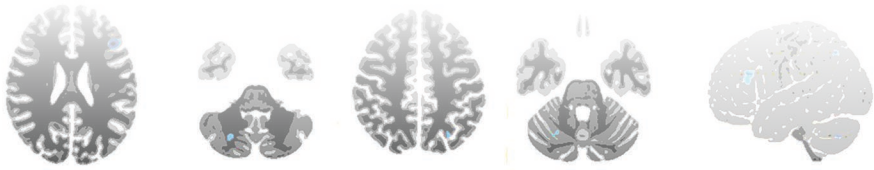
Sleepwalking Towards Bethlehem: On What We Know but Do Not Believe

David J. Frost

The Missouri Review, Volume 49, Number 2, Summer 2026, pp. 166-185 (Article)

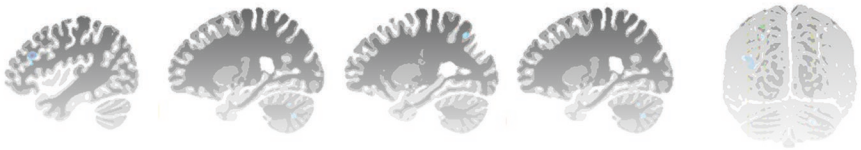
Published by University of Missouri
DOI: https://doi.org/10.1353/mis.2026.a994626



SLEEPWALKING TOWARDS BETHLEHEM

On What We Know but Do Not Believe

David J. Frost



Scans from peerj.com

Two years ago, my fiancée told me that I'd been sleepwalking and carrying on full-fledged conversations with her each night. When Kelly first relayed this, her brow twitched slightly with what I took to be suppressed irritation. These mornings, I never remembered anything about whatever it was I'd done overnight. After all, somnambulism is a classic example of automatic behavior. Yet the archetypal image of the sleepwalker—nightcap with drooping tassel, eyes closed, arms outstretched in front, wrists bent down—is wrong. Sleepwalkers engage in complex behavior: they might cook a meal, drive a car, carry on a conversation, murder a relative—all without the accompaniment of consciousness. It raises the question: if I'm not conscious, how am I seemingly able to adapt to

changing circumstances? Who was talking when I was conversing? As Kelly and I sat up in bed that first morning, her pent-up prickliness made me acutely aware of the score between us. What was the regrettable conduct I unfreely enacted each night but for which I must nevertheless accept responsibility? Evidently, I had been peeing in the closet.

During the first night, Kelly says, my rustling woke her up, and she asked me what I was doing. "Nothing," I said. "Going to the bathroom." But she could see that I was standing in front of the open closet and said, "That's not the toilet." "Yes, it is," I said, sounding offended, while gravely continuing to conduct my business. We would learn later why it's said not to wake up a sleepwalker. Although it's like a preprogrammed routine is being run, the person's capability to react in complex ways *can* remain, but often only in limited form. I could interact with Kelly, for instance, but likely only to counteract her attempts to thwart me. If you attempt to impede a sleepwalker, they might instinctively find it aversive and—one article suggested—react impulsively, yet they'd be without the benefit of consciousness to rein in their anger. That was the theory, anyway.

Luckily for Kelly and me, my nonconscious self was more likely to be cavalier and arrogant than angered by impediment. While following the going-back-to-bed algorithm, I responded to Kelly's protestations, she reports, by uttering "an aloof grunt of bemused recognition." During the night of my second episode, when she again informed me that I was standing at the closet, not the commode, I let go another self-satisfied, haughty grunt. But I apparently persevered in my task, because she could hear the unmistakable evidence—the damp applause—of my misguided micturition. A faux-fancy language affords me distance from disgusting deeds.

The third night Kelly screamed at me, "ARE YOU AWAKE?" "Yes," I said. In that moment, she struggled with what to conclude about my inner mental world. For I was passing the Turing test, so to speak, yet she didn't know if I was actually conscious. As we found out in the morning, I hadn't been.

When I asked her to recount the events for me once more (in furtherance of writing this essay), I noticed that her brow involuntarily furrowed again. The biggest irritation, she said, was having to clean up the soggy mess after I'd already roboted back to bed.

"That sucks," I replied dutifully. Yet, moments later, I caught myself feeling uninvolved again, like I hadn't done it. I *knew* I had done

it; I took what she said for the truth. But I did not *feel* that belief in my body. Usually when I've done something wrong, my self-criticism is more despairing. The knowledge of the injured party's disapprobation is more—I don't know—*piercing*. I get a nauseating knot in my stomach. But this time, my understanding that I was responsible for Kelly's aggravation was entirely intellectual and bloodless.

Because I have so tenuous a connection to what I've done while sleepwalking, I find it easy to be glad of my inopportune nocturnal emissions. If I had not sleep-peed in our closet, then I would not possess such a good example of an action for which I take responsibility but which I did not do of my own free will. It's been a fascinating experience. It's like when, at the dentist, I cannot keep my tongue from pressing against the hygienist's elongated mirror.

Above all, sleepwalking gave me a concrete path into wondering what it'd be like *as an experience* to deny free will, as a small number of public intellectuals have recently done. It was weird enough to think that I had carried on a conversation without free will one, two, or three times. What would it be like to think *every* action you undertake lacks free will? Not just outbursts but even conscious decisions made after careful deliberation. What would it be like to believe that each thought in one's stream of consciousness was an effect of previous causes?

For Stanford biology professor Robert Sapolsky, for instance, what is the experience of denying his own free will like? Sapolsky is convinced we do not have free will, so from his perspective, he *knows* it. But he's admitted he catches himself reverting to the default setting of free will *belief* every time he nonchalantly accepts responsibility. "Someone says, 'You're dressed nice today,' and I say, 'Thanks,' as if I had anything to do with it," Sapolsky joked during a guest appearance on NPR's *Radiolab*.

When Sapolsky manages to believe what he knows—when he is actively remembering that he doesn't have free will—how does his day-to-day self-conception change? Is denying free will bad for you? Or does it increase well-being? I'm talking about *psychological* effects. I became much less interested in the classic philosophical arguments for and against free will. Released from responsibility, does a free-will denier run amok, his face frozen in a gleeful rictus? Or does he lose all motivation and succumb to fatalism, thinking it doesn't matter what he does? Would denial, as Sartre worried, serve as an excuse for bad behavior? Perhaps, as Sapolsky imagines, pervasive denial would bring about a fairer, more just, less cruel society, "a giant leap forward in our moral

progress," as he's said. How would our thinking change if we cleansed our vocabulary of the term "free will"?

When Sapolsky and others deny free will, they don't mean that everything is destined, that what will be will be, no matter what we do. It's not fatalism necessarily. The precise logical implications of denial are, however, not a matter of what Sapolsky might want them to be; they are a matter of science and logic and will depend on arguments and experimental evidence. And, as it turns out, there is already an empirically informed debate going on.

My research into the psychological implications of free-will denial led me to novel places on the intellectual landscape: to a confrontation with bodies of knowledge called "informational existential threats," to the conceptual affinity between progressive political opinion and free-will skepticism, to the idea of an Area 51 for anti-free-will research, and finally, to the existence of brain-scanning technology that can read the inner contents of your mind and know what you are going to do before you know.

When I sat down in my office to start my research, the first thing I noticed about online pictures of Sapolsky was the compassionate twinkle in his eyes. Framed by long, curly gray hair and a gray Garibaldi beard, Sapolsky's eyes sparkled like quartz in the Tolstoian rocky crag of his sun-wrinkled face. A specialist in neuroendocrinology, he's spent thirty sunny summers in Africa studying simians. (Sorry! I couldn't control myself.) Returning to the same troop of baboons each year, he'd observe their behavior, tranquilize them, and take blood samples. Sapolsky found, among many other discoveries, that baboons ranking low in the troop's dominance hierarchy had chronically elevated levels of stress hormones (glucocorticoids) with concomitant negative effects on health, like hypertension, and on behavior, like an increase in isolated sulking and displacement aggression. Sapolsky estimates that nearly 50 percent of all violence among his troop of baboons is displacement aggression. Shit, as we know, flows downhill. "A high-ranking male loses a fight and so chases a subadult male, who promptly bites a female, who then lunges at an infant," Sapolsky has written. The more a baboon displaces aggression after losing a fight for rank, Sapolsky says, the lower his glucocorticoid levels are when the good doctor, sans sun hat, tests his blood.

That explanation—a diminishment in rank causes stress hormones, which cause the loser to punch down to gain relief—is only one example

from Sapolsky's immense storehouse of explanations of primate behavior. Sapolsky's 2018 best-selling book, *Behave: The Biology of Humans at Our Best and Worst*, exhaustively explores the myriad causal factors relevant to the conduct of your behavior: from your mix of neurotransmitters nanoseconds before a given action to the moral emotions bequeathed to all of us by our highly cooperative hunter-gatherer forebears; from your stress hormones thirty minutes ago to those of your prenatal environment; from the sexual selection processes at work in your grandparents' mating rituals to the organizing principles—whether individualistic, communitarian, or honor-based—of your great-great-great-grandparents' culture. Perpetually harried in his speech and even apparently in his ellipses-filled emails, Sapolsky—a 1987 MacArthur "Genius Grant" recipient—is charming and often approaches silliness. He turns serious, however, when he addresses what he calls the "appalling moral affront" of a society that continues to believe in free will. "The injustice and harm perpetrated by our justice system's antiscientific belief in free will should keep us up at night screaming," he said on another episode of *Radiolab*. Sapolsky has staked out the most uncompromising position against free will in the public sphere. "We are nothing more or less than the cumulative biological and environmental luck, over which we had no control, that has brought us to [where we are]," he writes. In October 2023, Penguin published Sapolsky's latest book, which is about living day-to-day with free-will denial, titled *Determined*: *A Science of Life without Free Will*. Having established in *Behave* that every human behavior has a cause, Sapolsky writes in *Determined* about the potential consequences of society-wide free-will denial and how to integrate those implications into our self-understanding.

After my sleepwalking episodes began, Kelly cracked open my copy of the eight-hundred-page *Behave*, and I started reading a galley copy of *Determined*. We would compare notes at our ritual of Sunday three o'clock tea, recommended by our premarital counselor. I also read the eighty-page *Free Will* by Sam Harris, another denier. Harris is a multi-topic public intellectual, unaffiliated with any university and independent-minded. The famously militant New Atheist, best-selling author, advocate of Buddhist meditation, neuroscience PhD, podcaster, app developer, and much-criticized critic of radical Islam agrees with Sapolsky that jettisoning belief in free will would be good for both society and the individual. Harris—who looks like a cross between Spock and a dour, karate-chopping Ben Stiller—writes that "losing the sense

of free will has only improved my ethics—by increasing my feelings of compassion and forgiveness, and diminishing my sense of entitlement to the fruits of my own good luck."

While Sapolsky's and Harris's predictions about a world after denial are quasi-utopian, the science does not appear to back up their sunny speculations. "Believe it or not," I said to Kelly at tea, "scientists have run experiments in this domain"—lab experiments to discover the psychological effect of denying free will on the person denying it. It's not long been the subject of empirical study, however. In one of the first papers, published in 2008, behavioral economist Kathleen Vohs, of the University of Minnesota's business school, and Jonathan Schooler, a psychologist at the University of British Columbia, found that when test subjects are primed with anti-free-will belief (by reading a passage denying free will in Francis Crick's book *The Astonishing Hypothesis*), they cheat and steal during real-money lab games in greater percentages than subjects in the control group.

Nearly all the other results in this line of research run counter to Sapolsky's and Harris's rosy predictions. Reading dozens of journal articles, I learned that disbelief in free will causes dishonest, selfish, aggressive, and conforming behavior; reduces helpfulness; hampers learning from one's misdeeds; curtails thinking for oneself; diminishes recycling; lowers expectations for occupational success; and decreases quality of job performance.

"What the fuck's 'recycling' doing in there?" Kelly asked in a moment of apparent lalochezia. I wondered silently what she might have been so upset about that she needed the pressure release of swearing. "Well," I said slowly, accusingly, drawing out the vowel, "Wehhhhl, experimental moral psychology has an unavoidable dollhouse nature." The lab can only be a simulacrum of real life's moral dilemmas. While researchers may want to test a belief's contribution to bad behavior, universities' institutional review boards, thankfully, do not allow actual harm to come to the participants in these studies. Similarly, good, heroic, or praiseworthy behavior must have a stand-in, which is what recycling is doing in there. To test whether an experimentally manipulated condition contributes to aggressive behavior, a number of studies have adopted the strategy of giving participants the opportunity to clandestinely drench another person's food in hot sauce. And, you guessed it, participants primed with free-will disbelief administer more hot sauce, i.e., are more

aggressive, than controls. The vast majority of the science around the psychological effects of denying free will seems to militate against the optimists and in favor of the pessimists. Who's right?

There's actually a third position. In the recent intellectual conversation around free-will denial, three mutually inconsistent schools of thought have developed. On the one hand, there's Sapolsky and Harris, the optimists, who think denial would rid the world of moral harms that are presently dependent on free-will credence; on the other hand, empirical research seems to show that not believing in free will has negative effects on our behavior and psychology; while, on the *third* hand—and this position is as strange as a third hand—there are those who say free will *is* an illusion *but* that the effects of disbelief would be so detrimental that this knowledge should be *suppressed*. Yes, that's right; some philosophers and scientists who deny free will's existence don't think it would be wise to let everyone know the truth. As Vohs and Schooler write, "If exposure to deterministic messages increases the likelihood of unethical actions, then identifying approaches for insulating the public against this danger becomes imperative."

They're not just trying to justify their grant money with a practical upshot for their research. They categorize the falsity of free will as an "informational existential threat," like the knowledge of how to build a nuclear bomb or make smallpox into an aerosol. The knowledge that free will is an illusion must be possessed only—the thinking goes—by a small elite who, supposedly, will be strong enough to resist the bad effects. Suppressionists never say *nobody* should know this truth. How could they, since *they* know it? I imagine there must be a certain frisson to making an exception for oneself precisely here. The physicist Paul Davies, writing in an essay commissioned by—of all places—*Foreign Affairs*, says free will is "a fiction worth maintaining," and denial "is the world's most dangerous idea."

It's not new, this moral panic. Catastrophizing about free-will denial goes way back. But today, it is not mere idle speculation, which it was in 1524, when Erasmus argued for free will against Martin Luther's denial. Erasmus worried that the common man was too "weak, ignorant, and wicked" to handle what Luther was proposing. Therefore, Luther ought not to say it. But Erasmus did not seriously contemplate Luther's denial becoming widespread, so his worries were merely abstract pearl-clutching—and self-serving to boot. Erasmus should have felt the bite of intellectual conscience, the pull of a certain epistemological scruple.

Because it was in Erasmus's time—and still *should* be in ours—a principle recognized by all truth-seekers that one cannot silence one's interlocutor in a debate by asserting that merely uttering their conclusion would cause harm. It's akin to an ad hominem argument, as it does not address the issue on its merits (i.e., on the question of whether it's true or not). Furthermore, it would preclude the discovery of harmful truths, for, indeed, some truths are harmful.

Today's suppressionism is different because Erasmus *believed* in free will. These days, recommending suppression is most interesting (and most ethical) when it's a move made by *deniers*, i.e., the thinker in Luther's position. Suppressionist scientists are advocating for the sequestration of their own research. They argue that if we really and truly believe "the un-free will" (a translation of Luther's 1525 title) to be dangerous knowledge, then it should be "born classified," as is the case now for developmental research for new nuclear weaponry. "Perhaps the government could," I joked with Kelly at the next tea, "in the vein of UFOs, attempt to discredit as a kook anyone who proclaims the dangerous and secret truth that we don't have free will."

Thankfully, outright deception may not be necessary. "We are *fortunately* deceived about free will," writes Saul Smilansky, professor of philosophy at the University of Haifa, in Israel, and the foremost academic proponent of letting the illusion be. In Plato scholarship, an elite promulgating a falsehood to maintain social order is called a "noble lie." A noble lie affirming free will may not be necessary, since few people see their way to the arguments for denial, and most, when the arguments are brought to their attention, do not accept their implications anyway—even when they cannot gainsay the logic. Smilansky worries that graduate students, working under the conviction that free will is false, will become depressed or psychotic postdisillusionment. So he discourages PhD students from working with him. He writes, "We must protect humanity in the hothouse of illusion from the chill of the ultimate perspective." "I like that," I remarked to Kelly. "The icy chill of the ultimate perspective."

Adopting the ultimate perspective without illusions would be to cross one more threshold in the ongoing diminishment of human metaphysical dignity, which includes a familiar list: Copernicus decentering the Earth and us with it; Darwinism making us animal and accidental; Nietzsche and Freud dethroning self-transparent consciousness and foregrounding the unruly unconscious. It's another instance of

specialized scientific knowledge undermining our day-to-day, practical concepts of the self, the world, and our exalted, or at least meaningful, place in it. Threatening to "unweave the rainbow" (as Keats said of Newton), it moves us farther along in our existentially demeaning journey toward the ultimate perspective and the debunking explanations found there. With denial, we are indeed "slouching towards Bethlehem." This phrase, coming from a Yeats poem describing a dreadful second coming of Christ, was borrowed by Joan Didion to describe the ostensibly utopian world of 1968 Haight-Ashbury as, instead, an apocalyptic nightmare. Analogously, free-will denial is seen by some as utopian and by others as catastrophic. The latter think it would cause societal collapse, which must be prevented by a paternalistic noble lie. The former, like Sapolsky and Harris, think the ultimate perspective would yield an egalitarian, socially just, quasi-heavenly sublime, if only we could get people to adopt it. "No one would deserve to be treated any better or any worse than anyone else," as Sapolsky writes.

When I shared that quotation with Kelly at tea, it struck me that denial seems to rhyme with progressive politics. Because, for example, it is the positive belief in free will that allows some people (usually on the political right wing) to blame poverty on the freely made bad decisions of the poor rather than on structural causes, which, by comparison, play a deterministic role. Also, many of the societal changes that progressives want to see are logical implications of free-will denial, including prison reform and intervention to prevent childhood poverty, to name only those on Sapolsky's short list.

"The progressive-aligned cause Sapolsky is most passionate about in his books," Kelly replied, "is government intervention to prevent *all* childhood poverty." Why? Because what the stress of being poor does to the developing brain is unfair. Poor people, according to research by Sapolsky and others, have chronically high levels of glucocorticoids—"Like low-ranking baboons," I interrupted mischievously—with adverse effects on health, brain function, and behavior. Sapolsky points out that the way stress works in baboons is exactly how stress works in humans. As part of the fight-or-flight system, the stress response is incredibly ancient—it's shared by all primates and every mammal, bird, fish, and reptile. Same chemical compounds, same interactions.

When triggered, glucocorticoids prepare a creature's body for emergency action, shutting down systems not essential in emergencies,

including digestion, reproduction, neurogenesis, and growth. (I learned there is such a thing as "stress dwarfism"—when growth is markedly stunted by chronic stress in childhood.) In animals less social than baboons and humans, stress hormones are washed away as quickly as they are triggered, as Sapolsky cleverly captured in the title of his 1994 book, *Why Zebras Don't Get Ulcers*. By contrast, baboons hunt, gather, and eat all the food they need each day in only a few hours, "leaving the rest of their time to make each other miserable," Sapolsky writes. The result is chronic stress (only for some, it must be said). In adult humans, the chronic stress of poverty creates a detrimental cognitive load, diminishing the ability to focus attention, according to studies on the psychological effects of poverty. In a 2013 study by researchers at Princeton, rich and poor study participants performed equally well on standard tests of cognition while preoccupied with lab problems involving trivial amounts of money. But if the amounts were $2,000, a poor person, on average, displayed a drop in cognitive function analogous to 13 IQ points, or an entire night's sleep. Rich participants preoccupied by the $2,000 scenarios averaged no drop in performance on the cognitive test.

For a child, whose brain is still developing, poverty-caused stress can spell disaster, Sapolsky says. Stress hormones adversely affect the growth of the prefrontal cortex, generating measurable effects like poorer executive function, worse emotional regulation, and deficits in impulse control. Neuroscientist David Eagleman, author of many popular books on the brain, says that poor impulse control makes sense as a factor in criminogenesis. Studies show that prison populations tend to fill out surveys with the correct answers to moral dilemmas—that is, they know right from wrong. But prisoners score below average on tests of impulse control, like the marshmallow test and others. Thereby, nonviolent economic crime, like shoplifting or burglary, would *not* be caused by lack of economic resources coupled with an upbringing that failed to inculcate prosocial values. It would be caused, instead, by an above-average tendency toward impulsivity, itself caused by factors beyond one's control.

Sapolsky writes, "The socioeconomic status of a child's family predicts the size, volume, and gray matter content in the prefrontal cortex . . . *in kindergarteners*." Kindergarten is crazy early to have effects like this show up. But socioeconomic status, i.e., class—human primates' version of rank among baboons—determines important features of the brain even earlier. The effect is there "in toddlers. In six-month-olds. In four-week-olds. You want to scream at how unfair life can be," Sapolsky

writes. "Actually," Kelly said, "Sapolsky should be advocating to eradicate poverty not just for all children but for all *mothers*." "Oh, yeah," I said, "you're right." For a four-week-old to have a smaller brain, the glucocorticoids must have been present in utero.

"You know who *doesn't* believe in free will?" I said to Kelly at the next tea. "The tobacco companies, the food industry, the Sackler family, advertising agencies, smartphone manufacturers, and social media platforms such as Facebook and TikTok." Addictive cigarettes and booze; snacks high in sugar, salt and fat; Oxycontin; product placement in movies; the doomscroll: the corporations behind these phenomena know precisely how effective they are in the aggregate with respect to changing the behavior of individuals. The "attention economy," according to Scott Galloway, celebrity investor, popular podcaster, and marketing professor at New York University, is only a subset of the larger "addiction economy"—made up of companies that have figured out ways to control our behavior by manipulating our dopaminergic reward system. Galloway says—borrowing from E. O. Wilson—that we have neolithic instincts, medieval social institutions, and futuristic, godlike technology, which adroitly pushes our buttons and pulls our strings.

These companies deny free will, at least notionally—i.e., they count on determinism—in their manipulations of our behavior; could denial, instead, help us protect ourselves? Presently, in a culture that believes in free will, resistance to the addiction economy is more difficult than it would be after pervasive denial. Free-will credulity makes it possible to deny disingenuously that the phenomena of addiction and manipulation are even happening ("There's no system of control. People *freely choose* to eat our unhealthy food product!") and recommends methods of resistance that might be hopeless ("Just use your willpower!"). Right now, strategies to resist corporate manipulation often recommend increased optionality so individuals can ostensibly *choose* a different path rather than the maladaptive, addictive one. "Sure, there's lots of processed food, but shoppers are free to buy fresh vegetables and whole foods just a few aisles away." Sure, they bought the object of addiction's desire, but they could have done otherwise. Yet if you deny free will, it'd be obvious that the best strategy would be to eliminate or diminish the addictive stimuli in the first place. Research suggests that you would increase the probability of healthy choices if veggies were centrally prominent in a store and processed foods were hard to get to. "But, notice," I said to

Kelly, "Free will drops out of the picture when predicting, manipulating, incentivizing, or just explaining behavior."

"Denial seems to be approaching a kind of critical mass right now," Kelly said at the next tea. "I mean, what's the historical evolution of determinism's threat to free will?" Myself, I'd assumed that determinism began to take its contemporary shape with the new mechanistic explanations of the Scientific Revolution of the sixteenth and seventeenth centuries. But recently I have begun to wonder if determinism were not older and deeper in our psyches—a feature of explanation per se, not mechanistic explanation specifically. Rather than coming about because of early science's impressive discoveries of the clockwork-like laws of cause and effect, maybe the first threat to humanity's sense of autonomy and dignity came from deterministic *stories*—that is to say, from Fate. Narrative paths and story arcs, imposed on us by gods or divine forces, were taken to undermine something important to us: happiness or meaningfulness, if not also autonomy. Fate, oracles, curses, and the like preoccupied the ancient Athenians especially—as evidenced by their invention of the genre of tragic theater. At Oedipus's birth, a divine oracle prescribed the inescapable story of his life, its end, and its lasting meaning, as clear and unchangeable as if chiseled in the firmament.

Twenty-one hundred years before the Scientific Revolution, Democritus (ca. 400 BCE) said that determinism governed the interactions between instances of what he speculated was the basic unit of matter: an invisible, indivisible substance he called *atomos*. Reacting to Democritus a hundred years later, Epicurus postulated that for there to be leeway for free will, atoms would have to undergo a "swerve" (Latin: *clinamen*), straying from their predetermined path. In other words, something somewhere would have to diverge from the deterministic causal chain *in virtue of no cause at all* (otherwise its diverging swerve would *itself* be part of a deterministic causal chain).

The Epicureans' contracausal reaction to Democritus's determinism seems to presage the position that philosophers today call libertarianism. No association with the specious political philosophy of Ayn Rand, Rand Paul, Ron Paul, and Paul Ryan, libertarianism about free will holds that humans have the ability "to have done otherwise"—as its slogan has it—i.e., specifically, to act in a way *contradicting* determinism. You can do otherwise than outside forces would have you do, or act contrary to internal desires, or even your long-standing character. Any influence

can in principle be resisted. Libertarianism says we are always capable of a spontaneous choice between alternative possibilities. Yesterday, you chose soup over salad. According to libertarianism, we could rewind the universe and let it play again, and this time, you could choose salad without *any change in any antecedent conditions*—for example, without even desiring salad. This is the sense one has of "could have done otherwise" when one occupies the ultimate perspective. Libertarianism *does not* mean the weaker proposition that *if I had wanted* to do otherwise, I could have done so. It's not that. It means I could choose to do something even if I didn't want to, i.e., without there being any particular cause of my choice. It means my decision has causal power (for example, to move my arm) but is itself *un*caused. It would have to be uncaused to allow for free will, because causation always comes in chains.

This "uncaused causer" of libertarian free will is called the *causa sui,* which Nietzsche, the most eminent critic of free-will credence, lambasted as "the best self-contradiction that has been conceived so far; it is a sort of logical rape and perversion." After 1900, the indeterminacy inherent to the newly discovered quantum realm gave libertarianism a boost, to be sure. There was now a permission structure to endorse the square circle of an uncaused causer. Yet for close to one hundred years, many philosophers have been going blue in the face pointing out that quantum, libertarian, or any other kind of indeterminacy is no help to friends of free will. If there is a spot of randomness between my intention to lift my arm and the event of my arm lifting, that doesn't make my movement free. It makes it random. I might intend to move my arm and kick out my leg if indeterminism were involved. Or, if I intend to move my arm and my arm moves, it could be thought of as only a fluke.

And yet the vast majority of Americans say they believe in free will. Most philosophers do too. According to a 2020 survey of more than seventeen hundred philosophy professors, only 11 percent denied free will in the way Sapolsky and Harris do. Still, Kelly was right that free-will denial is coming into a kind of vogue. The deniers have learned that doing so makes some people furious. Philosopher Galen Strawson received death threats because of his denial, according to a 2021 article in *The Guardian*. "It seems to provoke some people to an existential crisis," he said. In my conversations, many people bristle when I rehearse the arguments for denial. On a cross-country flight, a woman in my row—after asking me what I was writing about—stammered, "You . . . should not . . . think . . . that," her mouth quivering with indignation. Then she

proceeded to tell me with composed, clipped speech that "denying my free will means you are denying my godhood." Suffice to say, the rest of the flight was uncomfortable. I stood for much of it, stretching my legs by the bathrooms. While deplaning, I realized the book I'd been reading was missing from the seat pocket in front of me. The Vohs and Schooler psychology study borne out in the wild!

This emotional reactivity is also present in the academic debate. I've observed breaches of scholarly decorum—including accusations of willful blindness, ideological thinking, bad faith, and even moral turpitude. The philosopher Dan Dennett, who was the most eminent defender of free will before his unfortunate passing in the spring of 2024, said that denying free will causes "real harms"—harms for which deniers are morally responsible. Like Erasmus, Dennett has told deniers to "shut up." Although, unlike Erasmus, Dennett did it in a YouTube video.

Sam Harris and Dennett were friends—part of the four horsemen of New Atheism, along with Christopher Hitchens and Richard Dawkins. But after Harris published his short book denying free will, Dennett wrote a nasty review. For all the heated disagreement, Dennett and Harris agree about much, but then they disagree as to the significance of their agreement. Both say libertarianism is false—nobody has the ability to act contracausally—but Dennett thinks that "doesn't matter much," whereas Harris thinks if he can get the general public to admit that libertarianism is false, then "that's the whole shooting match."

"But wait. What *is* Dennett's position?" Kelly asked. "In another unlovely neologism from philosophers," I said, "he's what's called a compatibilist." Compatibilists assert that you can't solve free will by getting rid of the causal chain. That would be to embrace an unworkable, magical libertarianism. So compatibilists redefine free will. To be free, they say, an action needs to be not uncaused but *caused in a specific way*, by the right *kind* of causes—those meeting certain criteria. For example, University of North Carolina philosopher Susan Wolfe claims that a free action must "stem from values determined by reason." Dennett, for his part, says you are morally responsible (and free) if you act while capable of imagining the consequences of your actions, possessed of basic rationality, and not compelled by forces the direction of which you would disavow.

"But that means compatibilists maintain we are free *and* determined?" Kelly asked.

"Yes." In rebuttal, Harris says that compatibilist freedom is akin to a puppet loving where his strings deterministically take him.

"Who's right?" Kelly asked.

When new scientific consensus conflicts with folk understanding, sometimes folk beliefs change; sometimes they do not. The scientific difference between mass and weight has enjoyed—I'd wager—less than 100 percent adoption by the folk. When oxygen was discovered, however, *everyone* stopped referencing phlogiston. Will free will go the way of phlogiston? In *Blackwell's Companion to Free Will* (2023), Shaun Nichols, a philosopher at the University of Arizona, suggests that the concept of *free will* will change in one of two possible ways: in the way of the concept *witch* (eventual widespread acceptance that there are no such things as witches) or in the way of *whale* (eventual widespread acceptance, in spite of Herman Melville, that whales are not fish but mammals).

Dennett wants us to continue using the linguistic expression *free will*, but such that any libertarian part of it will disappear, and it'll have to start to mean something nonlibertarian via a connotational revision. Harris wants us to accept that *free will*, like *witch*, has no denotational referent in the real world. And the self-same responsibility-conferring mental states that Dennett wants to label "free will," Harris wants to use to label our concept of "voluntary willing."

In the first season of HBO's *Westworld*, Thandiwe Newton plays Maeve, an android brothel madam who doesn't know she's an android. "Everything you do, it's because the programmers . . . programmed you to do it," a technician named Felix (Leonard Nam) tells Maeve in a pivotal, revelatory scene. He shows her a futuristic iPad on which are displayed the words she is about to say. "I've run a brothel for years, and if there's one thing I know, it's when I am being fucked with," she says with unintended sphexishness, not yet entirely focused on the tablet. Finally comprehending what she's looking at, Maeve says, "That's not possible," a fraction of a second *after* we, the television audience, see the words *That's*, *not*, and *possible*, boxed, highlighted, and linked together—making a flowchart-like path, like stepping stones, among a cloud of alternative words she might have uttered. Confronted by accurate predictions of her behavior or speech, Maeve must conclude that her behavior is indeed caused by her programming. With this realization, she freezes and her eyes glaze over, evidently having crashed

like Windows 95. If you or I were confronted by accurate predictions of this kind, we might faint. It would seem to be compelling evidence that free will is an illusion.

The most interesting recent empirical work that threatens free-will belief—and in a way similar to *Westworld*—involves using brain scanners to literally "read minds" and accurately predict what a test subject is about to do. There is not yet a generally applicable, universal brain decoder, but the in-principle possibility is coming closer and closer to actuality. Neuroscientist John-Dylan Haynes and colleagues can predict which button a test subject in an fMRI brain-scanning machine will press—a button under the left index finger or one under the right—more than six seconds before the test subject consciously decides which button. These predictions—published in a study from 2008—were accurate 60 percent of the time. But a 2011 replication, using a technique other than fMRI, achieved 80 percent accuracy seven hundred milliseconds before the subject consciously decided.

In one experiment, participants are shown two numbers and over a short period may decide to either add or subtract them. Scientists can predict the participants' decision up to four seconds before the participants consciously decide which they're going to do. A researcher on Haynes's team can be locked away in a windowless room and still accurately predict which button the test subject will press, based only on the test subject's brain-scan data and artificial intelligence software. The upshot for free will is that your conscious decisions have their causes in nonconscious regions of your brain, which the scan can see even though you can't, since they are nonconscious. The fact of accurate prediction implies, more generally, that consciousness is fed its contents by nonconscious processes—nonconscious content bubbles up, on its own, to populate conscious experience.

"But if we feel free, is that not enough?" Kelly asked. I quoted Samuel Johnson, who said, "All theory is against freedom of the will; all experience for it." Sam Harris, I told Kelly, disagrees with the last part. Harris thinks experience is against free will too. In this opinion, he is almost entirely alone. Feeling like the author of your thoughts, Harris argues, is not a feeling we ever actually have. We *say* we have it, but if we paid attention to our experience with sufficient care, we'd see that thoughts, intentions, and desires simply arise without consciousness calling them up. Harris says you can come to this insight through mindfulness

meditation. Try it. Try meditating. You'll see what he's talking about. A guided meditation will ask you to focus on your breath and, when you catch your mind wandering, to refocus your mental attention back on the breath. It's not necessarily the goal to have a clear mind but rather to strengthen your ability to pay attention. In your early attempts, doing this is extremely difficult, and you immediately discover the wildness of your mind. Random thoughts pop up seemingly from nowhere; they take over your experience, leading you down tangents, pulling you away from paying attention to your breath. You'll be thinking about something else for a good long while before you realize that you are no longer focusing on your breath.

Harris says this is you identifying with the thought provided to your consciousness by nonconscious parts of your brain. You do not even notice that the thought came from nowhere, nor do you notice that you've identified with it automatically. You are *lost in thought*. If you are a novice meditator, and you're honest with yourself, the longest you can attend to your breath before your train of thought goes off in its own direction is probably two to three seconds. And the time spent *not* focusing on the breath without realizing it, i.e., the time spent identifying with your thoughts, can be much longer than two seconds before you realize it. For Harris, since being lost in thought is the same as identifying with your thoughts, enlightenment, or at least mindfulness, might be thought of as reflective distance from (nonidentification with) every thought, including seemingly willed intentions and decisions.

Harris's upshot for free will is exactly the same as that of Haynes's experiments: namely, that the contents of consciousness are provided by nonconscious processes in the brain. Feelings, ideas, desires, evaluations, decisions—they all come to consciousness, as it were, out of nowhere. Thoughts appear on their own, "unbidden," as Nietzsche put it. Thoughts come out of the darkness, entering the spotlighted stage of the mind's Cartesian theater on their own. As long as they do not conflict with our character or otherwise rock the boat, we immediately identify with them, and if asked, we will have, post hoc, filed them under "freely willed." Harris and Sapolsky both think (and Haynes's research seems to show) that when we make decisions, even momentous ones, the contents of consciousness are provided by nonconscious processes that we have no access to. Harris avows that he no longer feels like he could have done otherwise—a denial of libertarian experience, i.e., phenomenology.

To end, I return to my original question: what does denying your own free will *feel* like? Let's find out: there's no free will. There, I said it. And you read it, which means you thought it. Now try to believe it. The interesting thing is, even when assumed for the sake of argument, disbelief cannot be maintained for very long. One's tacit assumption of free will—one's tacit assumption of an event that Harris thinks we never actually experience—is nevertheless still so natural that disbelief evaporates quickly and usually unnoticed. Even when the denial of free will is based on the strongest evidence and most compelling arguments, it fades away (Harris's point notwithstanding). Indeed, without deliberately and effortfully keeping denial at the forefront of one's mind, one's mind reverts to its natural, implicit credence. For deniers, the falsity of free will is something they know but do not (always) believe.

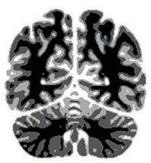
In those brief moments when we *do* actively deny our own free will, we get an intellectual dizziness; we start to freak out about the end of moral responsibility, and our actions start to seem not our own. But the discombobulation of the ultimate perspective is transient. Since it comes from thinking too much, I've taken to calling it "armchair vertigo." Once you get out of the armchair of philosophical reflection and engage again with life, it goes away. We get armchair vertigo only when, taking the ultimate perspective, we actively force ourselves to believe what we know (again, *pace* Harris). In day-to-day, normal life—that is, while not suffering from self-induced armchair vertigo—lost-in-thought deniers *know* they don't have free will, but they don't *believe* they don't.

Knowing without believing is, furthermore, the prescription offered by that other group of deniers who wish to suppress the knowledge that free will is false. They say, "Yeah, we know it's false, but it's better not to believe that." Maintaining that a nonveridical illusion is true amounts to not believing what you know. Of course, you cannot benightedly labor under an illusion that you know is an illusion. The self-same person, or a transparent, open society, cannot both know that X is an illusion and be taken in by the illusion. To not believe what you know requires rifts, internal divisions, impermeable walls between parts, self-deception, or forgetting.

We live in an age where much of what we know is not believed. Since the Scientific Revolution that started five hundred years ago, we have lived with this gap between the scientific image of the universe and the image tacitly animating how we experience and act in the world. Recall the last affecting sunset you saw. That fiery orb of nuclear fusion,

the sun, moving down through bands of luridly colored clouds, getting closer and closer to the horizon. It is easy to forget that the sun is not really moving in the way it seems to be. It's the Earth that is moving—rotating—so that, from our perspective situated on the sphere, the horizon eventually blocks our view of the sun. We are rotated away from the sight of it. We know this—if we take the time to think about it. Yet for our language to reflect our belief in it, we'd have to say, "What a beautiful Earth-turning tonight!"

What will be our fate once a critical mass of us finally knows the falsity of contracausal free will? Should we try harder to believe what we know? Or should we perhaps be grateful for temporary ignorance of the icy knowledge gleaned from the ultimate perspective? If you know something but do not believe it, you are in the position Oedipus was in immediately after he learned that he had killed his father and married his mother. For a split second, or maybe longer, he knew it, but it had not sunk in. He did not yet grok the new paradigm he now disastrously occupied. When his time of sleepwalking through life was at its end, when, finally, Oedipus truly believed what he knew, when he was pierced by the knowledge that he was a motherfucking patricide . . . he took a knife and pierced both his eyeballs.



David J. Frost

David J. Frost has published book reviews and essays in *SLAB literary magazine*, *Ruminate*, *The Smart Set*, *Philosophy Now*, and *As It Ought to Be* magazine. He wrote the first chapter in *Girls and Philosophy*, an anthology about Lena Dunham's HBO show *Girls*. He wrote—under a pen name—a self-help book that applies dual-process psychology to procrastination. He has been engaged as a speaker on productivity, the best science-backed strategies to fight procrastination, and how to set up one's writer's nook for a sustainable writing practice. With a BA in English literature from Columbia University, an MA in philosophy from University of California, Irvine, and a PhD in philosophy from the University of North Carolina, Chapel Hill, Frost teaches online for Alamance Community College in North Carolina and lives on the Oregon coast with his dogs Fritz and Lou Salomé, three thousand miles from Brooklyn. He writes about what we know but do not believe in his Substack newsletter, *Armchair Vertigo*. Follow him at davidjfrost.com.