Mark Balaguer, *Free Will as an Open Scientific Problem*, MA: MIT Press, Cambridge 2010. Pp. 202.

Mark Balaguer's recent book offers a defense of libertarianism (of its plausibility, not of its truth) from a deflationary view of the metaphysics of free will. On his account, libertarianism is the empirical hypothesis that decisions of a certain kind are undetermined in a way that enhances their freedom. He defends that the question of whether libertarianism is true is, at present, a wide open scientific question and the main problem in the metaphysics of free will. In chapters 1 and 2, he argues that, among the issues that settle whether free will is compatible with determinism, only one is a metaphysical issue: the empirical problem of which are the kinds of freedom that we have, whose most important subquestion is the (empirical) question of whether we have libertarian free will. In chapter 3, Balaguer develops his version of libertarianism, and argues that the question of whether this hypothesis is true comes down to the question of whether a sufficient amount of our *torn* decisions is undetermined in a specific way. Finally, in chapter 4, he argues that this is an entirely open scientific problem, and so that we have no good reason to believe either that we have libertarian freedom or that we don't have it.

According to Balaguer, questions like "What is free will?" and "Which is the freedom-requirement of moral responsibility?" are not metaphysical questions: they don't concern the nature of human beings and human decision-making processes, but the ordinary-language meaning of "free will", "moral responsibility", and related expressions. Answering these questions is obviously necessary (and also sufficient, since a complete notion of free will has compatibilism or incompatibilism built into it) for knowing whether determinism is compatible with free will and moral responsibility. But these questions are not part of the metaphysical component of the problem of free will and determinism. The metaphysical problem of free will concerns the nature of the freedom we exercise in our actual decision-making processes. It reduces to the empirical question "Which are the kinds of freedom that we have?" or, equivalently, "Which notions of free will apply to us?". And this question boils down largely to the most pressing of its subquestions: "Do we have a kind of free will that is enhanced by indeterminism?"

This metaphysical issue is independent from the semantic component of the problem of free will: whatever free will turns out to be (whatever "free will" turns out to mean), it may be that we don't have it ("free will" is not a natural kind term like "water"), and that we have capacities related to free will that are not denoted by "free will". For example, it may turn out that we have the capacities pictured by

some libertarian notion of freedom but that "free will" as we ordinarily use it refers to a capacity that is compatible with determinism. So claiming that humans have libertarian *freedom* (suspending judgment about whether it is correct to use "freedom" for referring to that capacity) does not commit one to defending that free will is libertarian free will. In a nutshell, conceptual analysis of the notion of free will is essentially irrelevant to the metaphysics of free will: our use of "free will" does not reflect anything important or nontrivial about the nature of our decisions (about how we author and control our decisions, the degree to which our reasons causally influence our decisions, the role determinism and indeterminism have in these processes, and so on).

Balaguer argues that the metaphysical component of the problem of free will -that is, the question of which are the kinds of freedom that we have- reduces, largely, to the question of whether we have libertarian free will, because this is the most important subquestion of the former, given that it is the most controversial one. (Balaguer takes it to be quite obvious that we have the capacities pictured by most compatibilist notions of freedom. For him, the next most important metaphysical question related to free will is the empirical question of whether our actions are sufficiently reasons responsive to meet the conditions set by any compatibilist account of freedom focused on reasons responsiveness. He thinks that, despite all the psychological studies which motivate pessimism about the issue, it is clear that at least some of our actions and decisions have the compatibilist freedom that is achieved through reasons responsiveness).

In chapter 3, Balaguer develops a variety of event-causal libertarianism that assumes mind-brain materialism and locates the required indeterminism at the moment of *torn* decisions. He argues that, if the world were indeterministic in the way hypothesized by this version of libertarianism, we would have more freedom than we would have if the world were not indeterministic in this way. Thus, he defends libertarianism from the criticism that indeterminism cannot in any way enhance free will.

We make a *torn* decision when we have reasons for more than one option but no conscious belief as to which option is best, and we choose one of our options without having consciously resolved this question. What determines whether a decision is a torn decision is its phenomenology: during deliberation, we feel torn as to which option is best, and then we have the experience of "just choosing" one of our live options without having formed the judgment that the option we choose (or any other option) is best. Given that what makes a decision be a torn decision is the way we experience the decision-making process, it is obvious and uncontroversial that we make many torn decisions in our everyday lives, both important and little: deciding whether to accept a job offer in one's town or move to another city, choosing between tiramisu and fruit for dessert, and so on.

Balaguer argues that a torn decision is as free as it can be only if it is *TDW-undetermined,* and so that full-blown freedom for a torn decision is libertarian freedom (a freedom enhanced by indeterminism). A torn decision is TDW-undetermined if and only if, until the choice is made, all the options favored to the same degree by the agent's reasons have roughly the same probability of being

chosen, and nothing outside the agent's conscious reasons and deliberation has a significant causal influence on the choice. (Note that the only thing that has to be undetermined at the moment of choice is which option will be chosen; the agent's reasons may be such that they causally determine that she will make a torn decision at a certain time, and even that she will choose from among a limited set of options). The key idea is that the scenario in which an agent exercises the highest degree of authorship and control over her decision is one in which the conscious reasons and intentions of the agent cause the decision, with no external factor *significantly* influencing the process. In a torn decision, there is a set of options that the agent takes to be equally good given her reasons, so only if it is *roughly* equally likely, at the moment of choice, that she will choose any option in that set, will her conscious reasons be the main cause of her choice. Only TDW-indeterminism ensures absence of hidden causal factors intervening in the decision-making process that are external to the agent's conscious reasons and thought. If the torn decision is partially or wholly determined, it must be so by virtue of subconscious reasons, nonmental physical events in the agent's brain, alien interveners, or some other thing. The damage will be more or less undesirable depending on which external factor is influencing the decision and on the degree to which it does; but any deviation from TDW-indeterminism will at least undermine, and at worst preclude, the authorship and control the agent exercises over her torn decision.

Balaguer argues that the truth of libertarianism depends on the question of whether our torn decisions (or at least most of them) are TDW-undetermined: first, because the full-blown, libertarian freedom of our torn decisions is worth wanting and important enough in itself, since we make a huge amount of torn decisions, and the idea that they are, contrary to what it seems to us, determined, is depressing; and second, because other kinds of decisions whose freedom would be enhanced by indeterminism require the same kind of indeterminacy for being as free as they can be, and the only plausible assumption is that indeterminism enters the same way in all these kinds of decisions, so if torn decisions are TDW-undetermined, these decisions are undetermined in a freedom-enhancing way. These decisions include Buridan's ass decisions – where the agent has two or more options that are equally good by her own lights, but her reasons for all of them are the same reasons- and decisions involving *leanings* – where the agent leans towards one or more of her options, but decides without having made a judgment as to which option is best. In torn decisions, any deviation from TDW-indeterminism means external influence to the agent's conscious reasons and thought; and, in general, any mismatch between the strength one assigns to one's reasons for an option and the probability of one's choosing that option means external influence.

So libertarianism, the thesis that the freedom of at least some of our free decisions is enhanced by indeterminism, boils down to the empirical hypothesis that the neural events that are our torn decisions are TDW-undetermined. Science is far from telling us whether this hypothesis is true, but the most controversial libertarian idea – that indeterminism brings or at least increases free will in a significant degree – is right.

In chapter 4, Balaguer argues that there are no good reasons to believe in the truth or falsity of TDW-indeterminism. He evaluates and rejects two a priori arguments, one for universal determinism, the other for the indeterminism required by libertarian freedom, and various empirical arguments for different kinds of determinism that are incompatible with TDW-indeterminism. I will briefly mention Balaguer's position in each case:

The first of the two a priori arguments he considers says that uncaused events, or things which "just happen", are inconceivable and so impossible. Balaguer's reply is that we can conceive of things just happening.

The second argument Balaguer rejects is supposed to show a priori that we have libertarian free will and that the world is suitably indeterministic. It says that we are morally responsible for at least some of our actions, and that this requires libertarian freedom which in turn requires indeterminism, so we have libertarian freedom and the world is indeterministic. Balaguer responds that, even granting that the premises are a priori, the argument doesn't provide good reason to believe the conclusion (despite the fact that the conclusion follows from the premises), because we are not justified in endorsing the conjunction of the premises, since arriving from one premise to the other requires holding that we are L-free and so that TDW-indeterminism is true (which we are far from knowing).

Moreover, Balaguer argues, no a priori argument can establish the truth of determinism or indeterminism, because these are contingent theses about the physical world.

Then, during most of chapter 4, Balaguer considers and rejects various empirical arguments in favor of various kinds of determinism which are incompatible with TDW-indeterminism:

Universal determinism (the thesis that all events are determined) is a wide open empirical hypothesis because we don't have compelling evidence for a deterministic or for an indeterministic interpretation of quantum mechanics.

Inductive arguments for macro-level determinism (the thesis that all macro-level events are determined) and for what Balaguer calls *virtual macro-level determinism* (the thesis that all macro-level events are made overwhelmingly likely by prior events together with physical laws) don't work because the premise they rely on (i.e., that our experience tells us that all macro-level events are determined or virtually determined) is unjustified: judging from the way they appear to us, many macro-level events could be undetermined, even if people often assume they are determined. And the argument that the success of science requires macro-level determinism, or at least virtual macro-level determinism, doesn't work mainly because universal causal laws connecting *kinds* of events are compatible with indeterministic causal connections between *particular* events of those kinds.

As for neural determinism and virtual neural determinism, presently they cannot be justified through any theory used, assumption made or evidence found in neurosciences.

Finally, Balaguer considers three sets of empirical data that might be used in arguments for torn-decision determinism or virtual torn-decision determinism, or directly against TDW-indeterminism: First, Tegmark's argument shows that quan-

tum indeterminacies *of a particular kind* cannot be relevant to neural processes, but this conclusion, Balaguer notes, is compatible with TDW-indeterminism; for all Tegmark argues, there may be other kinds of quantum indeterminacies relevant to neural events (in particular, relevant to mental events like torn-decisions). Secondly, Libet's results don't justify the hypothesis that the nonconscious brain processes that precede conscious intentions to act causally influence (much less determine or virtually determine) those intentions, so they don't give reason to think that torn decisions are not TDW-undetermined. Even granting that torn decisions are preceded by the kind of nonconscious electrical shifts in the brain that Libet calls *readiness potential*, these brain processes might causally influence only the fact that a torn decision will occur, or the fact that a torn decision will happen in one among a limited set of ways, and not which particular option will be chosen. For all Libet's data tell us, there are many functions the readiness potential may, equally likely, have; so we have no good reason for taking those data to show that, in a torn decision, there is always some factor external to conscious thought causally influencing that the agent chooses precisely the option she chooses. Third and last, Balaguer considers data from empirical psychology which seem to show that our actions are causally influenced by factors we are not aware of, that we are often mistaken about the reasons why we act, and the like; he argues that the data in question cannot be used to show that torn decisions are not TDW-undetermined, because they either don't say anything about torn decisions, or they don't say anything universalizable about them.

After rejecting a priori arguments for determinism or indeterminism and the just mentioned empirical arguments for different kinds of determinism, and assuming there are no good empirical arguments in favor of any kind of indeterminism, Balaguer concludes, in the end of chapter 4, that the question of whether TDW-indeterminism is true is a wide open empirical question to which we presently have no justified answer, and thus (given the arguments of chapter 3), that we have no good reason to believe that we have or lack L-freedom. Since (as was argued in chapter 2) the metaphysical problem of free will reduces largely to the question of whether we are L-free or not, this means that the metaphysical problem of free will is a wide open scientific problem, as the title of the book announces.

I think this book is a very interesting contribution to the libertarian literature. Its style is clear and ordered and its arguments well grounded. It raises interesting questions and gives a lot to think about, so it is a good text to generate discussion (for example, in a graduate seminar about free will), especially given the author's controversial positions in the meta-metaphysics of free will. Once this said, I would like to make one critical comment:

I am not convinced that the truth or falsity of TDW-indeterminism makes an important difference for freedom, understanding freedom as the capacity to guide one's behavior in light of one's reasons, or as an authorship and control over one's decisions that is based on the fact that one's conscious reasons are the main cause of one's decisions, or something like that – which seems to be roughly what Balaguer takes freedom to be. (From now on, I will use "freedom" and "free will" with this sense). I think Balaguer's version of the hypothesis that we have an inde-

terministic or libertarian free will reduces to the hypothesis that we exercise the highest degree of freedom even in situations where this requires indeterminism: we have the ability of determining what we do in light of our reasons, not only when our reasons pick out one option as our best choice (so we act as freely as we can only if we antecedently determine that we will choose our best option), but also when they don't (so our conscious reasons are the main cause of our decision only if our decision is suitably undetermined). And my worry is that most of the authorship and control Balaguer's L-free agent exercises is compatibilist authorship and control, and that agents in the best possible worlds where TDW-indeterminism was false would not lack much of the L-free agent's freedom. This inclines me to believe that, on Balaguer's version of libertarianism, the freedom-enhancing role of indeterminism is not central for the freedom of a full-blown free agent.

The role of indeterminism in Balaguer's picture is that of blocking those influences to decision-making processes which are freedom-undermining because they are external to the agent's conscious reasons and thought. (The argument for the thesis that TDW-indeterminism is freedom-enhancing is that, in some cases, the only possible scenario where these external influences are absent is one involving TDW-indeterminism.) But these influences threaten freedom for the compatibilist too, and for the same reasons. Balaguer points out that one's feeling of authoring and controlling a torn decision could be illusory only if it was causally influenced by a factor external to one's conscious reasons and thought (pp. 90-1). He takes it that, if a decision is caused mainly by the agent's conscious reasons and thought, it is sufficiently authored and controlled to count as free in the context of the problem of free will and determinism (and as authored and controlled as a libertarian should wish), as long as the decision is not influenced by alien intervention, nonmental brain events, and the like. That is, it seems Balaguer shares the compatibilist's idea that determinism in itself is not a threat to freedom; it is a threat only when it implies external influence. And this influence is freedom-undermining for compatibilists and incompatibilists alike. (One could even argue that Balaguer's libertarian hypothesis does not capture the incompatibilists' worries about determinism: the incompatibilist intuition that determinism seriously threatens freedom does not seem to concern only the freedom of decisions that agents make without having set their wills one way or another. It seems that it is central for incompatibilists to worry about how free the acquisition and maintenance of reasons, traits of character and the like could be in a deterministic world).

In Balaguer's picture, indeterminism has no role in the processes by which we get control and authorship over our character, values, beliefs, long- and short-terms goals, habits, and the like. So (assuming that, in order to have the capacity of acting freely by virtue of acting in light of our reasons, we must author and control our reasons) most of the freedom we have if Balaguer's libertarian hypothesis is true is compatibilist freedom. An L-free agent is free mainly because she has freedom over things she would have freedom over even if TDW-indeterminism was false.

Relatedly, I think that agents in the best possible worlds where TDW-indeterminism is false do not lack much of L-free agents' freedom. If a compatibilist imagined the best possible deterministic scenario for exercising the capacity of

guiding our decisions in light of our reasons, she would be able to get into her picture most of the freedom of Balaguer's full-blown L-free agent. She could give a deterministic picture where the lack of a match between real and experienced causal power of conscious reasons in torn (and similar) decisions were always due to the hidden influence of some unconscious reason strongly favoring one of the agent's live options. These unconscious reasons (or reasons that remain unconscious at least during deliberation, or reasons that the agent considers but whose influence in the decision-making process is not the one she thinks they have) would be acquired and maintained in a free (or non freedom-undermining) way, their influence on decisions would not go against the long-term values or interests that the agent most identified with, and so on. I believe that the only difference between this deterministic agent and Balaguer's full-blown L-free agent is that the L-free agent acts as freely as possible in more occasions. When the two agents act with as much freedom as possible, both do it by virtue of the same facts (presumably, authorship and control, free from external freedom-undermining influences, over reasons, deliberation, and decision). And, I think, when the L-free agent acts as freely as possible partly thanks to TDW-indeterminism, the determined agent acts quite freely too. Moreover: the falsity of TDW-indeterminism does not require determinism, so there are worlds where Balaguer's libertarian hypothesis is false just because our torn (and similar) decisions are only partially influenced by the above kind of unconscious reasons – that is, where these unconscious reasons fall short of functioning as hidden variables. These worlds are even less different from a TDW-indeterministic world than the best possible deterministic world. Given all this, and granting that the possibility that our torn decisions are causally influenced by unconscious reasons (or by conscious reasons in an unconscious way) is much more likely than the possibility that they are influenced by nonmental events or alien interventions or the like, the question of whether TDW-indeterminism is true doesn't appear so pressing.

To sum up, I am inclined to believe that the indeterminism introduced by Balaguer's libertarian hypothesis does not significantly increase the capacity it is supposed to enhance, and that the question of whether that hypothesis is true cannot be the most important metaphysical question related to free will. Nonetheless, I do think the best possible scenario for freedom (understood as Balaguer does) includes TDW-indeterminism, so I believe the book succeeds in the very difficult task of showing that indeterminism can be freedom-enhancing.

Mirja Pérez de Calleja
University of Girona
mirjaperezdecalleja@gmail.com