Criminal Law and Philosophy (2023) 17:235–242
https://doi.org/10.1007/s11572-022-09637-5

BOOK REVIEW

Review of Stephen P. Garvey, *Guilty Acts, Guilty Minds* (Oxford: Oxford University Press, 2020)

Christopher Bennett[1]

Accepted: 24 May 2022 / Published online: 25 June 2022

This is a book about culpability in criminal law, and one that sets out to take a whole new approach to the topic. Whereas some theorists have approached culpability in criminal law as an issue in moral philosophy, and others as an issue in political philosophy, Garvey thinks that the general assumption has been that what we are looking to explain is how criminal justice can be compatible with standards of *justice*. By contrast, Garvey suggests that we should be concerned with how criminal justice can be *legitimate*. The book.

> 'asks not if the rules comprising the criminal law – rules criminalizing conduct, providing affirmative defenses, and authorizing types and degrees of punishment – are just. Instead, it zeroes in on legitimacy. It first offers reasons to believe a democratic state is (or can be) legitimate, with the power to morally obligate conformity to its rules. It then tries to identify limits, captured in the twin requirements of *actus reus* and *mens rea*, on a democratic state's authority to ascribe culpability to a citizen who's chosen to do something the state has declared to be a crime. The focus is thus on one second-order moral constraint on a democracy's authority, to which a democratic state must adhere if the punishments it imposes are to be legitimate.' (p. 12)

1 Overview of the argument

Let's unpack this a bit. First of all, why legitimacy rather than justice? Garvey accepts arguments akin to those of Jeremy Waldron that a theory of legitimate authority has priority over the theory of justice, given that 'experts and non-experts' disagree about

✉ Christopher Bennett
c.bennett@sheffield.ac.uk

[1] University of Sheffield, Sheffield, England

the values and rights that would inform the latter (Waldron 1999), and that we therefore need an impartial procedure for decision-making that all parties can settle on despite their disagreements (p. 35). Garvey further accepts the idea that democratic procedures are the fairest and most appropriate way to deal with such cases of disagreement, and thus he takes it that the theory of legitimacy is basically the theory of democracy. The argument for the authority of democracy has two steps. First of all, there is a claim that any state, democratic or not, has the right that its subjects conform to legal *mala in se* standards that are necessary for social peace, and a moral permission (or liberty right) to coerce conformity to its rules, providing that: 'such coercion is necessary to assure conformity to its rules; conformity is necessary to achieve some good for all those subject to its coercion; and the good is one everyone, or most everyone, recognizes and wants but isn't able to secure without general conformity.' (p. 47) Secondly, there is an argument that respecting one's fellow citizens as an equal ('endowed with the same freedom to form beliefs about what justice requires when it comes to the rules governing his relations with his fellow citizens' [p. 51]) grounds an obligation 'to conform [one's] conduct to the rules resulting from the only real mechanism available to resolve their disagreements consistent with the status of each as equally free' (p. 52). That mechanism, Garvey thinks, is democracy. A democratic state, he claims, not only has a right to conformity; it also has a 'moral power to change the normative status of its citizens' (p. 53).

> 'With the coming of democracy, [democratic citizens], collectively, acquire the moral power to impose on themselves a moral obligation to conform to any rule they together agree upon through majority vote. This obligation binds them all, even those who believe that this or that particular rule is unjust.' (p. 44).

How does this relate to the issue of culpability? If democratic states have a moral power to change citizens' rights and duties, then presumably this implies that such states can impose obligations (and permissions) on their citizens that depart from standards of justice. In this respect, we might say that Garvey argues for the priority of legitimacy to justice. But Garvey does not think that such discretionary authority is unlimited. Rather, he accepts that justice imposes limitations on the legitimate exercise of authority. When it comes to criminal justice, democratic states face certain rights-based limits. Thus, in this further respect – at the margins, as it were – Garvey might be said to affirm the priority of justice to legitimacy. Garvey sees his task as 'to state, in an abstract and general way, the rights we believe citizens have limiting democratic authority over the rules comprising the criminal law. Those rules come in three sets – rules defining crimes, rules ascribing culpability, and rules prescribing punishments – and democratic citizens have, we believe, rights associated with each set.' (p. 61–2) With that general manifesto set out, the inquiry of this book focuses on culpability.

Of course, it might be said that 'experts and non-experts' also disagree about where the limits of legitimate authority lie, and therefore that the logical conclusion of Garvey's argument from disagreement is that there are no such limits, and everything must be decided by democratic procedures. Garvey is aware of this concern, but argues that such limits should be understood as 'self-evident truths,' such that,

'if you fail to acknowledge the right declared, your reason has, at least on this occasion, failed you' (p. 61). The basic idea is thus that one such set of limiting rights is expressed in the requirement that an agent has met the *actus reus* and *mens rea* conditions, and he develops a 'rights theory' of culpability conditions that explain where justice limits legitimate authority. (Actually, the book is not solely concerned with 'guilty acts' and 'guilty minds,' narrowly construed. Chapter Four looks at the insanity defence as resting on 'loss of agency,' while Chapter Five looks at issues of standing to punish, and the defence of 'rotten social background.' However, I will confine my discussion to the issues set out above.)

I agree with Garvey that we need a normative theory of culpability conditions. His book breaks new ground in setting out how to think about such a theory and will, I hope, stimulate further debate in this important area. I am also grateful to him for connecting the debates about culpability with the question of legitimacy. He deserves credit for shining a light on this comparatively neglected issue, and his book will be a key reference points for future work on this topic. Having said that, there are two critical questions that I would like to raise about his project. The first question is whether Garvey is right in the story I have just sketched about why we need a normative theory of culpability. As we have seen, Garvey's claim is that we need a normative theory of culpability conditions because of the limits justice places on legitimate authority. But is that explanation, which stakes its claims on the priority of legitimacy to justice, and the special legitimacy of democracy, the right one? The second question is whether the normative theory of culpability that Garvey proceeds to develop is the right one. Let us now look at these questions in more depth.

2 The case for legitimacy

The first question is whether Garvey is right that we need a normative theory of culpability that shows how justice limits legitimacy. My starting thought about this is that, while Garvey's framework *might* turn out to be the correct way of motivating the question, much of his 'rights theory' of culpability conditions could in fact be accepted by a theorist who rejects the importance of the distinction between justice and legitimacy. We can start to see this point by asking what difference the argument about the priority of legitimacy to justice makes to the argument of Garvey's rights theory. While Garvey does remind us periodically that the role of the rights theory is to find limits to democratic authority that cannot reasonably be doubted, I do not think that he needs the legitimacy framework sketched above in order to argue for the rights theory. Neither is it the case that the two arguments are mutually reinforcing, such that (1) accepting the rights theory strengthens one's reason to accept the arguments about legitimacy and democracy, and (2) accepting the arguments about legitimacy and democracy strengthens one's reasons to accept the rights theory.

Now, one might think that I am jumping the gun here, as I have not yet explained what Garvey's rights theory of culpability conditions commits him to. However, I think it can already be seen that one who rejects legitimacy could accept the rights theory. This is because, on Garvey's view, the constraints that are the concern of the rights theory are required for legitimacy *because their validity as considerations of*

justice cannot reasonably be doubted. This is of course compatible with Garvey's suggestion that they are constraints on legitimacy. But someone else, less convinced of the importance of legitimacy, might simply accept these constraints as constraints on and of *justice*. Such a theorist might simply say that our reason to accept these constraints is that they are fundamental constraints of justice, and that states are constrained by fundamental justice; and would thus be able to accept the main thrust of Garvey's rights theory of culpability while dispensing with legitimacy.

This suggest that the rights theory in itself is neutral among various views as to why a decent state needs culpability constraints on its right to punish. One need not sign up to the legitimacy framework to accept this part of Garvey's theory. And thus Garvey does not need to argue for the legitimacy framework in order to show that the rights theory is correct. This makes the rights theory stronger, to this extent, in that its success does not depend on our accepting Garvey's Waldron-esque approach. (And this is important given that Waldron's approach is controversial – see e.g. Barry and Øverland 2011.) In this sense Garvey might have done better, in demonstrating the potential attractiveness of his rights theory, to have made clear its compatibility with a wide range of views about the correct way to think about the right of the state to punish. The legitimacy framework allows us to understand the rights theory in a new light, but it is not required for it. Nevertheless, if Garvey's ambition was to show that a concern for legitimacy will make a big difference to how we view the substance of culpability conditions, I am not sure that he has succeeded.

The same point can be made about Garvey's commitment to democracy. Why does Garvey need to commit himself to democracy? To what extent does the rights theory that he defends depend on the truth of democratic theory as the answer to questions about legitimacy? If my interpretation is correct, it doesn't, on the grounds that Garvey's argument for the rights theory is that there are constraints of justice that cannot reasonably be doubted, and one can agree to that without agreeing to what Garvey believes about democratic authority. Furthermore, the democracy argument is controversial, and Garvey does not defend it against some well-known objections (some of which are considered by Waldron himself). One such criticism would be that citizens may reasonably disagree about the view of the requirements of treating people as free and equal on which Garvey's argument for democracy rests. Perhaps Garvey might argue that citizens cannot reasonably disagree with the proposition that people are free and equal. However, even if that is correct, it looks as though citizens could reasonably disagree with the claim that the only way to respect people as free and equal, in conditions of disagreement, is to institute a binding procedure of majority voting. But if such disagreement is reasonable, how can this controversial democratic procedure help us solve the problem of reasonable disagreement that motivated the turn to legitimacy? I am not unsympathetic to Garvey's democratic leanings, but the problems facing his view are not confronted here; neither is the democratic view needed as a foundation for the rights theory of culpability to follow. As with legitimacy, the democratic approach puts the rights theory in a new and interesting light, but this approach is not required to accept the rights theory.

3 The rights theory of culpability conditions

Let us now turn to the detail of Garvey's rights theory of culpability conditions. I think the issue that Garvey has raised by this point in the book is: why do the *actus reus* and *mens rea* requirements represent necessary conditions of legitimate punishment? Garvey avowedly does not present an argument for the claim that they do represent such necessary conditions – indeed at one point he claims not to know what such an argument would look like (p. 89). But he does helpfully explain how the normative basis of the *actus reus* requirement should be understood. He distinguishes two cases: one in which the agent knew he was committing a crime; and one in which he did not. In the former case, he says, an agent cannot legitimately be punished, even if they knew they were committing a crime, unless they had the capacity to conform their actions to the law: 'justice does not permit a democratic state to punish a citizen who lacked the capacity to conform to the law' (p. 89). In the latter case, Garvey claims that justice does not permit a democratic state to punish an agent who does not realise that they are committing a crime, unless they could have realised that they were committing a crime (p. 91). These formulations use 'capacity' or 'could have,' which can be a bit of a slippery fish. One question, as I have mentioned, is why such a counterfactual capacity is normatively relevant here. Garvey does not give an answer, but one traditional possibility – that he seems to be aware of – is that ability to do otherwise gives the agent a fair opportunity to avoid punishment (though that account will not be straightforwardly applicable to 'could have known'). Another question is how we should understand 'could have done otherwise.' Garvey's account of capacity to conform to the law is the 'Stephen Test' (p. 127) that an agent could have done otherwise unless 'the impulse to commit a crime [was] so violent that the offender would not be prevented from doing the act by knowing that the greatest punishment permitted by law for the offence would be instantly inflicted.'

As regards the *mens rea* condition, Garvey adopts the view that justice only permits the state to punish if the agent's action manifested insufficient concern for the law (p. 102). Insufficient concern is analysed roughly as the idea that the considerations the agent took to count in favour of the action were not compatible with the appropriate concern for the law. Garvey explains this idea with a psychological theory of the 'conative will,' according to which choice is the result of a struggle of competing 'motivating reasons' with causal power. However, I am not sure that this psychological theory is plausible, or that Garvey has any need to commit himself to it to establish the correctness of his core view. To decide whether the action is compatible with appropriate concern for the law, Garvey proposes the 'Jekyll Test' (p. 108) on which we replace the actual agent with an agent who has appropriate concern, and then ask how they would have acted.

Garvey develops these accounts in more detail than I can convey here, and with expert reference to a fascinating range of legal cases. He discusses many issues – from compatibilism to self-control to culpable ignorance and duress to delusion to provocation – that shows his adeptness with philosophical discussion. However, I do want to raise a general query about his approach. He contrasts his normative 'rights theory' of culpability conditions to the descriptive 'conventional' theory, according to which *actus reus* and *mens rea* do nothing more than 'identify, or refer to, the

elements of a statutorily defined crime' (p. 77). Garvey says that criminal lawyers probably have the conventional theory in mind when they talk about *actus reus* and *mens rea* (p. 76); and that is all very well given that the authority of the statute is the end-point of the inquiry as far as the lawyer's day-job is concerned. But the conventional theory is clearly not offered as an alternative normative theory of culpability to Garvey's own. This means that the 'rights theory' developed by Garvey is the only normative theory of culpability considered in these pages. The fact that he does not review alternative normative theories before arguing that his is the most adequate might mean that some readers find his account unpersuasive. Garvey explains that his method is to present a picture of how things might hang together rather than a full-scale argument that his is the only reasonable view (p. 17). While I agree that that can be a valid method, it would sometimes have been helpful to get a better sense of the considerations that underpin Garvey's theoretical choices.

To raise just one alternative theoretical possibility, it is not clear to me why Garvey doesn't consider that insufficient concern could be relevant to *actus reus* as well as *mens rea* conditions. Indeed, it is not entirely clear to me why Garvey chooses to analyse *actus reus* and *mens rea* separately, and to ascribe to them separate tests. In relation to both mind and action, there is an issue about what it is to have the relevant sort of freedom – what it means to have had the capacity to get it right in a situation in which one got it wrong – and then an issue about what it is to get it right – that is, the normative standard against which we are to be judged. An account of the capacity to conform is relevant to the first issue, but it is not only relevant to action, since we can also form false beliefs in situations where we should have got it right. Thus, an account of capacity to conform (an example of such an account might be the notion of 'reasons-responsiveness') is relevant to mental issues such as knowledge and insanity as much as it is to action. The Jekyll Test – or more broadly, an account of what sufficient concern involves – is relevant to the second issue regarding the appropriate normative standard. But it is again artificial to say that this applies only to mind rather than action, given that we can get our actions right as much as our states of mind. Indeed, a crucial point here is that reference to states of mind is often essential to the individuation of action (i.e., to saying which action has been performed). The law's distinction between *actus reus* and *mens rea* may be helpful for some purposes, but when taken as a guide to the metaphysics of mind and action, it can sow confusion. Garvey assumes this distinction, and treats action and mind as if quite different tests applied to them. I think this assumption needs to be examined in more detail before it is adopted (if it is to be adopted at all).

I also have a concern about the implications for 'insufficient concern' that Garvey draws from his legitimacy framework. Someone who commits a *malum in se*, he says, thereby commits three wrongs (p. 48): a wrong (which exists independently of any legal rule) directly against the victim; a wrong against those of one's fellow citizens who play fair and follow the rule; and a wrong against the state, whose authority has been transgressed. Now, if these are three separate wrongs, what are the relations among them, and how does each of them relate to the question of legitimate punishment? Are each of these wrongs individually sufficient for legitimate punishment, or is one or more of them necessary? For instance, does Garvey think it necessary that, for any case of legitimate punishment, there must be a wrong of types two and

three, as it is only in such a case that there is a 'public wrong'? Theories that start with a concern about authority often emphasise wrongs of type two and three, on the assumption that it is wrongs of those types that the state has a special right to punish. But such theories are thereby at danger of distorting the meaning of punishment by losing the focus on the wrong of type one. It seems essential to criminal censure that the wrong against the individual victim is, at least in serious cases, recognised to be of a far greater seriousness than the wrong against fair play or against the authority of the state. Garvey is not entirely clear on these questions, but it seems to me that he falls into this trap when he discusses the insufficient concern that underpins *mens rea*. He talks about the key issue being insufficient concern *for the law* (pp. 107-8), rather than concern for the victim – or, as on Alexander and Ferzan's formulation, concern for legally protected interests (Alexander and Ferzan 2009). He doesn't consider the possibility that the proper concern against which defendants should be judged is concern for the values that underpin the law – including the values of respect for persons and rights against harm – rather than concern for the law. I diagnose this problem as being a result of his concern with authority and legitimacy. But is it possible to have a view of criminal law that retains the focus on the wrong against the victim, while recognising the role of authority in the state's right to punish? In recent work, I have suggested that we can understand this possibility if we see the state as possessing an 'authority of moral oversight' that makes it legitimate for it to require citizens to answer for wrongs that of type one, and where their identity as wrongs of types two and three need not enter into the content of the censure (Bennett 2019).

Furthermore, while I have some sympathy with the overall picture being developed, Garvey's method sometimes prevents him from tackling the questions of principle that his discussion raises. One great virtue of Garvey's approach is that it raises a deep question about why meeting these culpability conditions is central to liability to punishment. To put it in a slightly different idiom, Garvey's thought could be expressed as the idea that one who meets these culpability conditions thereby forfeits a right against punishment (Wellman 2012). But – if we accept that idiom for present purposes – what is the explanation for the fact that acting in such a way thereby constitutes forfeiture? And how do the *actus reus* and *mens rea* conditions need to be understood in order to do this normative work? Garvey's approach takes it that what we need to do is explain the conditions under which punishment is (self-evidently) not justified. But the questions he has uncovered cannot be satisfactorily addressed until we have a theory that links (1) an account of the right of the state to punish with (2) an account of the conditions under which an agent becomes liable to punishment and then (3) an account of the conditions under which agents escape liability. It is no discredit to Garvey that he has not fully answered these questions; he has given us valuable light for the journey.

References

Larry A. Alexander and Kimberley Kessler Ferzan, *Crime and Culpability: A Theory of Criminal Law* (Cambridge: Cambridge University Press, 2009)

Christian Barry and Gerhard Øverland, 'Do Democratic Societies Have a Right to Do Wrong?' Journal of Social Philosophy 42 (2011), pp. 111–131

Christopher Bennett, 'The Authority of Moral Oversight: On the Legitimacy of Criminal Law,' Legal Theory 25 (2019), pp. 153–177

Jeremy Waldron, *Law and Disagreement* (Oxford: Oxford University Press, 1999)

Christopher Heath Wellman, 'The Rights-Forfeiture Theory of Punishment,' *Ethics* 122 (2012), 371–393