

Note: This Work has been made available by the authority of the copyright owner solely for the purpose of private study and research and may not be copied or reproduced except as permitted by the copyright laws of Canada without the written authority from the copyright owner.

Davis, Richard Brian. "Reply to Critics," In *Explaining Evil: Four Views*, edited by W. Paul Franks, 41-47. London: Bloomsbury Academic, Bloomsbury Publishing, Plc, 2019.

Reply to Critics

Richard Brian Davis

I am delighted to reply to Paul Helm, Michael Ruse, and Erik J. Wielenberg. Their challenging comments have given me much to think about, but not (I believe) anything substantial to repent of. In what follows, I take up some of the more pressing issues.

Paul Helm

Paul Helm's comments fall roughly into two parts. First, there is the claim that, as far as explaining evil goes, my Agent-Causal Theism (ACT) is theoretically deficient. It lacks an "account of choice" (31), and it isn't responsibly theistic. Secondly, Helm claims that his brand of theistic determinism is "perfectly consistent" with the "empirical evidence" (30) I cite in favor of agent-causal freedom (apart from which, I say, evil cannot arise).

Let's examine these claims in turn. Note first that ACT involves both an account of evil and of choice. It says, first of all, that a thought, desire, decision, or action (TDDA) counts as evil just in case it is (1) immoral, and (2) freely and deliberately entertained or undertaken, in order to (3) cause or permit significant harm to be done to oneself or others for an unjustifiable end. Thus, contra Helm, I don't think evil "is impossible to identify in terms of necessary and sufficient conditions" (29). I give the conditions—albeit tentatively.

So that's my aerial view of evil. Appended to it is also an account of (free) choice. I say that a TDDA is freely entertained or undertaken only if I am its author or originating cause. I produce it in light of

my reasons for doing so. If the cause of my TDDA is something *besides me*—say, some prior determining cause acting upon me to necessitate that thought, desire, decision, or action—then I am an agent in name only; I am Clarke’s sentient clock (Lumière): capable of neither good nor evil.

Now according to Helm, in saying all this, I haven’t actually given an “account of choice.” But why not? What, precisely, is it to do *that*? I must provide a narrative, he says, explaining the universal human “propensity or likelihood” to evil. “If we are talking of human choice and moral evil,” this must be cashed out either in terms of “something about human beings that *makes* each one liable to evil,” namely, some “feature of human nature, or of a combination of such a feature plus circumstance” (31; emphasis added).

I have two comments. First, if this is what an “account of choice” comes to, Helm is in big conceptual trouble. For then he can account for the blameworthiness of DISOBEY—that original choice to disobey God—only by appeal to the natures and circumstances of those original humans. But then if God is responsible for *that* arrangement, as Helm and I agree, things weren’t in fact “created good” and we have every reason to believe that God is the author of evil.

Second, Helm’s requested narrative confuses the question of evil’s *origin* (How does it arise?) with its *scope* (“Why is evil a universal phenomenon?” (29)). That my account of choice doesn’t address the scope question is no reason to think it doesn’t answer the origin question. In any event, if an explanation for evil’s scope is needed, the proponent of ACT could always adopt an ancillary hypothesis like (C2) above: that evil-inclining human natures are an effect of “the Fall.” That would do the trick nicely. The point is that ancillary questions of this sort can be handled by ancillary hypotheses. The same thing goes, I submit, for Helm’s faulting ACT for failing to tell us what God is like, or why he bestows agent-causal freedom upon us, knowing that we will misuse it and bring about evil.

Turning to Helm’s theistic determinism: while for the most part neglecting my criticisms, he still wants to say that his position is more resilient than I make out. For one thing, it “is perfectly consistent with” (30) the evidence I cite in favor of (the presumption of) our having agent-causal freedom. For a

compatibilist can also say that in making a decision she takes to be free, she experiences herself as deciding to act on one set of reasons, while being aware that she is perfectly able *on that very occasion* to act upon another.

Now Helm thinks that the compatibilist can agree to this “with a good conscience” (30–31). But is that really true? I don’t think so. Granted: she would have been able to decide to act on those different reasons, if she had been determined to do so by a *different* set of causes (on what would then have been some *other* occasion). But this is strictly beside the point. What I’m talking about is our being aware that we could have decided to act on those different reasons under the *same* causes on *one and the same* occasion. If Helm thinks he has experiences like that, he should dismiss them as illusory (given his theory), or admit there is presumptive evidence for agent-causal freedom and against his determinism.

Michael Ruse

Michael Ruse’s comments begin on a happy note (“I have great sympathy for Richard Brian Davis’s thinking about freedom” (32–33)), but then sadly take a precipitous turn to his final conclusion: “Philosophically, he is out to lunch” (36). Ruse proceeds mainly by way of story and assertion. Still, I think we can piece together a kind of argument.

An evil TDDA, I argued, must be one that is *freely* entertained or undertaken. Now this creates a severe problem for what Ruse calls in his own chapter “the supreme theory” (92): his Darwinian Naturalism (DN). For on DN, Ruse’s decision to have that adulterous affair with his colleague’s wife was no more than a physical event-structure in his brain, governed perhaps by “certain genetically determined, strategic rules or directives” (Ruse 2012b, 60). In fact, *he* didn’t make that “decision”; it’s something that happened *to* him—the inevitable outcome of a chain of prior determining causes and effects. He could no more have refrained from his “choice” than Lumière could have decided to stop its hands from moving. Not only was it not a *free* decision, it wasn’t a decision at all. But in that case, it seems to me, Ruse is in the clear. He shouldn’t reproach his 1970s self. What he did was neither blameworthy nor evil.

For as Clarke says, no man “can be angry with his clock for going wrong” (2011, 276).

Alas, however, Ruse is made of sterner stuff. What he wants to say, I take it, is that his decision was both determined and free. Like Helm, he finds himself “pushed toward compatibilism” (34). Now why is that? Because, he says, “the other option just doesn’t work” (ibid)—that option being the one I defended: agent-causation. The problem seems to be that if I am an agent (or first) cause of an evil choice, then my making that choice is uncaused. But if so, that TDDA occurs for no reason: “It just happens.” That “is not decision making,” says Ruse. It is “just craziness” (ibid).

I’m afraid this is simply a non sequitur. It is indeed true that if I am the first cause of a TDDA, then nothing is causing me to cause it. It hardly follows that I don’t have a reason for causing it. What Ruse is (falsely) assuming here is that reasons are causes, so that if I do something for a reason, that reason is the physical, efficient cause (apparently along with various hormones in Ruse’s case) of my doing it. He says we’ll come to see this once “we have reached the [appropriate] level of philosophical maturity,” where we reject the “thinking that reasons take us out of the real world—meaning the natural world” (36). So there it is. When everything is said and done, the reason to think that reasons are causes is that “the supreme theory” demands it.

But that’s not much of a reason. Indeed, it gives rise to a serious category mistake. Reasons are the sorts of things that can serve as premises in arguments. They are *of* or *about* things; brain events and materials (BEMs) are not. Reasons *represent* the world as being this way or that; BEMs do not. Reasons can have a *truth value* and be truth functionally connected (e.g., negated, conjoined, disjoined). Not so for BEMs. Reasons can *logically entail* propositions; BEMs can do no such thing.¹ The fact is: in most crucial respects, reasons are nothing at all like neural event-structures. They aren’t physical, efficient causes. Rather, they are abstract final causes that occasion and inform our decisions and actions.

According to Ruse, what “makes us different” from Lumière and the Mars Rover is “all of the thought business. Reasons.” Thus “I can certainly be judged for drawing a conclusion however determined I am” (ibid.). But how does that follow? If reasons are physical bits of (brain) hardware, subject to prior necessitating causes and operating according to “genetically determined” directives, we can

hardly deny that the Rover also acts on reasons. But then if our reasoning can be judged morally good or evil, so can the Rover's. Nothing could be more absurd.

Sometimes we have to stand up to our cherished theories—at least in certain respects. I would suggest that for Ruse and his all-purpose “supreme theory,” that time has come.

Erik J. Wielenberg

I turn finally to Erik J. Wielenberg's challenging and inventive remarks. Wielenberg doesn't dispute my claim that evil TDDAs require agent-causal freedom and consciousness. Nor does he disagree that the existence of evil is incompatible with both Darwinian Naturalism (as I defined it) and Calvinistic Theism. But he thinks I have overlooked “a fourth option” which he styles *Naturalistic Emergentism* (NE)—in brief, a naturalized version of ACT. “To the extent that such a view is plausible,” we're told, “it weakens Davis's case that the existence of evil requires the truth of Agent-Causal Theism” (37).

Well, to what extent *is* it plausible? Answer: to the same extent that it is plausible to claim that a material substance can be an agent-cause. This claim constitutes the beating heart of NE. In defense of NE's plausibility, Wielenberg draws upon Timothy O'Connor's work on agent-causation. Remarkably, however, O'Connor himself readily concedes: “Taking the agency theory seriously within a basic materialist framework brings forth a whole host of theoretical problems and issues” (1995, 180). Indeed, it does; here is one such problem.

I defined an agent-cause in terms of the power of self-motion: the power to originate volitions in the light of reasons, but without those reasons or anything else causing one to do so. However, since matter is essentially passive—that is, since it acts only if acted upon—it is conceptually impossible for a material object to be an agent-cause; in which case it is maximally *implausible* to suppose that brains are agent-causes or that they possess agent-causal freedom. Perhaps a case can be mustered up along these lines, but I don't see anything like one (or even the hint of one) in Wielenberg's remarks.

Now what about my other condition for evil: that there be *conscious* agents to bring it about? Here I argued that since consciousness is simple and unified, it cannot belong to a material

object composed of parts. Where “B” stands for my brain—a material whole consisting of parts $p_1, p_2, \dots p_n$ —Wielenberg and O’Connor unite as one man to oppose this conclusion. We can consistently affirm, they say, each of the following propositions:

- (1) B is identical with the sum of its parts: $p_1, p_2, \dots p_n$.
- (2) B is conscious.

And yet

- (3) None of $p_1, p_2, \dots p_n$ is conscious.

But how is that possible? If (1) is true, B isn’t a whole “over and above” $p_1, p_2, \dots p_n$. It’s not a *second* object in addition to those parts. But then surely if (2) is true, (3) is false. For my consciousness to belong to B *just is* for it to belong to $p_1, p_2, \dots p_n$. So it seems to me that (1)–(3) comprise an inconsistent set.

Wielenberg apparently disagrees. To see that these propositions are consistent, we need only embrace O’Connor’s conjecture:

CONJECTURE: Consciousness, though radically categorially distinct from the properties of matter, can be caused to emerge at the macro-level of the brain by the “micro-physical goings-on”² of its unconscious parts.

The basic idea, then, is that consciousness is an emergent property of the brain that “*supervenes* on its underlying base properties” (2000, 112)—that is, the properties of $p_1, p_2, \dots p_n$. O’Connor tells us that consciousness is “simple” or “partless” (2000, 111), and qualifies as a “radically new” feature of the world, “‘transcending’ the lower level properties from which it emerges” (2000, 112).

CONJECTURE strikes me as a blunt metaphysical impossibility. Let $R[p_1, p_2, \dots p_n]$ stand for the relational complex consisting of the properties of, and relations between B’s parts. Consciousness cannot be *generated by* $R[p_1, p_2, \dots p_n]$ unless it is in some sense *generated out of* $R[p_1, p_2, \dots p_n]$. Otherwise, we have a case of something coming to be out of nothing. Wielenberg doesn’t “feel the force of the intuition” (39). But it’s not clear to me that this is warranted. There are a couple of things to note.

First, although O'Connor speaks of emergent properties as being "exemplified by objects or systems" (2000, 111), this language is apt to be misleading. It suggests that the relationship between my brain and my consciousness is that of exemplification. But it isn't. O'Connor's properties are universals understood as "immanent constituents of the physical world" (2000, 73). They are tropes: concrete, particularized property instances. Now tropes can't be exemplified; that's just a category mistake. Rather, they are *constituents* of their associated objects or systems.³ So contra O'Connor, the relation obtaining between my consciousness trope (I dubbed it "C") and B is just this: C is a constituent of B. Hence, my brain *does* have conscious parts—precisely the thing CONJECTURE was engineered to avoid.

Secondly, I think O'Connor may well feel the force of the something-from-nothing intuition even if Wielenberg does not. In the end, O'Connor isn't prepared to say that C can be generated by $R[p_1, p_2, \dots p_n]$ without (in any sense) being generated *out of* it. In terms of "generating an emergent," he says, there is a "tendency, present in each microparticle, to jointly achieve its characteristic effect, which is the generation of a property" (2000, 114). So each of $p_1, p_2, \dots p_n$ has an internal *tendency to generate consciousness*. Now if a "tendency to consciousness" is a proto-consciousness, we're saddled with panpsychism. If it isn't a consciousness at all, then what is the plausible story about how such a "tendency" (whatever that means) is grounded in the material properties of the parts of my brain? Here, I'm afraid, we get no answer at all—not even the whisper of a conjecture.

According to Wielenberg, "As long as a view like Naturalistic Emergentism remains plausible, Davis has not established his conclusion that 'evil can only exist if God does'" (40). But that's just it. It isn't plausible.

Notes

- 1 On these points, see Lewis (1960, 21), Plantinga (1993, 116–17), and Gould and Davis (2014, 52–53).
- 2 The expression is Timothy O'Connor's. See O'Connor (2000, 109).
- 3 For details, see Moreland (2013).