survey, where the inquirers are not available to field such questions), but with fairly easy “yes”es: I suspect that what happens, at least with many respondents (but perhaps not all, given Roskies and Nichols’ results, reported at the start of this section), is that we just grasp after the hope mentioned at the end of the previous paragraph and make sense of the scenarios in the only straightforward way available to us: by supposing that determinism, and compatibilism, is true of our actual world, so that there is no bewildering issue of how to imaginatively get from here (how things actually are, when we do things like steal necklaces) to there (how things are in the deterministic scenario, when beings which seem a lot like us at least seem to do such things). But if that’s how some of the positive responses (that Jill and her like are free and/or morally responsible) are obtained, it’s not happening through any compatibilistic tendencies residing in the respondents’ souls, but through their willingness to suppose compatibilism is true in order to make sense of what would otherwise be a baffling question about a wildly far-out hypothetical situation. And if that’s what’s going on, then even though these scenarios are presented as alternative universes, not presumed to be our own, they do, at least sometimes, prompt a positive response on issues of freedom in the same way (at least broadly construed) as do the cases we considered in the previous section: by leading respondents to provisionally suppose compatibilism is actually true, whether or not they tend to think it’s actually true.

46. The Experimental Philosophy Battle over Intuitions: Attempts to Explain Away the Apparent Intuitions of the Other Side

Speculation aside, the project of discerning whether and to what extent compatibilism or incompatibilism is intuitively correct by testing our intuitions about cases like those we’ve been considering in these past two sections seems to me a mess, but I should perhaps give at least something of a summary report on the discussion that has taken place and is taking place in experimental philosophy, focusing on the portion of the discussion we’ve been considering, but saying how that part fits into the bigger picture. To do so, I will use the nicely succinct report in a recent *Stanford Encyclopedia of Philosophy* article by Joshua Knobe and Shaun Nichols. Below, Knobe and Nichols start their account of experimental philosophy’s empirical inquiry into the matter of whether incompatibilism is intuitive where we started, with NMNT’s study, and the passage I quote ends with the Roskies and Nichols paper we have discussed in this section:

One of the first experimental studies on free will found that people seemed to have compatibilist intuitions. Participants were presented with a scenario describing a deterministic universe, and then asked whether a person in the scenario was free and morally responsible (Nahmias et al. 2006). In one case, participants were asked to imagine a future scenario in which there is a supercomputer that is capable of predicting all future human behavior when provided with a complete description of the universe along with the laws of nature. In this scenario, a man robs a bank, and participants are asked whether the man is morally responsible for his action. Somewhat surprisingly, most participants gave compatibilist answers, saying that the person was morally blameworthy. This basic finding held across a number of scenarios.

In these early studies on intuitions about free will and moral responsibility, the description of determinism focused on the fact that in a deterministic universe, every event is in principle predictable from the past and the laws. In addition, the scenarios involved particular agents in our world doing bad things. Later studies emphasized the causal nature of determinism—that what happens at a given point is completely caused by what happened previously—and stressed that what happens in a deterministic universe is inevitable given the past. Even with this description of determinism, participants still tend to say that a specific concrete individual in such a universe who commits a heinous crime is free and responsible (Nichols & Knobe 2007; Roskies & Nichols 2008). However, when asked a more abstract question about whether it is possible in general for people in such a deterministic universe to be free and responsible, participants tend to say that moral responsibility is not possible in a deterministic universe. This incompatibilist response was also found in a cross cultural sample with participants from India, Hong Kong, Colombia, and the United States (Sarkissian et al. 2010). In addition to the abstract nature of the question, another important element seems to be whether one is considering an alternate deterministic universe or contemplating the possibility that our own universe is deterministic. When led to consider our own universe as deterministic, participants were more likely to say that people would still be morally responsible (Roskies & Nichols 2008).¹⁶⁴

Knobe and Nichols conclude:

¹⁶⁴ This is all from section 2.2, “Free Will and Moral Responsibility” of Knobe and Nichols’s *Stanford Encyclopedia of Philosophy* article on “Experimental Philosophy,” published 2017, at: <https://plato.stanford.edu/entries/experimental-philosophy>. The papers other than those by Nahmias et al. (NMNT) and Roskies & Nichols that Knobe here refers to are: Nichols, Shaun & Joshua Knobe, 2007: “Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions,” *Nous* 41: 663–685; and Sarkissian, Hagop, Amita Chatterjee, Felipe De Brigard, Joshua Knobe, Shaun Nichols, & Smita Sirker, 2010: “Is Belief in Free Will a Cultural Universal?”, *Mind & Language*, 25: 346–358.

Thus, it seems like people give compatibilist responses under some conditions and incompatibilist responses under others. One reaction to this apparent inconsistency is to treat one set of responses as defective.

And Knobe and Nichols go on to describe the efforts on each side to argue that the responses that seem to support the other side are defective, and/or that those that support their own side are not, before drawing the very balanced conclusion: “Thus, the state of the evidence currently suggests that people do have both incompatibilist and compatibilist intuitions.”

While I certainly think that the intuitions about characters in these deterministic scenarios are deeply defective measures of the intuitive appeal of incompatibilism, I quite unmoved by the particular complaints each side uses to attempt to undermine their opponents, for reasons I should fairly quickly give.

Nicholls and Knobe propose that the supposedly compatibilist judgments subjects make that characters in deterministic scenarios are free or morally responsible are distorted by emotional reactions those subjects have to extremely bad actions, and N&K thus propose classifying these responses as “performance errors,” rather than indications of respondents’ true intuitions.¹⁶⁵ There are some problems with supposing that a difference in responses to highly disturbing vs. “low affect” cases can account for the supposedly compatibilist intuitions it is supposed to undermine, and we will see them briefly in the next section. But here I want to register the worry that even if we are more likely to judge a character free if their action would appear to be egregiously wrong than if not, and even if this difference could play a major role in explaining where the supposedly compatibilist “yes” answers will be generated about cases, this wouldn’t seem to me properly thought of as detecting some “performance error.” Nicholls and Knobe entertain a rival “affective competence” explanation, on which it is subjects’ responses to the emotion-affecting cases that reveal their true intuitions, while their responses to the not-so-emotionally-engaging cases that represent a performance error (2007: pp. 672-673), and argue that the “affective performance error model” that best explains their data (pp. 676-677). But if we are trying to determine whether people find compatibilism or incompatibilism intuitive, and if we are going to do this (unwisely, I think) by considering their judgments about characters in deterministic scenarios, then if we found it depends on how the characters’ apparent actions affect our emotions, I think we’d just have to say that our intuitions vary depending on which kind of case we’re considering. The question wouldn’t be which position is intuitive, but which intuitions should be deemed credible, and that we should seek to verify in our theories. (As you saw in the “Taking our freedom seriously” portion of [section 37](#), I am myself not above appealing to “high-affect” cases (about our world as it is, with respect to whether determinism is true, rather than scenarios stipulated to be deterministic) to

¹⁶⁵ Nichols and Knobe 2007: The “affective performance error” proposal is laid out at pp. 671-672.

generate particularly strong intuitions of freedom, though I do advise that inquirers consider whether they think the highly affecting nature of the case might be a distorting, or rather a clarifying, factor, in deciding how much credence to give the relevant intuitions.)

On the other side, defenders of intuitive compatibilism, particularly Eddy Nahmias and Dylan Murray, have sought to undermine apparent incompatibilist intuitions that there is no freedom or moral responsibility in deterministic scenarios by claiming that those intuitions arise from a confusion over what determinism is. Finding that those who respond that there is no moral responsibility in scenarios described in deterministic ways also tend to agree with statements like that in those same scenarios, a person's decisions, and what they want, and what they believe, have "no effect on" what they do, Nahmias and Murray conclude that these respondents think of determinism as including the thought that such things as an agent's decisions, wants, and beliefs are "bypassed" in the process by which actions are produced, and since determinism does not include that thought, these respondents are simply misunderstanding what determinism is.¹⁶⁶ Continuing this project, in a new paper, "Intuitions About Free Will and the Failure to Comprehend Determinism" (*Erkenntnis*, forthcoming), Thomas Nadelhoffer, Samuel Murray, and Elise Murray (NMM), find that the misunderstanding of determinism as containing aspects that are in fact no part of the thesis runs alarmingly deep indeed—to the extent that it undermines the whole attempt to determine whether ordinary intuitive thought is compatibilistic or incompatibilistic, until this problem of misunderstanding can be solved.

I too take a dim view of trying to discern the intuitive status of incompatibilism by means of surveys like those taken in the experimental philosophy literature, but for me the problem lies in the whole project of asking subjects about the actions of agents in deterministic scenarios. So far from worrying that *respondents* to these surveys are giving incompatibilist answers because they don't understand determinism, I instead suspect that the survey *designers* are misunderstanding how genuine incompatibilists would react to their questions. These philosophers ask participants what things would be like in a deterministic world or scenario. The researchers think the clearly correct answer is that if everything were determined, human beings could still do all of the stuff we normally associate with agency (deciding what to do, considering reasons, etc.); it's just that this entire process would be deterministic. But they then find that ordinary folks do not think that. Instead, ordinary folks tend to think that if everything was completely deterministic, human agency would be completely different from the way it is in the actual world—to the extent that it becomes

¹⁶⁶ See Nahmias and Murray, "Experimental Philosophy on Free Will: An Error Theory for Incompatibilist Intuitions," in Aguilar, Buckareff, and Frankish, ed., *New Waves in Philosophy of Action* (London: Palgrave Macmillan UK, 2011), pp. 189–216; and Murray and Nahmias, "Explaining Away Incompatibilist Intuitions," *Philosophy and Phenomenological Research* 88 (2014): pp. 434–467.

problematic to think of them as even making decisions about what to do, at least in anything like the way we do. In my view, ordinary, natural Libertarians are wise to at least suspect all this—and probably even to outright think it. Their responses don't indicate that they misunderstand what determinism is, but rather show that the incompatibilism of ordinary people runs deeper than the researchers suppose, affecting not just whether actions can be free (and things for which the agent can be responsible) in deterministic scenarios, but also whether anything like the process by which we end up performing free actions can happen in such worlds. But the compatibilist researchers conclude that the ordinary folks are making such a glaring error about the implications of determinism as to show that they are misunderstanding what it even means for things to be deterministic.

In taking the real lesson here to concern how deep incompatibilism can run in ordinary thought, I am following the perceptive reaction of David Rose and Shaun Nichols, in their paper, "The Lesson of Bypassing."¹⁶⁷ This, from Rose and Nichols's abstract, gives a good summary of their conclusions, and the studies they used to support them:

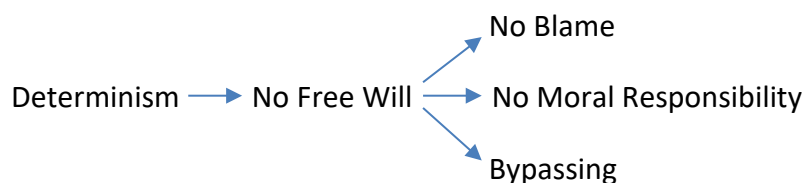
And a number of recent studies find that people give apparently incompatibilist responses in vignette studies. When participants are presented with a description of a causal deterministic universe, they tend to deny that people are morally responsible in that universe. Although this suggests that people are intuitive incompatibilists, Eddy Nahmias and Dylan Murray, in a recent series of important papers, have developed an important challenge to this interpretation. They argue that people confuse determinism with bypassing, the idea that one's mental states lack causal efficacy. Murray and Nahmias present new experiments that seem to confirm the bypassing hypothesis. In this paper, we use structural equation modeling to reexamine the issue. We find support instead for an incompatibilist explanation of the bypassing results, i.e., incompatibilist judgments seem to cause bypassing judgments. We hypothesize that this phenomenon occurs because people think of decisions as essentially indeterministic; thus, when confronted with a description of determinism they tend to think that decisions do not even occur. We provide evidence for this in three subsequent studies which show that many participants deny that people make decisions in a deterministic universe; by contrast, most participants tend to allow that people add numbers in a deterministic universe. Together, these studies suggest that bypassing results don't reflect a confusion, but rather the depth of the incompatibilist intuition. (p. 599)

How will those taking a survey respond when you tell them that the world you are describing is deterministic, but then you go on to describe agents in that world acting away, seemingly just as we do, as if this is all completely straightforward and unproblematic? As I've been urging, it seems the Libertarians among them should at least suspect that the world you

¹⁶⁷ David Rose, Shaun Nichols, "The lesson of Bypassing," *Review of Philosophy and Psychology* 4 (2013): pp. 599–619.

are describing may be impossible, and should at least wonder how it is happening that these fully determined beings seem to acting much as we do, though when we act freely, it is essential to our so acting that we are not determined to act as we do. (Truth be told, even those who tend toward compatibilism should be a bit worried about what they're being asked, insofar as they allow the Libertarianism even might be true. What's strange is just putting together, as if this can be done straightforwardly and unproblematically, things that can't so straightforwardly be put together on some of the views being tested for.) But that's just how it seems subjects in some sense *should* respond. But how will they actually respond, especially in the setting of taking a survey, where there's no mechanism for complaining about a bewildering question, or asking for clarification? Well, it seems that sometimes respondents will suppose determinism is actually true, in order to help them make decent sense of the question, and they'll then give supposedly "compatibilist" responses—or so I've speculated. But it turns out you can also describe the scenarios and ask the questions in a way as to get supposedly incompatibilist answers that deny free will and/or moral responsibility to the inhabitants of these deterministic scenarios. I must advise not to make too much out of these responses.

However, some of what Rose and Nichols found, though I suppose it has to be deemed preliminary, was so fascinating and suggestive about the shape of ordinary thinking about free will that I will quickly convey it, without commentary (see Rose and Nichols's paper for more information), even though this all does just trace thought about what should be bewildering questions about potentially wild scenarios. For where respondents' reactions go in the incompatibilist direction, there seems to be an order to the "cascade of effects" that the supposition of determinism has on ordinary judgments. Namely, determinism seems to directly produce judgments of No Free Will, that then, in turn, result in judgments of No Blame for the "actions" of the determined beings; in judgments of No Moral Responsibility for those "actions"; and for judgments of "Bypassing": i.e., judgments that things like the agents' decisions, wants, and beliefs "have no effect" on what they do:



47. Asking an Abstract Question

If we were to take positive responses to questions about the freedom (or moral responsibility) of characters in deterministic scenarios to be compatibilist judgments, we would be faced with a jarring disparity between those “compatibilist” results (at least where the responses are indeed largely positive, as in studies like NMNT’s) and the incompatibilism that many philosophy teachers report in very strong terms to be the views of their students and others they talk to about the matter. As Shaun Nichols and Knobe ask in their (2007) (one of the studies that they report on in the passage quoted above):

But if people so consistently give compatibilist responses on experimental questionnaires, why have some philosophers concluded that ordinary people are incompatibilists? Have these philosophers simply been failing to listen to their own undergraduate students? (p. 668)

I’ve been urging that the whole method of asking about particular characters in deterministic scenarios is a problem here, but Nichols and Knobe’s study is of particular interest to us, as their attempt to understand the disparity led them to at least take a step away from this methodology, and toward the obvious alternative of just straightforwardly asking whether free will is compatible with determinism: They still asked about a particular deterministic scenario they described, but, in their key study, they did not ask about the particular actions of a character in that scenario, but abstractly asked whether acting in a morally responsible way is possible in that universe. And this returned very strong results.

Nichols and Knobe started by giving participants in their study, not just a description of a deterministic universe, but of both a deterministic and an indeterministic one:

Imagine a universe (Universe A) in which everything that happens is completely caused by whatever happened before it. This is true from the very beginning of the universe, so what happened in the beginning of the universe caused what happened next, and so on right up until the present. For example one day John decided to have French Fries at lunch. Like everything else, this decision was completely caused by what happened before it. So, if everything in this universe was exactly the same up until John made his decision, then it *had to happen* that John would decide to have French Fries.

Now imagine a universe (Universe B) in which *almost* everything that happens is completely caused by whatever happened before it. The one exception is human decision making. For example, one day Mary decided to have French Fries at lunch. Since a person’s decision in this universe is not completely caused by what happened before it, even if everything in the universe was exactly the same up until Mary made her decision, it *did not have to happen* that

Mary would decide to have French Fries. She could have decided to have something different.

The key difference, then, is that in Universe A every decision is completely caused by what happened before the decision—given the past, each decision *has to happen* the way that it does. By contrast, in Universe B, decisions are not completely caused by the past, and each human decision *does not have to happen* the way that it does. (p. 669)

Before getting to questions of moral responsibility, N&K started with this interesting comparative question:

1. Which of these universes do you think is most like ours?

And received even more interesting answers:

Across conditions, nearly all participants (over 90%) judged that the indeterministic universe is most similar to our own. (p. 669)

That is a fascinatingly strong result, but though N&K report that “the purpose of this initial question was simply to see whether subjects believe that our own universe is deterministic or indeterministic,” I wonder how good a test their answers are for that, as what respondents were comparing were perhaps rather odd examples of universes of the two types.¹⁶⁸

At any rate, N&K then prepared for asking the question we are most interested in by dividing respondents into two groups:

After answering the initial question, subjects received a question designed to test intuitions about compatibilism and incompatibilism. Subjects were randomly assigned either to the concrete condition or to the abstract condition. (p. 669)

¹⁶⁸ Though I’m an indeterminist, N&K’s Universe B looks rather odd to me. I suspect that if human decisions to act are not causally determined by what happens before them, then there are probably lots of other events that are also not causally determined by prior events. I think of Universe A as more like our actual universe in that, though it is deterministic while I think our universe is not, human decisions to act are at least not in it (as they also are not in our universe) a weird and lonely exception to what the universe is otherwise like with respect to whether it’s deterministic—though I of course think Universe B is more like our actual universe in that human decisions to act are not determined. I suppose in the end I’d give the same response to N&K’s question that almost all of their respondents gave, but it is a bit messy. And if, as I think, indeterminism is true in our actual universe, then Universe A is perhaps quite odd among deterministic universes in containing what seem to be ordinary people seeming to be doing things like eating French fries.

We are most interested in the abstract condition. Here, though N&K's description of Universe A had featured french-fry-eating John, they did not ask about that or any character's particular actions, but instead abstractly asked:

In Universe A, is it possible for a person to be fully morally responsible for their actions?

YES NO

In this condition, most subjects (86%) gave the incompatibilist response! (p. 670)

"NO," one presumes. And that's a whole lot of "NO," especially against the background of results like NMNTs',¹⁶⁹ and well-deserving of the exclamation mark. And a lot more "NO" than N&K got from respondents in their concrete conditions, in all of which respondents got the same descriptions of the two universes that were given to those in the abstract conditions (and that we have already seen), but were then asked about the actions of particular characters in the deterministic universe, which characters were not french-fry-eating John, but were new

¹⁶⁹ Of course, many factors, going beyond N&K's asking an abstract question, are likely involved in the great difference here. And indeed, in a new paper, "Intuitions About Free Will and the Failure to Comprehend Determinism" (*Erkenntnis*, forthcoming: page numbers to the online prepublication preview), Thomas Nadelhoffer, Samuel Murray, and Elise Murray (NMM) report getting very different results indeed when they presented survey-takers with N&K's descriptions of Universes A and B and abstractly asked whether respondents agreed that "It is possible for a person in Universe A to be morally blameworthy when they do something wrong" (p. 8). NMM's results: "While 86% of the participants from the study by Nichols and Knobe judged that people could not be fully morally responsible in a deterministic universe, we found instead that 58% (170) of our participants judged that people *could* be responsible whereas 30% (87) judged that people could not be morally responsible. These are markedly different results from Nichols and Knobe (2007)" (p. 10). However, despite using N&K's descriptions of Universes A and B, and despite, like N&K, asking an abstract question about the possibility of moral responsibility in Universe A, there are many differences in how NMM conducted their survey (going beyond the one readers here are already in a position to spot: the significantly different abstract question about moral responsibility), one of which very much caught my attention: After presenting the N&K's descriptions of the Universes, and after (like N&K) asking which was most like our own universe, NMM then asked whether Universe A is possible, and asked a question checking for comprehension, and then added this notable bit of direction: "Regardless of how you answered the previous questions, imagine that it is possible that there is a universe where everything that happens is completely caused by what happened before it, such that, given the past, everything that happens has to happen..." (p. 7). In a note, they report: "This type of 'suspension of disbelief' paradigm was originally used in Nahmias et al., (2005, 2006)" (n. 16, p. 7), and, indeed, it is the feature of NMMT's survey that I was fretting over in section 44—and that I worry might function like a "suspension of incompatibilism" direction. Now that I see NMM's results, I am even more interested in seeing how much such directions influence results. At any rate, given all the different variables in play here, in judging whether asking an abstract question can make a difference in the results, it's best to rely on the difference N&K found between the significantly different answers they got to their abstract question as opposed to their concrete questions, because here other factors (beyond what questions are asked) are held constant.

characters introduced in the asking of the “concrete” question (so, going beyond how Universe A was described to those in the abstract conditions).

Nichols and Knobe were testing their own guess as to what was behind the great disparity. Their suspicion:

On our view, most people (at least in our culture) really do hold incompatibilist theories of moral responsibility, and these theories can easily be brought out in the kinds of philosophical discussions that arise, e.g., in university seminars. It’s just that, in addition to these theories of moral responsibility, people also have immediate affective reactions to stories about immoral behaviors. What we see in the results of the experiments by Nahmias and colleagues is, in part, the effect of these affective reactions. (p. 668)

Nichols and Knobe’s hypothesis is actually a bit tricky to state squarely. I don’t think we can just say, as they do, that it’s “immoral behaviors” that respondents are affectively reacting to, since these are “behaviors” that are occurring in deterministic worlds, and so that aren’t immoral at all on one of the views being tested. I think the idea is that these are “behaviors” that would be immoral if the agents involved were responsible for them, or that would appear to us to be immoral, if we were to observe them, or something like that. And that would appear to us to be *grossly* immoral, or at least immoral in a way that we would have a particularly strong, negative emotional reaction against: Nichols and Knobe construed their “affective performance error” hypothesis so that it predicted that these distorting emotional responses would not arise when the actions being judged would have appeared to us to be immoral, but not in a particularly upsetting way. That this prediction seems to have turned out to be wrong is what Knobe and Nichols are admitting when they report, in the history quoted in the previous section, that “a meta-analysis indicates that there is very little evidence that emotions play a critical role in generating compatibilist judgments (Feltz & Cova 2014).”¹⁷⁰

¹⁷⁰ This hypothesis seems a bit strained to me, even before learning of any results of a meta-analysis, at least if it is meant to account for very much of tendency toward positive results in studies like NMNT’s original. Nichols and Knobe’s first concrete case was indeed one that could provoke an emotional response: “In Universe A, a man named Bill has become attracted to his secretary, and he decides that the only way to be with her is to kill his wife and 3 children. He knows that it is impossible to escape from his house in the event of a fire. Before he leaves on a business trip, he sets up a device in his basement that burns down the house and kills his family.” But is the bank robbery so abstractly presented in NMNT’s original case so affecting? This is not to deny that, say, being a teller involved in a robbery, and being threatened at gunpoint, would be traumatic. But NMNT’s presentation of the case didn’t specify that it happened in a way that would involve such trauma for anyone involved, and I wonder whether respondents would get so worked up about the case as presented. When Nichols and Knobe formulated a “low affect” case to test, it involved a character cheating on their taxes. While that does indeed seem quite “low affect,” and sharply lower than the case of the murderer of his own family, I wonder if it’s really all that much of a contrast with an abstractly presented bank robbery?

Since I think the whole methodology of testing intuitions about deterministic scenarios is mistaken, my interest in Nichols and Knobe's results is due to their asking, in their key case, an abstract question, as opposed to a concrete question about a particular action (whether particularly upsetting, or not). And it should be noted that the meta-analysis which showed that high vs. low affect difference may make little or no difference to the results also concluded (admitted) that "the abstract/concrete asymmetry seems a robust result" (Feltz & Cova, p. 238).

48. **Surveying Our Freedom**

If asking about the "actions" of characters in deterministic scenarios is such a ruinous mistake, as I claim, how then *should* we test the intuitiveness of incompatibilism? What questions should we ask—ask of others, where we want to see what their intuitions are, or of ourselves, where we're just thinking through the matter on our own?

Nichols and Knobe's abstract question is a step in the right direction—and I find their strong negative result (that moral responsibility is not possible in their deterministic universe) quite telling (and encouraging). But while Nichols and Knobe didn't ask about the particular "actions" of a character in a deterministic scenario, their presentation of their deterministic Universe A did include French-fry-eating John, a character who seemed much like us, acting in a way that seems much as we act. And their question, while abstract, does ask about the "actions" of "people" in Universe A, as if we can with no further guidance straightforwardly imagine what these people and their actions are like, which seems an easy and straightforward task only on the hope that our actual universe is itself deterministic (since then, they can be supposed to be just like us).

If you want to drive home the point that *everything* in your described universe is determined, with no exceptions for decisions to act, if there are any such decisions, I would think such decisions and such actions would have to be added to your presentation of the universe conditionally, in something like this manner (which, admittedly, might be quite confusing for respondents rushing through a survey): "If beings like us, deciding to perform and performing actions much like ours, are possible in such a deterministic universe, then the decisions to act made by these beings, like everything else in the universe, would be completely determined by what happened earlier in the universe." And if you want to include a particular character like John, that too would be best done conditionally: "So if, say, ordering French fries for lunch, in much the way we would do so, is possible in such a deterministic universe, then in

much the way that we do, John orders French fries for lunch in Universe A.” I’d think you should then add something like: “And if beings who are not much like us, but who can still decide in their own way to perform what can still be actions, are possible, then their decisions to perform actions are, like everything else in the universe, completely determined by what happened earlier in the universe.” And then, instead of asking whether the actions of people in your deterministic universe can be free, or whether these people can be morally responsible for their actions, it would be better just to ask whether there can be free actions in this universe (or whether there can be actions for which agents are morally responsible).

But having come this far, I wonder why we wouldn’t just straightforwardly ask whether free will (or actions for which an agent is morally responsible) is compatible with determinism, after having explained what determinism is, if that’s needed. Indeed, how did we end up asking about the freedom or responsibility of characters in deterministic scenarios, rather than asking the obvious, straightforward question that just is the question whose answer we’re ultimately interested in? As I have guessed, this move may have been motivated by quite general thoughts to the effect that our best intuitions involve how philosophically important concepts do or do not apply to particular cases. For instance, in epistemology, some of our best known arguments are driven by premises that characters in quickly described cases do or do not know something to be the case, and so are the contents of such intuitions about cases. So the thought may be that, just as in epistemology we ask whether Henry knows he is seeing a barn in the philosophically famous case in which he is a character traveling through the Land of Fake Barns,¹⁷¹ so in the metaphysics of free will we should ask such things as whether Jeremy acts freely when he robs a bank in a deterministic scenario we quickly construct. But this method seems a disaster in the latter case, and the cases we have been considering here, because the situation we end up asking respondents to imagine—as if this can just be easily and straightforwardly done—is potentially impossible on one of the positions whose intuitive standing we are investigating, and is in any case potentially extremely different from the actual situations in which we have learned to apply the relevant concept, and greatly different in ways very relevant to the application of the concept.

Or the thought may be that respondents are not able to comprehend the obvious question well enough to respond competently to it, and are better able to process simple

¹⁷¹ Here I allude to a philosophically well-known example, due to Carl Ginet, which made its first appearance in published writing in Alvin Goldman’s “Discrimination and Perceptual Knowledge” (*Journal of Philosophy* 73 (1976): 771–91; pp. 772–3). While I do think that applications of concepts to particular cases comprise some of our most solid philosophical intuitions, I am among the epistemologists who follow Ruth Millikan (see her “Naturalist Reflections on Knowledge,” *Pacific Philosophical Quarterly* 65 (1984): 315–34) in saying that I find this particular intuition that Henry does not know in this case, at least as the case is described in Goldman’s paper, to be quite murky: See DeRose, *The Case for Contextualism*, p. 49, n. 2, for discussion.

questions about the particular actions of characters in described cases. This may be closely related to the thought we just previously considered: Perhaps it is (in part) because we are in general better able to understand questions about the application of concepts to particular cases that those applications are thought to be our best intuitions? Those whose procedure is driven by such concerns about comprehensibility may have been worried about the complicated additions I was suggesting two paragraphs above, thinking these would ruin the advantage in clarity that questions about particular cases often enjoy. If their concern is specifically about whether survey-takers can adequately understand what determinism is, they may reject my suggestion to “just explain what it is on the survey” for reasons of survey methodology: Perhaps no abstract description of what determinism is that is suitable for inclusion on a survey can give survey-takers an adequate grasp of the notion? That actually seems quite plausible to me, as my understanding is that these surveys are taken by respondents who are rushing through them in order to claim a small reward, hoping to earn enough such rewards to make the time they spend on answering surveys worthwhile. And perhaps for other reasons as well (beyond the difficulty of grasping what determinism is), our obvious question is just too difficult to be itself suitable for experimental-philosophy-style surveys, as perhaps are questions about particular deterministic scenarios when they include difficult complications like those I suggest two paragraphs above.

If so, and if the potential “remedy” is to instead ask simpler questions about the actions of characters in deterministic scenarios, then I think we’re dealing with a matter whose intuitive status cannot be ascertained by means of the kind of mass surveys used in experimental philosophy, because, for the reasons we have seen, this “remedy” is a bad one here. Perhaps we are then dealing with a question whose intuitive answer can’t be discerned by such surveys (though I certainly don’t rule out the possibility that those much better than me at designing such surveys might come up with a good way to survey this matter), but has to be ascertained by a longer and perhaps interactive in-person explanation of what determinism is, followed by the key question of whether one’s respondents think such determinism is compatible with freedom. And I’m guessing that that may be pretty much what happens in the classrooms of the many philosophy teachers who report that their students seem very much to start out as incompatibilists.

49. The Great “Upshot” Argument against Free Will

I think of the main argument for Pessimism (no free will) to be another of the great, simple, intuitively powerful arguments of philosophy, though in this case the intuition is one born of a

certain scientifically derived way of construing human thought. We previewed this argument in the third paragraph of [sect. 33](#), and the presentation here actually won't go into greater detail in terms of the steps that would make it into a premises-key-steps-and-conclusion outline of the argument. It really just is that all our decisions to act, and then our actions, including my raising of my hand (our example of a purportedly free action), are all just the upshot of mindless, tiny physical events, mostly occurring in our brains, over which we have no control, in some way (in some good sense of "just the upshot of...") that renders us really no more responsible for any of them than for any others, nor more responsible than is the victim of our mad scientists (our example of a subject who is not acting freely). Here, we will look at the intuitive view of human decision-making that lies behind that premise and will begin a look at the meaning of the key terms of the argument which makes it intuitively powerful both that the argument's premise is true and that it implies the argument's ruinous conclusion – which look will be continued in the next section, as we also then explore how to escape this argument. And I will also in this section explain why the argument is best construed as having the simple basic structure I claim for it, rather than the structure it is often construed as having.

The great intuitive case for Pessimism is often presented as a dilemma argument in its basic structure: Either determinism or indeterminism (where this just means the denial of determinism) is true; and either way, we cannot be free in our actions; so, no free will. Our choices and actions are either determined, or else—to get just a bit into how the second horn of the dilemma is often developed—those choices are not determined, and so are just random, or just a matter of chance, or something to that effect; and, either way, we don't act freely. Taking the argument to have this dilemma structure would handily allow us to view this great argument for Pessimism as an extension of the case for Incompatibilism that we have already spent considerable energy on, since the first half of the dilemma can be rested either on the intuitive incompatibility of determinism with free will, or on the intuitive power of the great Consequence argument for Incompatibilism, both of which we have already discussed. The key to this argument against free will, insofar as it is something new for us, going beyond the old case for just Incompatibilism, would then lie in the case made for the dilemma's second half: Why we are not free in the case where determinism is false.

But employing an insight I believe I picked up many years ago from Unger,¹⁷² would it not be a surprise if determinism, and then also determinism's bare denial, themselves both

¹⁷² I recall this as something Unger told me (though I suppose it's possible that the telling went the other way) some time while I was at NYU (1990-1993) when we discussed the ideas that he later wrote up in his paper "Free Will and Scientiphicalism," *Philosophy and Phenomenological Research* 65 (2002): pp. 1-25, and then in the corresponding chapter of *All the Power in the World* (Oxford UP, 2006). I was a bit surprised to find that he presents the argument as a dilemma one in these write-ups, and also, relatedly, that "upshot," which I remember as playing the key role I'm giving it here in our discussions, does not

militated against our freedom? Would it not be more reasonable to suppose that, deeper down, some other thought is at work here, which thought challenges our freedom whether or not determinism is true, though it perhaps works out in somewhat different forms in the deterministic vs. indeterministic cases? I think when we look at what lies at the root of our supposed dilemma's second half, we find just such a thought—indeed, that thought is the premise of the argument as I think it is best and simply construed: that our decisions to act, and then our actions, are all just the upshot of mindless, tiny physical events.

At any rate, the picture of human decision-making that we can find so intuitively compelling but that we also find to be a grave threat to our freedom is effectively presented by van Inwagen, unsurprisingly (at least I find this unsurprising) in a textbook, *Metaphysics*.¹⁷³ Though van Inwagen is a Libertarian, and so a believer in human freedom, he does in a way endorse the dilemma argument against free will: not, of course, as compelling, nor even as winning the day, but as at least being strong enough to face us with a mystery until we can, as van Inwagen thinks has not yet been done, answer this argument¹⁷⁴—which answer, he thinks, would have to be an explanation of how the argument's second (indeterministic) half fails: an explanation of how we can be free if indeterminism is true. Van Inwagen is using his description of human decision-making in making the case for the indeterministic half of the dilemma. In presenting my somewhat streamlined (eliding over some of van Inwagen's text) version, I here italicize and put in brackets the material where van Inwagen is being explicit about his assumption of indeterminism:

[*Let us look carefully at the consequences of supposing human behavior to be undetermined.*] Suppose Jane is in an agony of indecision; if her deliberations go one way, she will in a moment speak the

play a key role in Unger's rendering of the argument. Perhaps these are features of our discussions that he had forgotten by the time he wrote up his ideas—or perhaps he for some reason decided against presenting the argument as I do here. But I happily note a use of “upshot” by Unger in his book (but not in the parallel portion of his article, at p. 16) that is along the lines I recall, and that does concern van Inwagen's deliberating character, Jane, whom we are about to meet in my text, though it does not occur where Unger explains the argument against free will, but later in the chapter: “Being just the upshot of how it is with the manifesting of those reciprocal probabilistic Propensities, our Scientifical Jane's speaking her mind, even if objectively quite likely, isn't something that, from among real options for her own activity, she really chooses to do” (*All the Power in the World*, p. 331).

¹⁷³ Using this bit of van Inwagen's *Metaphysics* for this purpose is also something I picked up from Unger, who does the same in both of the works cited in the note immediately above. I am not surprised to find this effective presentation in a textbook, as opposed to a work intended primarily for other professional philosophers, as it is in textbooks, and other works for general audiences, that philosophers can be expected to work the hardest to make their arguments intuitively compelling.

¹⁷⁴ See especially van Inwagen's aptly titled “Free Will Remains a Mystery,” *Philosophical Perspectives* 14 (2000): pp. 1-19. As van Inwagen puts it in the opening of the paper: “I conclude that free will remains a mystery—that is, that free will undeniably exists and that there is a strong and unanswered *prima facie* case for its impossibility” (pp. 1-2).

words, “John, I lied to you about Alice,” and if her deliberations go the other way, she will bite her tongue and remain silent. [*We have supposed there to be physically possible continuations of the present in which each of these things happens. Given the whole state of the physical world at the present moment, and given the laws of nature, both these things are possible; either might equally well happen.*]

Each contemplated action will, of course, have antecedents in. . . Jane’s cerebral cortex. . . . Let us make a fanciful assumption about these antecedents, since it will make no real difference to our argument what they are. (It will help us to focus our thoughts if we have some sort of mental picture of what goes on inside Jane at the moment of decision.) Let us suppose that a certain current-pulse is proceeding along one of the neural pathways in Jane’s brain and that it is about to come to a fork. And let us suppose that if it goes to the left, she will make her confession, and that if it goes to the right, she will remain silent. [*And let us suppose that it is undetermined which way the pulse will go when it comes to the fork. . . .*]

Van Inwagen thinks this look at deliberation makes Jane’s choice, and then her resulting action, seem like something that is not up to her, since it is the result of a physical happening—the direction the current-pulse takes at the fork in the neural pathway—that is not up to Jane, but is something that, to use the term van Inwagen chooses, “just happens.” This is just van Inwagen’s “fanciful assumption” of how decision making goes. But the serious idea is that whatever the actual details of the story are, our decisions are just the upshot of little (indeterministic, van Inwagen is currently supposing) physical events which will seem to “just happen” in the same way that the direction that the pulse takes in the example seems to “just happen.” In what immediately follows the passage quoted above, van Inwagen moves (on behalf of the argument) quickly to his key “just happens” claim:

Now let us ask: Is it *up to Jane* whether the pulse goes to the left or to the right? If we think about this question for a moment, we shall see that it is very hard to see how this could be up to her. Nothing in the way things are at the instant before the pulse makes its “decision” to go one way or the other makes it happen that the pulse goes one way or goes the other. If it goes to the left, *that just happens*. If it goes to the right, *that just happens*.

I think what van Inwagen means by saying that the direction the pulse takes, and then Jane’s decision and resulting action are things that “just happen” is very close to what I mean by saying her decision and resulting action are just the upshot of little physical events over which she has no control, or that are not at all up to her. Both are meant to exclude Jane’s being at all responsible for what happens, or what happens being at all up to her.¹⁷⁵ (The likes of “just a matter of chance” and “just random”—the terms in which the indeterministic half of

¹⁷⁵ In the case of both terms, they can be used to exclude *anyone’s* being responsible for what happens, or it being up to *anyone* (and not just the subject in question). Thus, one could say, in the relevant sense, “Though what she decided wasn’t up to Jane, it turned out not to be something that just happened, for Jill was behind the scene calling the shots.”

the dilemma are often put—are closely related, I think: In the relevant sense they more specifically convey both that a happening “just happened” in the relevant sense and also that it could have not happened.¹⁷⁶) I think my way of putting things is a bit more specific than van Inwagen’s: It is because these things seem just the upshot of little physical events over which Jane had no control that they seem to be things that just happen, though such things could in the relevant sense “just happen” in other ways: by, say, being the upshot of minor spiritual, non-physical events over which Jane has no control, or by not being the upshot of anything—of being things that in a more thorough-going way, and not just from the standpoint of Jane’s responsibility, “just happen.” So, though we phrase the premise of the argument for no free will (or, in van Inwagen’s case, the premise for the indeterministic half of the argument for no free will) a bit differently, we both think there is a way of understanding that premise, whichever terms we use, so that it both does seem to undermine Jane’s—and then our—freedom, and also so that it is rendered plausible by a good look at what human decision making is like. Indeed, I think it is as well motivated by van Inwagen’s description of human decision making without the assumption of indeterminism—so, when omitting the portions of van Inwagen’s description that I put in brackets above.

For van Inwagen, as you can see as we pick up his presentation of the case from where we left it above (in fact, repeating the last two sentences of above), the assumption of indeterminism plays a key role as he brings the argument home:

If [the pulse] goes to the left, *that just happens*. If it goes to the right, *that just happens*. There is no way for Jane to *influence* the pulse. There is no way for her to make it go one way rather than the other. Or at least there is no way for her to make it go one way rather than the other and leave the “choice” it makes an undetermined event. If Jane did something to *make* the pulse go to the left, then obviously its going to the left would *not* be an undetermined event. It is a plausible idea that it is up to an agent what the outcome of a process will be only if the agent is able to arrange things in a way that would make the occurrence of *this* outcome inevitable *and* able to arrange things in a way that would make the occurrence of *that* outcome inevitable. If this plausible idea is right, there would seem to be no possibility of its being up to Jane (or to anyone else) what the outcome of an *indeterministic* process would be. And it seems to follow that if, when one is trying to decide what to do, it is truly

¹⁷⁶ Thus, I think that Unger is wrong when he compares the key term he uses in the argument with van Inwagen’s, writing: “Where I talk of what’s a *matter of chance*, van Inwagen writes of what *just happens*; but the suggested thought’s the same” (*All the Power in the World*, p. 322), if he is suggesting those key phrases mean the same thing. I do think Unger is very much on to the same idea as van Inwagen here, but Unger uses a term that, as part of its meaning, adds indeterminism on to that idea. That van Inwagen doesn’t use “just happens” with that add-on, I think, comes out in the longish passage we’re about to look at, where van Inwagen argues that it is a consequence of indeterminism that the direction the pulse goes is something that “just happens”: He does not argue as if he takes this to be a simple matter of “just happens” having indeterminism built into its very meaning.

undetermined what the outcome of one's deliberations will be, it cannot be up to one what the outcome of one's deliberations will be. It is therefore far from clear whether incompatibilism is a tenable position. The incompatibilist who believes in free will must say this: it is possible, despite the above argument, for it to be up to an agent what the outcome of an indeterministic process will be. But how is the argument to be met?

We will be figuring out how to meet this great argument in the next section, but it is important that we get the argument we respond to in its best form. I think van Inwagen is wrong about why it seems the pulse's path cannot be up to Jane. The real reason has nothing to do with indeterminism.

To see this, consider this case of bearing some responsibility for an indeterministic event. A billionaire keeps a lot of money in an advanced high-tech safe that sends out a daily signal that notifies the billionaire's various personal devices that all is well with the safe. I figure out a high-tech way to steal one million dollars from the safe, but my action may change the nature of the first signal the safe sends out after the theft—which, as the timing works out, will be sent several hours after the theft. Even if the signal is altered, it will have the same effect as the usual signal does on the billionaire's devices, so the billionaire will as usual be informed (though mis-informed, in this case) that all is well with their safe. But, unfortunately, the altered signal will instantly kill you (due to some signal-receiving device that was implanted in your brain for some reason or other, we may suppose). Now, my action doesn't inevitably change the signal: We suppose it's truly undetermined whether my action will change that first signal. Let's say there's a 90% chance that it will (with the result that you are killed), but a 10% chance that it won't (leaving you safe as ever, in that case). Knowing all this,¹⁷⁷ I steal the money, making myself one million dollars richer, but, as was likely to happen, you do end up being killed by the signal, which was indeed altered. Here, I cannot say in the sense that relieves me of all responsibility that your death is something that "just happened," or was not at all "up to me," or was "just the upshot of" little physical events over which I had no control (or that it was "just a matter of chance"). And this does not depend on the particularly high chance of

¹⁷⁷ We are dealing with a causal, pre-moral sense of "responsibility" here, which may* not require that I know about the chances of harm that I am causing for me to be responsible for it. However, this kind of causal responsibility is required for me to be morally responsible for it, so we here give me the knowledge that gives me some moral responsibility for your death, in order to make it clear that in the relevant way I bear some causal responsibility for your death. (*Alternatively, we may be using an intermediate, intentional, but still pre-moral, sense, where I need to know what I'm doing to count as "responsible" for it, but is still pre-moral in that it is silent about whether I'm justified in creating this chance of your death. In that case, we are configuring the circumstances so that I am clearly in the wrong in order to make it clear that in this sense, where being "responsible for" what happens is still required for (but not sufficient for) moral responsibility, I am intermediately responsible for what happens.)

disaster that we have supposed: It all remains the case if we flip the probabilities so that the chances of the signal being changed are only 10% if I steal the money. If you are unlucky and are killed, I bear some responsibility for your death, and cannot say in the sense that relieves me of all responsibility that the altered signal was something that “just happened.”

Low probability or high, there does seem to be a related but narrower use of “just happened” and the other phrases (including “just a matter of chance”) that may be applied (and that I may be anxious to apply, to make it sound as if I am not responsible for what happens), and that will be helpful for us to recognize, but these don’t relieve me of all responsibility for your death. Suppose I have stolen the money, and, during the time between the theft and the sending of the signal, I tell a friend how I got one million dollars, but explain that I have created a 10% chance of your being killed. They and I then figure out how, by fairly quickly taking another action before the signal is sent out, I can get an additional one million dollars from the safe, but only by making it determined and inevitable that the altered signal will be sent out, killing you. The friend leaves me to my decision; I decide to just be happy with the one million dollars I already have; but you are killed, anyway, as the 10% chance of the signal being altered that I have already created is realized. As I later explain to my friend that, no, I didn’t take the second million dollars, I *think* there is some narrow sense tied to my responsibility being somewhat diminished, in which I can say that the altering of the signal is something that “just happened.” This limited sense seems here to just rule out that I bear any *further* responsibility for the altering of the signal beyond what I bear for creating the small chance that it might happen. While I may be anxious to find some good responsibility-related sense in which the altered signal “just happened,” my friend might be quick to remind me that the deadly altering of the signal did not “just happen” in the sense that relieves me of all responsibility, for I am partly responsible for it happening in virtue of my creating the chance that it would happen.

So, the idea that

it is up to an agent what the outcome of a process will be only if the agent is able to arrange things in a way that would make the occurrence of *this* outcome inevitable *and* able to arrange things in a way that would make the occurrence of *that* outcome inevitable

is a plausible one, as van Inwagen claims, only if it expresses a requirement for the outcome being *completely* up to the agent; it seems wrong as a statement of what’s needed for the outcome being *at all* up to them, or for them to be *at all* responsible for it.

Suppose that Jane has done many wrong things in the past that were completely up to her. (If you think this is a problematic assumption here, hold that thought: we’ll be all over it in the next paragraph.) And suppose that the result of the pulse going to the right at the fork—

Jane's remaining silent and not confessing the lie—is in Jane's circumstances seriously wrong in a way similar to those past wrong actions of Jane's. And suppose those past actions have had an effect on the configuration of the neural pathway, affecting the probabilities of the pulse's direction: the chances of the pulse going to the right would have been very low were it not for those past bad actions, but because of those actions, they are in fact very high—though it is still not completely inevitable that the pulse will go to the right. Now, if we suppose that the pulse does what it was likely to do because of Jane's past actions, the result is “not up to Jane” and was something that “just happened” only in the limited way that rules out any *further* responsibility for Jane that goes beyond what responsibility she bears for the creating the high likelihood of her sin. If informed of what's going on in her head, Jane might be anxious to avail herself of such descriptions of the process, to make it sound as if she isn't responsible for what happened. But if we are to be honest with her, we will have to insist that the result is not something that “just happened” in the sense that would absolve her of all responsibility, since she was responsible for the result having the high chance of happening that it had. (And in fact, we'd have to tell her something fairly similar even if the probability of the pulse going in the bad direction were low, so long as she was responsible for it being as high as it was.)

The truly plausible reason for thinking that Jane can't be at all responsible for what happens, and that the result is something that in the relevant sense “just happens,” then, is not that Jane would have to have done something for which she is responsible and that has rendered the outcome of the current process *certain* in order for that outcome to be at all up to her. Rather, it's that Jane could not have done anything in the past for which she was truly responsible and in virtue of which even the chances of the pulse going to the right are as high as they are, because anything she might have done in the past would also be just the upshot of little physical events over which she had no ultimate control. *All* of Jane's, and all of our, deliberations and resulting actions are just the upshot of little, mindless physical events, deterministic or indeterministic, over which we ultimately have no control. Or so we intuit, based on a scientific picture of human thought that we have come to hold. *That's* the great problem with supposing we ever act freely.

50. Freedom's Place in the Causal Structure of the World: The Shape of the Libertarian Hope

The way to resist this great argument for Pessimism is to deny its key (and its only) premise. Our hope for being free then depends on our deliberations and our resulting actions not being just the upshot of tiny physical events over which we have no control. But how could that be?

Haven't we learned that that is how our world works: That everything is the upshot of tiny physical events?

Our guide on how to fit human agency into a picture of the causal structure of the world that is friendly to our freedom will be the already-mentioned Timothy O'Connor—though we will only be taking a quick version of his trip. What I like about O'Connor's presentation of this picture, beyond the ableness of his defense of it from objections (and most of his defensive work won't be conveyed here) is how he frames the picture he is painting.¹⁷⁸ O'Connor realizes that the scientifically inspired way of construing of human thought that he opposes is very deeply ingrained, but insists that the contrary picture he defends deserves to be called the "commonsense picture of agency," and does reflect how things seem to go when we perform deliberate actions:

[I]t seems for all the world to be *up to me* to decide which particular action I will undertake. The decision I make is no mere vector sum of internal and external forces acting upon me during the process of deliberation (if, indeed, I deliberate at all). Rather, I bring it about—directly, you might say—in response to the various considerations: I am the source of my own activity, not merely in a relative sense as the most proximate and salient locus of an unbroken chain of causal transactions leading up to this event, but fundamentally, in a way not prefigured by what has gone before. Or, again, so it seems.¹⁷⁹

When we briefly looked at the phenomenology of deliberate action with O'Connor's help at the end of [section 35](#), I was focused negatively on how we seem not to be causally determined to act as we do. But the appearance is more specific than that—as it needs to be, if it is to be an appearance that validates our freedom in the face of the argument for Pessimism. If the threat to our freedom were just causal determinism, and especially if that threat were paired with one on the other side of our actions being completely uncaused, random happenings, the natural remedy would be to suppose that our volitions (and then resulting actions) are caused by prior states (our reasons and desires being good candidates for these

¹⁷⁸ The way O'Connor frames the issue that I love is particularly how he does this in his fairly early paper, "Agent Causation" (in O'Connor, ed., *Agents, Causes, and Events: Essays on Indeterminism and Free Will*, Oxford UP, 1995, pp. 173–200). The argumentative details of his defense get further developed in subsequent work of O'Connor's (especially his book, *Persons and Causes: The Metaphysics of Free Will* (2002), but I find the material in the first 7 pages (through half-way down p. 180) and the last 2 pages (from the beginning of section VI toward the top of p. 195 to the end of the text) of "Agent Causation" to be a nicely concise and effective framing of the issues. The material in between those blocks responds to objections. Actually, those last two pages at the end is also presented as a response to yet another objection, but given the nature of the objection being addressed, and how O'Connor answers it, his answer serves as a nice presentation of the epistemology of our situation with respect to our own agency.

¹⁷⁹ "Agent Causation," p. 173.

states), but in an indeterministic way. But the mere absence of determinism is not enough, for reasons we are now familiar with.¹⁸⁰ And what positively seems to be the case is that I bring about my decision to act, in such a way that that decision is not caused, deterministically or indeterministically, “by what has gone before.” I seem to do it. I am, or at least seem to be, a “fundamental causal agent” (to use a phrase from a passage quoted below). This is the “agent causation” of the tradition that O’Connor is working in¹⁸¹: an agent producing an effect that in a way that is not reducible to event-event causal relations.

I may be influenced by what precedes my decision to act. (Thus, O’Connor writes that rather than calling us “unmoved movers,” as many refer to what we are on the agent causation view when we act freely, it would be more accurate to say we are “not wholly moved movers.”) We can certainly be free in cases where past events and states of ours satisfy the Leibnizian slogan of “inclining without necessitating” us to act as we do.¹⁸² But we need, and seem to have, more than a lack of necessitation in cases where we act freely. A mere lack of necessitation can be secured by prior events and states causing our decisions to act, but in an indeterministic or probabilistic way, so that our causing those decisions were nothing over and above that indeterministic causing of our decision by those prior events. We need not just the room to make a fundamental (not reducible to event-event causal relations) causal contribution to what happens; we, as agents, need to make that contribution.

¹⁸⁰ For O’Connor’s criticism of the “indeterminist event-causal” view, I here switch to his later paper, “Agent-Causal Power” (in T. Handfield, ed., *Dispositions and Causes*, Oxford University Press, 2009, pp. 189-214), at pp. 193-193: “However, according to many critics (myself among them), indeterminist event-causal approaches falter just here, in the fact that the free control they posit is secured by an *absence*, a removal of a condition (causal determination) suggested by the manifestly inadequate varieties of compatibilism. If there is no means by which I can take advantage of this looser connectivity in the flow of events, its presence can’t confer a greater kind of control, one that *inter alia* grounds moral responsibility for the action and its consequences. Given the causal indeterminist view, if I am faced with a choice between selfish and generous courses of action, each of which has some significant chance of being chosen, it would seem to be a matter of luck, good or bad, whichever way I choose, since I have no means directly to settle which of the indeterministic propensities gets manifested.”

¹⁸¹ In Chapter 3 of *Persons and Causes*, O’Connor discusses and criticizes the agent causation accounts of his most important predecessors: the eighteenth-century Scottish philosopher, Thomas Reid, and the twentieth-century Americans, Roderick Chisholm and Richard Taylor.

¹⁸² As Robert Adams notes, “One of the things Leibniz never tired of saying about free choices is that their causes, motives, or reasons ‘incline’ but do not ‘necessitate’.” A main issue with Leibniz’s account(s) of freedom is how, and, more urgently, whether, Leibniz can make good on this slogan, given various of his views. See Adams, esp. pp. 34-36, for discussion of Leibniz’s attempts to make good on the “not necessitating” part of the slogan, though Adams’s whole first chapter—a reworking of his great paper “Leibniz’s Theories of Contingency” (*Rice University Studies* 63 (1977): pp. 1-41)—can be seen as devoted to the issue.

Well, or at the very least we need to *be able to* make that fundamental contribution. *Perhaps* we can act freely in some cases where prior events cause our decisions to act and then our subsequent actions, despite our making no fundamental causal contribution to the process as agents, so long as we *could have* redirected the process, but opted not to. Some matters of how event- and agent-causal processes interact seem negotiable, so that holders of the agent causal view can take different positions—and we may be at a point where I am inclined to do things a bit differently from O'Connor. He seems to think that in cases where we are conscious of our acting, but make no conscious decision to act, or at least in many such cases, our actions are just event-caused by prior events, and we do not act freely.¹⁸³ I lean toward thinking that when we intentionally act, we agent-cause a volition to act, but we sometimes do this in an automatic way, with no conscious thought given to deliberation.¹⁸⁴ In such cases, we do not get the strong, positive appearance of our acting freely, as we do in the cases involving conscious deliberation that we (following O'Connor) have focused on in making the case for our freedom. But nor is it open to simple observation that we are *not* acting freely in such cases, and I tend to think that we are acting freely. But your mileage (or conditional mileage, if this is for you a hypothetical exercise in which way you would take the agent-causal view) may vary.

What is essential to the view is that in the clearest cases of free action, where we consciously decide what we will do, we are fundamental causal agents, agent-causing our decision to act in a way that is not reducible to event-event causal relations. If that is right, *as it seems to be*, our decisions to act are not just the upshot of tiny little physical events taking

¹⁸³ O'Connor, *Persons and Causes*, pp. 122-123: "Another issue an emergentist version of the agency theory faces is a result of the fact that if there are agent-causal events, there is no neat and simple way of dividing them from event-causal ones. It surely must be allowed that some human behavior, even conscious behavior, is directly brought about by event-causal factors. (Not all action is free action.) As we noted, this is likely to be true of behavior governed by unconscious factors and highly routinized actions. Precisely to what extent, then, is an ordinary human's behavior directly regulated by the agent himself, and to what extent is it controlled by microdeterministic processes? More generally, how do event- and agent-causal processes interact? Even when I act freely, I am usually not even trying to control directly the precise degree of muscle contraction, limb trajectory, and so forth. This makes it plausible to hold that our memory system stores action sequences that we simply activate through conscious choice. It also explains the facility of an experienced performer in carrying out complex movements, such as a sequence of dance steps. It may be that these choices are at times even brought about event-causally, while we simply monitor the result and retain the capacity to agent-causally redirect things as need be."

¹⁸⁴ I wouldn't count the attaining of "the precise degree of muscle contraction, limb trajectory, and so forth" (from the O'Connor passage quoted in the immediately previous note) as intentional actions at all in the typical cases O'Connor has in mind, so I don't think we will to attain them. However, I do think there are many intentional actions that we perform without consciously deliberating over whether to perform them, and often without consciously deciding to perform them.

place in our central nervous systems. The great argument for pessimism is answered by how things seem to go in deliberate action.

But the problem is that this is just one of the seemings involved here at this crucial point in how we see human thought fitting into the world. We face a clash in appearances.

[rest of section and Part under construction]

[more sections and Parts to come...]