THE SCIENTIFIC PLAUSIBILITY OF LIBERTARIANISM

Christopher Evan Franklin

Free Will and Moral Responsibility, (eds.), Ishtiyaque Haji and Justin Caouette, pp. 123-141. (Newcastle upon Tyne: Cambridge Scholars Publishing, 2013).

ABSTRACT: Many have worried that libertarianism is scientifically implausible. Libertarians, given plausible assumptions about the mind and brain, are committed to neuronal indeterminism. Call this 'the libertarian hypothesis.' Recently, Manuel Vargas has argued that this hypothesis renders libertarianism naturalistically implausible. He contends that we ought to prefer compatibilism to libertarianism since compatibilism has fewer empirical commitments *and* libertarianism's commitment to neuronal indeterminism is dubious—because it both lacks any empirical evidence and in fact has empirical evidence against it. I show that Vargas's assessment of libertarianism is too sanguine. First, I argue that weighing a theory's empirical commitments is not a mere matter of weighing the number of its commitments, but also the quality of such commitments: in particular how demanding they are. I go on to argue that the libertarian hypothesis is relatively undemanding. I defend this claim by arguing, contra Vargas, that we lack any good evidence against the libertarian hypothesis, that many of our observations concerning the workings of the brain fit with the libertarian hypothesis, and that the main changes that libertarianism requires concerns not what we observe, but how we assess or interpret our observations.

Free will—moral responsibility—libertarianism—indeterminism

1 Introduction

Libertarians claim that free will and moral responsibility exist, and that free will and moral responsibility are incompatible with determinism.[1] The conjunction of these theses entails that indeterminism obtains—that at least some events are undetermined. However, libertarians are clear that not any location of indeterminism will help. Agents in worlds where all the events surrounding the moment of action are deterministically caused by prior events precludes the possibility of freedom and responsibility (cf. van Inwagen 1983; Kane 1996). Thus, libertarians are committed to indeterminism being relevant to the genesis of action: some free *actions* must be undetermined. It is

[1] Beyond this libertarians disagree about what else is required for free will and moral responsibility. In general, libertarians can be divided into three camps, all of which admit of further subdivisions. There are non-causal libertarians (Ginet 1990; McCann 1998), event-causal libertarians (Kane 1996), and agent-causal libertarians (O'Connor 2000). These accounts differ concerning what causal requirements they place on exercises of free will. Non-causal libertarians do not require (and in fact may require the absence of) causation in order for an action to be free. Event-causal libertarians believe that free action requires causation, but that the causal relata involved need only be events causing other events. Agent-causal libertarians also require causation for free action, but maintain that the cause of such actions cannot be solely events (and maybe cannot be events at all), but also (or in place of the events) must be a substance that is identical to the agent. My aim in this paper is only to assess libertarianism as a generic thesis, and not any of these specific varieties of libertarianism.

also widely accepted that the mental supervenes[2] on the physical, and on the brain in particular. Assuming this is true, mental actions, such as choice and attention, must supervene on brain events. Therefore, if choice is undermined, so also must the brain event on which it supervenes. Libertarians, if this argument is sound, are committed to neuronal indeterminism.[3] Let us refer to this commitment, that the brain is indeterministic, as 'the libertarian hypothesis.' Many have worried that the libertarian hypothesis renders libertarianism, in one way or another, scientifically implausible. According to what I will call the 'scientific plausibility objection,' while libertarianism may at first glance appear to be a compelling philosophical analysis of freedom and responsibility, its empirical commitments render it untenable (cf. Honderich 1988; Churchland 2002). This worry often remains inchoate, rarely taking the form of a well-developed argument. Manual Vargas, however, has recently sought to take some steps toward filling in this argument by arguing that libertarianism violates his "standard of naturalistic plausibility" (Vargas 2004). Vargas argues both that libertarianism is comparatively less plausible than compatibilism and that the libertarian hypothesis is dubious. In sections 2 and 3, I consider and respond to Vargas's particular version of the scientific plausibility objection against libertarianism. In section 4 I broaden my line of defense to survey the neuroscientific evidence (or lack thereof) for the libertarian hypothesis. I respond to Vargas's contention that our best models of the brain tend to be deterministic, showing this to be false. In section 5 I turn to the issue of how demanding the libertarian hypothesis is, and argue that

[2] The supervenience relation need only be nomological supervenience in order to generate the commitment to neuronal indeterminism, and thus is a relatively weak commitment. Even many dualists accept nomological supervenience of the mental on the physical.

[3] I am moving very quickly here. First, notice that libertarianism itself has *no* empirical commitments. It is committed to free action being undetermined, but it has no commitment about whether free actions are part of or dependent on the physical world. Libertarian's commitment to neuronal indeterminism only arises when it is wedded to theses about the mind and brain, theses that are not themselves commitments of libertarianism. This is important to emphasize. Nevertheless, upon widely held assumptions about the nature of the mind and brain, libertarians are committed to neuronal indeterminism, and thus it becomes important to assess the impact of this commitment for the overall plausibility of libertarianism. I will often write of "libertarians" having this or that empirical commitment, when, strictly speaking, it is libertarianism in conjunction with independent philosophical and scientific theses that generate these commitments.

it is relatively undemanding, requiring little to no changes in what observe, but rather changes concerning how we interpret these observations.

2 Vargas's Challenge

Vargas (2004) raises an important and distinctive challenge to libertarianism. It is often remarked that there is something unscientific about libertarianism; that libertarianism, despite all its attempts to domesticate itself, remains indelibly at odds with a scientific picture of the world. Unfortunately this is usually the extent of such arguments. Vargas, however, intends to "take some steps toward filling in this argument" (2004, p. 403). He is only taking some steps toward filling it in because he is attempting to show that "there is little reason to believe that extant libertarian theories satisfy a standard of naturalistic *plausibility*…." (2004, p. 403), and apparently there are questions concerning the relationship between naturalistic plausibility and *scientific* plausibility. Below I will argue that Vargas's standard of naturalistic plausibility is not an apt standard for assessing a theory's overall plausibility: a theory might be overall plausible and yet not satisfy the standard of naturalistic plausibility. I will then draw out some of the insights in Vargas's discussion and offer a more promising line along which to assess the scientific plausibility of libertarianism.

According to Vargas, philosophy "is something like the search for probable truths in domains in which we lack any reliable methodology for determining what the truth is" (2004, p. 404). Initially, a proponent of a theory might be satisfied with showing that some theory, which up until now was thought impossible, is possible. Showing how such a theory could be true is a genuine and helpful advance. However, as advancements in the theory develop we will likely raise the bar of assessment. If a theory is shown to be possible, then we will want to know if it is

plausible; if plausible, then whether probable; and so on. Vargas concedes that libertarians have not only shown that libertarianism is possible, but also how it might be true given our conception of the world, including ourselves, derived from science. Libertarianism satisfies the standard of naturalistic *compatibility*: "compatibility with an independently acquired, broadly scientific conception of the universe, especially the parts we might inhabit" (2004, p. 405).[4] The reason the standard of naturalistic compatibility is important, Vargas argues, is because a theory that is in tension with our best science is a theory that is in tension with our "best pieces of knowledge" (2004, p. 405).

The libertarian's achievement, although indeed an advance over earlier libertarian accounts, is short lived: for we will now want to know whether libertarianism satisfies more demanding standards of assessment. Vargas cautions that we should not raise the bar of assessment too high too quickly, but rather should prefer an incremental approach. Rather than moving in one stride from possibility to probability, Vargas suggests that we slow down and consider naturalistic *plausibility*. A theory satisfies this standard just in case it is "both compatible with a scientific worldview *and* that the balance of known truth-relevant considerations would be sufficient to lead a group of informed, well-reasoning, and disinterested persons to think the theory is plausible" (2004, p. 406). The sense of plausibility that Vargas is interested in "is one where the balance of known truth-relevant considerations could, via an appropriate deliberative path, lead a group of informed, well-reasoning, and disinterested persons to accept the theory" (2004, p. 406). Plausibility does not imply that it is likely to be true or even more likely to be true than competing theories, but concerns the likelihood of acceptance among informed, well-reasoning, and disinterested persons.

The standard of naturalistic plausibility raises "serious difficulty" for libertarianism since "we are being asked to take on commitments for which we lack independent evidential support" (2004,

[4] Vargas's use of 'naturalism' and its cognates is thin: it is only meant to indicate a generally pro-scientific attitude (Vargas 2004, p. 405). It seems to me that it would be best to replace his uses of 'naturalistic' with 'scientific' in order to avoid confusion. But then again, Vargas's standard seems intended to capture something different from scientific plausibility and thus necessitates a different title.

p. 409). And any theory that has unsupported theoretical commitments, commitments for which we lack independent support, will be rightly considered implausible. Vargas's claim is a comparative one: we ought to prefer theories that give up some feature of common sense while preserving plausibility over theories that preserve some feature of common sense at the expense of implausibility. So although both compatibilism and libertarianism will to some degree outstrip our evidential basis, libertarianism is less plausible than compatibilism insofar as it will always have an additional unsupported theoretical commitment: namely that indeterminism occurs somewhere in the genesis of action.[5]

But the problems for libertarians get worse. As we saw above, given widely held assumptions about the mind and brain, libertarians must locate indeterminism in the brain.[6] This, according to Vargas, generates the comparative point: as libertarianism has this additional, unsupported theoretical commitment, it is comparatively less plausible than compatibilism. But there is also evidence that this theoretical commitment is false. Vargas argues that the libertarian hypothesis suffers both from the fact that there is no evidence that the brain is indeterministic and that there is evidence for thinking that the brain is deterministic. In defense of these claims, Vargas cites the widespread rejection of S.R. Hameroff and R. Penrose's (1996) hypothesis about how quantum indeterminacies magnify to the neuronal level. Moreover, he cites neuroscientist and philosopher Henrik Walter (2001), who claims that there is indeed good reason to think that the brain is not indeterministic. Vargas concludes, "Given that there are no credible scientific models of indeterminacies in the brain, and given that there are reasons internal to neuroscience for thinking

[5] In defending this charge of implausibility, Vargas considers that libertarians might appeal to introspection, maintaining that the evidence from introspection supports libertarianism over compatibilism. He rightly rejects this response, arguing that libertarians owe us a reason to trust introspection as a reliable guide to the metaphysics of agency.
[6] Again, libertarianism itself does not generate this commitment, although I think any plausible version of libertarianism is committed to neuronal indeterminism.

that the brain is not so organized, libertarianism…will fail to meet a standard of naturalistic plausibility" (2004, p. 413).

3 Scientific and Naturalistic Plausibility

Vargas's challenge raises deep methodological and epistemological questions. First, it raises questions concerning the relationship between scientific plausibility, naturalistic plausibility, and (simply) overall plausibility. How does the standard of naturalistic plausibility relate to the standard of scientific plausibility? Can a theory be scientifically plausible and yet fail to satisfy the standard of naturalistic plausibility? Moreover, how do the notions of scientific and naturalistic plausibility relate to overall plausibility? Does conceptual space allow for the possibility that a theory is overall plausible and yet neither scientifically nor naturalistically plausible? These are important questions to consider when assessing the import of Vargas's argument. I doubt that the relationship between, on the one hand, naturalistic plausibility and, on the other hand, scientific plausibility and overall plausibility is a tight one. But in the present section I will restrict my attention to raising some general worries about the suitability and relevance of Vargas's standard of naturalistic plausibility for assessing either the scientific or overall plausibility of libertarianism—arguing that although his standard is a helpful start for addressing how to assess the status of a philosophical theory in light of our scientific worldview, sufficient problems remain to undercut its relevance and suitability for such a task. I will not offer an alternative standard in its place, but will instead suggest some considerations that bear on assessing the overall plausibility of libertarianism in light of its commitment to neuronal indeterminism. A key insight of Vargas's standard is its directing us to the importance of assessing the relative *demandingness* of a theory's empirical commitments in order to

assess the theory's scientific and overall plausibility. In sections 4 and 5 we will turn to the scientific data itself in order to assess exactly how demanding the libertarian hypothesis is. There I will argue that the demandingness of the hypothesis is negligible.

Returning to the standard of naturalistic plausibility, Vargas sometimes claims that he is arguing that libertarianism is implausible and at other times that he is making the more qualified point that libertarianism is comparatively less plausible than compatibilism. Vargas retreats to this qualified conclusion since he realizes that every philosophical theory may well violate the standard of naturalistic plausibility, as (almost?) every philosophical theory outstrips our evidential base. Consider for example standard compatibilist theories that make rational reflection and guidance of behavior central to freedom and responsibility. Often these theories are committed to reasons playing a central causal role in the etiology of free and responsible action. But as we lack independent support for this commitment, it is unclear that a group of informed, well-reasoning, and disinterested persons could be led to accept this theory—especially in light of the countervailing empirical evidence that calls into question the relevance of reason for our actions (cf. Doris 2002; Wegner 2002). But Vargas rightly notes that libertarians will always have one more empirical commitment than compatibilists (who neither require the truth nor falsity of determinism),[7] and to the degree that a theory has unsupported commitments theorists will, for that reason, find the theory implausible. Consequently, theorists will conclude that libertarianism fails to satisfy the standard of naturalistic plausibility to a greater degree than compatibilism, and in this way that compatibilism is more plausible than libertarianism.

[7] This claim actually needs qualification: it is certainly conceivable that a compatibilist theory has more empirical commitments than a libertarian theory. It is possible for a compatibilists to construct a theory that makes very precise predictions about the neural mechanisms underlying free action and in this way have more empirical commitments than standard libertarian theories that refrain from making such precise predictions. Presumably what Vargas had in mind was that libertarianism qua libertarianism will always have one more commitment than compatibilism qua compatibilism.

In response, I contend that things are much more complicated than Vargas's discussion suggests. When comparing libertarianism and compatibilism, one must consider all their strengths and weaknesses, not just the number of empirical commitments. Compatibilism may have fewer empirical commitments than libertarianism, but we must also compare these theories' philosophical merits.[8] If the degree to which a theory satisfies the standard of naturalistic plausibility is solely a function of the quantity of empirical commitments, then we have reason to think that the range of theories that are overall plausible is not co-extensive with the range of theories that are naturalistically plausible. For example, if we have compelling arguments for libertarianism, then this fact would outweigh the relevance of empirical commitments that lack independent support: better to have a theory with unsupported empirical commitments than one that is false. Now, of course, compatibilists will hardly be moved by this, but they need not be. Putting aside which theory has the strongest philosophical arguments in its favor, if Vargas's standard of naturalistic plausibility is to have a hope of tracking overall plausibility, then more must be relevant to assessing a theory with respect to this standard than the number of empirical commitments. We have two comparative points: we must compare both the number of empirical commitments of each theory and also each theory's philosophical virtues. In this way, we cannot immediately infer that libertarianism is comparatively less plausible than compatibilism because it has more empirical commitments than compatibilism. Before any such conclusion can be reached we must also assess each theory's philosophical merits. These comparative points are not altogether independent. For example, one might take the latter comparative point to militate against the former: the fact that compatibilism has fewer empirical commitments shows that it is a better account of freedom and responsibility. This is of course too strong; a more plausible claim would be that the fact that compatibilism has fewer empirical commitments is a point in is favor—albeit a defeasible one. But we must weigh each

[8] Although Vargas is aware of this point (2004, p. 409-410), it plays no role whatsoever in his assessment of the comparative plausibility of libertarianism.

feature in order to determine which theory is overall more plausible. The structure, then, of assessing comparative plausibility is more complicated than Vargas lets on. Merely having less empirical commitments than a rival theory cannot establish overall comparative plausibility. Rather this feature must be weighed in light of the theory's other virtues: for example, the theory's ability to properly conceptualize the kinds of powers necessary for freedom and responsibility.

In addition to weighing a theory's virtues and vices and the number of empirical commitments, we must take into account the relative plausibility of the empirical commitments that lack independent support. If a theory has a commitment that requires a radical departure from current and widely accepted scientific theories, then this will count more heavily against the theory's overall plausibility than if the commitment is one that is consonant with, even though not demonstrated by, what we currently take science to show. That is, in addition to considering the *quantity* of empirical commitments that a theory has, we must also consider the *quality* of such commitments—specifically how demanding they are. The demandingness of a commitment is partly a function of how radically things must change from what we currently take ourselves to know in order for the commitment to be satisfied. If a commitment requires a major change in central and well-confirmed aspects of our scientific understanding, then such a commitment will be very demanding. Alternatively, a commitment will be less demanding that contravenes nothing we take ourselves to know, but merely requires that some things turn out to be the case for which we currently lack evidence. Not all unconfirmed empirical hypotheses are created equal: *ceretis paribus*, the more demanding the hypothesis the less plausible the theory. It is possible for a theory to have more empirical commitments than another theory and yet be overall more plausible, since each of its commitments is undemanding while the other theory's single empirical commitment is extremely demanding. Therefore, when assessing a theory's overall plausibility, in addition to weighing its

philosophical merits and number of empirical commitments, we must also weigh the demandingness of these commitments.

The scientific plausibility objection that Vargas and others have raised against libertarianism is intended as an *independent* argument. That is, this argument is supposed to undermine the plausibility of libertarianism independent of an assessment of the putative success of standard philosophical arguments for libertarianism. The import of the argument is that while libertarianism may well seem plausible when we limit our inquiry to philosophical analysis, once we broaden our perspective to include scientific data, libertarianism founders. Consequently, I will assume that there are strong philosophical reasons for preferring libertarianism to compatibilism.[9] In this respect, theorists would rank libertarianism as better satisfying the standard of naturalistic plausibility. Crucial, then, for determining which theory is overall comparatively more plausible, is to assess the demandingness of the libertarian hypothesis: is the cost high enough to outweigh libertarianism's philosophical merits. In the following sections I will argue that the cost is not high. This argument will not only serve as a response to Vargas's contention that compatibilism is comparatively more plausible than libertarianism, but should also serve to silence more general worries about the libertarian hypothesis.

Ever after we consider the complicated nature of comparing the plausibility of theories, many questions remain about Vargas's standard. First, Vargas does not always state the standard in the same terms. Sometimes he words the standard so that the group of theorists must be led to *accept* the theory (p. 406) in order for it to satisfy the standard, and in other places he claims that the

[9] One might object: why would any non-libertarian accept this assumption? First, because of the structure of the objection I am responding to: again it is often admitted that libertarianism, when restricted to purely philosophical concerns, such as its analysis of human powers, is more plausible than compatibilism. But, the objection goes on, the empirical commitments engendered by such a theory force us to revise our original analysis of power. (Vargas (2004) uses this argument to argue for a kind of compatibilist-revisionism.) Thus, it is not objectionable for me to assume a claim that the objection I am responding to assumes. Second, given that we are involved in a weighing task—weighing libertarianism's philosophical virtues against its empirical commitments—it is helpful to assign its philosophical virtues a fixed weight. My arguments below are concerned to show that libertarian's empirical commitments are not weighty. One can accept this conclusion regardless of one's assessment of libertarianism's philosophical virtues.

theorists only need to think the theory is *plausible* (p. 406). But accepting a theory and judging it to be plausible are different cognitive states: I have often rejected theories that I think are plausible. Which then does Vargas have in mind? I will assume that he has the second reading in mind since this, after all, is the standard of naturalistic plausibility.

Epistemological questions loom large once we begin to consider how we can know whether a theory satisfies the standard. We cannot answer these questions until we hear more about who counts as informed, disinterested, and well-reasoning. Must such a person know all the relevant facts? Merely all the relevant facts that *we* know? Or all the relevant facts we *think* we know? Even when these points are clarified, questions remain since what all the relevant facts that we know are—let alone what all the facts are—is a substantive philosophical and scientific issue. Lastly, how many members of the group must be led to judge the theory plausible in order for it to satisfy the standard: All? Some? A majority? In light of these unanswered questions and the complicated nature of comparing the respective aspects of libertarianism and compatibilism, I find Vargas's conclusion that libertarianism violates this standard to a greater degree than compatibilism too sanguine.

All this suggests to me that his standard is too ill-defined to do the required work. I believe that the spirit of Vargas's standard is correct and does contain important insights. Libertarians should not rest content with naturalistic compatibility, but should strive to meet more demanding standards, and if libertarians are committed to the brain operating in a specific way, then they ought to take seriously the judgments concerning the plausibility of this commitment by leading neuroscientists. They ought to take these judgments seriously, but they are not the last word. What really should concern them is not what neuroscientists believe about the brain, but the reasons neuroscientists offer in support of these beliefs about the brain. This is another respect in which Vargas's standard is defective: it focuses too much on what experts would say and too little on why

they would say it. The question before us should concern less what some hypothetical or real theorists would think—even informed, disinterested, and well-reasoning people can endorse poor arguments—and should be concerned more with the data and its significance.

Worries about the standard of naturalistic plausibility notwithstanding, Vargas's assessment of the neuroscientific literature ought to trouble libertarians. He claims "there are no credible scientific models of indeterminacies in the brain, and…there are reasons internal to neuroscience for thinking that the brain is not so organized" (2004, p. 413). These contentions return us to our question above when comparing the plausibility of libertarianism and compatibilism: how demanding is the libertarian hypothesis? In determining the demandingness of this hypothesis we need clarity concerning the current state of neuroscientific knowledge. In the next section I will offer a summary of this. As we will see, Vargas is incorrect in his assessment of the evidence: not only is there no evidence for determinism in the brain, but much of how we currently observe the brain operating is what we would expect were the libertarian hypothesis true. Indeed, it seems that the libertarian's demand requires little to no change in what we observe about the brain, but merely a change in how we interpret these observations.

4 Indeterminism and the Brain

In this and the following section I seek to show that the libertarian hypothesis is relatively undemanding, and, more specifically, that its degree of demandingness does not outweigh the theory's philosophical virtues (assuming that libertarianism does indeed have these virtues). Whereas Vargas's original argument that the libertarian hypothesis is at odds with current findings in neuroscience was intended as additional evidence that libertarianism fails to satisfy the standard of naturalistic plausibility, his contention has force outside this context. Even if we reject the relevance of Vargas's standard, if libertarianism is indeed at odds with our best theories about the brain, then

this may well be a powerful reason to reject libertarianism, either in favor of compatibilism or even skepticism about free will and moral responsibility. I will now argue that neither conclusion is warranted because the libertarian hypothesis is relatively undemanding.

I begin my defense of this contention by considering whether there is evidence that the brain is deterministic. Vargas draws on the work of Henrik Walter (2001) and Patricia Churchland (2002) to defend the claim that our best models of the brain "tend" to be deterministic (Vargas 2004, p. 411). However, both Walter and Churchland conflate determinism with something like the thesis of universal causation: the thesis that all events are caused (Walter 2001, pp. 155-157; Churchland 2002, p. 207). They then utilize evidence that the brain is a thoroughgoing causal system to establish that the brain is deterministic. But libertarianism is consistent with the brain's being deterministic in *this* sense. While Walter and Churchland's argument has force against non-causal libertarianism—at least those versions that maintain that free action *cannot* be caused—it has no import for event-causal and agent-causal libertarianism, which maintain that free actions must have a causal etiology. What these libertarians deny is that all causation is deterministic. Hence, Walter and Churchland's argument leave these positions untouched.

There are two further problems with Walter's arguments. First, he fails to consider the accounts offered by Frederick Beck and John C. Eccles (1992) and Henry Stapp (1993) concerning how quantum physics is relevant to the workings of the brain, neither of which is subject to the criticisms he raises against other accounts. The second problem is that Walter believes that "Quantum physics gives us the only means with which we can establish an indeterministic version of being able to do otherwise" (2001, p. 154). But this does not seem to be the case. It may well be that neuronal indeterminism requires the truth of physical indeterminism, but it does not follow that the only *evidence* for neuronal indeterminism stems from quantum physics. The difference I have in mind can be understood as follows. One might seek to establish that the brain is an indeterministic

system by showing how the brain is significantly affected by micro-level indeterministic events. For example, some argue that certain brain processes, such as the release of vesicles, might be affected by the behavior of a quantum event in such a way as to inherit micro-level indeterminacies. A calcium ion might be in the superposition of states *A*&*B*. If the ion collapses to *A*, then vesicles will be released. If the ion collapses to *B*, then vesicles will not be released. Kane (1996, pp. 128-130) seems to defend the presence of indeterminism in the brain along these lines. However, one might think this is an overly restrictive account of the evidence for neuronal indeterminism. On a more liberal account, we allow for the possibility of gaining evidence for the existence of neuronal indeterminism simply by observing and experimenting on the brain. On this account it is possible to discover evidence for indeterminism without having to first map out how indeterminism in the brain bubbles up from indeterminism in micro-particles. Thus, Walter is mistaken to restrict the possibility of evidence for neuronal indeterminism to the first picture. As we consider the evidence for neuronal indeterminism we will want to keep these different models in mind.

Rather than continuing to respond to arguments that the brain is a deterministic system, let us turn to consider observations about how the brain works. As I understand them, these observations leave open the question of whether the brain is deterministic, but they certainly undermine the idea that, as we observe the workings of the brain, we are ineluctably led to postulate deterministic laws to explain its behavior. They also undermine the claim that the best models of the brain tend to be deterministic. But before I can defend these claims, we need a little neuroscience under our belt.

A possible location of indeterminism in the brain, given what we know about physics and neuroscience, is in the neuronal processes at the synaptic gap (cf. Beck and Eccles 1992; Stapp 1993; Shadlen and Newsome 1998; Atmanspacher 2006). A neuron consists of a cell body (the soma) and numerous branches that connect to other neurons forming synapses. The axon of a neuron is a

single branch that sends information coded in electrical signals to other neurons. The dendrites, the remaining, thinner branches of the soma, are poised to receive such information from other neurons. The synaptic gap, a one-millionth of a centimeter gap between the axon of the transmitting neuron (the presynaptic neuron) and the dendrite of the receiving neuron (the postsynaptic neuron), is the *locus* of information transaction in the brain. The following is a brief description of how this process of communication takes place.[10] Rapid changes in the flux of ions (which, depending on the direction and nature of flow, is known as depolarization or hyperpolarization) in the presynaptic neuron generate a moving impulse of electrical charge known as an action potential. When an action potential reaches the synaptic gap it stimulates the influx of calcium ions, which in turn causes the presynaptic neuron to release neurotransmitters. These neurotransmitters dock on the postsynaptic neuron and a flurry of electrochemical activity takes place. Once the postsynaptic neuron reaches a certain level of electrical activity as a result of the docking neurotransmitter, it fires its own action potential and the process repeats. This process is known as exocytosis (cf. Kandel *et al.* 2000, part III; Dayan and Abbott 2001).

One cannot over-emphasize the centrality of this exchange of information: it underlies all major psychological processes, such as cognition, emotion, and action (cf. Churchland 1996, p. 4). What is significant for our purposes is that the process of exocytosis admits only of stochastic or probabilistic modeling. Consider the following passages from Peter Dayan and L.F. Abbott's widely used neuroscience textbook:

> Neural responses can vary from trial to trial even when the same stimulus is presented repeatedly. There are many potential sources of this variability, including variable levels of arousal and attention, randomness associated with various biophysical processes that affect neuronal firing, and the effects of other cognitive processes taking place during a trial. The complexity and trial-to-trial variability of action potential sequences make it unlikely that we can describe and predict the timing of each spike deterministically. Instead, we seek a model

[10] Actually, it is only a description of how the process takes place *chemically*. Information exchange can also take place electrically, but for the sake of simplicity I ignore this complication.

> that can account for the probabilities that different spike sequences are evoked by specific stimulus. (Dayan and Abbott, 2001, p. 7)
>
> While the map from stimulus to *average* responses may be described deterministically, it is likely that single-trial responses such as spike-count rates can be modeled only in a probabilistic manner. (Dayan and Abbott, 2001, p. 16; emphasis mine)
>
> This reflects the fact that transmitter release is a stochastic process. Release of transmitter at a presynaptic terminal does not necessarily occur every time an action potential arrives and, conversely, spontaneous release can occur even in the absence of the depolarization due to an action potential.[11] (Dayan and Abbott, 2001 p. 179)

A couple of comments are in order. First, one of the major obstacles to showing that the brain is deterministic is the sheer complexity of neuronal events and, consequently, the difficulty of isolating the influences affecting the evolution of the system for any given trial. Testing for determinism requires an ability to set up experiments in which we can be confident that the initial conditions of the various trials are the same and that identical causal factors are present in each trial. The complexity of the brain makes it extraordinarily difficult to accomplish either task, destroying our confidence that each trial is "identical." Moreover, the brain is subject to many influences so that even if one could be sure that the initial conditions of the brain in two different trials are identical, it is highly difficult to guarantee the same causal factors are at work in the processes under consideration. These twin obstacles make testing for neuronal (in)determinism very difficult. Finally, and in light of these comments, we need to firmly distinguish a model's being stochastic from the brain's being indeterministic. Stochasticity is a property of models; indeterminism is a property of systems (such as the brain). A model of the brain can be stochastic and yet the brain itself deterministic. Therefore, to discover that our best models of the brain are stochastic is not tantamount to the discovery that brain is indeterministic.

However, given our best abilities to place constraints on trials to ensure sameness of initial conditions and sameness of causal factors, we observe apparent, non-negligible variability. The

[11] I am indebted to Balaguer (2009) for these references.

processes surrounding events at the synaptic gap admit only of stochastic modeling and explanation. Adina Roskies confirms Dayan and Abbott's conclusions:

> The picture that neuroscience has yielded so far is one of mechanisms infused with indeterministic or stochastic (random or probabilistic) processes. Whether or not a neuron will fire, what pattern of action potentials it generates, or how many synaptic vesicles are released have all been characterized as stochastic phenomena in our current best models. (Roskies 2006, p. 420; cf. Dean 1981; Tolhurst *et al.* 1983).

These considerations militate against Vargas's claim that the best models of the brain tend to be deterministic. Importantly though, this offers no evidence that the brain is indeterministic *in the way* libertarianism requires. Recall that choice must be undetermined in order for us to be free and that this means that brain states identical to or underlying free choice must themselves be undetermined. We know that such brain states will not consist in a single neuron's activity, but rather the activity of large assemblies of neurons (cf. Churchland 1996; Kandel *et al.* 2000; Dayan and Abbott 2001). Consequently, one might argue that none of the above observations suggest that the brain is indeterministic in the way required by libertarians—namely near or at the moment of choice.

This rejoinder is correct, but the above observations do undermine the idea that the best brain models are deterministic and it is against this contention that I offered them. This is relevant to determining the demandingness of the libertarian hypothesis. If the best models of the brain are deterministic, then the libertarian hypothesis requires that we come up with new brain models. This would increase the demandingness of the hypothesis. However, since the best brain models are stochastic, they are consistent with the libertarian hypothesis, rendering the hypothesis less demanding.

But this is not the end of the story. Observed variability is not limited to single neuron firing. The stochasticity of brain models extends to the brain on a larger level. Neuroscientists and psychologists are constructing models of perceptual and actional decision that are stochastic. Michael Platt *et al.* reports, "Most decisions of interest from a psychological, economic, or

evolutionary point of view appear stochastic" (Platt, *et al.* 2008, p. 130; cf. Shadlen and Newsome 1998; Smith and Ratcliff 2004; Gold and Shadlen 2007; Shadlen *et al.* 2008). Platt *et al.* goes on to construct an elaborate theory of decision-making that builds in this stochasticity. Variability is thought to affect the gathering and weighing of evidence, as well as the decision based on the evidence (Platt, *et al.* 2008). Variability in the brain is, therefore, not considered a mere aberration, something we can ignore in our theorizing, but is accorded a central place in our understanding of the brain. Not only, then, do we have evidence that the brain is indeterministic, we also have evidence that it is indeterministic in the way libertarianism requires.

Notice that none of this evidence against neuronal determinism turns on mapping out connections between micro- and macro-events. One way to argue against the hypothesis that the brain is deterministic is to show, first, that some micro-events are indeterministic and, then, that the indeterminism of these micro-events has non-negligible effects on the brain.[12] The evidence I have catalogued so far is derived solely from observations of how the brain works. However, once quantum mechanics is brought to bear, more evidence is accumulated against Vargas's contention that the brain is deterministic, since the behavior of calcium ions, the chemicals that determine whether the ion channels that release the vesicles open, are known to be subject to quantum effects (Gribbin 1995, p. 7ff). Therefore, we have reasons based in both physics and neuroscience against the contention that the brain is deterministic.

In light of these findings, we should conclude that the libertarian hypothesis requires little to no change in what we observe concerning the workings of the brain. We have two distinct sources of evidence against Vargas's contention that the brain is deterministic. The first comes from neuroscientific observations in which the neuronal behavior of both single neurons and large

[12] Kane employs this method. He argues that brain is a chaotic system in which tiny changes in initial conditions can have large effects. In this way he believes it is possible that micro-indeterminism may be magnified to the macro-level, resulting in non-negligible effects on the brain. See Kane (1996, pp. 128-131).

assembles of neurons (underpinning decision-making) varies from trial to trial. An additional source of evidence comes from quantum mechanics that suggests that the central processes involved in neuron firing are governed by indeterministic laws. The relationship between these may turn out to be tight: we observe variability because the brain is affected, in non-negligible ways, by quantum indeterminacies.

5 Libertarianism's Demandingness and the Goal of Science

What does all this show? What it does *not* show is that the brain is indeterministic.[13] For all we know, the appearance of indeterminism is just that: an appearance. One could try to explain the observed neuronal variability from trial to trial either by appealing to underlying phenomena that are completely deterministic but give rise to processes that appear indeterministic,[14] or by supposing the observed variability from trial to trial is due not to fundamental indeterminism but current limitations concerning our ability to secure truly identical initial conditions and causal factors.[15]

These responses remain real, genuine possibilities. Thus, I do not take myself to have *established* that the brain is indeterministic, or is even likely to be indeterministic. Here is what I take the forgoing to have shown. First, it shows that Vargas is mistaken: we do not have neuroscientific evidence against the libertarian hypothesis and the best brain models do not tend to be deterministic. Second, it shows that there is some evidence for the libertarian hypothesis. The fact that we observe variability from trial to trial raises the probability of the libertarian hypothesis: for this is what the hypothesis predicts we will observe. To say that this observation raises the probability of the libertarian hypothesis is not to say that it makes the hypothesis more probable

[13] However see Maye *et al.* (2007) for an argument that the brain is fundamentally indeterministic.
[14] It is important to realize that this option may well require one to endorse a deterministic interpretation of quantum mechanics, since once we have reached the level of exocytosis the only lower-levels appear to be ones known to be subject to quantum effects. Bohm (1952) offers the most plausible and well-developed attempt to understand the quantum phenomena in a deterministic way.
[15] Obviously these options are not exclusive.

than not, nor is it even to say it raises its probability by a significant degree. It is just to say that it raises its probability by some amount. And this suffices to show that these observations offer *some* evidence for the libertarian hypothesis.

Third, the libertarian hypothesis is relatively undemanding. Recall that demandingness of a hypothesis is a function of how radically things must change from what we currently take ourselves to know in order for the hypothesis to be satisfied. So how radically must things change from what we know in order for the libertarian hypothesis to be satisfied? The answer seems clear: not much at all. First, our current observations of the brain cohere well with the requirements of libertarianism. The libertarian hypothesis requires that the brain is indeterministic and thus that our best models would be stochastic. And this is what we find: our best models of the brain are indeed stochastic. Libertarianism also requires that indeterminism affects the decision-making process, and thus that our best models of decision-making would be stochastic. And this is what we find: we observe variability in trials concerning decisions-making—just as we would expect if the libertarian hypothesis were true. But why then is there such widespread resistance to the idea that the brain is indeterministic?

Paul Glimcher (2005) contends that much of the resistance to the libertarian hypothesis traces back to the early beginnings of the enlightenment when the idea that the goal of science was to reduce uncertainty was widely embraced. Glimcher argues that Karl Popper (1934/2002) solidified this conception of science in the early twentieth-century. Popper argued that the goal of modern science was to falsify extant theories through empirical observation. According to Popper, theories could not be confirmed by empirical observation, only disconfirmed. Theories that build indeterminism into them do not seem amenable to this conception of science. For if we assign a probability of 0.5 to a coin landing heads up, then no empirical outcome could falsify it: the coin landing tails up, on its side, and floating in mid-air are all consistent with this prediction. This led

Popper to argue that the proposition that the world is fundamentally indeterministic is an unscientific one.

Glimcher claims that even though many contemporary scientists dispute various aspects of Popper's philosophy of science, the importance of seeking a deterministic theory within the neurobiological and behavioral sciences has retained its grip (2005, pp. 51-52). Many have assumed that indeterminism at the quantum level has negligible effect on macroscopic biological organisms, and hence leaves untouched the enlightenment conception of science for these sciences. Glimcher offers a systematic critique of this assumption, arguing that there is evidence for indeterminism at the social, individual behavioral, neurophysiological, cellular, and sub-cellular levels. Glimcher takes this evidence to be a serious challenge to the enlightenment conception of science and argues that taking irreducible indeterministic hypotheses seriously may require no less than a re-envisioning of the scientific goal and methods for testing hypotheses. It is no wonder that many are loath to accept indeterminism.

But Glimcher suggests that the acceptance of indeterminism is not as radical as it may seem. Although many scientists aspire to the Popperian ideal of falsification, this is rarely attained in practice. Rather, "We often test theories against each other. We ask which theory provides more explanatory power, which yields a smaller [level of error in our predications], and we then discard the less efficient theory" (Glimcher 2005, p. 53). Behavioral theories are falsified, not by single observations, but with statistical generalizations. Returning to the coin case, although this prediction is compatible with any single case observation, if one were to flip the coin a thousand times and the coin landed tails in over nine hundred trials, then the prediction that there is a 0.5 chance of the coin landing heads up is falsified. So although an acceptance of irreducible indeterminism would call for a modification in our *ideal* of science, it would not change much of the practice. The main difference would concern how we evaluate theories that fall short of deterministic predictions:

sometimes this would be not a fault of the theory, but a reflection of the indeterministic nature of the world.

We are now in a position to understand what I meant earlier when I claimed that the libertarian hypothesis does not require so much of a change in what we observe, but more a change in how we interpret what we observe. For many, observing variability from experiment to experiment suggests that we do not yet understand all the causal factors at play in neuronal behavior, or perhaps that we have failed to secure sameness of initial conditions. Variability is interpreted as a shortcoming of the experiment or theory. For such theorists, that our best brain models are stochastic is indicative of a kind of failure and should move us to continue to try to come up with better models. If libertarianism is true, however, then we should expect to find limitations in our models, for such limitations are imposed on us by nature. Therefore, libertarianism requires us to reassess our criteria for successful scientific theories and experiments: observed variability need not be indicative of a failed theory or explanation. But this, again, requires a change in our interpretation of what we observe and not what we actually observe, and thus does not render the libertarian hypothesis particularly demanding.

6 Conclusion

In this essay I have sought to allay worries about the scientific plausibility of libertarianism. It is often argued that while libertarianism better accommodates our intuitive conception of ourselves as free and morally responsible agents, it must be rejected in light of its moral, metaphysical, and scientific problems. I have been concerned with defending libertarianism on the scientific front, arguing specifically that the libertarian hypothesis does not outweigh libertarianism's other virtues. In defending this claim I have considered Manuel Vargas's important challenge to libertarianism. In response to Vargas I argued that his standard of naturalistic plausibility is too vague to serve as a

guide in determining libertarianism's scientific plausibility. I then suggested that we focus instead on the actual neuroscientific data and the arguments given for (and against) the libertarian hypothesis. Vargas's main argument that libertarianism violates the standard of naturalistic plausibility was that there is no evidence that the brain is indeterministic and that there is evidence internal to neuroscience for thinking that the brain is actually deterministic. I have called both the claims into question. First, we saw that, at the present time, we can only model central processes in neuron firing stochastically. This militates against both Vargas's claim that we our best brain models tend to be deterministic and his claim that we have *no* evidence for the libertarian hypothesis. Second, we saw that the libertarian hypothesis is relatively undemanding, requiring not a departure in what we observe concerning the workings of the brain but rather a change in our assessment of these observations. While we lack anything like strong or decisive evidence for the libertarian hypothesis, its demandingness is not sufficient, by itself, to cast doubt on the tenability of libertarianism.[16]

[16] I am grateful to Manual Vargas and the members of the University of California, Riverside Dissertation Workshop, especially Mark Wrathall, for their helpful comments.

Bibliography

Atmanspacher, Harald. 2006. 'Quantum Approaches to Consciousness', *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/qt-consciousness/. Accessed July 2010.

Balaguer, Mark. 2009. 'Why There Are No Good Arguments for any Interesting Version of Determinism', *Synthese* 168: pp. 1-12.

Beck, Friedrich and Eccles, John C. 1992. 'Quantum Aspects of Brain Activity and the Role of Consciousness', *Proceedings of the National Academy of Sciences of the USA* 89: pp. 11357-11361.

Bohm, David. 1952. 'A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables, I and II', *Physical Review* 85: pp. 166-193

Churchland, Patricia. 2002. *Brain Wise: Studies in Neurophilosophy*. Cambridge: MIT Press.

Churchland, Paul. 1996. *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*. Cambridge: MIT Press.

Dayan, Peter and Abbott, L.F. 2001. *Theoretical Neuroscience*. Cambridge: MIT Press.

Dean, A.F. 1981. 'The Variability of Discharge of Simple Cells in the Cat Striate Corext', *Experimental Brain Research* 44: pp. 437-440.

Doris, John M. 2002. *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.

Fischer, John Martin. 2006. *My Way: Essays on Moral Responsibility*. New York: Oxford University Press.

Ginet, Carl. 1990. *On Action*. New York: Cambridge University Press.

Glimcher, Paul. 2005. 'Indeterminacy in Brain and Behavior', *Annual Review of Psychology* 56: pp. 25-56.

Gold, Joshua I. and Shadlen, Michael N. 2007. 'The Neural Basis of Decision Making', *The Annual Review of Neuroscience* 30: pp. 535-574.

Gribbin, John. 1995. *Schrödinger's Kittens and the Search for Reality: Solving the Quantum Mysteries*. New York: Little Brown

Hameroff, S.R. and Penrose, R. 1996. 'Conscious Events as Orchestrated Spacetime Selections', *Journal of Consciousness Studies* 2: pp. 98-111.

Honderich, Ted. 1988. *A Theory of Determinism: The Mind, Neuroscience, and Life-Hopes*. Oxford: Oxford University Press.

Kandel, Eric R., Schwartz, James H., and Jessell, Thomas M. 2000. *Principles of Neuroscience*, 4th edition. New York: McGraw-Hill Medical.

Kane, Robert. 1996. *The Significance of Free Will*. New York: Oxford University Press.

Maye, Alexander, Chih-hao Hsieh, George Sugihara, Björn Brembs. 2007. 'Order in Spontaneous Behavior', *PLoS One* 5(e443): pp. 1-14.

McCann, Hugh J. 1998. *The Works of Agency: On Human Action, Will, and Freedom*. Ithaca: Cornell University Press.

O'Connor, Timothy. 2000. *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press.

Platt, Michael, Dayan, Peter and Dehaene, Stanislas, *et al.* 2008. 'Neuronal Correlates of Decision Making', in *Better than Conscious?: Decision Making, The Human Mind, and Implications for Institutions*, eds. Christoph Engel and Wolf Singer. Cambridge: MIT Press, pp. 125-154.

Popper, Karl. 1934/2002. *The Logic of Scientific Discover*. New York: Routledge.

Roskies, Adina. 2006. 'Neuroscientific Challenges to Free Will and Responsibility', *Trends in Cognitive Science* 10: pp. 419-423.

Shadlen, Michael N., Kiani, Roozbeh, Hanks, Timothy D., and Churchland, Anne K. 2008. 'Neurobiology of Decision Making: An Intentional Framework', in *Better than Conscious?: Decision Making, The Human Mind, and Implications for Institutions*, eds. Christoph Engel and Wolf Singer. Cambridge: MIT Press, pp. 71-101.

Shadlen, Michael N. and Newsome, William T. 1998. 'The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding', *The Journal of Neuroscience* 18: pp. 3870-3896.

Stapp, Henry.1993. *Mind, Matter, and Quantum Mechanics*. New York: Springer-Verlag.

Tolhurst, D.J., Movshon, J.A. and Dean, A.F. 1983. 'The Statistical Reliability of Signals in Single Neurons in Cat and Monkey Visual Cortex', *Vision Research* 23: pp. 775-785.

Inwagen, Peter. 1983. *Essay on Free Will*. Oxford: Clarendon Press.

Vargas, Manuel. 2004. 'Libertarianism and Skepticism About Free Will: Some Arguments Against Both', *Philosophical Topics* 32: pp. 403-426.

Walter, Henrik. 2001. *Neurophilosophy of Free Will: From Libertarian Illusions to a Concept of Natural Autonomy*. Cambridge: MIT Press.

Wegner, Daniel. 2003. *The Illusion of the Conscious Will*. Cambridge: MIT Press.