The Mental Basis of Responsibility

Walter Glannon
(Aldershot, Ashgate, 2002)

Given the apparently never-ending stream of publications on responsibility and free will, any new book on the subject must do something to distinguish itself. Glannon seeks to develop a distinctive 'capacity' theory of moral responsibility, according to which moral responsibility depends on the possession of certain relevant capacities that together give us the capacity for 'reflective self-control' (and does not require that we have genuinely 'alternative possibilities of choice and action' (5)). He tell us that there are 'six features of the book that distinguish it from other works on moral responsibility', including other capacity-based theories –

> (1) attention to the agent's epistemic capacities, especially beliefs about the foreseeable consequences of his actions and omissions; (2) attention to the essential role of emotions in prudential and moral reasoning, and the idea that emotion and cognition are interdependent, interacting mental faculties necessary for rational and moral agency; (3) a conception of personal identity that can justify holding persons responsible at later times for actions they performed at earlier times, a conception that accords with and is shaped by our normative practices; (4) a compatibilist theory of responsibility for actions, omissions and consequences whose requirements are less strict and broader than those in standard compatibilist theories; (5) an emphasis on neurobiology rather than physics as the science that should inform our thinking about free will and responsibility; and (6) the melding of literature on free will and responsibility in contemporary analytic philosophy with legal cases, abnormal psychology, neurology, and psychiatry, which gives greater nuance and a richer texture to the general debate on the relevant issues (5-6, 144).

To do all this, in a mere 144 pages of text, would indeed be a significant achievement. I fear, however, that the book falls well short of fulfilling this promise.

The general outlines of a capacity theory of responsibility are by now familiar enough. We explain what it is to be a morally responsible agent, or to have 'free will', in terms of the various capacities that characterise responsible agents. We might talk initially of capacities for autonomous action or, as Glannon talks, for 'reflective self-control'; but we must then go on to identify the more specific capacities on which these general capacities depend. Glannon identifies six such capacities: the ability to 'form and respond to desires'; the abilities 'to form and respond to beliefs about the circumstances and consequences of action' (theoretical reason) and 'to form and respond to reasons for or against actions, and to form intentions to act' (practical reason); 'the ability to have and respond to emotions'; the 'ability to execute desires, reasons and intentions in choices and actions'; and 'the

ability to perform voluntary bodily movements' (14-5). Much of the book consists in the more detailed analysis of these capacities and of their significance for moral responsibility. Whilst Glannon's account is clearly, as he notes, a close relative to other capacity-based theories, in particular those that emphasise the importance of the agent's responsiveness to reasons, he argues that others have paid insufficient attention to the roles of the capacity for theoretical reason and of emotion – two of the features that make his book distinctive. However, both these charges seem at least overstated.

Theorists have typically included the capacity to form and respond to reasonable beliefs about the circumstances and consequences of my actions as a necessary condition of moral responsibility: that is why ignorance of relevant facts, or lack of foresight of consequences, is typically an excuse – at least if it was not reasonably avoidable; and that is why someone in whom that capacity is seriously lacking or impaired is not treated as a responsible agent. The failure that most concerns Glannon is, it seems the failure of those who defend some kind of principle of alternative possibilities to recognise the ways in which an agent who foresaw the situation in which he would lack alternative possibilities could properly be held responsible for what he does, or fails to do, in that situation (ch 5).

What Glannon appeals to here (and makes rather heavy weather of explaining) is the familiar doctrine that conditions that could otherwise exculpate an agent will not do so if he culpably brought them about, or culpably failed to prevent them arising. If, before going to an exhibition of fragile glass, I take a drug that I know will (or suspect might) cause me to have a violent fit during the exhibition, I can hardly excuse myself for the damage I cause by pointing out that my violent movements were involuntary; if I take a sleep-inducing drug knowing (or even intending) that it will cause me to be asleep at the time that I should be attending a meeting, I can hardly excuse my failure to attend the meeting by pleading that I was asleep and thus unable to attend. Glannon offers us more complicated examples than these, but that seems to be the essential point. However, it is not clear that they do constitute counter-examples to the principles he is attacking. The first is PAP': 'A person is not morally responsible for what he has done if he did it only because he could not have done otherwise' (97). But it is not clear that, in my example, I damage the exhibits '*only* because I could not have done otherwise'; it seems natural to say that I damage them because I took the drug – and a similar point applies, I think, to Glannon's examples.

The second principle, as specified by Glannon, is that 'A person is morally responsible for failing to perform some act A at some time T' only if he could at some time T not later than T' have performed A at T'' (107): I find this somewhat opaque, since it seems to suggest that I could 'perform

A at T'' at some time earlier than T'; but it is meant to capture the thought in van Inwagen's clearer principle that 'An agent cannot be blamed for a state of affairs unless there was a time at which he could have arranged matters such that that state of affairs not obtain' (108).[1] Now this clearly allows me to be blamed for missing the meeting when I do so because I took a soporific knowing that this would be its effect: but Glannon thinks it cannot deal with the case of a non-swimmer who impersonates a lifeguard, and who is then unable to save a swimmer who gets into difficulties from drowning. He is, Glannon argues, not only 'morally responsible ... for impersonating a lifeguard and for the consequence of the swimmer drowning', but also 'for failing to perform the lifesaving act, even though at no time is he physically able to perform that act' (108).

There are various problems with this argument, which I think are symptomatic of more general problems with the book. First, the example is too sketchy: in particular, it is not clear whether others (other lifeguards, the swimmer) rely on this supposed lifeguard (he persuades the other lifeguards that he is qualified, and sits in the lifeguard's chair, which suggests that others might be relying on him; but '[h]is presence or absence makes no difference to what occurs within the causal sequence once it is underway' (108)). If others did rely on him, he could be held responsible for the swimmer's death insofar as it is likely that the swimmer would not have entered the water, or that other lifeguards would have been on watch there, had it not been for the impersonator's deception. If they did not rely on him, it is hard to see how the swimmer drowning is a 'consequence' of his impersonation, or how he is responsible for *that* (despite Glannon's comments at 110). Second, suppose he had not impersonated a lifeguard, but had just been innocently on the beach, and had watched with helpless horror as the swimmer drowned. We presumably would not in that case say that he 'failed to perform the lifesaving act' – in which case he could have so arranged matters that he did not fail to perform the lifesaving act, by not engaging in the impersonation (or by leaving the beach, for that matter). Thus the example does not threaten van Inwagen's principle, once we get clear about what the relevant 'state of affairs' is.

I have spent some time on this argument because I think it exemplifies the way in which Glannon is too often prone to over-complicate matters, and to argue too hastily and sketchily (a result, no doubt, of trying to fit so much material into so short a book). The latter defect is also evident in his comments on the role of emotion in practical reasoning – another aspect of the book that is supposed to distinguish it from others, by correcting other theorists' failure to pay enough attention to that role. That charge seems strange, given the rich literature on the emotions of the last

[1] Quoting van Inwagen, 'When Is the Will Free?' in *Philosophical Perspectives III* (1989) 399, 419.

few decades – a literature that has included quite a lot of attention both to the character of emotions (in particular the extent to which they involve a cognitive aspect), and to their role in moral thought and action. It is also strange to be told, twice, that we need to look to neurobiology to discover the role that emotions play in practical reasoning (33, 38): there is of course room for serious discussion about the relationship between philosophy and such sciences as neurobiology, but philosophers from Aristotle onwards have managed to say quite a lot that is useful and illuminating on this topic (as well as, inevitably, a lot that is confused or misleading) without the benefit of expertise in neurobiology. The real problem with Glannon's discussion, however, is that whilst he often tells us that the emotions play a crucial role in practical reasoning, he never tells us in enough detail either how we should understand and analyse emotions (beyond making clear that he rejects accounts that portray emotions in overly cognitive terms), or what their role is – beyond a number of rather sketchy examples, and a discussion of the by now well worn example of psychopathy. The discussion of psychopathy is puzzling. Glannon argues that psychopaths are partly responsible for their actions since they have 'the capacity to recognise moral reasons for and against performing certain actions', though they lack 'the capacity to respond to these reasons in the appropriate way' (61). But, first, it seems strange to hold a person even partly responsible for wrongful actions if he lacked the capacity to respond to or be moved by the moral reasons against such actions – unless we could claim that he was responsible for that lack of capacity. Second, it is not clear that one who lacks the capacity to respond to reasons, and the emotional capacities that psychopaths are said to lack, can really have the capacity to recognise them *as reasons*: for surely to recognise something as a reason involves at the least an understanding of how people could be moved to act by it, if not seeing it as something that could move me; and in virtue of the role that emotions play in such understanding, it is not clear that it is available to the psychopath as Glannon describes him. Perhaps Glannon could say more to explain and defend his claim: but he does not say nearly enough here to render it plausible.

I have focused on some (though not all) of the features of this book that Glannon claims makes it distinctive. The book is useful as a detailed statement of a capacity-based theory of moral responsibility, and as bringing together themes from contemporary discussions of the conditions of responsibility, of free will, and of personal identity: but its arguments are too often sketchy and underdeveloped, and its claims too often insufficiently explained to render them clear or plausible.

R A Duff
(Department of Philosophy
University of Stirling)