

Neil Levy

Consciousness and Moral Responsibility.

Oxford University Press 2014.

176 pages

\$45.00 (Hardback ISBN 9780198704638)

Sleep walking, sleep driving and even sleep murder, are not one's responsibility. According to the Canadian Supreme court and common opinion, somnambulism is an excuse for an agent's behavior. But why? In his latest book Neil Levy argues that consciousness is a necessary condition on moral responsibility: "moral responsibility requires that agents be conscious of the moral significance of their actions" (vi). Levy urges: "only when our actions are expressions of our selves can we be appropriately identified with them, such that we can be assessed on their basis" (x).

The book is divided into six chapters. In chapter one Levy quotes from Nomy Arpaly and Angela Smith to illustrate that there are apparent cases where people are held responsible for actions in which their motivations and reasons are not conscious to them, and cases in which we hold people responsible for being forgetful (like forgetting a child in a car that results in the child's death). Such cases, acknowledges Levy, appear to counter the consciousness thesis, so his task in the remainder of the book is to illustrate why they are mistaken.

Chapter two is devoted to sorting out the details of what is meant by the thesis. Levy notes that he is concerned with (i) information-bearing states, such as beliefs (29), that (ii) play a role in guiding behavior (32), that are (iii) retrievable, easily and effortlessly (33), and thus are personal, rather than subpersonal (31). Such easily accessible states Levy calls "conscious" even though they might not be occurrent and claims that this is what "thinkers like Freud seem to have in mind when they deny that agents are conscious of some of their motivations for their behavior" (35). Levy then narrows the set of states to only those that contribute to the positive or negative "valence" of the agent's behavior. Levy is rather equivocal on this matter, however, writing: "Typically facts that make the action bad play this privileged role in explaining why the responsibility is valenced negatively" (36). But he then continues: "agents are not morally praiseworthy for donations to good causes if they mistook the ... collection box for 'People for cannibalism.' Valence is explained not by facts pertaining, but the facts the agent takes to pertain" (37). Chapter three is devoted to defending the functional role of awareness as having an integrative role that allows for "flexible, reasons-responsive online adjustments of behavior." Levy offers a breezy discussion of the global workspace view of consciousness and offers some remarks against those theses that appear to contradict it, focusing in particular on the work of Jesse Prinz.

In chapter four Levy argues that consciousness is required for flexibility of responses to situations. He examines the somnambulist behavior of Ken Parks, who drove across town and killed his mother-in-law. He argues plausibly that the best explanation for Parks' behavior is a series of learned action scripts. So although the behavior is most certainly complex, in fact Parks' behavior was not really flexible: Parks was unresponsive to the cries of his mother-in-law when he attacked her. Integration of content, allowing for flexible responses, is the key function of consciousness, Levy claims.

Having concluded that consciousness integrates actions scripts and allows for flexible reasoning and behavior, Levy turns to normative issues, addressing the evaluative agency account of moral responsibility. Levy is keen to reject Arpaly and Smith's claim that an agent can be

responsible for actions in which they lack awareness of their reasons for those actions, so he begins with the generally agreed upon cases of non-responsible somnambulist behavior. He notes that somnambulists are not sensitive to stimuli that conscious folk are and thus are not flexible in their behavior or reasoning.

Smith and Arpaly did not have inflexible non-conscious behavior in mind, of course—only cases in which the agent's reasons were non-conscious. So Levy turns to these cases and baldly contradicts their thesis saying that such reasons are not expressive of the agent's evaluative agency. Perhaps an agent discovers that the reasons they give for their choice (e.g., deciding between a woman and a man, based on the candidates' qualifications, who is best qualified to be a police-officer) are confabulated, and that, in fact, they have an unconscious sexist preference driving their judgment. Had they been conscious or aware of such tendencies they might have done their best to combat them and they would have certainly rejected these preferences, for they express unjust attitudes. According to Levy, "while the choice had a sexist content, this content was not an expression of agents' personal-level attitudes. [By definition] *For that reason*, it is not plausibly taken to be an expression of their evaluative agency, their deliberative and evaluative perspective upon the world. This perspective must be a relatively coherent, relatively stable, set of attitudes ...When an action has a significance that conflicts with a range of attitudes central to an agent's evaluative perspective, and which, were the agent able to detect this inconsistency she would prevent, there are no plausible grounds for holding that the action expresses the agent's evaluative perspective" (97, italics added).

In the final chapter Levy examines the notion of guidance control, commonly taken to be a necessary condition on an agent's responsibility for their behavior. He claims that we cannot satisfy the guidance control requirement when the mental states causing our behavior are not conscious. Levy writes, however, not of guidance control over choices or behavior but of guidance control over the mental states that cause or guide our behavior: "because these agents were conscious neither of the implicit attitude that caused the confabulations, nor of the moral significance of the decision they made, they could exercise guidance control over neither" (123). But this leaves it open whether the behavior is under the guidance control of the agent in virtue of the agent's unconscious reasons and unconscious (sexist) attitudes. The agents in such tests are conscious and aware of what is going on around them and they are responsive to reasons when they are offered—indeed, they are asked to give reasons for their decisions. But they respond to the request for reasons that justify their choice with something confabulated, something that reveals they do not understand their own motivating mental states. Their behavior seems to be reason-responsive. They offer reasons to explain their behavior and their unconscious reasons and preferences also guide their choices; giving false reasons is not being unresponsive to reasons. More convincingly, however, Levy takes it that because these mental states are non-conscious our behavior is not under our reason-responsive guidance control: "Insofar as our behavior is shaped by facts of which we are unaware [non-conscious mental states] we cannot respond to these facts [non-conscious mental states], nor to the conflict or consistency between these facts [non-conscious mental states] and other [conscious] reasons" (115). Thus the unavailability of these reasons to other reasons, so to speak, is what discounts the action as being under our guidance control. We are unable to consciously assess, integrate and evaluate these non-conscious mental states. Levy ends the chapter, almost as an afterthought, with a discussion and rejection of George Sher's views from Sher's *Who Knew? Responsibility without Awareness*.

Consciousness and Moral Responsibility has the advantage of being a short book and the reader could, I think, limit their labour further by skipping to chapters five and six to find the core

of the argument. Levy complains about intuitions and their uncertain causes when responding to Sher's examples but he helps himself to "intuitions" regarding the non-responsibility of sleepwalkers. Some of his comments about central concepts puzzled me. In a footnote he writes: "Mental states are personal when they can rightly be predicated of the person herself ('Jenni believes that p', for instance); and subpersonal when they can only be predicated of some lower-level components of the mind ('edge detection occurs in V1', for instance)" (31). But I don't understand why my subpersonal states aren't properly predicated or attributed to me, a person, just as my non-conscious physical processes are appropriately predicated of me (e.g., "Brian digests his dinner"). But perhaps this definition was driving the show? If Levy takes it that the person doesn't really include "their" subpersonal mental states then these states do not express the agent's evaluative agency (they cannot even be predicated of the person) and, not being properly predicated of the person, the agent has no internal guidance control over them. With that assumption in hand the consciousness thesis is well defended, but it is an awfully controversial assumption and is, I would think, at the core of the debate. I am the person who has these attitudes and they are not due to any pathology or abnormal psychology. Levy writes "Our evaluative agency is caused or constituted by our personal-level attitudes" as if this were almost a premise of an argument rather than one of the substantial conclusions he is urging us to accept.

Brian Jonathan Garrett
Kwantlen Polytechnic University