

# ON THREE ARGUMENTS AGAINST METAPHYSICAL LIBERTARIANISM

KEN M. LEVY

## I

**T**OWARD THE END OF MY RECENT DEFENSE of metaphysical libertarianism, the theory that genuine responsibility requires “full self-determinism,”<sup>1</sup> I write:

The Self itself—an irreducible, self-creating effort that determines for itself how much it wants to align with its values—will make this determination [whether to reach a certain threshold of effort]. The Self’s self-determining effort is determined neither by factors outside its control, including its previous non-self-determined self, nor by pure chance. It is determined entirely by itself. And this self-determining is itself self-determining. And this self-determining self-determining is itself self-determining. And so on. It is self-determining all the way down, an infinite and instantaneous feedback loop that begins and ends with the Self.<sup>2</sup>

In this rather colorful passage, I was trying to capture metaphysical libertarians’ idea that what lies at the very beginning of our free decisions and actions is a self rather than nothingness and therefore full self-determinism rather than mere indeterminism. But I have since had some misgivings about these words. I now realize that they do not really

---

Correspondence to: klevy@lsu.edu.

<sup>1</sup> I use “genuine responsibility” as shorthand for “basic desert” (Derk Pereboom, *Free Will, Agency, and Meaning in Life* [New York: Oxford University Press, 2014], 2, 127), “true desert-entailing responsibility” (Galen Strawson, *Freedom and Belief* [New York: Oxford University Press, 1986], 3), and “the sort of responsibility that is desert entailing, the kind that makes blaming and punishing as well as promising and rewarding justified” (Michael McKenna, “Compatibilism and Desert: Critical Comments on Four View of Free Will,” *Philosophical Studies* 144, no. 1 [2009]: 12). Putting these different formulations together, genuine responsibility is a property of humans that genuinely justifies or warrants either positive attitudes toward and treatment of them (for example, gratitude and praise) or negative attitudes toward and treatment of them (for example, anger and blame). I will explicate “full self-determinism” in section 2.

<sup>2</sup> Ken Levy, *Free Will, Responsibility, and Crime: An Introduction* (New York: Routledge, 2020), 73.

help the metaphysical-libertarian “cause.” On the contrary, far from showing that metaphysical libertarianism is possible or true, resorting to an arguably nonsensical regress of self-determinings may help to explain why metaphysical libertarianism generally gets a bad rap.<sup>3</sup> Many philosophers assume it is so hopeless that the few who still defend it are stubbornly ignorant, if not worse.

I hope to show, however, that, stubbornness aside, metaphysical libertarians like myself are not ignorant (or worse). Specifically, I will contend that the three strongest arguments against metaphysical libertarianism—what I will refer to as the “randomness objection,” the “constitutive luck objection,” and the “physicalist objection”—are actually unsuccessful and therefore that metaphysical libertarianism is more plausible than the common philosophical wisdom allows. My more positive thesis, which I will refer to as “agent exceptionalism,” will be that, when making decisions and performing actions, human beings can indeed satisfy the four conditions of metaphysical libertarianism: the control condition, the rationality condition, the ultimacy condition, and the physicalism condition.

As the extensive citations in the footnotes indicate, the issue of metaphysical libertarianism is well-trodden ground, but I still hope to make at least three original contributions. In section 2, I will offer what I take to be a uniquely clear, concise, and comprehensive overview of the theory. In section 4, I will introduce the concept of “relational origination” to address the constitutive luck objection. And in section 5, I will argue that metaphysical libertarians need not, as they generally do, resort to event-causal libertarianism in order to make metaphysical libertarianism compatible with physicalism; instead, they may maintain the compatibility of agent-causal libertarianism and physicalism simply by adopting the highly plausible assumption that the brain is special, unique among physical objects in its ability to initiate bodily motion.

---

<sup>3</sup> Robert Kane, “Two Kinds of Incompatibilism,” in *Agents, Causes, Events: Essays on Indeterminism and Free Will*, ed. Timothy O’Connor (New York: Oxford University Press, 1995), 115, and Timothy O’Connor, *Persons & Causes: The Metaphysics of Free Will* (New York: Oxford University Press, 2000), 61, make a similar point.

## II

Suppose that Archie is deciding between calling his neighbor a racial slur (*S*-ing) and refraining from calling her a racial slur (*R*-ing). Archie has reasons for both courses of action. In favor of *S*-ing, Archie is an angry and hateful racist. Call this  $S_{\text{Reason}}$ . Against *S*-ing and in favor of *R*-ing, Archie is aware that racism is morally wrong, name-calling is morally wrong, his neighbor's feelings will be hurt, and his own reputation might suffer. Call these  $R_{\text{Reasons}}$ . In the end, after deliberating between  $S_{\text{Reason}}$  and  $R_{\text{Reasons}}$ , Archie decides to *S*.<sup>4</sup>

Metaphysical libertarians claim that Archie is genuinely responsible and therefore genuinely blameworthy for *S*-ing as long as (a) the minimal compatibilist conditions required for moral responsibility are satisfied;<sup>5</sup> (b) he was not predetermined to *S*—that is, not determined to *S* by causal forces ultimately outside his control;<sup>6</sup> and (c) the indeterminism behind his *S*-ing was located in the “right” place.<sup>7</sup>

---

<sup>4</sup> This is a quintessentially akratic action; Archie knowingly and willingly does what he believes to be the wrong or less valued thing. See Christopher Evan Franklin, “Event-Causal Libertarianism, Functional Reduction, and the Disappearing Agent Argument,” *Philosophical Studies* 170, no. 3 (2014): 423; Kane, “Two Kinds of Incompatibilism,” 129–37; Gideon Rosen, “Skepticism about Moral Responsibility,” *Philosophical Perspectives* 18 (2004): 1307; Scott Sehon, *Free Will and Action Explanation: A Non-Causal, Compatibilist Account* (New York: Oxford University Press, 2016), 136–37, 152–54.

<sup>5</sup> See Neil Levy, *Hard Luck: How Luck Undermines Free Will and Moral Responsibility* (New York: Oxford University Press, 2011), 44.

<sup>6</sup> See Mark Balaguer, *Free Will as an Open Scientific Problem* (Cambridge, Mass.: The MIT Press, 2010), 3; Justin A. Capes, “What the Consequence Argument Is an Argument For,” *Thought* 8, no. 1 (2019): 51–55; Franklin, “Event-Causal Libertarianism,” 416, 424; Robert Kane, *The Significance of Free Will* (New York: Oxford University Press, 1998), 8; John Lemos, *A Pragmatic Approach to Libertarian Free Will* (New York: Routledge, 2018), 47–48; K. Levy, *Free Will, Responsibility, and Crime*, 26, 60–61; Theodore Sider, “Free Will and Determinism,” in *Riddles of Existence: A Guided Tour of Metaphysics*, ed. Earl Conee and Theodore Sider (New York: Clarendon Press, 2005), 115–17; Richard Taylor, *Metaphysics*, 4th ed. (Englewood Cliffs, N.J.: Prentice Hall, 1992), 45–48.

<sup>7</sup> For discussions of where to locate indeterminism, see Balaguer, *Free Will as an Open Scientific Problem*, 16, 67, 132; Mark Balaguer, *Free Will* (Cambridge, Mass.: MIT Press, 2014), 58–63, 67–72; Joseph Keim Campbell, *Free Will* (Malden, Mass.: Polity Press, 2011), 25, 53, 78–79; Robert Kane, *A Contemporary Introduction to Free Will* (New York: Oxford University Press, 2005), 133; K. Levy, *Free Will, Responsibility, and Crime*, 12–13, 68, 72–77; Pereboom, *Free Will, Agency*, 32–33; Manuel Vargas, *Building Better*

Regarding (a), it must be the case that Archie was not compelled or coerced to  $S$ ,<sup>8</sup> he was not subconsciously manipulated or defrauded into  $S$ ,<sup>9</sup> and he exhibited at least minimal rationality or “normative competence” or “moderate reasons-responsiveness.”<sup>10</sup> A fourth

---

*Beings: A Theory of Moral Responsibility* (New York: Oxford University Press, 2013), 65.

<sup>8</sup> See Kane, “Two Kinds of Incompatibilism,” 131–32; Sider, “Free Will and Determinism,” 127–31; Taylor, *Metaphysics*, 43–45.

<sup>9</sup> For discussions of the potential threat to responsibility posed by manipulation, see Campbell, *Free Will*, 66–69; Sam Harris, *Free Will* (New York: Free Press, 2012), 24–25; Kane, *A Contemporary Introduction*, 2–3, 101, 113–15, 118; Lemos, *A Pragmatic Approach*, 29–56; Neil Levy, *Consciousness and Moral Responsibility* (New York: Oxford University Press, 2014), 131–34; Alfred R. Mele, “Free Will and Luck: Reply to Critics,” *Philosophical Explorations* 10, no. 2 (2007): 204–05; Eddy Nahmias, “Your Brain as the Source of Free Will Worth Wanting,” in *Neuroexistentialism: Meaning, Morals, and Purpose in the Age of Neuroscience*, ed. Gregg D. Caruso and Owen Flanagan (New York: Oxford University Press, 2018), 264; Dana Kay Nelkin, *Making Sense of Freedom and Responsibility* (New York: Oxford University Press, 2011), 57–60; O’Connor, *Persons & Causes*, 49; Derk Pereboom, *Living without Free Will* (New York: Cambridge University Press, 2001), 112–16; Pereboom, *Free Will, Agency*, 74–77; Matthew Talbert, *Moral Responsibility* (Malden, Mass.: Polity, 2016), 88–99; Taylor, *Metaphysics*, 46–47; Bruce Waller, *The Stubborn System of Moral Responsibility* (Cambridge, Mass.: The MIT Press, 2015), 212, 247.

<sup>10</sup> See John Martin Fischer and Mark Ravizza, *Responsibility and Control: A Theory of Moral Responsibility* (New York: Cambridge University Press, 1998), 62–91; Joshua Greene and Jonathan Cohen, “For the Law, Neuroscience Changes Nothing and Everything,” *Philosophical Transactions of the Royal Society of London* (2004): 1778; K. Levy, *Free Will, Responsibility, and Crime*, 17–18, 21–25, 60–61, 150–51; N. Levy, *Hard Luck*, 64, 73–74; N. Levy, *Consciousness and Moral Responsibility*, xi, 96–97, 112–14, 126–27; Stephen J. Morse, “Reason, Results, and Criminal Responsibility,” *University of Illinois Law Review* 363 (2004): 382–83, 440–43; Thomas Nadelhoffer and Eddy Nahmias, “Neuroscience, Free Will, Folk Intuitions, and the Criminal Law,” *Thurgood Marshall Law Review* 36 (2011): 159–60; Pereboom, *Living without Free Will*, 139; Derk Pereboom, *Wrongdoing and the Moral Emotions* (New York: Oxford University Press, 2021), 38–39, 127; Christian Perring, “Mental Disorder, Free Will, and Personal Autonomy,” in *The Bloomsbury Companion to Philosophy of Psychiatry*, ed. Şerife Tekin and Robyn Bluhm (London: Bloomsbury Academic, 2019), 437–38, 442–43; Saul Smilansky, “Free Will and Respect for Persons,” *Midwest Studies in Philosophy* 29 (2005): 250–55; compare Pereboom, *Free Will, Agency*, 135–36; Andrea C. Westlund, “Selflessness and Responsibility for Self: Is Deference Compatible with Autonomy?” *Philosophical Review* 112, no. 4 (2003): 484–85, 491–515. Matthew Talbert, “Moral Competence, Moral Blame, and Protest,” *Journal of Ethics* 16, no. 1 (2012): 89–101, argues that normative competence is *not* necessary for moral responsibility.

condition that many compatibilists support is Archie's deciding to *S* on the basis of a reflectively endorsed identification with a specific part of his psychological makeup—his character, his values, his higher-order desires, his rational faculties, or his long-term plans.<sup>11</sup>

Regarding (b), metaphysical libertarians believe that indeterminism is necessary for two conditions that are themselves necessary for genuine responsibility: the ability to do otherwise and “full self-determinism”—that is, Archie's being the first, uncaused cause of his decision to *S*.<sup>12</sup> If Archie's decision to *S* had been predetermined, then it is difficult to see how he could be genuinely responsible for *S*-ing. It was inevitable, “in the cards”—as if Archie had been “set up” to *S* since before he was even born.<sup>13</sup> Conversely, Archie does seem

---

<sup>11</sup> See Franklin, “Event-Causal Libertarianism,” 419–27; Kane, *The Significance*, 90; Lemos, *A Pragmatic Approach*, 71; Nadelhoffer and Nahmias, “Neuroscience,” 162; Perring, “Mental Disorder,” 437; Kenneth A. Richman and Raya Bidshahri, “Autism, Theory of Mind, and the Reactive Attitudes,” *Bioethics* 32, no. 1 (2018); Sider, “Free Will and Determinism,” 130–31; Tamler Sommers, *Relative Justice: Cultural Diversity, Free Will, and Moral Responsibility* (Princeton, N.J.: Princeton University Press, 2012), 17–24; Taylor, *Metaphysics*, 45; George Vucoso, “Background, Responsibility, and Excuse,” *Yale Law Journal* 96 (1987): 1678–81; Westlund, “Selflessness,” 484, 489–91.

<sup>12</sup> See Kane, “Two Kinds of Incompatibilism,” 119–21, 125, 129–32, 137, 142–44, 146; Kane, *The Significance*, 44–78; Robert Kane, “Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism,” *Journal of Philosophy* 96, no. 5 (1999): 224–26; Lemos, *A Pragmatic Approach*, 4, 11–12, 31–33; K. Levy, *Free Will, Responsibility, and Crime*, 9; Pereboom, *Free Will, Agency*, 4, 9. Martine Nida-Rümelin, “Freedom and the Phenomenology of Agency,” *Erkenntnis* 83, no. 1 (2018): 61–86, argues that free action is active behavior, and active behavior requires only the “causal power” (or general capacity) to act otherwise, not the actual ability to do otherwise at that particular time.

<sup>13</sup> See Greene and Cohen, “For the Law,” 1780; Pereboom, *Free Will, Agency*, 73, 82–103; Taylor, *Metaphysics*, 47, 49–50. Put another way, Archie's lack of responsibility for the ultimate cause of his *S*-ing “transfers” to his *S*-ing itself. This is known as the “Transfer of Non-Responsibility Principle” (TNR) in the literature. See Fischer and Ravizza, *Responsibility and Control*, 24, 152–69, 249–51. Mele's “Zygote Argument,” Pereboom's “Four-Case Manipulation Argument,” and Van Inwagen's “Direct Argument” are all designed to support TNR. See Alfred R. Mele, *Free Will and Luck* (New York: Oxford University Press, 2006), 188–95; Pereboom, *Living without Free Will*, 112–17; Pereboom, *Free Will, Agency*, 74–82; Derk Pereboom, “A Defense of Free Will Skepticism: Replies to Commentaries by Victor Tadros, Saul Smilansky, Michael McKenna, and Alfred R. Mele on *Free Will, Agency, and Meaning in Life*,” *Criminal Law and Philosophy* 11, no. 3 (2017): 627–29; Peter van Inwagen, *An Essay*

genuinely responsible for *S*ing if the causal chain behind his *S*ing started with *him* rather than with some event in the distant past and therefore he was at least somewhat independent of his past and the laws of nature.<sup>14</sup>

Regarding (c), the indeterminism behind Archie's decision to *S* rather than *R* must have been located in the "right" place—that is, somewhere in Archie's history where the absence of determinism made possible, and possibly even contributed to, Archie's responsibility for *S*ing. The strongest candidate seems to be the point at which Archie was deliberating between his competing sets of reasons,  $S_{\text{Reason}}$  and  $R_{\text{Reasons}}$ , because this is the point where Archie's conscious self seems to be maximally involved in determining his immediate future.<sup>15</sup>

Putting (a), (b), and (c) together, metaphysical libertarianism captures everything that compatibilism does plus something that compatibilism does not: the notion of self-origination.<sup>16</sup> Compatibilists believe that determinism and moral responsibility are compatible and therefore that Archie may be responsible for *S*ing even if his decision to *S* did not originate with him but, rather, with the Big Bang. By contrast, metaphysical libertarians believe that Archie cannot possibly

---

*on Free Will* (New York: Clarendon Press, 1983), 183–88. See also Thomas Nadelhoffer, "Moral Responsibility Has a Past: Has It a Future?" *Journal of Information Ethics* 28, no. 1 (2019): 25–26.

<sup>14</sup> See Luis E. Chiesa, "Punishing without Free Will," *Utah Law Review* 2011 (2011): 1439–40; Christopher Evan Franklin, "Farewell to the Luck (and Mind) Argument," *Philosophical Studies* 156, no. 2 (2011): 200, 214; Lemos, *A Pragmatic Approach*, 22–27, 80–81, 84–112.

<sup>15</sup> See Franklin, "Farewell," 202, 205–08, 214, 228; K. Levy, *Free Will, Responsibility, and Crime*, 13–14; Lemos, *A Pragmatic Approach*, 106–12; Robert Nozick, *Philosophical Explanations* (Cambridge, Mass.: Belknap Press, 1981), 294–301, 304–05, 314–15. Neil Levy argues that an undetermined weighting of reasons is irrational and arbitrary; see N. Levy, *Consciousness and Moral Responsibility*, 23; N. Levy, *Hard Luck*, 70–71.

<sup>16</sup> See Randolph Clarke, "Toward a Credible Agent-Causal Account of Free Will," in *Agents, Causes, Events*, 203, 211; Timothy O'Connor, "Agent Causation," in *Agents, Causes, Events*, 185, 190; O'Connor, *Persons & Causes*, 51. According to Nida-Rümelin, "Freedom and the Phenomenology," 72–86, the essence of free action is not self-origination but activity or active behavior (as opposed to passivity or "mere happenings"). Nida-Rümelin's theory is still libertarian to the extent that it regards active behavior, and therefore freedom, as incompatible with "microphysical determinism." But it is also compatibilist to the extent that it regards active behavior, and therefore freedom, as perfectly compatible not only with psychological determination but also with even the metaphysical impossibility of counterfactual actions.

be genuinely responsible for *S*ing if he is merely a “middleman”—merely an intermediate, embedded link in the causal chain—between the distant past and his *S*ing.

Whence this fundamental disagreement? Metaphysical libertarians oppose compatibilists’ middleman picture and prefer the self-origination picture for three reasons. First, the middleman picture does not line up well with our experience.<sup>17</sup> When Archie is deliberating between *S*ing and *R*-ing, it certainly *seems* to him that his final decision is entirely up to him, not up to a causal chain stretching back to the Big Bang. And Archie is hardly alone; most if not all of us tend to think of our futures, both short-term and long-term, as genuinely open and therefore genuinely ours to “make happen.”<sup>18</sup>

Second, motivated belief. (This is less an argument for metaphysical libertarianism than it is a motivation, whether good or bad, for subscribing to it.<sup>19</sup>) Metaphysical libertarians tend to believe that genuine self-origination is a necessary condition of one or more things that are “worth wanting”—that is, that we strongly value: accountability, autonomy, dignity, humanity, meaning, meaningful relationships, morality, personhood, just punishment, respect, responsibility, or self-pride.<sup>20</sup> More generally, metaphysical libertarians

<sup>17</sup> See Farah Focquaert, Andrea L. Glenn, and Adrian Raine, “Free Will Skepticism, Freedom, and Criminal Behavior,” in *Neuroexistentialism*, 236–37; Greene and Cohen, “For the Law,” 1781; O’Connor, “Agent Causation,” 173, 196; Pereboom, *Free Will, Agency*, 45–47; Taylor, *Metaphysics*, 53. Compare Kane, “Two Kinds of Incompatibilism,” 132, 143. Martine Nida-Rümelin, “Doings and Subject Causation,” *Erkenntnis* 67, no. 2 (2007): 255–60, 262–64. Nida-Rümelin, “Freedom and the Phenomenology,” 62–64, 71–72, offers a detailed phenomenology of action (“doings”).

<sup>18</sup> See Gregg D. Caruso, “Free Will Eliminativism: Reference, Error, and Phenomenology,” *Philosophical Studies* 172, no. 10 (2015): 2828–30; Kane, *A Contemporary Introduction*, 17.

<sup>19</sup> Daniel Speak, “Toward an Axiological Defense of Libertarianism,” *Philosophical Topics* 3, nos. 1–2 (2004): 358–59, 364–67, argues that values can serve as reasons for belief.

<sup>20</sup> See Kane, *The Significance*, 79–101; Lemos, *A Pragmatic Approach*, 6, 114–16, 131, 157–63; C. D. Meyers, “Automatic Behavior and Moral Agency: Defending the Concept of Personhood from Empirically Based Skepticism,” *Acta Analytica* 30, no. 2 (2015): 194, 206; Morse, “Reason, Results,” 443; Nadelhoffer, “Moral Responsibility Has a Past,” 21–22, 28; Nozick, *Philosophical Explanations*, 2; Saul Smilansky, *Free Will and Illusion* (New York: Oxford University Press, 2000), 161–69; Smilansky, “Free Will and Respect for Persons,” 248–50, 253, 256–61; Speak, “Toward an Axiological

are threatened by determinism in a way that compatibilists are not. They tend to feel that the inevitability of all their decisions and actions is (somewhat) degrading, dehumanizing, and demoralizing. It likens us to rats in mazes, our lives to games that have already been won or lost before we even start playing. Yes, we may not know what will happen, but this mere “epistemic freedom”<sup>21</sup> seems much less valuable than the “real deal,” genuine metaphysical freedom, the difference between watching a recorded game (without knowing the outcome) and actually playing in the game. Compatibilists are happy (enough) with the former; metaphysical libertarians are not.

Third, another possible motivation (not argument) for metaphysical libertarianism is that its conception of human agency is much deeper—much richer, more intriguing, more mysterious—than it is on the compatibilist view. For compatibilists, human agents are just physical objects acting in accordance with their intrinsic natures, immediate circumstances, and the laws of nature, no different in principle from computers and robots. But on the metaphysical-libertarian view, human agents can do what no computer or robot can: reflect, deliberate, and decide to carve out a path not fully determined by their intrinsic natures, immediate circumstances, and the laws of nature. There is still some “wiggle room” left for agents themselves to consider their reasons<sup>22</sup> and make the final determination which way to go—in Archie’s case, whether to *S* or to *R*. And this wiggle room, this tiny “black box” at the root of all our (free) decisions and actions, is one of the great wonders of the universe.<sup>23</sup>

---

Defense,” 366–67. But see Pereboom, *Living without Free Will*, 135, 141–43, 151–52; Pereboom, *Free Will, Agency*, 4, 104–99; Pereboom, “A Defense of Free Will Skepticism,” 625.

<sup>21</sup> See Pereboom, *Free Will, Agency*, 107–26, 194; Pereboom, *Wrongdoing*, 24–26.

<sup>22</sup> More precisely, the agent’s reasons can influence her decision; they just cannot necessitate it. See Franklin, “Farewell,” 208–09; K. Levy, *Free Will, Responsibility, and Crime*, 14.

<sup>23</sup> See Viktor E. Frankl, *Man’s Search for Meaning* (Boston: Beacon Press, 1959), 148: “I would like to depart from a theological concept that has always been fascinating to me. I refer to what is called *mysterium iniquitatis*, meaning, as I see it, that a crime in the final analysis remains inexplicable inasmuch as it cannot be fully traced back to biological, psychological, and/or sociological factors. Totally explaining one’s crime would be tantamount . . . to seeing in him or her not a free and responsible human being but a machine to be repaired.”

## III

By definition, compatibilists believe that indeterminism is not necessary for (genuine) responsibility. Many compatibilists also believe that indeterminism is actually incompatible with genuine responsibility. This conclusion follows from two premises:

1. Responsibility requires a threshold level of agency or self-determinism. If parts of my body are fully determined to move as they do by something other than me—for example, another person such as an evil neurosurgeon remotely controlling my brain, an external object such as a falling rock or strong wind, or an autonomous event in my body such as a twitch or seizure—then I am not responsible for these bodily motions.<sup>24</sup>

2. Indeterminism is not self-determinism; an *undetermined* decision is not even a minimally *self*-determined decision.<sup>25</sup>

Both metaphysical libertarians and antilibertarians accept 1. So the success or failure of this argument, the randomness objection, depends entirely on 2.

2 claims that, logically speaking, the absence of determinism cannot contribute to its very opposite: determinism—specifically, determinism by a certain kind of cause (the self).<sup>26</sup> What self-determinism requires is a self doing the determining. But if Archie's decision to *S* rather than *R* was ultimately undetermined, then it was not ultimately Archie-determined. What ultimately determined the final

---

<sup>24</sup> See Franklin, "Event-Causal Libertarianism," 417; Kane, "Two Kinds of Incompatibilism," 125; K. Levy, *Free Will, Responsibility, and Crime*, 11, 32–33, 108–09.

<sup>25</sup> For different versions of this argument, see Campbell, *Free Will*, 55–56, 71–72, 78–80, 85, 91–92; Gregg D. Caruso, *Free Will and Consciousness: A Determinist Account of the Illusion of Free Will* (Lanham, Md.: Lexington Books, 2012), 15–55; Franklin, "Farewell," 200–02, 209–28; Harris, *Free Will*, 15–16; K. Levy, *Free Will, Responsibility, and Crime*, 11, 66, 76; N. Levy, *Hard Luck*, 34–35 n. 15, 41, 43–44, 53–54, 64–76; Nelkin, *Making Sense*, 86–89; O'Connor, *Persons & Causes*, 59, 76; Pereboom, *Living without Free Will*, 128; Pereboom, *Free Will, Agency*, 41, 47–49; Smilansky, "Free Will and Respect for Persons," 250, 255; Talbert, *Moral Responsibility*, 83; Vargas, *Building Better Beings*, 13–14, 54–58; Vuoso, "Background, Responsibility," 1667–78; Susan Wolf, *Freedom within Reason* (New York: Oxford University Press, 1990), 57–66, 70–73, 78, 92–93. Compare Franklin, "Event-Causal Libertarianism," 429.

<sup>26</sup> See Talbert, *Moral Responsibility*, 84; Taylor, *Metaphysics*, 51; Vargas, *Building Better Beings*, 79–84.

outcome—*S*ing or *R*ing—was not Archie; it was pure chance. Archie's decision to *S* rather than *R* ultimately came from nothing, not from Archie.<sup>27</sup>

The problem may be more clearly framed in terms of so-called cross-world luck: (a) indeterminism entails luck and (b) luck is incompatible with control and responsibility.<sup>28</sup> As we saw in section 2, metaphysical libertarians require indeterminism for responsibility; it must be the case that Archie's decision to *S* was fully determined by *him* and therefore not fully determined by anything outside his control—specifically, the past history of the universe in conjunction with the laws of nature. This means that at time *t*, when Archie decides to *S* rather than to *R*, it must be the case that he could have decided to *R* rather than *S*; that in another possible world that is identical in every respect (internal, external, and nomological) to the actual world up to *t*—call it *PW*—Archie decides to *R* rather than to *S*. But because both worlds are identical in every respect up to *t*, nothing can explain this difference in decisions. Therefore Archie's actual decision to *S* is random, a matter of luck, outside his control. And because control is necessary for responsibility, Archie is not responsible, not blameworthy, for his decision to *S* or his consequent *S*ing (in the actual world).

Metaphysical libertarians, however, can overcome the randomness objection by framing decisions in terms of not only indeterminism and

---

<sup>27</sup> For discussions of this “settling” problem, see Franklin, “Event-Causal Libertarianism,” 424–26; Alfred R. Mele, “On Pereboom's Disappearing Agent Argument,” *Criminal Law and Philosophy* 11, no. 3 (2017): 562–63, 565–66; Pereboom, *Free Will, Agency*, 32; Pereboom, “A Defense of Free Will Skepticism,” 630–34; Pereboom, *Wrongdoing*, 19.

<sup>28</sup> For discussions of the cross-world-luck formulation of the Randomness Argument, see Caruso, *Free Will and Consciousness*, 16–18, 42–51; Franklin, “Farewell,” 202, 217–24; Greene and Cohen, “For the Law,” 1777; Lemos, *A Pragmatic Approach*, 15–20, 58–59, 88; Kane, “Responsibility, Luck,” 217–23, 226, 228–29, 239; K. Levy, *Free Will, Responsibility, and Crime*, 12–16, 70–73; N. Levy, *Hard Luck*, 46–47, 50–57, 64–76; Mele, *Free Will and Luck*, 49–78; Mele, “Free Will and Luck: Reply to Critics,” 195; Mele, “On Pereboom's Disappearing Agent Argument,” 563–66; Nida-Rümelin, “Freedom and the Phenomenology,” 73–74; Pereboom, *Free Will, Agency*, 31–32, 51–54, 65; Sider, “Free Will and Determinism,” 120–21; Taylor, *Metaphysics*, 47, 49–51; Vuoso, “Background, Responsibility,” 1669–70. Mele, “Free Will and Luck: Reply to Critics,” 196, argues that libertarians should respond to the problem of cross-world luck by “ceasing to focus exclusively on the moment of action and the agent's here-and-now powers and taking a broader temporal perspective on agents.” See also Mele, *Free Will and Luck*, 117–23.

self-determinism but also effort.<sup>29</sup> Instead of considering Archie's decision between *S*-ing and *R*-ing to be a matter of which set of reasons he calmly decides to prioritize, suppose we consider it to be a matter of Archie struggling to do what he knows is right (*R*-ing) rather than giving in and doing what he knows is wrong (*S*-ing). Archie is trying to do the right thing, trying to resist the pull of his very strong desire to *S*. In this framing, there *is* one thing that accounts for the different decisions in the actual world and *PW*: Archie. In the actual world, Archie decides to go with his reasons for *S*-ing; in *PW*, Archie decides to go with his reasons for *R*-ing. So Archie's decision was ultimately and entirely up to Archie and Archie alone. Far from being purely random, this kind of ultimate self-causation is exactly what control and genuine responsibility are all about, at least for metaphysical libertarians.<sup>30</sup>

Indeterminism seems to contribute to Archie's responsibility here in a way that it does not in section 2. When we consider Archie's deliberative process as merely undetermined, Archie's ultimately "swiveling" toward *S* rather than toward *R* seems to be a matter of chance. But when we consider Archie's effort to decide in favor of *R*, the random-swiveling imagery is unsuitable because it implies the very opposite of effort: both indifference and mere mechanism. Which "wins"—Archie's better self or Archie's worse self—will depend entirely on his level of effort, entirely on how much energy he exerts on behalf of success (*R*-ing) over failure (*S*-ing). And this level of effort will be entirely up to Archie, not up to the causal chain leading from the Big Bang to this moment. Archie will make all the difference here. And such ultimate difference-making is the essence of control and therefore of genuine responsibility.<sup>31</sup>

---

<sup>29</sup> Robert Kane relies on this framing as well. See Kane, "Two Kinds of Incompatibilism," 133–35; Kane, *The Significance*, 27–28, 126–70, 193–94; Kane, "Responsibility, Luck," 225, 227, 231–35; Kane, *A Contemporary Introduction*, 135–45. See also Lemos, *A Pragmatic Approach*, 57–81; N. Levy, *Hard Luck*, 58–63. Some metaphysical libertarians frame genuine responsibility in terms other than effort. For example, Balaguer, *Free Will*, 63–78, uses "torn decisions," and Lemos, *A Pragmatic Approach*, 106–12, uses the weighting of reasons.

<sup>30</sup> See Franklin, "Farewell," 223–24.

<sup>31</sup> See Lemos, *A Pragmatic Approach*, 14–22, 26–27, 60–62, 65–66, 77–81, 85; N. Levy, *Hard Luck*, 69; Nozick, *Philosophical Explanations*, 310–16. See also Pereboom, *Free Will, Agency*, 28, 40–43; Vargas, *Building Better Beings*, 83–84.

The antilibertarian will likely reiterate the randomness objection: If Archie's level of effort is *undetermined*, then it is not Archie-determined. Here is the argument in more detail. Suppose three things: Archie failed, he would have succeeded if he had tried harder, and he could have tried harder—that is, he had not reached the upper limit of expendable effort. The question is: Why didn't he try hard enough? The most likely reason is that he felt the additional effort would have been too uncomfortable or unpleasant. Call this  $U_{\text{Reason}}$ .  $U_{\text{Reason}}$  was clearly competing with all his reasons for wanting to prevail—call them  $P_{\text{Reasons}}$ . If Archie's deliberation between these reasons was undetermined, then whichever reason(s) prevailed was just a matter of chance.

Fortunately, metaphysical libertarians now have available to them a stronger response: An undetermined level of effort *can* be fully self-determined, at least more self-determined than if it had been predetermined.<sup>32</sup> Archie was not randomly swiveling between  $U_{\text{Reason}}$  and  $P_{\text{Reasons}}$ . Even if he “landed on”  $U_{\text{Reason}}$ , it was not because *chance* landed him there. Instead, it was because *he* landed *himself* there. When Archie was oscillating between  $U_{\text{Reason}}$  and  $P_{\text{Reasons}}$ , the struggle was not between two independent sets of reasons but, rather, between Archie on the one hand and  $U_{\text{Reason}}$  on the other.<sup>33</sup> As a result, calling both (a) the process by which Archie decided whether he “beat”  $U_{\text{Reason}}$  or  $U_{\text{Reason}}$  “beat” him and (b) the result of this process random or a matter of chance is simply inaccurate. Instead, the process and result were purely matters of will—strength of will. And strength of will, whether successful or not, is fully attributable to Archie. Archie, not his competing reasons, was the ultimate difference-maker.<sup>34</sup>

In this way, Archie's effort can be both undetermined and up to him without contradiction. The three conditions that need to be satisfied: Archie could have tried harder to  $R$ ; had he tried harder to  $R$ , he would

<sup>32</sup> See Kane, “Two Kinds of Incompatibilism,” 140; Kane, “Responsibility, Luck,” 232, 236–37.

<sup>33</sup> See Brendan de Kenessey, “People Are Dying Because We Misunderstand How Those with Addiction Think: A Philosopher Explains Why Addiction Isn't a Moral Failure,” *Vox* (March 16, 2018).

<sup>34</sup> See Balaguer, *Free Will as an Open Scientific Problem*, 93; Robert Kane, “Free Will, bound and unbound: reflections on Shaun Nichols' bound,” *Philosophical Studies* 174, no. 10 (2017): 2482; Lemos, *A Pragmatic Approach*, 98–112; Pereboom, “A Defense of Free Will Skepticism,” 630, 633–34. Azim F. Shariff and Kathleen D. Vohs, “The World without Free Will,” *Scientific American* 310, no. 6 (2014): 79, argue that belief in free will actually strengthens will power (and disbelief weakens it).

have succeeded; and his level of effort to *R* rather than *S* was not predetermined. Only if all three of these conditions are satisfied may we consider Archie to be genuinely responsible, genuinely blameworthy, for *S*-ing rather than *R*-ing—and for all future actions that trace back to his decision to *S*.

Of course, the antilibertarian may still insist that whether Archie's will is strong enough, whether his effort to *R* succeeds, is still ultimately a matter of chance. Yes, he may have a reason for *S*-ing ( $S_{\text{Reason}}$ ) and he may have reasons for *R*-ing ( $R_{\text{Reasons}}$ ); to this extent, whichever decision he makes, *S* or *R*, is not random in the sense of baseless.<sup>35</sup> But it is still random in the sense of *ultimately* baseless; Archie has no *deeper* or *contrastive* reason for *S*-ing over *R*-ing (or vice versa), no "metareason" for siding with one set of reasons over the other set of reasons. And this contrastive or metareason is necessary to defeat the randomness objection.<sup>36</sup>

The metaphysical libertarian might respond that Archie's *S*-ing rather than *R*-ing does have a deeper reason—call it  $S_{\text{DeeperReason}}$ —but then the antilibertarian would simply ask why  $S_{\text{DeeperReason}}$  prevailed. Either  $S_{\text{DeeperReason}}$  was opposed or it was not. If it was opposed—by  $R_{\text{DeeperReason}}$ —then it remains unexplained why, and therefore a matter of chance that,  $S_{\text{DeeperReason}}$  prevailed over  $R_{\text{DeeperReason}}$ .<sup>37</sup> If it was not opposed, then  $S_{\text{DeeperReason}}$  was predetermined to prevail, which is contrary to indeterminist hypothesis.

The more successful response for the metaphysical libertarian is to simply hold the line and argue that a contrastive reason is not necessary for genuine responsibility.<sup>38</sup> Instead, mere "dual" or "plural"

<sup>35</sup> See Kane, "Two Kinds of Incompatibilism," 126–27, 133–34, 139.

<sup>36</sup> See Franklin, "Farewell," 219–24; Mele, *Free Will and Luck*, 72–73; Mele, "Free Will and Luck: Reply to Critics," 195–98.

<sup>37</sup> See Campbell, *Free Will*, 79–82; Thomas Nagel, *The View from Nowhere* (New York: Oxford University Press, 1986), 115–17.

<sup>38</sup> See Roderick M. Chisholm, "Agents, Causes, and Events: The Problem of Free Will," in *Agents, Causes*, 99; Clarke, "Toward a Credible," 206; Kane, "Two Kinds of Incompatibilism," 128, 136–39; K. Levy, *Free Will, Responsibility, and Crime*, 15–16, 71, 77; Nida-Rümelin, "Freedom and the Phenomenology," 84–85; Nozick, *Philosophical Explanations*, 302, 304–05; O'Connor, "Agent Causation," 188–89, 194; O'Connor, *Persons & Causes*, 75–76, 91–95. Franklin, "Farewell," 221–24, 228–29, argues that metaphysical libertarians *can* provide a contrastive explanation: Whichever reasons

rationality—that is, an agent’s having a reason for whichever choice she makes, even if she does not have a deeper reason for making one choice rather than the other—is sufficient.<sup>39</sup> If the antilibertarian is still not satisfied with this response, the burden is on her to explain why it is wrong. And it is not clear to me that this burden has been, no less can be, satisfied.<sup>40</sup>

#### IV

The constitutive luck objection says that Archie’s effort cannot be fully up to him because this “him”—Archie—was never itself up to him.<sup>41</sup> Archie cannot be the ultimate cause of his effort because he is not the ultimate cause of the self that made (or caused or created) this effort. Instead, Archie was created by his two biological parents. And while this self may have caused itself to change in certain ways over the course of its development, none of these self-changes were ultimately self-caused because they were all made by a self that was not itself ultimately self-caused.<sup>42</sup>

---

ultimately prevail helped raise the probability of the action they succeeded in motivating.

<sup>39</sup> See Kane, *The Significance*, 107–09, 142–44; Kane, “Responsibility, Luck,” 237–38; Kane, *A Contemporary Introduction*, 128–30. Kane supplements dual rationality (sometimes dual control) with “dual willings,” the idea that deliberation generally involves simultaneously trying to make both choices. See Kane, “Responsibility, Luck,” 231–35; Kane, *A Contemporary Introduction*, 135–45. For helpful discussions of Kane’s dual-willings approach, see Caruso, *Free Will and Consciousness*, 45–51; Lemos, *A Pragmatic Approach*, 8–27, 37–39, 57–81; N. Levy, *Hard Luck*, 57–63; Talbert, *Moral Responsibility*, 83–85. For a discussion of plausible metaphysical-libertarian alternatives to Kane’s dual-willings approach, see Lemos, *A Pragmatic Approach*, 83–112.

<sup>40</sup> See Franklin, “Farewell,” 223–24; K. Levy, *Free Will, Responsibility, and Crime*, 71.

<sup>41</sup> See Greene and Cohen, “For the Law,” 1779–80; Lemos, *A Pragmatic Approach*, 85; K. Levy, *Free Will, Responsibility, and Crime*, 67–69, 86; Nahmias, “Your Brain,” 265; Talbert, *Moral Responsibility*, 70–79; Wolf, *Freedom within Reason*, 377–78.

<sup>42</sup> See Campbell, *Free Will*, 55–56, 71–72, 85, 91–92; Franklin, “Farewell,” 213; Harris, *Free Will*, 61–62; Kane, *A Contemporary Introduction*, 16, 169; Lemos, *A Pragmatic Approach*, 46–53, 75–76; K. Levy, *Free Will, Responsibility, and Crime*, 67–69, 86–87; N. Levy, *Hard Luck*, 4–5; Nozick, *Philosophical Explanations*, 354–57; Strawson, *Freedom and Belief*, 25–60;

Suppose, for example, that (a) Archie was brought up in a white-racist family and community; (b) at the age of fifteen, he was first exposed to the opposing point of view—that racism is seriously wrong; and (c) after struggling to adopt this more enlightened view, Archie gave in and stuck with bigotry. The constitutive luck objection says that even if it can be said that Archie recreates himself here, the Archie who recreates himself is something that Archie had not himself created or recreated. If we think of Archie as “doubling down” on his racism, the Archie who engages in this doubling down is a self that, up to this moment, had been entirely out of his hands. Archie did not create his body, brain, or environment; they were all “given” to him—fully shaped by external causes. So Archie’s doubling down on racism is really just a combination of external forces causing Archie to double down on racism. Even though it may seem to Archie like he is finally taking charge, the decision that results from this taking charge is just as much the effect of these external forces as all of his previous, *ex hypothesi* non-ultimately-self-caused decisions.

While this objection is powerful, it is not decisive. The self certainly cannot create itself *ex nihilo* or, therefore, anything it ever causes, such as its decisions. But the metaphysical impossibility of (human) creation *ex nihilo* is not fatal to full self-determinism, genuine responsibility, or (therefore) metaphysical libertarianism. The fact that the self does not create *ex nihilo* the “raw materials” of any decision does not mean that it cannot create *ex nihilo* the relationships among these raw materials. In Archie’s case, the raw materials are primarily his deliberation between *S*-ing and *R*-ing (that is, between  $S_{\text{Reason}}$  and  $R_{\text{Reasons}}$ ); his beliefs about objective moral truth; and his values, including his valuation of what he believes to be objective moral truth.

Again, Archie is not the ultimate originator of these raw materials, but he can still fully originate the relationships among these raw materials. And this power, the power of genuine, fully self-caused

---

Galen Strawson, “The Unhelpfulness of Indeterminism,” *Philosophy & Phenomenological Research* 60, no. 1 (2000): 150–52; van Inwagen, *An Essay on Free Will*, 146; Waller, *The Stubborn System*, 56–61, 72–73, 83; Wolf, *Freedom within Reason*, 14. Bruce Waller, *The Injustice of Punishment* (New York: Routledge, 2018), 115, puts it well: “The skill and fortitude and optimism and confidence which you ‘play the cards that were dealt you’ are ultimately *among* the cards that were dealt you (dealt either by one’s genetic legacy or one’s early conditioning).”

“relational origination,” is all that genuine responsibility requires.<sup>43</sup> So, yes, Archie did not create *ex nihilo* the self that engaged in his struggle between *S*ing and *R*ing, but the non-ultimately-self-created self that engaged in this struggle effectively recreated itself, or part of itself, at that moment.<sup>44</sup> For the first time in his life, Archie *owned* the racism; he made it *his*—without being predetermined to make it his. Going forward, then, Archie is fully responsible for the racist behavior that traces back to this ultimately self-caused decision.<sup>45</sup>

At this point, the antilibertarian may yet again try to renew her initial argument against metaphysical libertarianism, the randomness objection. She may argue that even if Archie is the ultimate cause of his decision to *S* rather than *R*, the fact that, under the very same circumstances, he equally might have ended up deciding to *R* rather than *S* makes his decision to *S* random, merely undetermined rather than ultimately self-determined and, therefore, not something for which Archie is responsible.

But there is nothing random about Archie’s chosen relationship here, given that one of the relata is Archie’s beliefs about objective moral truth. To the extent that Archie chooses or tries to choose in accordance with these beliefs, a nonarbitrary reason can be given for whichever choice he makes. If, contrary to hypothesis, Archie had chosen to *R*, this would have been because he chose to act in accordance with his beliefs about objective moral truth, which is the very opposite of random. And even though Archie did not actually choose to act in accordance with his beliefs about objective moral truth, his choice was still not random; he made a deliberate value-choice—specifically, to prioritize his base desire to *S* over his beliefs about objective moral truth. If the antilibertarian thinks that a deeper, contrastive reason for prioritizing his base desire to *S* over his beliefs about objective moral truth is necessary for genuine responsibility, she

---

<sup>43</sup> For a similar argument, see K. Levy, *Free Will, Responsibility, and Crime*, 66–77.

<sup>44</sup> See Kane, *The Significance*, 96; compare N. Levy, *Consciousness and Moral Responsibility*, 104–08.

<sup>45</sup> See Kane, “Free Will, bound and unbound,” 2484–85; Lemos, *A Pragmatic Approach*, 34–35, 69–73; K. Levy, *Free Will, Responsibility, and Crime*, 64, 68–77; Mele, “Free Will and Luck: Reply to Critics,” 199.

must offer an argument for this conclusion. And as I suggested at the end of section 3, I am not aware of any such argument.

## V

The physicalist objection says that even if full self-determinism were metaphysically possible, it is not clear that it is physically possible.<sup>46</sup> Specifically, it remains difficult to give a fully physicalist (or scientific or naturalist) account of this “godlike” power, of how (full) self-determinism can be realized by entirely physical objects governed entirely by the laws of physics.<sup>47</sup>

I offer two responses to the Physicalist Objection. First, physicalism is arguably false because neither philosophers nor scientists have been able to explain either (a) how the brain, a complex combination of carbon-based molecules, gives rise to seemingly non-carbon-constituted consciousness or (b) how this seemingly nonphysical consciousness can make a causal difference in the physical world—for example, by causing these very words to be typed out on my computer.<sup>48</sup> Consciousness, mental causation, and agency are realized

---

<sup>46</sup> The question of what exactly “physical” means is very difficult to answer. See Daniel Stoljar, *Physicalism* (New York: Routledge, 2010), 5, 13–162. Fortunately, for the purposes of this article, I need not commit to any particular definition. My only commitment here is that, on any reasonable definition, human brains and human bodies are indeed physical.

<sup>47</sup> For discussions of the tension between metaphysical libertarianism and physicalism (or science), see Patricia E. Erickson and Steven K. Erickson, *Crime, Punishment, and Mental Illness: Law and the Behavioral Sciences in Conflict* (New Brunswick, N.J.: Rutgers University Press, 2008), x, 10–12, 14, 21, 85–86, 111–12, 133, 180, 183; Michael S. Gazzaniga, *Who’s in Charge? Free Will and the Science of the Brain* (New York: HarperCollins Publishers, 2011), 218–19; Greene and Cohen, “For the Law,” 1779; K. Levy, *Free Will, Responsibility, and Crime*, 3–5; Colin McGinn, *The Mysterious Flame: Conscious Minds in a Material World* (New York: Basic Books, 1999), 167–69; Alfred R. Mele, *Free: Why Science Hasn’t Disproved Free Will* (New York: Oxford University Press, 2014), 3, 91; Nagel, *The View*, 110–11; Nelkin, *Making Sense*, 87; O’Connor, *Persons & Causes*, 54; Pereboom, *Free Will, Agency*, 3; Sider, “Free Will and Determinism,” 118–22; Stoljar, *Physicalism*, 2, 13, 15–16; Vargas, *Building Better Beings*, 7; Vuoso, “Background, Responsibility,” 1677–78; Waller, *The Stubborn System*, 174–76.

<sup>48</sup> Stoljar, *Physicalism*, argues that there is no version of physicalism that is “both true and deserving of the name.” Philosophers who discuss the “explanatory gap” between consciousness (and sometimes agency) and brain

in the brain, the most complicated physical object in the known universe,<sup>49</sup> but how the brain realizes these wonderful things may just be beyond its own ability to understand.<sup>50</sup> Given these deep explanatory gaps at the center of physicalism, it is not entirely clear why any theory of responsibility, including metaphysical libertarianism, must be consistent with it in order to be considered acceptable.<sup>51</sup>

Second, even if we assume that physicalism is true, it does not necessarily follow that we must then reject metaphysical libertarianism; conversely, metaphysical libertarianism does not necessarily entail nonphysicalism. There are two possibilities: Either the agent reduces to the brain (or part of the brain) or the agent is irreducible to the brain (or part of the brain). I will argue that there is a version of the first possibility—the “reducibility thesis”—that is compatible with metaphysical libertarianism.<sup>52</sup>

---

include David J. Chalmers, *The Conscious Mind: In Search of a Fundamental Theory* (New York: Oxford University Press, 1996), 47, 93–122, 244–45; Sam Coleman, “Being Realistic: Why Physicalism May Entail Panexperientialism,” in *Consciousness and Its Place in Nature: Does Physicalism Entail Panpsychism?* ed. Anthony Freeman (Charlottesville, Va.: Imprint Academic, 2006), 51; Philip Goff, “Experiences Don’t Sum,” in *Consciousness and Its Place*, 56–57; Jaegwon Kim, *Physicalism, Or Something Near Enough* (Princeton, N.J.: Princeton University Press, 2005), 6, 170–74; K. Levy, *Free Will, Responsibility, and Crime*, 2–3; N. Levy, *Consciousness and Moral Responsibility*, 27–28; William G. Lycan, “Resisting ?-ism,” in *Consciousness and Its Place*, 68; Fiona Macpherson, “Property Dualism and the Merits of Solutions to the Mind–Body Problem,” in *Consciousness and Its Place*, 86–87; Colin McGinn, *The Mysterious Flame*, 8–18; Nagel, *The View*, 13–53; Thomas Nagel, *Mind & Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False* (New York: Oxford University Press, 2012), 13–69; O’Connor, *Persons & Causes*, 110–25; William Seager, “The ‘Intrinsic Nature’ Argument for Panpsychism,” in *Consciousness and Its Place*, 133–34, 139–40, 143–44; Peter Simons, “The Seeds of Experience,” in *Consciousness and Its Place*, 148–50; Galen Strawson, “Realistic Monism: Why Physicalism Entails Panpsychism,” in *Consciousness and Its Place*, 15, 17–18, 21–24, 28–29; Catherine Wilson, “Commentary on Galen Strawson,” in *Consciousness and Its Place*, 181–82.

<sup>49</sup> See Simons, “The Seeds of Experience,” 148.

<sup>50</sup> See Kim, *Physicalism*, 170–74; McGinn, *The Mysterious Flame*.

<sup>51</sup> See Randolph Clarke, “Are We Free to Obey the Laws?” *American Philosophical Quarterly* 47, no. 4 (2010); Nida-Rümelin, “Freedom and the Phenomenology,” 86; O’Connor, “Agent Causation,” 174.

<sup>52</sup> I am assuming here that physicalism requires either reduction of the mental to the physical or reductive explanation of the mental by the physical. Philosophers who are sympathetic to reductive physicalism in one form or

The latter—“irreducibility thesis”—is generally referred to as “agent causation” (or “nonoccurrent causation”), which is the idea that agents are special, nonphysical entities that can somehow cause events in the physical world without themselves being either caused by events in the physical world or dictated by the laws of nature.<sup>53</sup> One common objection to agent causation is that it lacks explanatory power. This objection takes two forms. The first regards rationality: Agent causation fails to explain why an agent (like Archie) arrived at a particular decision (at a particular time). All it says is that Archie decided to *S* (at *t*) because Archie himself made (or caused) this decision; it says nothing

---

another include Kim, *Physicalism*; Lycan, “Resisting ?-ism,” 65–68; Kevin Morris, *Physicalism Deconstructed: Levels of Reality and the Mind-Body Problem* (New York: Cambridge University Press, 2019); David Papineau, “Comments on Galen Strawson[’s] ‘Realistic Monism: Why Physicalism Entails Panpsychism,’” in *Consciousness and Its Place*; Richard Rorty, *Philosophy and the Mirror of Nature* (Princeton, N.J.: Princeton University Press, 1979), 17–127; David M. Rosenthal, “Experience and the Physical,” in *Consciousness and Its Place*, 121–27; J. J. C. Smart, “Ockhamist Comments on Strawson,” in *Consciousness and Its Place*; compare Nahmias, “Your Brain,” 252–54. Importantly, however, there are three theories that reject my assumption about reduction/reductive explanation here. First, nonreductive physicalism says that physicalism can be compatible with mental (qua mental) causation. See Coleman, “Being Realistic,” 41–42, 44; Kim, *Physicalism*, 33–35. (For powerful critiques of nonreductive physicalism, see Kim, *Physicalism*, 32–69; Morris, *Physicalism Deconstructed*, 1–155.) Second, panpsychism says that even the smallest bits of matter enjoy some minimal level of experience. See David Skrbina, “Realistic Panpsychism: Commentary on Strawson,” in *Consciousness and Its Place*, 151–57; Strawson, “Realistic Monism,” 24–29; Galen Strawson, “Panpsychism? Reply to Commentators with a Celebration of Descartes,” in *Consciousness and Its Place*, 256–62; compare Coleman, “Being Realistic,” 48–52. Third, John Campbell, “Interventionism, Control Variables and the Qualitative World,” *Philosophical Issues* 18 (2008): 437–42, argues that physicalism does not entail that “every physical outcome has a physical cause.”

<sup>53</sup> For discussions of agent causation, see Clarke, “Are We Free,” 392, 395–96; Greene and Cohen, “For the Law,” 1779; Kane, “Two Kinds of Incompatibilism,” 117–24; Lemos, *A Pragmatic Approach*, 8; Nida-Rümelin, “Doings,” 263; Nida-Rümelin, “Freedom and the Phenomenology,” 70–72; Pereboom, *Free Will, Agency*, 30–31, 50–51, 57; Pereboom, *Wrongdoing*, 20–21; Taylor, *Metaphysics*, 51–53. But Pereboom, *Free Will, Agency*, 105, argues that indeterminism may not be essential to agent-causation. Nida-Rümelin, “Doings,” 260–61, 264–67, proposes “weak dualism” or “subject-body dualism,” which is agent causation without substance dualism: (a) agency is essential to explaining “doings” (of which actions are a subset), (b) agency cannot be reduced to the brain, and (c) agency still at least nomologically and possibly metaphysically supervenes on neural events.

about *why* Archie decided to *S*, no less why Archie decided to *S* rather than to *R*. Archie's decision is therefore random, a matter of chance, not in his control. (See section 3.) The second form of the explanatory objection regards physicalism: Agent causation fails to explain how an *ex hypothesi* nonphysical agent acts on physical objects like the brain and body. The very idea seems too "magical" and unscientific.<sup>54</sup>

Because of these two problems, metaphysical libertarians generally adopt the former—reducibility—thesis, which is commonly referred to as "event-causal" libertarianism.<sup>55</sup> According to event-causal libertarianism, decisions and actions are, like everything else in the

---

<sup>54</sup> See Greene and Cohen, "For the Law," 1780, 1782; Kane, "Two Kinds of Incompatibilism," 115–16, 124, 141, 145; Pereboom, *Free Will, Agency*, 50, 61–62, 65–69; Pereboom, *Wrongdoing*, 20–21; Talbert, *Moral Responsibility*, 82; Taylor, *Metaphysics*, 53. While "many philosophers who discuss [metaphysical libertarianism] seem to simply assume that its adherents are 'dualists'" (O'Connor, *Persons & Causes*, 121), modern self-designated agent-causal libertarians tend to repudiate dualism. See, for example, Clarke, "Toward a Credible,"; O'Connor, "Agent Causation," 178–80, 196; O'Connor, *Persons & Causes*, 43–66, 71–84. Still, I find it difficult to understand modern agent-causal libertarians' attempts to naturalize agent causation without reducing it to event causation. Compare Nagel, *Mind & Cosmos*.

<sup>55</sup> For discussions of event-causal libertarianism, see Balaguer, *Free Will as an Open Scientific Problem*, 67, 152; Balaguer, *Free Will*, 2–3; Franklin, "Farewell," 203–04, 219; Franklin, "Event-Causal Libertarianism," 415–24; Kane, "Two Kinds of Incompatibilism," 124–25, 141–42; Kane, *The Significance*, 17, 94, 192–95; Kane, "Responsibility, Luck," 223; Kane, *A Contemporary Introduction*, 4; Lemos, *A Pragmatic Approach*, 8–13, 22–25; O'Connor, *Persons & Causes*, 68–71, 108–25; Pereboom, *Wrongdoing*, 18–19. Some libertarians have responded to the problems with agent-causal libertarianism by adopting noncausalism. Noncausalists generally maintain that (a) causality takes place only between events or states of affairs; therefore (b) the relationship between agents and their decisions/actions is not causal; therefore (c) whatever problems arise from viewing the relationship between agents and their decisions/actions as causal automatically evaporate. For discussions of noncausalism, see Campbell, *Free Will*, 73–76; O'Connor, *Persons & Causes*, xiii, 24–27, 34, 59; Pereboom, *Free Will, Agency*, 5, 31, 39–43; Pereboom, *Wrongdoing*, 21–22. For discussions of the distinction between agent-causal and event-causal libertarianism, see Campbell, *Free Will*, 80–82; Caruso, *Free Will and Consciousness*, 19–29; Clarke, "Toward a Credible," 202–10; Clarke, "Are We Free," 392, 396; Randolph Clarke, "Free Will, Agent Causation, and 'Disappearing Agents,'" *Noûs* 53, no. 1 (2019): 76–80; Franklin, "Event-Causal Libertarianism," 413–14, 426; Kane, "Two Kinds of Incompatibilism," 144–45; Lemos, *A Pragmatic Approach*, 8–10; O'Connor, *Persons & Causes*, 68–101, 175–80, 184, 186; Pereboom, *Living without Free Will*, 135–36; Pereboom, *Free Will, Agency*, 3, 55–58, 67–69; Sehon, *Free Will and Action Explanation*, 3–4, 74–126; Sider, "Free Will and Determinism," 120.

world, caused by events—specifically, mental events that are themselves assumed to be reducible to neural events. In contrast to agent-causal libertarians, event-causal libertarians suggest that what causes Archie to decide to *S* is not Archie qua nonphysical agent but rather Archie qua events in Archie’s mind/brain.

The main difficulty with the event-causal approach is reconciling mental/neural causation with the self-origination that is central to metaphysical libertarianism. (See section 2.) Either decision- and action-causing mental/neural events are ultimately caused by the agent or they are not. If they are, then it is unclear how event-causal libertarianism differs from, and thereby avoids the problems with, agent causation. If they are not, then they are either entirely uncaused or caused by earlier mental/neural events. And neither of these disjuncts seems compatible with self-origination.<sup>56</sup> As we saw in section 3, being entirely uncaused is not being self-caused. And being caused by earlier mental/neural events, which are themselves either entirely uncaused or caused by earlier mental/neural events (and so on), is not being either ultimately *self*-caused or *ultimately* self-caused.

My proposal is that there is one, and only one, way to satisfy the four conditions that most, if not all, modern metaphysical libertarians require for genuine responsibility: (a) the agent has control over her decision and action (the control condition), (b) the agent decides and acts as she does for a reason (the rationality condition), (c) the agent is not predetermined to decide or act on this reason (the ultimacy condition), and (d) the agent’s reason, decision, and action are fully reducible to brain states or events (the physicalism condition). According to my Agent Exceptionalism, the only way to reconcile conditions a through d is to assume that the laws of nature are “open” or “lenient” enough<sup>57</sup> to allow agents qua brains to do something that no other physical object can: initiate physical events—that is, cause certain

---

<sup>56</sup> This is known as the “disappearing agent” argument, which says that event-causal libertarianism leaves no causal role for the agent herself in decision-making or action. See Clarke, “Free Will, Agent Causation”; Franklin, “Event-Causal Libertarianism,” 413–15, 424–29; Mele, “On Pereboom’s Disappearing Agent Argument”; Pereboom, *Free Will, Agency*, 3, 5, 32–43, 48–51, 55; Pereboom, “A Defense of Free Will Skepticism,” 629–34; Pereboom, *Wrongdoing*, 19–20.

<sup>57</sup> See Michael Tooley, *Time, Tense, and Causation* (New York: Clarendon Press, 1997), 101–03.

bodily motions for reasons that are reducible to neural states/events without these causings being predetermined.<sup>58</sup>

Admittedly, agent exceptionalism is somewhat ad hoc and likely violates the “causal closure of the physical”—the widespread physicalist assumption that the behavior of every physical object, including every human brain and body, must be fully explainable by the immediately preceding physical state of that object, its immediate physical environment, and the laws of physics.<sup>59</sup> But I see no reason, a priori or empirical, why this metaphysical possibility is not also a physical possibility in the actual world.<sup>60</sup> First, it is not at all clear that physicalism entails causal closure of the physical. For example, two physicalist theories, nonreductive physicalism and panpsychism, are arguably compatible with nonphysical causation.<sup>61</sup> Second, while it is difficult to understand both how agents can be reducible to brains and how brains, physical objects, can initiate motion in themselves and in

---

<sup>58</sup> Compare Clarke, “Are We Free”; K. Levy, *Free Will, Responsibility, and Crime*, 76–77; Nida-Rümelin, “Doings,” 264; O’Connor, *Persons & Causes*, 65, 76–77; Taylor, *Metaphysics*, 51–53. While Agent Exceptionalism may seem similar to Donald Davidson’s anomalous monism (Donald Davidson, “Mental Events,” in *Essays on Actions and Events*, ed. Donald Davidson [New York: Clarendon Press, 1980]), I believe that the two theories differ in several important respects, including their positions on whether natural laws help uniquely determine (that is, dictate) brain-events.

<sup>59</sup> For discussions of causal closure, see Caruso, *Free Will and Consciousness*, 16–17, 24–27, 31–42; Coleman, “Being Realistic,” 44; K. Levy, *Free Will, Responsibility, and Crime*, 12, 76; Kim, *Physicalism*, 15–16, 42–52; Lycan, “Resisting ?-ism,” 69; Pereboom, *Living without Free Will*, 128; Stoljar, *Physicalism*, 211–14; Taylor, *Metaphysics*, 51–53; Vargas, *Building Better Beings*, 76; Waller, *The Stubborn System*, 140–41, 147, 208–09, 237.

<sup>60</sup> See Stoljar, *Physicalism*, 220–22. But see Greene and Cohen, “For the Law,” 1777: “What if, somewhere, deep in the brain, there are mysterious events that operate independently of the ordinary laws of physics and that are somehow tied to the will of the brain’s owner? In light of the available evidence, this is highly unlikely. . . . [T]here is not a shred of scientific evidence to support the existence of *causally effective* processes in the mind or brain that violate the laws of physics.” See also Pereboom, *Wrongdoing*, 21.

<sup>61</sup> See n. 52 above. Nida-Rümelin, “Doings,” 269–71, offers a *reductio ad absurdum* against causal closure of the physical: (a) If causal closure is true, then we suffer from “massive, permanent and fundamental error” (that is, “massive illusion”) about agency, about what causes us and others (including animals) to behave as we do; and (b) it is at least implausible, if not absurd, to think that we are all massively, permanently, and fundamentally wrong about something for which we have overwhelming subjective and observational evidence. See also Nida-Rümelin, “Freedom and the Phenomenology,” 65–69.

other physical objects, our inability to understand a particular phenomenon *P*, given either current limits on our scientific knowledge or permanent cognitive limits, does not mean that *P* is physically impossible—any more than our current inability to understand the consciousness–brain relation means that consciousness does not exist or is not caused or constituted in some way by the brain.<sup>62</sup> Epistemology does not dictate metaphysics. What we can and cannot understand does not determine the intrinsic nature of things, what they are in themselves.<sup>63</sup>

## VI

Whether physicalism is true or false, it is indisputable that the human agent is both natural and exceptional. It is natural in the sense that it is created by a clearly natural process (reproduction) and causally impacts the natural world, both the body it “inhabits” and the objects that surround this body.<sup>64</sup> But it is exceptional in the sense that it has distinctive cognitive abilities: “writing, rewriting, thinking, reasoning, creating, imagining, deliberating, reflecting, judging, concentrating, exerting, pausing, daydreaming.”<sup>65</sup> (I would add joking and being amused to the list.) These abilities, which are not possessed either at all or nearly to the same degree by any other object or organism, make the human agent unique, not to mention fascinating.<sup>66</sup> And it is this unique (and fascinating) status, this agent exceptionalism,

---

<sup>62</sup> See Coleman, “Being Realistic,” 44–48; Goff, “Experiences Don’t Sum,” 56–57; Lycan, “Resisting ?-ism,” 68; Macpherson, “Property Dualism,” 74; McGinn, *The Mysterious Flame*, 65–68, 84–85; McGinn, “Hard Questions,” in *Consciousness and Its Place*, 98; Georges Rey, “Better to Study Human Than World Psychology: Commentary on Galen Strawson’s ‘Realistic Monism: Why Physicalism Entails Panpsychism,’” in *Consciousness and Its Place*, 111; Seager, “The ‘Intrinsic Nature’ Argument,” 139, 143–44; Simons, “The Seeds of Experience,” 148–49; Daniel Stoljar, “Comments on Galen Strawson[’s] ‘Realistic Monism: Why Physicalism Entails Panpsychism,’” in *Consciousness and Its Place*, 174–76; Stoljar, *Physicalism*, 193–95, 210; Strawson, “Realistic Monism,” 4, 7, 28; Strawson, “Panpsychism?” 274–75; Wilson, “Commentary on Galen Strawson,” 182. But see Strawson, “Realistic Monism,” 15, 18.

<sup>63</sup> See Taylor, *Metaphysics*, 1–8.

<sup>64</sup> See Kane, *The Significance*, 92–95.

<sup>65</sup> K. Levy, *Free Will, Responsibility, and Crime*, 4.

<sup>66</sup> See *ibid.*, 12, 70–71, 76.

that makes metaphysical libertarianism at least a viable, if not the most plausible, theory of free will and responsibility.

*Louisiana State University*