

Review Essay: Sam Harris's *Free Will*

Eyal Mozes
Bethesda, MD

1. Introduction

In philosophical discussion of the relation of mind and body, the most heated debate throughout the history of philosophy, which continues today as strong as ever, has been that of free will versus determinism. Sam Harris's *Free Will*¹ is a recent, highly acclaimed defense of determinism.

The case for free will is clear; it is a self-evident, directly perceived fact. Every reader of this article can directly perceive that the amount of mental effort he spends on considering and trying to understand it—and then whether he agrees with me or not—is under his own control. The same is true every time any one of us is engaged in any thought process of any difficulty or makes a decision of any significance in his actions.

In contrast, in reading the writings of determinists, it is often unclear just what their case is for accepting determinism. It is common for determinists to tout determinism as scientifically proven, or declare that we must accept determinism in order to be scientific, without ever stating precisely what evidence or arguments they believe they have in support of determinism. When their arguments are identified and examined, they always turn out to be very weak.²

Sam Harris's book demonstrates the worst qualities of writings by determinists. He writes in a supercilious tone, full of pronouncements declaring the case for determinism to be conclusive and declaring free will to be an illusion; but he never gives a clear statement of what precisely he thinks the basis is for accepting determinism, requiring the reader to piece together statements from various parts of the book to figure out just what Harris's case is. And when his case is identified and examined, it turns out not to consist of any actual evidence, but only of the dogmatic acceptance of certain philosophical premises about causality.

¹ Sam Harris, *Free Will* (New York: Free Press, 2012).

² For a survey and examination of the various arguments that have been used to support determinism, see Eyal Mozes, "Is There a Rational Basis for Determinism?" available online at: <https://sites.google.com/site/eyalmozesonobjectivism/determinism>.

Harris advocates determinism as part of a more general model of the universe, which involves two basic principles:

- (1) The universe is built out of physical particles whose movements are determined by their previous movements and their physical impact on each other. (A principle commonly referred to as *mechanism*.)
- (2) Human beings are complex systems of these physical particles, and causal laws governing those particles completely determine the actions of the system. (A principle commonly referred to as *reductionism*.)

The model of the universe based on these two principles was originated by the Greek atomists, and is associated in modern times with the physicist Pierre-Simon Laplace. It is held by Harris, Daniel Dennett,³ and many other contemporary determinists.

What is Harris's case for accepting this model, and for accepting determinism? The book's loose structure, and the fact that Harris leaves his central line of argument to implication and never states it explicitly, make his case difficult to identify, but when we piece together Harris's various statements, we find that his central line of argument consists of accepting two unstated, unsupported assumptions about the nature of causality. His two dogmas of causality are that it requires mechanism and determinism and that it is a relation between events.

Harris also presents two lines of alleged observational evidence for determinism:

- (1) The claim that introspective experience, when seen with "serious self-scrutiny," demonstrates that we do not control our actions.
- (2) The claim that the Libet experiments, and similar subsequent experiments, demonstrate that we do not control our actions.

When we examine these lines of evidence, however, we find that in both cases Harris twists the evidence to fit his pre-conceived assumptions about causality. In both cases, when the evidence is considered without such assumptions, it provides no support at all for Harris's claims.

³ See Daniel Dennett, *Freedom Evolves* (New York: Viking, 2003). For a detailed analysis of Dennett's defense of determinism, see my review of *Freedom Evolves* in *Navigator* (December 2003), available online at: http://www.atlassociety.org/daniel_dennett_freedom_evolves.

2. Harris's Two Dogmas of Causality

a. Causality as requiring mechanism and determinism

Harris's most basic reason for accepting determinism and the mechanist/reductionist model, is his assumption that causality requires it. This is an assumption that runs throughout Harris's discussion. He presents no argument to support this assumption, never states it explicitly, and it is not clear whether he is fully aware of it. However, the entire book is written with this assumption accepted as self-evident and unquestioned.

In stating the possible positions regarding free will and determinism, Harris's summary of the libertarian position is: "human agency must magically rise above the plane of physical causation."⁴ Harris provides no citation to any libertarian writers who describe human agency as "magical"; I very much doubt he can find even one. Even if such writers could be found, they are rare exceptions. What Harris is describing is not the libertarian view, but rather his own view: that any violation of reductionism would have to be magical, that is, be a violation of causality. He takes this idea for granted so completely that he states it not as an argument against libertarianism, but as an alleged summary of it.

A consequence of the assumption that causality requires determinism is that the only alternative to determinism is randomness. In his introduction, Harris writes: "Free will . . . cannot be made conceptually coherent. Either our wills are determined by prior causes and we are not responsible for them, or they are the product of chance and we are not responsible for them."⁵ Later, in a chapter titled "Cause and Effect,"⁶ he discusses the possibility that free will could be based on quantum indeterminism in brain processes; having correctly dismissed that possibility, he then believes he has refuted any possible alternative to determinism. Harris takes it as a given that the view actually held by libertarians—that our actions are neither determined by prior causes nor the product of random events, but are under our own control—is "magical," violates causality, and therefore requires no discussion. When he says that free will "cannot be made conceptually coherent," what he actually means is that free will cannot be made consistent with his first dogma of causality.

⁴ Harris, *Free Will*, pp. 15-16. Note that Harris is here following the common usage of referring to the view that rejects determinism and affirms the existence of free will as *libertarianism*. I regard the use of this term, with the potential confusion with the unrelated political meaning of the same word, as unfortunate, but since I don't know a better alternative, I will follow the same usage.

⁵ *Ibid.*, p. 5.

⁶ *Ibid.*, pp. 27-30.

b. The event-event model of causality

Harris's second basic assumption about causality is that causality is a relation between events. He assumes the law of causality to mean that for every event there has to be some prior event which is its cause.

An alternative view of causality not discussed by Harris, dating back to Aristotle, is that causality is a relationship not between one event and another, but between an entity and its actions: the way an entity acts, including the way it reacts to the actions of other entities, is a function of its nature. While it is often convenient to refer to some action as the "cause" of a subsequent action, such usage is derivative; primarily, an action's cause is the nature of the acting entity. For example, the motions of atoms or ions are caused by their mass, electric charge, etc., which determine how the forces operating on them affect their movement. If the nature of these entities were different, they would act differently in response to the same external forces.

In the case of living things—for example, the contraction of a muscle, caused by the nature of the animal's muscular and nervous systems—the action's direction and energy come from sources internal to the acting entity. This special type of entity causation is referred to as *agent causation*.

Entity causation and agent causation are compatible with determinism in specific cases; there are many entities whose nature allows only one possible action in any given situation. Whether an entity's nature is deterministic or not is a question that has to be answered based on the evidence. For inanimate objects (above the level of subatomic particles) and for vegetative biological processes (i.e., all processes in bacteria and plants, and those processes in an animal's body that do not involve consciousness), many deterministic laws have been discovered and verified by the scientific method, precisely predicting the actions of all of these entities given the situation. These verified deterministic laws, and the observations supporting these laws, are the evidence justifying the conclusion that such entities behave deterministically.

However, unlike the event-event model of causality, the entity-action model does not *a priori* mandate determinism. It does not forbid the nature of an entity from including the ability to weigh alternative courses of action and deliberate about them, and consequently the capacity for genuine choice; such entities also act in accordance with causality, not in any way in contradiction to it. Once we get rid of the assumption of event-event causality, the question of whether human nature includes this capacity becomes a question that has to be answered based on the evidence, not on *a priori* requirements of causality.⁷

Similarly to his first dogma, Harris assumes the event-event model of causality without ever presenting any argument to support it, or even stating it

⁷ For more discussion of the event-event model of causality, and its role as an unstated, unsupported assumption at the base of defenses of determinism, see my "Is There a Rational Basis for Determinism?"

explicitly. As we will see below, he accepts it so completely that it becomes a filter through which all experience is forced so as to fit his pre-conceived theories.

3. Harris's Analysis of Introspective Evidence

Harris's first line of alleged observational evidence for determinism is his claim that it is supported by introspective experience. He writes:

Seeming acts of volition merely arise spontaneously . . . and cannot be traced to a point of origin in our conscious minds. A moment or two of serious self-scrutiny, and you might observe that you no more decide the next thought you think than the next thought I write.⁸

The examples Harris presents in support of this claim are of two types: examples of arbitrary or frivolous decisions, and one example of a serious, consequential decision.

a. Examples of arbitrary or frivolous decisions

The bulk of Harris's examples to support his claim about introspective evidence, are examples of arbitrary, inconsequential, or frivolous decisions. As his first example of a human decision, Harris writes: "I generally start each day with a cup of coffee or tea—sometimes two. This morning, it was coffee (two). Why not tea?"⁹ As his final example, he offers us: "In fact, I will now perform an experiment in free will for all to see: I will write anything I want for the rest of this book. . . . I can be ungrammatical if I pleased. And if I want to put a rabbit in this sentence, I am free to do so."¹⁰

It is common practice for determinists, when providing examples of human choices, to use these types of examples. The implication of using such examples is that if free will exists at all, it can only be applicable to choices that are arbitrary and frivolous or of no consequence.

To libertarians, in contrast, such examples seem of little relevance. The significance of free will is in our ability to deliberate on the reasons for and against a decision; examples in which no deliberation is possible, because there is nothing to deliberate on, are a distraction from the relevant issue. In these types of arbitrary decisions, we also don't have a clear introspective experience of self-control, and free will cannot be regarded as a self-evident, clearly perceived fact. Free will is self-evident when we make decisions based on deliberation.

⁸ Harris, *Free Will*, p. 6.

⁹ *Ibid.*, p. 7.

¹⁰ *Ibid.*, p. 65.

When we make arbitrary decisions, in situations in which there is no reason for one choice rather than another, is our action in fact fully determined by prior causes? This is a question that cannot be answered either way by simple everyday observations or by any philosophical arguments; it would have to be answered by scientific experiment. The answer is unknown at this time (as I discuss below, the Libet experiments don't provide an answer), but we can expect it to be known eventually. If the answer turns out to be that such decisions are entirely the product of prior causes, I doubt that any libertarians would find that to be disturbing news.

The irrelevance of such examples is made even clearer by Ayn Rand's crucial insight that the center of free will is in man's ability to direct his mental focus.¹¹ Man's basic choice is in focusing his mind: whether to focus it, to what level, and what to focus it on. Man is free to keep his mind in full focus, to drift automatically without focus, or actively to evade and refuse to think. Man is also free to focus on all relevant facts and considerations, and make a deliberate, conscious effort to think of additional factors and find anything that might have been missed so far; to limit his thinking to the factors that he notices easily and think no further; or actively to refuse to consider some of the facts. All other choices man makes are results of this basic choice. Man does not directly make a free choice on what ideas to accept; his freedom consists in controlling what facts and arguments his mind focuses on, and this selection of facts and arguments determines what ideas he then accepts as truth. Man does not directly make a free choice on what action to take; his freedom consists in controlling which considerations relevant to his decision his mind focuses on, and this selection of facts and considerations determines his actions.¹²

Rand's insight provides a clear criterion for what kind of human choices are relevant to the issue. If the center of free will is in directing your mental focus, then it is relevant when your choice is based on some reason that you had to think about. Mental focus requires something to focus on. The choice to focus is not itself the result of deliberation; it is one of the pre-conditions for deliberation about one's more derivative choices, specifically about one's choice of action; another pre-condition for deliberation is the existence of considerations to deliberate about. In Harris's examples of decisions without reason, this second pre-condition is not met; whatever choice you make in focusing your thinking, in Harris's examples you cannot choose to focus it on facts relevant to the decision, since there are no such

¹¹ See Leonard Peikoff, *Objectivism: The Philosophy of Ayn Rand* (New York: Dutton, 1991), pp. 55-69.

¹² For a more detailed discussion of Rand's identification of mental focus as the center of free will, and of how this insight answers the remaining objections to free will that were not adequately answered before, see my "Is There a Rational Basis for Determinism?"

facts. This makes such decisions, and introspective experience that accompanies them, irrelevant to the question of free will.

b. A more serious example

Harris makes one more serious attempt to support his claims about introspective experience. He tells a hypothetical story of an overweight person who, after several half-hearted and failed attempts to go on a diet, finally commits himself seriously, successfully loses weight, and improves his life in various other ways. Harris then considers what would be the person's introspective view of the causes of this change, if he engaged in serious introspection:

If you pay attention to your inner life, you will see that the emergence of choices, efforts, and intentions is a fundamentally mysterious process. Yes, you can decide to go on a diet—and we know a lot about the variables that will enable you to stick to it—but you cannot know why you were finally able to adhere to this discipline when all your previous attempts failed. . . . Yes, you can do what you want—but you cannot account for the fact that your wants are effective in one case and not in another. . . . You wanted to lose weight for years. Then you *really* wanted to. What's the difference? Whatever it is, it's not a difference that *you* brought into being.¹³

Harris presents this as an account of introspective experience; but the experience he describes has no similarity at all to the actual experience of people who have made life-changing decisions of this kind. As it happens, I can speak to this issue from direct personal experience; I was overweight for many years as a child and a teenager, and then successfully lost weight. I can clearly remember the difference between talking about wanting to lose weight without any genuine commitment to doing something about it, and later focusing on the importance of losing weight and committing myself to the effort. There is nothing fundamentally mysterious, or mysterious at all, about the difference. Anyone who has ever made any important decision involving a serious effort can see this by examining his own experience.

Why does Harris make a claim about introspective experience that is so obviously contrary to fact? The reason becomes clear when we look at his statement two pages later:

Choices, efforts, intentions, and reasoning influence our behavior—but they are themselves part of a chain of causes that precede conscious awareness and over which we exert no ultimate control. My choices matter—and there are paths toward making wiser ones—

¹³ Harris, *Free Will*, p. 37.

but I cannot choose what I choose. And if it ever appears that I do—for instance, after going back and forth between two options—I do not *choose* to choose what I choose. There is a regress here that always ends in darkness. I must take a first step, or a last one, for reasons that are bound to remain inscrutable.¹⁴

This statement makes clear that Harris's claims are based not on introspective experience, but on what his view of causality requires. Harris assumes the event-event model of causality, requiring that any choice he makes must be caused by some prior event. He assumes *a priori* that agent causation cannot exist; therefore, if he experiences himself as the cause of his choice, without any event that acts as the cause, this experience cannot be accepted. His choice must have been caused by some prior event, and so the only experience he is willing to accept as an explanation is the experience of some prior event. If he can find a prior event in his experience, such as some prior choice, providing an explanation, then this event itself requires an explanation through some third, even earlier event. It is logically inevitable that this regress will end at some point, with some choice or thought or intention that he cannot explain by any prior event in his experience, making it by his assumptions "fundamentally mysterious" and "inscrutable."

Harris claims to have introspective evidence against free will. When his evidence is examined, however, we find that his "serious self-scrutiny" is merely the filtering of introspective experience so as to fit his philosophical dogmas.

4. The Libet Experiments

Benjamin Libet performed a series of experiments in the early 1980s, in which subjects were asked occasionally to move their hands at arbitrary intervals, and to note and report the time at which they made the decision to move their hand. An EEG measurement, taken during the experiment, showed that the brain waves preceding the hand movement started some fraction of a second before the time the subjects reported as the time they made the decision.¹⁵ Several other researchers have since conducted similar experiments, all following the same pattern: the subject is asked to make some arbitrary decision, noting and reporting the time at which he made the decision, and EEG or fMRI measurements detect the brain activity containing information about the decision some time before the subject reported making

¹⁴ Ibid., p. 39; internal footnote omitted.

¹⁵ Benjamin Libet, Curtis A. Gleason, Elwood W. Wright, and Dennis K. Pearl, "Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential): The Unconscious Initiation of a Freely Voluntary Act," *Brain* 106 (1983), pp. 623-42.

it.¹⁶ Harris, following several other determinists of the past few decades, trumpets these results as proof of determinism, allegedly demonstrating that decisions we apparently make by our free will are the result of neural processes that happen before we become conscious of the decision.

In fairness to Libet, it should be noted that he is not a determinist, and this is not his own interpretation of his results. Libet's own interpretation, rather, is that his results demonstrate that free will is purely negative. Libet's theory is that our actions are the result of urges created by neural processes that are outside our control, but that we are able consciously to override these urges and decline to act on them; this conscious "veto power" is Libet's view of free will. As I discuss below, Libet's interpretation does not follow from his experiments either, but it is not nearly as blatant a *non sequitur* as taking the results to be evidence of determinism.

The Libet experiments in fact do not have any interesting implications regarding free will, for two basic reasons. First, the experiments created a situation in which the subject's decision (at what time to move his hand) is necessarily arbitrary; there is no possible reason for the subject to move his hand at one time rather than another. The same is true for all of the later experiments that found similar results; all involve asking the subject to make a decision without any reason to regard one alternative as better than the others. As I discuss above, arbitrary decisions concerning options such as "Tea or coffee?" or "Should I put a rabbit in the sentence?" are irrelevant to understanding free will; the decisions studied in the Libet experiments, and in subsequent experiments, are an even more extreme case. Such situations are fundamentally different from real-life situations in which people make decisions, and it seems likely that that difference would completely change how free will operates. Even if these experiments proved anything about the decision-making process in the laboratory situations they created—situations in which a decision is completely arbitrary, without reasons—it would be impossible to draw any conclusions from that about the decision-making process in real-life situations, in which a person makes a decision by considering reasons for and against a course of action.¹⁷

Second, the Libet experiments don't prove anything even about the decision-making process in the laboratory situation they created; both the determinist interpretation of the experiments and Libet's own interpretation ignore the fact that perception takes time. Libet asked subjects to report the time at which they decided to move their hands, with accuracy down to a fraction of a second, by watching a clock-face with a fast-moving dot, and noting the position of the dot at the moment they made the decision. But the

¹⁶ Harris, *Free Will*, p. 73, nn. 3 and 4, provides references to several such experiments.

¹⁷ David Kelley makes this point in the Q&A period of his lecture series *The Nature of Free Will*, presented at the Portland Institute, 1986.

perception of the clock-face—like all visual perception—is a process that takes time; it is therefore likely that the dot position reported by the subjects was the position not at the moment they became conscious of the decision, but some fraction of a second later. Furthermore, generally when we intentionally perform a movement, we monitor that movement with our vision, and so our visual processes when intentionally performing a movement (such as the hand-movement performed in the experiment) will naturally be alert to the moment at which the movement occurs, which will be some fraction of a second after the movement is consciously initiated. This again makes it likely that the dot's position on the clock-face, perceived by the subject as being at the same time he became conscious of the decision, was in fact at a time some fraction of a second later. Daniel Dennett makes this point in his *Freedom Evolves*.¹⁸ Given the fact that Harris is clearly familiar with Dennett's book, citing it several times, and given the central importance Harris claims to attach to the Libet experiments, it would be natural to expect him to address Dennett's analysis of these experiments. However, Harris never addresses Dennett's arguments on this—indeed, never even acknowledges them.

The same problem applies to all but one of the later experiments: all of these experiments relied on visual cues to help the subjects note when they made their conscious decision, and it is thus very likely that the time reported by the subjects was some fraction of a second later than the time they actually became conscious of their decision. (The only exception is one experiment in which the delay measured was several seconds, which cannot be explained by delay in perception.¹⁹ This experiment was similar to all of the other ones in that the subjects were asked to make an arbitrary decision with no reasons, and so the previous point still fully applies.)

It is very doubtful whether anyone has ever been convinced of the truth of determinism by Libet's results; the fallacies are too obvious. Those who cite the Libet experiments or later similar experiments as support for determinism, are taking them as confirmation for a view they have accepted *a priori*. Harris's discussion of the experiments makes clear that that is precisely what he is doing. After describing the experiments, Harris writes:

These findings are difficult to reconcile with the sense that we are the conscious authors of our actions. One fact now seems indisputable: Some moments before you are aware of what you will do next—a time in which you subjectively appear to have complete freedom to

¹⁸ Daniel Dennett, *Freedom Evolves*, pp. 227-42. While in general I have a very low opinion of *Freedom Evolves*, the section on the Libet experiments is the one section of the book that contains useful and interesting information. The fact that Dennett is himself an avid advocate of determinism, but took the time to discuss and expose the problems with bad arguments presented in support of his own position, is the one aspect of his book that deserves respect.

¹⁹ Cited by Harris, *Free Will*, p, 73, n. 3.

behave however you please—your brain has already determined what you will do. . . . There will always be some delay between the first neurophysiological events that kindle my next conscious thought and the thought itself.

However, he then adds:

And even if there weren't—even if all mental states were truly coincident with their underlying brain states—I cannot decide what I will next think or intend until a thought or intention arises.²⁰

Having cited Libet's results as evidence for determinism, Harris then admits that these results are not actually relevant to his advocacy of determinism, and that he would have stuck to the same view had Libet's results been the opposite.

Describing Libet's interpretation of his own results, as consistent with a form of free will which is purely negative, Harris writes: "This suggestion has always seemed absurd on its face—for surely the neural events that inhibit a planned action arise unconsciously as well."²¹ In rejecting Libet's alternative interpretation, which is equally consistent with the experimental results, Harris admits that his claim that all neural processes underlying human choices are controlled by unconscious causes, is not based on the experimental results; it is based on his *a priori* conviction that this is "surely" the case and that any suggestion it might not be is "absurd on its face."

Harris's alleged experimental evidence for determinism thus turns out to be the same as his introspective evidence; it does not actually provide any support for determinism, except to interpreters who have already accepted Harris's pre-conceived assumptions.

5. Conclusion

Harris repeatedly heaps scorn on anyone who would disagree with determinism, and blatantly uses arguments from intimidation. In a chapter on the political implications of the debate, Harris states that the defense of free will is motivated by the "religious fetish of individualism," which leads political conservatives to want to give people credit for their achievements, and writes:

[O]ne gets the distinct sense that if certain conservatives were asked why they weren't born with club feet or orphaned before the age of

²⁰ Ibid., p. 9.

²¹ Ibid., p. 73, n. 2.

five, they would not hesitate to take credit for these accomplishments.²²

In his introduction, Harris describes a gruesome murder, claims that the murderers had no freedom to choose not to commit this murder, and then writes: “There is simply no intellectually respectable position from which to deny this.”²³

After describing the idea of compatibilism,²⁴ Harris writes:

Today, the only philosophically respectable way to endorse free will is to be a compatibilist—because we know that determinism, in every sense relevant to human behavior, is true.²⁵

Harris is evidently hoping that the fear of being labeled “philosophically [un]respectable” would dissuade readers from too closely examining just what case he has presented for his claim that “we know” determinism to be true.

When we do examine Harris’s case, as remarked above, it does not consist of any scientific evidence or logical arguments, but only of the dogmatic acceptance of certain philosophical premises about the nature of causality. Both Harris’s alleged introspective evidence and his alleged experimental evidence turn out to be merely the filtering of observation in order to fit his *a priori* beliefs. Harris’s defense of determinism is an emperor who turns out not to be wearing any clothes.

²² Ibid., pp. 61 and 62.

²³ Ibid., p. 4.

²⁴ Compatibilism is the idea, advocated by Dennett and rejected by Harris, that people should continue to use the concept of free will, but redefine it so that it no longer involves choice among several possible alternatives, and can thus be made compatible with determinism.

²⁵ Harris, *Free Will*, p. 16.

