Reply to Critics (2007)

by John Gardner
Professor of Jurisprudence
University of Oxford
http://users.ox.ac.uk/~lawf0081

John Gardner, *Offences and Defences: Selected Essays in the Philosophy of Criminal Law* (Oxford: OUP 2007)
doi: 10.1093/acprof:oso/9780199239351.001.0001

12

Reply to Critics

1. The rule of law and the harm principle

The criminal law is concerned with wrongdoing. Of course, not all wrongs recognized by the law are crimes. There are also torts, breaches of contract, equitable wrongs, etc. On the other hand, all crimes are wrongs recognized by the law. The word 'recognized' here is deliberately ambiguous. It covers two different scenarios. In scenarios of the one type – *malum in se* – the wrong in question is a wrong anyway, quite apart from the law's recognition of it as a wrong. In scenarios of the other type – *malum prohibitum* – it is the law's recognition of the wrong that makes it into a wrong. The distinction is simple, but its application is far from clear-cut. Even murder and rape, paradigmatic *mala in se*, are sharpened up by the law near their borderlines. The law, and especially the criminal law, needs more determinacy in the demarcation of wrongs than is needed, or possible, in ordinary life apart from the law. This means that we should have modest expectations when debating the merits of different ways of demarcating particular wrongs. We should not, for example, expect there to be an independent answer to the question of which kinds of fraud, used to induce a sexual encounter, suffice to turn the inducer into a rapist. This is a line that the law itself has to draw. Within limits, it can reasonably be drawn by different legal systems in different places.

The need for added (and up to a point arbitrary) determinacy in the demarcation of criminal wrongs principally reflects the demands of the ideal known as the rule of law. Under the rule of law, *inter alia*, we must all be given adequate warning of what the

law requires of us, be informed of any charge against us and given an opportunity to answer it, and be judged in court in accordance with the law that we were charged under. In short, we must not be ambushed by the law. Towards the end of 'Rationality and the Rule of Law in Offences against the Person' I outlined one way in which the criminal law can improve its conformity with the rule of law, one way in which it can give us better warning before we blunder into committing a crime. It can do this, I said, by making crimes more 'action-specific'. A more action-specific crime is one that is more specific concerning the possible means of its commission. A crime of causing death by dangerous driving, for example, is more action-specific than a crime of causing death *simpliciter*.

Alan Bogg and John Stanton-Ife protest that the connection I am trying to make between action-specificity and the rule of law is unclear. It is unclear, they write,

> why exactly the rule of law's warnings, where they succeed in demanding advertence, should also require great specificity of action-description. It is not clear that is how this extra requirement would serve the rule of law values of avoiding arbitrariness and human indignity or promoting freedom or autonomy.[1]

I am not sure that I agree with the list of 'rule of law values' here. Sometimes, as we just saw, the rule of law positively requires an element of arbitrariness. To meet the demands of the rule of law, a criminal wrong sometimes needs to be sharply demarcated by arbitrary legal fiat. So perhaps there is a difference of opinion here over how to interpret the ideal of the rule of law.

Bogg's and Stanton-Ife's main challenge, however, lies in their suggestion that action-specificity does not aid conformity with the rule of law even as I interpret it. Action-specificity, they

[1] Bogg & Stanton-Ife, 'Protecting the vulnerable: legality, harm and theft', *Legal Studies* 23 (2003), 402.

say, adds nothing to the law's ability to warn us of our impending violations. My contrary line of thought goes like this. In countless everyday activities (switching on a light, opening a door, riding a bike, cooking a meal...) one may cause another's death. So an offence of causing death *simpliciter* gives one no simple and reliable way of avoiding its commission. Whereas one has a simple and reliable way of avoiding the commission of an offence of causing death by dangerous driving. One may simply give up driving. Bogg and Stanton-Ife counter, not implausibly, that the same protection against ambush by the law could equally be provided by adding an element of mens rea into the crime of causing death *simpliciter*. If it is turned into an offence of knowingly causing death (if, in their words, it is a crime that 'demands advertence') one also has a simple and reliable way of avoiding its commission. One can stop whatever one is doing as soon as one knows that one will thereby cause death. If the mens rea element is already in place, what extra protection against ambush is added by the action-specificity?

The answer is that the action-specificity adds a second and different protection against ambush. It enables one to organize one's life so as to avoid even having to worry about committing certain criminal offences, viz. those that can only be committed while driving. The same goes for criminal offences that can only be committed by threatening, trading, using a telephone, etc. Of course, the argument only applies where the action in question is one that cannot be performed without noticing that one is performing it. But so long as that condition is met, the specific action's inclusion in the definition of the offence makes a distinct contribution to one's freedom. It means that one need not have the existence of the offence in the back of one's mind whenever one does anything at all. One need only have it in the back of one's mind when one is performing the specific action of driving, betting, publishing, trading, etc. The offence of theft under the Theft Act 1968, on which Bogg and Stanton-Ife focus, is a good example. There are many reasons to regret the way in which the

element of 'appropriation' in the definition of theft was gradually emptied of content by the English courts after 1968. But one is that this has gradually deprived each of us of one reliable and simple way we would otherwise have had of steering clear of the law of theft, viz. by not appropriating anything belonging to another. It is some comfort, but not complete comfort, to discover that we can still avoid the law of theft by not being dishonest in anything we do, whether we appropriate anything or not. From the point of view of the rule of law, it is better to have the belt as well as the braces. *Pace* Bogg and Stanton-Ife, neither is superfluous.

The ideal of the rule of law regulates the way in which wrongs can be recognized by the criminal law. Some wrongs, such as betraying a friend, cannot be criminalized consistently with the rule of law. Others, such as interrupting or talking over other people in a conversation, cannot be criminalized consistently with the harm principle. The role of the harm principle in thinking about the criminal law is often exaggerated. Some people think that, thanks to the harm principle, they need to explain the wrongfulness of every criminal wrong in terms of its harmfulness. If there is no harm, there can be no wrong. This is a mistake. Many actions, such as rape and blackmail, are wrongful quite independently of their harmfulness. The role of the harm principle comes later, in regulating whether and how such wrongs are properly recognized by law. This is one theme of 'The Wrongness of Rape'. Arthur Ripstein misrepresents the theme somewhat when he accuses the paper of trying

> to explain the wrongness of such acts by appeal to the harm they do to the 'practice' of sexual autonomy, as if rape would not be wrongful in a society that had not 'adopted' such a practice.[2]

[2] Ripstein, 'Beyond the Harm Principle', *Philosophy and Public Affairs* 34 (2006) 215 at 227.

In fact 'The Wrongness of Rape' does not try to explain the wrongness of rape by an appeal to this harm, or to any other harm. It tries to explain the wrongness of rape by pointing to the rapist's sheer use of another person, a use which may be entirely harmless. The puzzle explored towards the end of the paper was this: How, if at all, can the criminalization of such wrongful but harmless rape be reconciled with the harm principle? The answer suggested in the paper is that the harm principle does not require of each legal prohibition that it (proportionately) prohibits harmful wrongdoing. Rather, it requires of each legal prohibition that it (proportionately) serves to prevent harm. The law may prevent harm by prohibiting a harmless wrong where the non-prohibition of such a harmless wrong would itself be harmful. If harmless rape were not prohibited, that would probably increase the incidence of harmful rape (because it would license men to regard women as less sexually autonomous, and hence encourage their use as sex objects). So far as the harm principle is concerned, such an indirect connection to harm is all that is required to justify a (proportionate) legal intervention. But it is not what makes rape wrongful. Rape is *malum in se* and remains wrong irrespective of the case for legal intervention.

Ripstein misrepresents this line of thought in a second way when he associates it with the view that the harm principle authorizes the use of the law to protect social practices. It is true that the harm principle authorizes the use of the law to protect social practices, where the loss of the practice in question would be harmful. This is how the crime of bigamy passes the test of the harm principle. It protects the social practice of marriage, on the footing that having this practice gives many people an option without which they would be worse off. If there were no such practice, then there should be no such crime. But the crime of rape is different. It attempts to protect sexual autonomy, especially that of women, irrespective of whether there is already a social practice of respecting their sexual autonomy. If there is no such practice there should be, for people are harmed by its

absence, and one task of the law of rape is to help bring it into existence. That is the position represented by 'The Wrongness of Rape'. Nowhere in the paper is it suggested that the practice must already exist for the harmfulness, let alone the wrongfulness, of rape to be established. Rape, the essay makes clear, is wrongful at all times and in all places. Barring times and places where the law is counterproductive, it ought also to be a crime in all times and in all places. At any rate, the harm principle yields no credible objection to its criminalization.

Peter Cane complains that this way of reading the harm principle strips it of its motivation and its power:

> We might say (as Gardner and Shute say in relation to rape) that a society in which the creation of certain risks was not a crime, or in which attempting and contemplating crimes were not themselves crimes, would be (in some sense) a worse society to live in than one in which they were. A worry about this sort of argument, however, is that it depends on the aggregate effect of many such acts, and does not seem to justify coercion of any individual agent. At the same time, classifying such 'diffuse effects' as harm 'seem[s] to reduce the significance of Mill's principle to vanishing point.'[3]

Cane is here forgetting, I think, that the would-be rapist is a would-be wrongdoer. This already picks him out as a suitable person to be threatened with punishment (coerced). It is not the job of the harm principle to pick him out again. The job of the harm principle is to regulate the wider purposes of the law that does the threatening. This law, and indeed every coercive law, must have and fulfil a harm-prevention purpose. The prevention of offence, distress, pain, vice, or indeed further wrongdoing is not sufficient warrant for coercion by law unless by such

[3] Cane, 'Taking Law Seriously: Starting Points of the Hart/Devlin Debate', *Journal of Ethics* 10 (2006), 21 at 33–4. The words quoted by Cane at the end of the passage are from N.E. Simmonds, 'Law and Morality' in E. Craig (ed), *Routledge Encyclopedia of Philosophy* (2004).

coercion the law also prevents harm. The law's threats, moreover, must be in proportion to the harm thereby prevented.

Even as interpreted in 'The Wrongness of Rape' this is a principle of much more than vanishing significance. True, it includes diffuse effects among those that can warrant proportionate coercive interventions. But the effects in question must be harmful, not just bad. A law might be proposed, for example, to criminalize the possession of material that contains photographic depictions of (simulated) torture or rape designed for the sexual arousal of its viewers. Such material is intrinsically abhorrent. All else being equal, the world is a better place if none of it exists, and thus if nobody possesses it. Perhaps a ban on possession can take us a long way towards that result. In which case the ban has a lot to be said in its favour. Nevertheless the ban is warranted, according to the harm principle, only to the extent that it also serves to prevent harm. Not everything that makes the world a better place prevents harm. To prevent harm is to improve someone's prospects in life. Without further inquiry, we cannot tell whether anyone's prospects are going to be improved, let alone proportionately improved, by a ban on possession of this kind of material. We need to know who is supposed to benefit, and by how much, and at what cost, and so on. It is no answer to say that we all benefit just by the bare fact that such material no longer exists. Such material exists right now and many people's prospects are unaffected either way by it. Sure, they are offended or upset to encounter the material or to hear of its existence. But afterwards their lives go on as before; the encounter is upsetting but harmless. To justify a ban on possession under the harm principle we need to find the people who are not so lucky, the ones whose life-prospects are being affected adversely by the non-existence of the ban. Probably there are such people. Probably the ban does indeed pass the test of the harm principle. Probably, for example, it makes a diffuse contribution to improving the way in which women generally are regarded, and hence treated, by men. Possibly it also makes a

less diffuse contribution to protecting some women somewhere from some men who are tempted to act out fantasies that they have seen depicted in pornography. All of this requires detailed empirical study. And that is the point. The point is that the government does not adequately defend the proposed ban until it explains which harms the ban is intended to prevent and why the ban is a proportionate measure to prevent them.

So the harm principle is far from redundant. Yet it cannot do all the work sometimes assigned to it. Satisfying the harm principle is a necessary but not a sufficient condition for the law to use coercion against would-be wrongdoers. The law must also respect the right to freedom of expression, the right to freedom of conscience, the right to sexual freedom, the right to political participation, the right to privacy, etc., each of which restricts the law in the kinds of harms it can legitimately prevent and the ways in which it can legitimately prevent them. Quite possibly – this also requires detailed thought - the mooted law against possession of violent pornography falls at the hurdle of the right to freedom of expression or the right to sexual freedom even if it passes the hurdle of the harm principle. In addition, as already outlined, it has to overcome the obstacles created by the ideal of the rule of law, which regulates the way in which laws are created, promulgated, and maintained. And on top of all that there is the basic principle of rationality by which nobody, not even the law, should take futile or counterproductive actions. If one expected all the work that is done by these various principles to be done by the harm principle, then inevitably one is disappointed by the principle's more humble contribution to settling the limits of legitimate law.

2. Differentiating wrongs

According to 'The Wrongness of Rape', those who are raped and know it tend to be traumatized by the rape because rape is wrong quite apart from the trauma of it. The pure case of rape,

on this view, is the case of rape that goes undetected and hence unexperienced by its victim. Victor Tadros wonders whether the point is supposed to extend beyond rape:

> Similarly, we might seek to identify what is wrong with domestic abuse in the absence of the psychological trauma that is suffered by victims. Perhaps it will be claimed that the 'pure' case of domestic abuse is the case in which the victim is not traumatized by the abuse ... But I think that would be a mistake. The fact that there may be cases of this sort ought not to incline us to think that psychological trauma is not central to what is wrong with domestic abuse.[4]

I do not share Tadros's confidence that we should think of domestic abuse (or recognize it in the criminal law) as a distinct wrong. I think it is better thought of as an insidious pattern of behaviour in which various different wrongs – they may include wrongs of manipulation, coercion, intimidation, exploitation, theft, harassment, humiliation, neglect, assault, battery, torture, rape and ultimately murder – are repeated episodically. Yet one characteristically wrongful aspect of domestic abuse that gives it some unity is the abuser's use of terror (the terror of episodic repetition) with a view to breaking the will of the abused person.[5] Now obviously one cannot use terror against a person without terrifying her. So I agree with Tadros that, in the case of domestic abuse, some of the wrongs of the wrongdoer are partly constituted, even in the purest case, by the reaction of the person who is wronged. The reaction in question, the terror, is a reaction to other actual and threatened wrongs that also form part

[4] Tadros, 'The Distinctiveness of Domestic Abuse' in R.A. Duff and Stuart Green (eds), *Defining Crimes: Essays on the Special Part of the Criminal Law* (2005), 119 at 136.

[5] To be more exact it is a regime of alternating terror and false hope that is designed by the abuser (albeit perhaps only subconsciously) to induce a condition of severe dependency, and hence a radical loss of personal autonomy, in the person abused.

of the pattern of abuse. Rape is different because in the case of rape the reaction under discussion is a reaction to the rape itself, not to some other actual or threatened wrong. This illustrates the point, also made in 'The Wrongness of Rape', that different wrongs are wrong in different ways, by virtue of different lines of argument, and each therefore calls for study in its own right. There is no general formula of wrongdoing.

Terrorizing, unlike rape, is a wrong partly constituted by its result, viz. the terror of the person terrorized. If recognized in criminal law this wrong would count, in criminal lawyer's terms, as a 'result crime'. 'Result crimes' are those that are committed by making a certain causal contribution to a certain outcome. Andrew Ashworth, who has some lingering doubts about result crimes in general, focuses his criticisms of my work on just one sub-category of result crimes. These are constructive crimes: result crimes that one commits by committing another (lesser) crime and thereby contributing to a certain outcome (which may or may not have been foreseen or foreseeable). In 'Rationality and the Rule of Law in Offences Against the Person' I suggested a possible way of thinking about constructive crimes. I said that by committing the lesser crime one 'changes one's normative position' such that a certain outcome that would not otherwise have counted against one now counts against one, and adds to one's crime. Ashworth complains that I do too little to explain this remark. It strikes him as cryptic. What he does manage to glean by way of explanation he finds unsatisfying:

> Gardner sets great store by the fact that people have been 'put on notice' by the law's clear statement ... that this higher liability may be imposed if the outcome is more serious than D anticipated. Fair warning and notice are important components of the rule of law, but they are not capable of supplying substantive moral justification for a

particular head of liability. Giving fair warning of an unfair rule does not turn it into a fair rule.[6]

I agree. I regret that my remark about 'changing one's normative position' was taken (and not only by Ashworth) to be an attempt at offering a 'substantive moral justification' for any constructive criminal liability. When I wrote the words I didn't really mean to justify anything. I only meant to analyze the law's own moral outlook. I meant, in other words, to set out the thing that needs to be justified rather than the justification. I meant to present the law as morally intelligible but not necessarily as morally acceptable. But I can see that I muddied the waters by casually mixing in some alien notes of approval and disapproval.

My 'fair warning' remarks, on the other hand, were meant to have some justificatory importance. They were meant to meet an important objection to constructive criminal liability, viz. that it breaches the rule of law. This struck me and still strikes me as an important objection to constructive criminal liability because the principle *actus non facit reum nisi mens sit rea* is itself a constituent principle of the rule of law. It should be interpreted accordingly. If constructive liability does not breach the rule of law then equally it does not breach the *actus non facit reum* principle, as correctly interpreted. So I thought it worth showing that constructive liability does not breach the rule of law.

But this, of course, is a mainly negative project. If successful it shows that constructive criminal liability does not fall foul of one particular moral objection. It does not provide anything resembling a positive moral case for constructive criminal liability. Indeed nothing in 'Rationality and the Rule of Law' provides or attempts to provide such a positive moral case. Inasmuch as I have tried to provide a positive moral case, I have set it out elsewhere. Elsewhere, to be more exact, I have argued

[6] Ashworth, 'A Change of Normative Position: Determining the Contours of Culpability in Criminal Law', forthcoming, [13].

(against Kant) that acting with bad results (and irrespective of fault) is the basic or elementary type of moral wrongdoing. Some of the relevant considerations are sketched in 'Fletcher on Offences and Defences' and in 'Crime – in Proportion and in Perspective'. However my main work on this topic is not included in this volume because it is not focused principally on the criminal law but rather on the law of torts.[7] It therefore requires some adaptation to fit the criminal law context. I am inclined to think that two main adaptations are required. First, unlike tort liability, criminal liability should not descend on those who commit moral wrongs of the basic kind unless they do it culpably. Justifications and excuses, in other words, should always be available to those who commit criminal wrongs (or else should be anticipated in the very definitions of those criminal wrongs). Second, the rule of law makes stricter demands of the criminal law than it does of the law of torts. This is where the *actus not facit reum* principle, and other requirements of fair warning, come in. I see nothing in these two adaptations that would cast suspicion on constructive criminal liability as such.

One can see in these remarks the explanation for a certain ambivalence that Ashworth detects in my attitude to constructive crimes. Sometimes I talk as if the basic moral wrong of a constructive ('unlawful act') manslaughterer is his unlawful act, the fatal consequence of that act being an aggravating factor. Sometimes, on the other hand, I talk as if the basic moral wrong is the killing, the unlawful act being a precondition for attaching criminal liability to the killing. Ashworth protests:

> [I]t is unclear whether Gardner is able properly to claim that his starting point is that D killed V, when the change of normative position on which he places such weight is surely an assault, and nothing more.

[7] Gardner, 'Obligations and Outcomes in the Law of Torts' in Peter Cane and John Gardner (eds), *Relating to Responsibility* (2001); Gardner, 'The Wrongdoing that Gets Results', *Philosophical Perspectives* 18 (2004), 53.

This, after all, is how he explains other examples – [on Gardner's analysis of a section 47 crime] the assault changes D's normative position so as to render him properly liable for the more serious offence of assault occasioning actual bodily harm.[8]

The answer is that there are two competing criteria of basicness. Morally, the killing is the basic wrong and the prior unlawful act is a precondition for its criminalization. However, in the criminal law's own perspective – with everything seen through the distorting lens of the rule of law - the unlawful act is the basic wrong which changes the defendant's normative position in such a way that its fatal consequences can be counted against him. That, at any rate, is how I see the matter.

Result crimes, as I said, are those that are committed by making a certain causal contribution to a certain outcome. We have just been scratching the surface of the question of whether result crimes should exist. But if result crimes should exist, what causal contribution must one make to the result before one commits them? Many courts and textbook writers seem to expect the answer to this question to belong to the general part of the criminal law, and hence to remain constant across all result crimes. In 'Rationality and the Rule of Law in Offences Against the Person' I challenged this expectation. I suggested that we can draw rational distinctions between wrongdoers not only according to the outcomes of what they do (death, grievous bodily harm, actual bodily harm, etc.) but also according to the causal contribution that they make to those outcomes (occasioning, causing, inflicting, etc.). In 'Causality and Complicity' I developed this idea, and fortified it in the process. I argued that such causal distinctions among wrongs not only *can* be drawn in rationally defensible ways, but to some extent *must* be: as rational beings, responsible for our actions, we cannot live without them. The distinction between principals and

[8] Ashworth, 'A Change of Normative Position', above note 6, [14].

accomplices, in particular, is a causal distinction central to moral life, which is part of rational life. Of course, it does not follow that the distinction need be institutionalized in the criminal law. But it is so institutionalized, and part of the aim of 'Complicity and Causality' was to show that this part of the criminal law has a moral foundation.

One possible reaction to this line of thought is to deny that a crime of complicity is a result crime – in other words to deny that it is a wrong committed by making a causal contribution to something. Responding to 'Complicity and Causality', Lindsay Farmer flirts with this possibility. But he ends up defending only the more modest thesis that a wrong of complicity is not *always* a wrong committed by making a casual contribution to something. He floats a number of possible counterexamples. One is the case in which I help to conceal the wrong of another, e.g. by hiding the body of a murder victim. He says that in this case 'while the action of the accessory does not cause the action of the principal (in Gardner's terms) it undoubtedly "contributes" to it.'[9] I am not sure what to make of the quotation marks around 'contributes' here. More importantly, I am not sure which action Farmer means by 'the action of the principal'. If he means the action of escaping detection or arrest, then concealing the body may certainly contribute, and in a straightforwardly causal way, to the action of the principal. But if he means the original action of killing, then I do not see how the concealer contributes to it at all. Unless, of course, there is something we are not being told. Did the concealer offer to conceal the body beforehand? In that case she may have made a contribution to the killing (as an encourager). In that case, as I explained in the penultimate section of 'Complicity and Causality', the contribution is still

[9] Farmer, 'Complicity beyond Causality: a Comment', *Criminal Law and Philosophy* 1 (2007), 151 at 153.

causal. So I see nothing here to suggest that the contributions made by some accomplices are of a non-causal type.

Another case that strikes Farmer as a counterexample to my thesis is the case in which one is morally complicit in wrongs committed by one's country. For countries, he says,

> do not act through [their citizens] but act (in different ways) in the name of their citizens. The forms of representation and collective action always already imply a notion of political and legal complicity that goes beyond the causal forms discussed by Gardner.[10]

This is the kind of case I mentioned in the final section of 'Complicity and Causality'. Thanks to the work of Christopher Kutz, I was astute to the difficulties of explaining how those human agents who are also constituent members of a collective agent come to be responsible for the actions of that agent.[11] I did not offer a general answer, because it seems to me that there is no general answer. Rather, there are two possibilities. The first is that an individual member of the collectivity is personally responsible as an accomplice because she failed to prevent the collectivity's wrong, failed to impede it, or made some other wrongful causal contribution to its commission. The second is that she is vicariously responsible, meaning that, just because she is a member of the collectivity, she is responsible for what the collectivity does irrespective of whether she made any causal contribution to its being done. It seems to me that between them these possibilities exhaust the possible modes of responsibility for the wrongs of another. That my country represents me and thereby 'acts in my name' is clearly not a *tertium quid*, for it leaves open precisely the question at issue, namely: What is the *basis* of my responsibility, if any, for the acts of those who represent me? Is it a personal responsibility as an accomplice, based on

[10] Ibid, 154-5.
[11] Kutz, *Complicity: Law and Ethics for a Collective Age* (2000).

something I do to contribute to their wrongdoing, or is it a vicarious responsibility based on my membership alone?

Farmer doubts whether this distinction between personal and vicarious responsibility can be sustained. His doubts centre, not on the joint exhaustiveness of the two modes of responsibility, but on their mutual exclusivity. He mentions two sources of doubt. One is that the law might choose a regime of vicarious responsibility as a way of giving legal effect to what is really (i.e. apart from the law) a kind of personal accomplice responsibility. Thus possibly the vicarious responsibility of an employer for his employee's torts is a 'way of institutionalizing in law the fact that [employers are] *always* participating in the wrongs of their subordinates because of their position.'[12] The second is that there seem to be some cases at the borderline that are indeterminate as between personal accomplice responsibility and vicarious responsibility. The position of being vicariously responsible 'can[not] always be distinguished from the position of failing to prevent wrongdoing.'[13] I agree with Farmer on both points. Some of the law's policy goals, in extending liability out beyond principal wrongdoers, can be served either by a regime of personal accomplice responsibility or by a regime of vicarious responsibility. The law therefore has, within limits, a choice of regimes, and different legal systems can reasonably choose differently. But a choice presupposes a distinction. So, for that matter, does a borderline. It is true that, in some cases there is no telling whether one's responsibility for the wrongs of another is personal or vicarious; it could plausibly be analyzed either way. But this does not mean that there is anything amiss with the distinction. Every distinction yields indeterminate cases at the borderline. Lawyers, driven by the ideal of the rule of law, are paid to get such cases decided one way or the other, to force

[12] 'Complicity beyond Causality', above note 9, 153.
[13] Ibid.

them off the borderline. Philosophers have no such duty. Their job is to understand the world, not to change it.

Like Farmer, Tatjana Hörnle doubts whether causality can do the work that I expect of it in explaining the nature of complicity. But her doubts are less far-reaching than Farmer's. In particular, unlike Farmer, she agrees with me that making a causal contribution to the wrong of a principal is a necessary but insufficient condition for being an accomplice. Where she thinks I err is in trying to draw a causal *distinction* between principals and accomplices. She writes:

> Not much is gained by invoking causality to decide whether someone is to be punished as principal or as an accomplice. Causation is a necessary condition for both, but different criteria are needed to discriminate between principals and accomplices.[14]

Hörnle's words 'punished as' draw attention to one respect in which she misreads me. She attributes to me the view, which she thinks a good rule of thumb, 'that accomplices should receive lighter sentences than principals.'[15] This is not my view and I can see nothing in 'Complicity and Causality' that supports it. Often the biggest fish, the one who deserves the most punishment, is an accomplice. I am not thinking about those unusual cases, mentioned by Hörnle, in which an infant or a mentally ill person is used as a tool to carry out a crime. Here, as Hörnle rightly points out, the user is a principal rather than an accomplice, because the infant or mentally ill person is not responsible for her actions. I am thinking, rather, of the everyday situation in which a gangster 'persuades' some petty criminal to carry out a 'hit' in lieu of repaying a debt. Here the gangster, although the prime mover, is an accomplice, while the debtor, although a pawn, is

[14] Hörnle, 'Commentary to "Complicity and Causality"', *Criminal Law and Philosophy* 1 (2007), 143 at 147.
[15] Ibid, 146.

the principal. It seems to me that in this story, all else being equal, the accomplice is the one who should be punished more severely by the criminal law. With this in mind I tend to favour the doctrine in the Accessories and Abettors Act 1861, under which the maximum sentence for the accomplice is identical to that for the principal, leaving the court free to determine which, in any given case, is the more egregious offender. This clearly sets me against Hörnle's rival proposal that principals should be distinguished from accomplices by their supposedly greater degree of control over the commission of the offence. For in the case I just sketched the accomplice, the one who makes the indirect causal contribution to the death through the principal, is the one who exercises greater control over the killing.

In 'Complicity and Causality' I focused on cases of killing, which I treated (for the sake of argument) as a wrong *per se*. I needed to illustrate the contrast between principalship and complicity using a simple wrong that is constituted by making a direct causal contribution to an outcome. I did not mean to suggest that the principalship/complicity distinction is always and only a contrast between direct causal contribution and indirect causal contribution. Some wrongs – such as rape and attempted murder - are not constituted by making a causal contribution to anything. Other wrongs – such as assault occasioning actual bodily harm or permitting premises to be used for the purposes of prostitution - are constituted by making a causal contribution that need not be direct. In the latter cases, some or all of the domain of complicity is inevitably absorbed into the domain of principalship. We have seen this development very starkly in the law of torts, especially in connection with the tort of negligence. As well as making extensive use of vicarious liability, the courts have overcome the absence of complicity liability in tort law by extending the class of principal wrongs to include many indirect causal contributions to damage. This tendency has been less marked in the criminal law, although the criminal courts are so bewildered by causal questions that it is often hard to tell.

Perhaps this is what encourages Farmer to say that often 'the distinction [between principals and accomplices] confuses the issue and might usefully be discarded.'[16] But the importance of a distinction is not a function of its ease of use. The distinction between principals and accomplices might perhaps be excised from the law (e.g. for rule-of-law reasons), but, as 'Complicity and Causality' tries to show, it cannot be excised from life.

3. Justifying wrongdoing

Is there such a thing as justified complicity? The question I have in mind is not a moral or legal one. It is a conceptual one. As I framed the question in 'Complicity and Causality': Is complicity the unjustified thing or the thing that calls for justification? The papers 'In Defence of Defences' and 'Justifications and Reasons' tackle the master question of which this question about complicity is one instance. Is there such a thing as justified wrongdoing? Is wrongdoing the unjustified thing or the thing that calls for justification? A convergence of Benthamite and Kantian influences has privileged the former view. How so? Wrongdoing is action in breach of duty. According to Bentham, one has a duty to perform that action which, on the balance of reasons, one ought to perform. According to Kant, reasons of duty defeat all other reasons and cannot be defeated even by other duties, for duties by their nature cannot conflict. So according to Kant and Bentham alike, once wrongdoing is established, there can be no further question of justification. The question is already closed. In 'In Defence of Defences' this is labelled the 'closure' view of wrongdoing. I argued that the closure view is mistaken. There is nothing contradictory (or even puzzling) in the proposal that, although I acted wrongfully (in breach of my duty), I was amply justified in doing so.

[16] Farmer, above note 9, 154.

Mitchell Berman professes to be agnostic about the closure view. Even so, he has doubts about my case for rejecting it. My case includes the thought that wrongs, even justified ones, leave trails of unfulfilled duty behind them. Justified wrongs are hence suitable occasions for such remedial actions as (depending on the circumstances) reparation, apology, and regret. Berman objects:

> [A]lthough Gardner is surely right that some duties – most especially the duty to show regret – arise from justified actions, ... this is not, I think, because the justified action remains *wrongful*. Rather, it's because the justified action has produced some unfortunate state of affairs (such as injury to another person) about which it is a mark of decency, and perhaps a duty of empathy, to feel regret. Tellingly, this duty to express regret arises even from conduct that (on most accounts) is not wrongful at all – such as nonnegligently causing injury to a negligent victim.[17]

Non-negligently causing injury to a negligent victim can indeed be wrongful, so its regrettability does not help to make Berman's point. Yet he has a good point. Remorse, on the one hand, is apposite only in respect of unjustified and unexcused wrongs. Regret, on the other hand, is an apt reaction to many things that are not wrongs at all (including some things that are not actions at all, but mere happenings). So where in between these attitudes, we may wonder, is the distinctive attitudinal trail that is appropriately left behind by justified wrongs?

An answer was famously offered by Bernard Williams. Williams identified a distinctive attitude: agent-regret, something more than regret, but less than remorse.[18] It is regret about an action that one had a reason, sometimes a duty, not to perform, regret that is not extinguished by the recognition that one's nonperformance was not one's fault. Does this help? The

[17] Berman, 'Justification and Excuse, Law and Morality', *Duke Law Journal* 53 (2004), 1.

[18] Williams, 'Moral Luck', *Proceedings of the Aristotelian Society Supplementary Volume* 50 (1976), 115 at 122–4.

problem is not that no such attitude exists. Clearly it does. The problem is that agent-regret is phenomenologically distinct from simple regret only inasmuch as it adds a judgment on the part of the regretter that (albeit faultlessly) she failed in some respect as an agent. In experiencing agent-regret as a distinct attitude, in other words, the regretter presupposes that the closure view is false. One therefore needs to rely on the falsity of the closure view in order to defend the regretter's experience of agent-regret as a distinct attitude. One cannot at the same time rely on the regretter's experience of agent-regret as a distinct attitude to defend the thesis that the closure view is false. This point can be generalized. In general, it is hard to make a positive case for the falsity of the closure view. The issues at stake are among the deepest in the philosophy of rationality. If one wishes to draw conclusions at this depth, it is hard to find any deeper truths that one can rely upon as premisses. The best one can do is bring out some ways in which the falsity of the closure view chimes with ordinary moral experience. One is therefore always vulnerable to the response - characteristic of tidy-minded Benthamites and Kantians - that ordinary moral experience is shot through with irrationality. This explains the sense that some may have, on reading 'In Defence of Defences' and 'Justifications and Reasons', that neither paper offers much in the way of positive argument. Each proceeds mainly by setting out a view and showing how it hangs together with itself, as well as resonating with various aspects of ordinary moral experience (some of which are reflected in English criminal law).

Responding to 'Justifications and Reasons', Alan Norrie reasserts the closure view. My response to that view, he says,

> is to argue that the prima facie reason is not merely provisional or evidential in character, so that even if it is defeated it is not undermined or cancelled. Yet how can this be the case? If it is defeated then it is indeed undermined, and, more strongly, cancelled. In the relevant

sense, the prima facie reason did not turn out to designate an actual wrong, but one that was only putative.[19]

'Justifications and Reasons' already showed how it *can* be the case that a prima facie reason is not cancelled, by showing that there is an intelligible sense of 'prima facie' other than the 'provisional or evidential' one preferred by Norrie. But is this intelligible sense the *relevant* sense when we come to explain the logic of justification? The question is complex. One thing that I attempted to do in 'Justifications and Reasons' and again in 'Fletcher on Offences and Defences' was to show how the closure view gets its false allure. Some of its allure, in my view, just comes of wishful thinking (if only life were simpler; if only there were no tragedy; if only there were never anything to rue in a justified action). But some of its allure comes of the fact that it contains a grain of truth. In the justificatory scenario there is indeed some cancellation of reasons going on. A justification for wrongdoing, as I explained it, is a cancelling permission. It does not cancel the reason not to commit the wrong – it only defeats that reason - but it does cancel the reason's mandatoriness. This proposal contradicts the closure view but preserves an important aspect of it, and this helps us to see why Norrie might be resistant to its abandonment. (Although, as I joked in a footnote to 'In Defence of Defences', it is curious to find Norrie's resistance to the abandonment of the Kantian closure view presented as part of his critique of Kantian thinking in criminal law theory!)

Norrie also has a lawyer's objection to my abandonment of the closure view. He objects to the 'contingency' or 'fluidity' of the resulting distinction between offences and defences, which means that the same legal doctrine (e.g. self-defence, consent, necessity) or the same case (e.g. that of a mercy killer or a soldier killing in battle) could be placed on opposite sides of the line by different legal systems, or indeed by the same legal system at

[19] Norrie, *Punishment, Responsibility, and Justice* (2000), 153.

different times or for different purposes or in different contexts. This is similar to Farmer's objection to my use of the distinction between personal and vicarious responsibility. Both Norrie and Farmer exhibit the traditional lawyer's interest in how to classify particular doctrines and cases, and they seem to expect philosophical analysis to settle such matters in advance of the law. But I have more modest ambitions. It is true that I often reveal my views about which cases fall on which side of the line in morality apart from the law. Nevertheless the law, as I have always tried to emphasise, faces many constraints, but also many choices, in settling how morality is to be institutionalised. I regard it as a merit of 'Justifications and Reasons', not a failing, that it explains what the difference is between denying an offence and offering a (justificatory) defence, and what makes that difference salient for rational beings, without attempting to pre-empt the decisions of particular legal systems about whether and where to draw the line in particular legal contexts.

4. From justification to excuse

Although I have not made it my main mission to classify them, I have frequently used legal doctrines and cases from particular legal systems to exemplify the categories under investigation. This exposes me to an oft-repeated line of criticism. The criticism is that the doctrines and cases are not examples of what I take them to be examples of. Peter Westen, for instance, objects to my use of duress cases in 'The Gist of Excuses':

> Gardner's problem with duress is this: Gardner argues that excuse comes into play only with respect to offenses that are unjustified; yet, given Gardner's definition of 'justification', offenses committed under

duress – as opposed to offenses committed under mistaken duress – ought to be regarded as offenses that are *justified*.[20]

We should leave aside for a moment cases of mistaken duress, which raise classificatory problems of their own. Is Westen right to think that, where the threat is real and the person threatened does not misperceive it, a duress defence is always a justificatory defence? There is no doubt that sometimes it is. Sometimes the person threatened was right, all things considered, to commit the wrong that he committed in response to the threat. But sometimes he was not so lucky. His fear of the threat's being carried out, itself justified, drove him to commit an unjustified wrong. This is not a case of mistake. The person threatened did not misinterpret the threat or misjudge the likelihood of its being carried out. He merely overreacted; he let his fear get the better of him. Usually the overreaction was exactly what the threatener banked on. The threatener selected a vulnerability of the person threatened: his love of his children, his loathing of rats, or just his anxiety about his own fate (excessive preoccupation with which is a widespread human limitation and therefore the simplest one to exploit without special knowledge of the person one is threatening). The person threatened exhibited this limitation in his action of surrendering to the threat. Then the allowing of the duress defence is, as the courts sometimes put it, a 'concession to human frailty'. The fortitude exhibited is suboptimal but not below acceptable limits. In this case the force of the duress defence is excusatory rather than justificatory.

What leads Westen to deny the existence of this class of cases (cases of excusatory duress without mistake of fact)? He writes:

A person who is correct in thinking that A's threat is genuine is a person who acts both in his mind and in actuality upon a balance

[20] Westen, 'An Attitudinal Theory of Excuse', *Law and Philosophy* 25 (2006), 289 at 348.

between self-interest and the interests of others that society regards as acceptable, all things considered – which is precisely the sort of person whom Gardner has said is 'justified'.[21]

The problem here is that 'act[ing] upon a balance ... that society regards as acceptable' is ambiguous, even when we restrict our attention to the actuality rather than A's perception. Even when she in no way misperceives her situation, it is possible for the way that A balances two considerations to live up to the relevant standard for thought without her action on the strength of that thought living up to the relevant standard for action. Of course there is something of the tragic about this possibility. That one thinks acceptably about how to act, and acts on the strength of that thinking, does not guarantee that one acts acceptably. What it does guarantee is that one is not at fault in acting unacceptably. For while one lacks a justification, one has a full excuse. Westen accepts this much in cases of misperception, and so classifies a reasonably mistaken belief in duress as excusatory, even though duress itself would be for him justificatory. His error lies in failing to see that misperceptions are not the only distortions in practical rationality that can drive a wedge between what one acceptably thinks about what to do, and what one acceptably does. Strong emotions too can supply the relevant distortion, for reasonable practical thought is characterised and constituted by reasonable affect as well as by reasonable cognition.

Hamish Stewart and Victor Tadros both take my account of excuses to task on the very point that Westen concedes. They both hold that a reasonable mistake as to an element of a justificatory defence preserves the justificatory character of the defence, rather than turning it into an excuse. I will come to Stewart's line of thought in section 7 below. Here let me focus on Tadros's criticism, which takes the form of a *reductio*:

[21] Ibid, 350.

> [On Gardner's view] it is difficult to see how one could ever be justified in taking a risk where things do not turn out as one hoped. If even reasonably held beliefs cannot ground a justification defence where the facts turn out unexpectedly to be different from those that D believes them to be, where are we to stop in our analysis of what the facts are? For it is always the case, where one takes a risk and things turn out badly, that there is some further fact about the world that one could have known that would make that risk not worth taking. This is true even if determinism is false: there are facts about what will happen in the future as well as facts about the world as it is now.[22]

Here Tadros raises another of the most difficult problems in the theory of reasons. Reasons are facts. If I make a mistake of fact (e.g. if I think I am being attacked when I am not) then I do not have the reason to act that, in the grip of that mistake, I take myself to have (e.g. to defend myself). I may have ample reason to believe that I have that reason to act, but this can only furnish me with an excuse, not a justification, for so acting. This is the view advanced in 'Justifications and Reasons' and assumed in 'The Gist of Excuses'. Tadros challenges it by asking us to shift from the present tense to the future tense, thinking about cases where one acts on the strength of beliefs about what is yet to take place. Suppose D thinks that, in all likelihood, he is going to be attacked later. On one view, the probabilist view, the risk that D is going to be attacked later is a fact in its own right and is capable of counting as a reason for D to run away. There are risks in the world, and not only in our incomplete *ex ante* beliefs about the world. On another view, the actualist view, D is either going to be attacked or he is not. D has a reason to run away only if he is indeed going to be attacked. If D is not going to be attacked, then the most D can possibly have is a reasonable *belief* that he has a reason to run away. Tadros thinks that my view of the present-tense cases commits me to taking an actualist view in the future-tense cases, such that all references to risks in practical argument

[22] Tadros, *Criminal Responsibility* (2005), 286.

should be read as references to epistemic uncertainties (i.e. as a shorthand reference to the existence of reasons to believe in the existence of reasons to act), and hence should be read as excusatory rather than justificatory. He takes this to be a *reductio* of my position on the present-tense cases.

To this objection I have various reactions. First, I am not convinced that my view on the present-tense cases commits me to the actualist view on the future-tense cases. I do not find the question of whether there are facts about what will happen in the future as easy to answer as Tadros seems to find it. Second, I do not find the advertised implication of the actualist view (viz. that reasonably mistaken gambles on the future can only excuse, and not justify, actions taken on the strength of them) as unpalatable as Tadros seems to expect. To the extent that I am sympathetic to the actualist view,[23] this is partly because of this implication, not in spite of it. Third, thanks to the existence of what (in 'Fletcher on Offences and Defences') I called 'fault-anticipating wrongs', what would otherwise excuse one's action is sometimes sufficient to make it the case that one commits no wrong in the first place, so that one has nothing to excuse. Many wrongs are defined in terms of recklessness or negligence and hence build a sensitivity to risk into their very definitions. More generally, many rules exist to regulate risk-taking, providing more concrete practical guidance on how to cope with epistemic uncertainty. In this way the rationality of belief can be indirectly relevant to the rationality of action.

Finally, it is salutary to note how much trouble Tadros himself has in escaping from an epistemic conceptualization when talking about the present-tense cases. He returns to a case

[23] In a paper called 'The Mysterious Case of the Reasonable Person', *University of Toronto Law Journal* 51 (2001), 273, I confessed to actualist sympathies. Tadros relies on this confession to corroborate his charges.

devised by Bernard Williams.[24] I have in front of me a glass of petrol. I have ample reason to believe it to be a glass of gin. Do I have any reason to drink it? Williams and I agree that the answer is no. There may be a complete and adequate rational explanation for my having drunk the contents of the glass if that is what I do. But the explanation includes (because there is) no reason for my having drunk the contents. It cites only the ample reasons for me to have *believed* that I had reason to drink the contents. Tadros says, by contrast, that there was also a reason for my having drunk the contents. His argument goes like this:

> I am still motivated by a fact about the world in this case: the fact that it appears [to be gin in my glass]. That appearance does not always correspond to reality does not deny the status of that latter as a fact about the world. It may be objected to this that ... talking of the fact of appearance collapses into talk of belief. But that is not the case. The fact of appearance is to be distinguished from belief by virtue of the fact that appearance is appearance to any believer, not just this [one].[25]

Tadros cannot literally mean that appearance is appearance to *any* believer. Believers with unreasonable beliefs are distinguished by their failure to see things as they appear to others to be. Who are these others? They are those with reasonable beliefs. So the standard of appearance is none other than the standard of reasonable belief, i.e. the standard of those with beliefs that are held for adequate reason. The relevant standard, in other words, remains epistemic. In a sense it is true, as Tadros adds, that 'appearance is objective whereas belief is subjective.' But the objective standard in question is an objective standard applicable to belief, viz. the standard of being a belief held for adequate reasons. This being so, Tadros has not identified a way to escape from my conclusion about the present-tense cases, whatever we

[24] Williams. 'Internal and External Reasons' in Ross Harrison (ed), *Rational Action* (1979) 17 at 18.
[25] Tadros, *Criminal Responsibility*, 284.

may want to say about the future-tense cases. Acting on what Tadros calls the appearances is reasonable action not in the literal sense of being justified action but in the elliptical sense of being action on the strength of reasonable beliefs, which is excused.

At one point Tadros seems to mount a more radical challenge to this thesis. He seems to deny that there is any difference between having adequate reason to believe that one has a reason to act and actually having that reason to act. 'It is difficult to see,' he says of Williams' petrol-drinking case, 'how to drive a wedge between having good reason to believe that the stuff is desirable to drink and having good reason to drink it.' This remark comes as a surprise. It seems to undermine many of Tadros's previous thoughts, which presuppose that there is logical space for people to make reasonable mistakes as to the existence of reasons for acting. If it turns out that there is no such logical space, because one always has all the reasons for acting that one reasonably believes oneself to have, then the question of whether reasonable mistakes regarding justificatory features of one's actions are excusatory does not arise, and there is no point in Tadros debating the answer with me. It seems to me that Tadros does not really mean to pursue this nuclear option. Perhaps all that he is trying to say is that Williams and I have exaggerated the *importance* of the contrast, within the rational explanation of human action, between reasons to believe in reasons to act and reasons to act themselves. If that is what he means then my answer is that the importance of this contrast would be very hard to exaggerate. It is absolutely central to a proper understanding of our predicament as rational beings.

5. Excuses and incapacities

Tadros's critique of my account of excuses continues with an attempt to re-establish a relationship between excuses and incapacities. In 'The Gist of Excuses' I argued that an incapacity to act better than one does is no excuse, and does not constitute

the gist of any excuse. Central to my argument was the thesis that courage and honesty and other moral virtues are none other than capacities to act in certain ways, viz. in the ways that exhibit those virtues. This has two implications. First, nobody ever has the capacity to exhibit more courage (etc.) than they do exhibit. So one's capacity to exhibit courage cannot possibly be used as a standard against which to judge one's exhibitions of courage. Since it cannot be used as a standard, it obviously cannot be used as an excusatory standard. Second, anyone who lacks the capacity to exhibit courage also lacks courage. So the assertion of an incapacity to act more courageously than one did, inasmuch as we can make sense of it, is not exculpating but inculpating. In making this assertion one is confessing to one's inadequate moral character, confirming one's fault. Since an excuse is one kind of denial of fault, such an assertion cannot be excusatory.

Tadros resists both of these lines of thought. Against the first he raises a complex and fascinating objection. At its heart lies the following supposed counterexample:

> I am in the pub with a few friends. We have time for two drinks. ... If I buy neither round of drinks, that may show that I am not generous, particularly if I am wealthier than my friends. But in not buying the first round of drinks I show nothing about my capacity for virtue. This shows that I may have the capacity for generosity but not exercise it. Although I have the inclination to be virtuous, I also allow other people to be virtuous by buying a round of drinks. In doing so I show that I possess a different virtue, sensitivity.[26]

The lesson of the example, says Tadros, can be generalized across the moral virtues, including courage: 'in failing to manifest a virtue at an appropriate time, I do not necessarily show that I do not possess that virtue, or the capacity for [that] virtue.'[27]

[26] Ibid, 312.
[27] Ibid, 313.

This, however, cannot conceivably be the lesson of Tadros's example. Why not? Because, according to Tadros himself, the occasion of the first round of drinks was not 'an appropriate time' for me to exhibit generosity. Given that I was about to buy the second round of drinks, buying the first would have been, says Tadros, 'beyond the realm of generosity':

> Not only am I not required to buy both rounds of drinks, I would also not be generous in doing so. This is what Aristotle means when he says that in acting virtuously we must avoid both excess and deficiency: the doctrine of the mean.[28]

The Aristotelian doctrine is correct. Insensitive (indiscriminate) generosity is not virtuous generosity. Rather it is generosity to a fault, a distinct moral vice. Possibly I have an unexercised capacity for generosity-to-a-fault when, in Tadros's example, I do not buy the first round of drinks but only the second. But I cannot have an unexercised capacity for *virtuous* generosity when I do not buy the first round of drinks but only the second. This is because a capacity counts as unexercised only if there is also an opportunity for its exercise. And *ex hypothesi* there is no such opportunity in Tadros's example. There was no occasion for me to do any better, from the point of view of virtuous generosity, than buy the one round of drinks that indeed I did buy. So Tadros fails to show, by this example, that anyone ever has the capacity to exhibit any moral virtue beyond the moral virtue that he does exhibit. He fails to show any gap between virtue and the capacity for it that would enable us to judge people's exhibitions of virtue relative to their capacity for such exhibitions.

Tadros resists my second line of thought in a very different way. He criticises my failure to note that there are some moral vices with which the criminal law should not concern itself:

[28] Ibid, 312.

> [T]hose that show themselves only to have the vice of cowardice, as opposed to showing that they have an insufficient regard for the criminal law and the values that it enshrines, ought to be entitled to a defence of duress. Those who could not have resisted had they held the criminal law in sufficiently high regard have displayed none of the vices required for criminal liability.[29]

I am taken aback by the suggestion that 'insufficient regard for the criminal law' ought to be thought of as a vice, even by the criminal law itself. It is one of the hallmarks of fascist law that it judges people not by their attitudes to authentically valuable things but by their attitudes to the law itself (which may or may not be valuable depending on its content, and which is already made distinctly less valuable by the very fact that it contains this narcissistic insistence on respect for itself). Perhaps Tadros means to emphasize not the criminal law as such, but only 'the values that it enshrines', as proper objects of regard; and perhaps he means to restrict attention to cases in which these values are valid (i.e. are not erroneously espoused). I certainly hope he does not mean to go further in making a false idol of the law.

But be all that as it may, I am not sure how Tadros's criticism in this passage is supposed to bite against my views. In 'The Gist of Excuses', I did not discuss the question of which failings of character should be recognized as such by the criminal law. I happen to think that the criminal law could legitimately take an interest in any of them, but that was not my point. My point was rather that once the criminal law has recognized that something is a failing of character (as it does with cowardice in the face of threats, intemperance in the face of taunts, inattentiveness in the face of risks, and various other shortcomings) it cannot allow people to argue that they lacked the capacity to avoid exhibiting that very same failing. For this lack of capacity *is* their failing. This applies as much to the failings to which Tadros would like

[29] Ibid, 317.

to see the criminal law attend (cruelty, callousness, prejudice, etc.) as it does to those failings that I discussed.

Barry Mitchell shares Tadros's sense that there must be a logical gap between virtue and the capacity to exhibit it. He accuses me of harbouring 'a singularly simplistic view of capacity according to which a person who has the capacity to behave in a particular way in the circumstances will behave in that way.'[30] This view is indeed simplistic, but the simplification is Mitchell's, not mine. I advanced the much narrower proposition that *some* capacities – virtues of character – are capacities that one does not possess unless one possesses the matching propensities. They are capacities that exist only in their exercise. That proposition does not commit me to any more general 'view of capacity'. But it seems that Mitchell would be unhappy with it all the same. For he uses my illustration of the narrower proposition that I advanced to bring out what he holds to be simplistic about the broader proposition that I did not advance. I had written:

> If one sees the world through genuinely courageous eyes one does not see the danger to oneself the way that more cowardly people see it, as a threat, but rather as a challenge, something which, up to a point, one inclines towards rather than away from.[31]

Mitchell has the following complaint:

> The truth is that we simply do not fully understand why some people are braver than others. Moreover, people may behave courageously when faced with a particular set of circumstances simply because at the critical moment they focused more on the need to (seek to) prevent some other harm than on the danger to himself. We do not know how far they [their?] focus on the harm to others rather than themselves is a matter of chance rather than a manifestation of some aspect of their

[30] Mitchell, 'The Minimum Culpability for Criminal Homicide', *European Journal of Crime, Criminal Law and Criminal Justice* 9 (2001), 193 at 196.
[31] Above, 173.

personality which enables them to be brave. Another person may, on that occasion have focused instead on the danger to himself, but on a subsequent occasion would have shown the same degree of fortitude as the courageous person. The second person does not necessarily have a lesser capacity for being courageous.[32]

This criticism seems to be at cross-purposes with the passage that it is supposed to criticize. I said nothing about why people are brave, or what brings it about that they react bravely today and without any bravery tomorrow, or similar questions of aetiology. I have no views on these matters which, as Mitchell points out, require empirical investigation.[33] I discussed only the conceptual question of *what counts as* bravery. This question must be answered first. One must always know what counts as an X before one can discuss any empirical questions about Xs.

Mitchell must have his own idea of what counts as courage, but I am not sure what it is. Inasmuch as I can work it out, it seems to have a starting point in common with my own. He seems to agree with me that the person who 'focused on the danger to himself' was not being courageous on the occasion on which he did so; only 'on a subsequent occasion' did he show 'the same degree of fortitude [=courage?] as the courageous person.' So a person is courageous only when he does not 'focus on the danger to himself'. One's courage (as I put it) depends on how one sees the world. But does Mitchell mean to add the extra qualification (against me) that such a focus does not show one to be courageous if it is 'a matter of chance rather than a manifestation of some aspect of [one's] personality'? Or is this extra variable relevant only to determining whether one had the 'capacity for being courageous' when one was not? If the latter, which way does it cut? Does one have the capacity to be

[32] Ibid, 196.

[33] Mitchell complains (ibid, 196) that in my paper 'there is a conspicuous absence of any attempt to address the scientific data available.'

courageous only if one's focus was or would have been a 'manifestation of some aspect of [one's] personality' or only if it was or would have been a matter of chance? If the former, then what aspect of one's personality are we looking for other than the courage itself? For it must clearly be something other than the courage to save us from vicious circularity; it must a further trait that 'enables [one] to be' courageous rather than simply constituting one's courage. But what could this further trait be? Could this trait (unlike one's courage) be a capacity that exists only in its exercise? If not – if it could be a capacity without propensity - then in what sense would it qualify as an 'aspect of one's personality'? And if this trait is a capacity that exists only in its exercise, then what is so wrong with the view that courage itself is also the same kind of capacity? All of these are obscurities that prevent us from crystallizing, and hence evaluating, Mitchell's rival account of courage (if rival it be).

Jeremy Horder is another critic who thinks that I underestimate the significance of incapacity in setting excusatory standards in the criminal law. At least he rejects my rival explanation for the variability of excusatory standards as between defendants. Some excusatory statements that are commonly interpreted as capacity-invoking ('she's a kid', 'he's a beginner') are better interpreted, in my view, as role-invoking. In a case of provocation, for example, the excusatory salience of being an adolescent is not that adolescents *cannot* be as even-tempered as adults but that they *should not* be. One should act one's age. Adult temperance would not befit an adolescent. This view is sketched in 'The Gist of Excuses' and elaborated, with more examples, in 'Provocation and Pluralism'. Horder objects:

> The problem with this view is that the very fact that I need an excuse, because I have engaged in wrongdoing, shows that I have failed in my role ... [T]he more serious the wrongdoing the more plausible it will be to say that one failed to meet the preconditions for successful role-

fulfilment, and so the normative expectations attached to one's role are no longer relevant.[34]

Horder seems to be making two distinct objections here. The first is *ad hominem*. I (Gardner) am the one who holds that wrongdoing should be understood as failure in a role. Suppose this is true, says Horder. How, having failed in a role, and hence having fallen short of the expectations for that role, can one still meet the expectations of that role and hence enjoy a role-specific excuse? I explained how this can be so in 'The Gist of Excuses', in a passage that is earlier quoted by Horder. I pointed out that there are two sets of standards (and hence two sets of expectations) that go to define a role. There are standards of success/failure and standards of fitness/unfitness. One can fail in a role without being unfit for it. The role's excusatory standards are among its standards of fitness/unfitness. So one can meet these excusatory standards even though one has committed a wrong and hence failed in the role. Which is just as well, since one needs an excuse precisely because of one's failure.

Horder's second point is that (serious) wrongdoing should not, in any case, be understood as failure in a role. Instead it should be understood as failure at an earlier hurdle, before one even gets to occupy, and hence be judged by the standards of, any role. What should we make of this idea? A wrong is no more and no less than a breach of duty. There are clearly many duties that are incidents of friendship, parenthood, citizenship, etc. One is subject to them only if one is in the role of friend, parent, citizen etc. It follows that there are some wrongs that one commits only as an occupant of these roles. Perhaps, according to Horder, these are not *serious* wrongs? Horder does not seem to honour this 'seriousness' restriction in the rest of his argument. He goes on to say that 'avoid[ing] wrongdoing' – serious or not – is 'avoid[ing] violating the conditions in which we and others can

[34] *Excusing Crime* (Oxford 2004), 115.

securely and confidently set about right-doing', this in turn being the point at which we assume our various roles.[35]

To make sense of this proposal, as Horder points out, we need to be able to distinguish between 'doing ... what is right' and 'avoiding ... what is wrong.'[36] But this distinction is obscure and Horder's examples do not help us to unpack it.[37] Avoiding wrongdoing is the same as rightdoing. The rightdoer fulfils her duty; the wrongdoer breaches it. There is no *tertium quid*. Of course there are many valuable things that one might do beyond the call of duty. Some of these may make one particularly successful in one's role. But one does not need to do any of these things to be an adequate occupant of one's role, i.e. to avoid failure in it. To avoid failure in one's role one need only perform its duties. So Horder gives us no credible reason to think that wrongdoing, serious or otherwise, is not failure in a role. Consequently he gives us to reason to think that excusatory standards cannot likewise be role-specific standards.

6. Adapting excuses to the criminal law

I already mentioned and doubted Tadros's view that the criminal law (in forging its excusatory doctrines) should not be concerned with deficiencies of character *tout court*, but should limit its attention to a subset of deficiencies which evince 'insufficient regard for the criminal law and the values that it enshrines'. William Wilson also argues that my view of excuses needs to be

[35] Ibid, 116.
[36] Ibid, 116.
[37] He cites some remarks by Raz in *The Authority of Law* (1979), at 224, as helping to unpack the distinction and show its import. But the cited passage from Raz deals with a different distinction, viz. the distinction between avoiding evil and doing further good beyond avoiding evil. Raz's distinction is axiological whereas Horder's is deontic.

modified before it is suitable for use in the criminal law. Focusing on the example of duress, he writes:

> [C]oercion may excuse in different ways. Sometimes ... reasonableness of reaction clearly grounds the defence and in such cases the presence or absence of subversion of the 'will' must be of peripheral concern. ... In other cases it is equally clear that the defence may be grounded not in the reasonableness of the reaction but simply in the unfairness of expecting conformity [to the criminal law]. Here the defendant may be acting upon fear alone where he just 'does it' regardless of the consequences because he is too terrified to do anything but obey. ... His claim to be excused is that we cannot always be expected to live up even to the standards which we set for ourselves. The criminal law is in place to punish us for our self-interestedness and the lack of concern we show to others not the fortuitous falls from grace which attend our unpracticed responses to crisis.[38]

I have a few preliminary comments about this passage. First, I am not sure why Wilson puts 'will' in inverted commas in the second quoted sentence, but with the inverted commas excised there is no reason to think that a focus on reasonableness of reaction marginalizes questions about the subversion (or overbearing) of the will. One's will is overborne, in the relevant sense, when someone else, by issuing a conditional threat, intentionally creates a reason (or what one reasonably takes to be a reason) for one to ϕ, and one ϕs for that reason, and one could not reasonably have been expected to do otherwise than to ϕ for that reason. The reasonableness of one's reaction is, in short, a necessary condition of the overbearing of one's will. Second, I am not sure what is supposed to be the extra significance, in the penultimate quoted sentence, of our falling short 'even [of] the standards which we set for ourselves'. Since it is part of being human that one always aims to be justified in what one does, the

[38] Wilson, 'The Filtering Role of Crisis in the Constitution of Criminal Excuses', *Canadian Journal of Law and Jurisprudence* 17 (2004), 387 at 405–6.

standards that we set for ourselves are always higher than excusatory standards. They are justificatory standards. One cannot be expected invariably to live up to these higher justificatory standards, but still one is held to (and holds oneself to) standards of character and skill. This is exactly what 'The Gist of Excuses' points out. So here Wilson seems to echo, not challenge, my views. Third, the final quoted sentence can be interpreted (and is presented by Wilson in his footnotes) as echoing Tadros's view that only some and not all standards of character should be institutionalized in the criminal law. But this point bears on what should qualify as a reasonable reaction for the criminal law's purposes, not on whether reasonableness of reaction is the right thing for the criminal law to care about. So this remark seems orthogonal to Wilson's main complaint, which is that reasonableness of reaction is not the right thing for the criminal law to care about, or rather not the only one.

So what is the other right thing for the criminal law to care about? While in some cases it should admittedly care about the reasonableness or otherwise of D's reaction, according to Wilson, in other cases the criminal law should care about the fairness or otherwise of expecting D to conform to its norms. This, for Wilson, is the gist of at least some criminal-law excuses. By way of example he mentions automatism and involuntary intoxication as well as duress. Now automatism is not an excuse but a denial of responsibility. And involuntary intoxication is not a defence at all (excusatory or otherwise) but merely a possible way of supporting one's denial of mens rea.[39] But never mind these quibbles about particular legal doctrines. Whichever legal doctrines it is supposed to apply to and illuminate, Wilson's proposal is unhelpful on different and deeper grounds.

[39] Strictly speaking, the involuntary intoxication doctrine is merely an exception to the general doctrine according to which evidence of intoxication cannot be used to support a denial of mens rea for crimes of 'basic intent' (and arguably may even be used to help establish such a mens rea).

It is unfair to expect conformity to criminal-law norms in a wide variety of circumstances and for a wide variety of reasons. It is unfair to expect conformity to a criminal-law norm when, in violation of the rule of law, no fair warning of its potential application is given. It is unfair to expect conformity to a criminal-law norm when the defendant is not responsible for her actions (e.g. is an infant or severely mentally ill or is sleepwalking). It is also unfair to expect conformity with a criminal law norm when breach of it is justified or excused. In each of these cases there is a different explanation of why it is unfair to expect conformity with the norm. In the last case, it is unfair to expect conformity with the norm because the breach of it is excused. So it cannot possibly be the case, *pace* Wilson, that the breach is excused because it is unfair to expect conformity with the norm. An explanation of the nature of criminal excuses must explain not only *that* it would be unfair of the criminal law to expect conformity with the norm when non-conformity is excused, but also *why*: what it is about an excuse that *makes* it unfair to expect such conformity. In all of the examples given by Wilson the explanation is the same. It is that the defendant's reaction to the crisis before him was within the bounds of reason. In the example given in the passage – of someone who gives in to a threatener regardless of the consequences – the explanation is no different. Was it reasonable for D to be this terrified of the threatener, so terrified that he obeyed in blind panic and ended up doing an unjustified (=unreasonable) thing? Was he justified in his terror, even though not in his terrified action?

7. *The hierarchy of defences*

As these remarks make clear, I subscribe to what Douglas Husak calls 'the logical priority thesis', according to which one needs to understand what a justification is in order to understand what an excuse is. Husak criticizes this thesis, and my reliance on it, by criticizing a thesis that he takes to be entailed by it:

> All theorists who hold the [logical] priority thesis are committed to the claim that justifications and excuses are mutually exclusive. Excuses are defined so that if a defendant has an excuse, he cannot also have a justification. This definition, of course, entails (as its contrapositive) [that] if a defendant has a justification, he cannot also have an excuse.[40]

Husak offers three main objections to this claim of mutual exclusivity. The first is that one may, on one and the same occasion, act with both justification and excuse. One may act in self-defence, for example, even though one also acted under grave provocation. The second, which turns out to be a variant of the first, is that there may be partial justifications and partial excuses, which might in principle add up, on a particular occasion, to yield a more complete defence. Using excessive force in self-defence is always partially justified (because *ex hypothesi* one is justified in using *some* of the force that one uses) but beyond that point one needs to carve out an excuse (e.g. that one was reasonable in one's misjudgment of the amount of force that was called for). The third, well-illustrated by the same example, is that there could be hybrid defences, 'shar[ing] characteristics of both justification and excuse.'[41]

I agree with Husak on all three points and follow him in thinking that, for these reasons among others, justifications and excuses cannot be mutually exclusive in the sense of never being available in tandem in respect of one and the same wrong. What I do not understand is why my endorsement of the logical priority thesis is supposed to commit me to the opposite view. It is true that in 'Justifications and Reasons' I characterized excused actions as unjustified actions on the strength of justified beliefs, emotions, attitudes, etc. The reference to 'unjustified actions' suggests mutual exclusivity. The characterization is, however, a

[40] Husak, 'On the Supposed Priority of Justification to Excuse', *Law and Philosophy* 24 (2005), 557 at 560.
[41] Ibid, 583.

bit rough and ready. All I meant was that excuses are offered *on the footing* that the action is unjustified. The action may in reality be fully justified. Yet one concedes otherwise (at least for the sake of argument) when one offers an excuse.

Thus – to develop an example of Husak's - when A says 'excuse me' as he squeezes past B to exit the bus, A need not be denying that his action is justified. A may well believe that squeezing past is fully justified and this belief may well be true. Quite possibly B is blocking A's way, time to exit the bus is short, A has an urgent reason to exit, B's reaction time is slow, and squeezing past is therefore the right thing to do all-things-considered. Nevertheless, by asking B to excuse him, A is conceding *arguendo* that he is incompletely justified. Making this concession is a convention of good manners that, by promptly revealing one's willingness to admit error, helps to avoid triggering conflict. If B is a person of good manners he will promptly reciprocate by saying something like 'no, excuse *me*!' and then the incident will be over. This example bears out Husak's view about the mutual exclusivity of justifications and excuses. A may have a valid excuse to offer even though he also has a valid justification that he declines to offer. He may equally have a valid excuse to make up for any incompleteness in his justification. He may also have a hybrid defence that is part justification and part excuse. But none of this is inconsistent with my version of the logical priority thesis, according to which making an excuse is asserting that one's action, even if unjustified, was taken on the strength of justified beliefs, emotions, attitudes, etc. This 'even if' formulation makes clear that, as between justification and excuse, there is no mutual exclusivity of the kind that Husak objects to.

My analysis of the example of A and B does, however, presuppose another thesis to which Husak takes exception. This is the 'normative priority thesis' according to which it is better, all else being equal, to be justified than to be excused. Husak resists this thesis by counterexample:

> Smith and Jones deliberately inflict a fairly serious but not life-threatening injury on White. Imagine the story is embellished in either of two ways. In the first scenario, Smith's behaviour is barely tolerable, although no one would regard it as commendable. Suppose White is a deranged thief who is escaping with Smith's television set. Smith's act, I assume, is permissible and thereby justified. In the second scenario, Jones is well below the age of criminal responsibility, and injures White in a schoolyard brawl. Even if Jones has not acted in self-defence, I assume he is excused. ... I believe that the former scenario casts the agent in a worse moral light. Had I seriously injured a person at some point in my life I would prefer to have done so under the circumstances described in the second scenario.[42]

Unlike Husak, I do not see any reason to regard Jones as excused. His defence, if he has one, is that he is not responsible for his actions. But let's leave this point on one side for a moment. There are two other problems with the way that Husak sets out his challenge. First, we may wonder whether he honours the 'all else being equal' proviso in the normative priority thesis as I defend it. Isn't Husak perhaps tempting us to think of Smith as an intentional injurer and Jones as an accidental injurer? And aren't we perhaps distracted by sympathy for the deranged White in the first scenario, which does not extend to the brawling White in the second? Don't we tend to think that Jones' childish mistakes should be forgiven and forgotten more readily than Smith's adult ones, forgiveness being a quite separate matter from justification and excuse? And doesn't the description of Smith's action as 'barely tolerable' make his supposed justification seem dubious in a way that Jones' supposed excuse is not? Toleration, after all, is an attitude that is normally called for in respect of unjustified actions. All of this makes me think that Husak is not setting a fair test for the normative priority thesis. He is subtly stacking the deck in favour of Jones and against Smith, and hence in favour of excuses and against justifications.

[42] Ibid, 573.

Secondly, Husak asks which of the two agents is 'cast[ ] ... in a worse moral light.' This is a misleading way to phrase the question. It suggests that, according to believers in the normative priority thesis, excused wrongdoers are more blameworthy or less creditworthy than their justified counterparts. But this cannot be the implication of the normative priority thesis. Most believers in the normative priority thesis, including me, hold that there can be complete excuses as well as complete justifications, and that what makes an excuse or justification complete is the fact that it renders the agent entirely blameless. And most believers in the normative priority thesis, including me, hold that justified wrongdoers need not be creditworthy at all, let alone more creditworthy than excused wrongdoers. So the respect in which it is better to be justified than excused, according to most believers in the normative priority thesis, cannot be that justification earns one less blame, or more credit, than excuse. The normative difference must lie elsewhere. On my version, the difference is that justification brings one closer than excuse does to perfect conformity with reasons, and hence closer to perfect success as a rational being. That one falls short of perfect success as a rational being need not be one's fault; nor need the degree to which one falls short reflect the degree of one's fault. Failure need not betoken unfitness. Yet failure is still a matter of regret, and it still leaves a blemish on one's life. Making an excuse, I claim, is admitting a failure (in conformity with reasons) that goes one step beyond the failure one already admits in offering a justification, and that remains true even though both equally constitute denials of one's fault. As I put the point in 'The Mark of Responsibility', excuses are 'second best' to justifications because they involve 'an admission of rational defeat.' This is all that there is to the normative priority thesis as I defend it. Inasmuch as Husak suggests otherwise by landing me with the view that excuses 'cast[ ] [one] in a worse moral light' than justifications, his objection does not hit home.

I have also expressed the normative priority thesis in terms of self-respect. I said this in 'The Gist of Excuses':

> The self-respecting person aspires to live up to the proper standards for success in and fitness for the life she leads, and holds herself out to be judged by those standards. ... She wants it to be the case that her actions were not truly wrongful, or if they were wrongful, that they were at any rate justified, or if they were not justified, that they were at any rate excused.[43]

This is where Hamish Stewart's criticisms come in. Recall that Stewart, like Tadros, is interested in why I classify as excused, not as justified, those who act on the strength of reasonable mistakes concerning the existence of justificatory facts. Some of his objections to this classification parallel Tadros's. But Stewart also has an extra objection. He objects to my classification because, in tandem with the claims I made about the self-respecting person, it suggests to him

> that the self-respecting individual ought to reproach himself or herself for acting on facts that he or she could not reasonably have known. ... In terms of the reasons that [a defendant] could reasonably have been expected to act upon ... there is no difference between [a case of self-defence and a case of reasonably mistaken self-defence], and [the defendant] ought therefore to respect himself equally in each case.[44]

Here Stewart makes a misinterpretation similar to Husak's. I am sorry that I failed adequately to forestall it. It is not my view that someone who has an excuse should reproach himself. To be fully excused (I repeat) is to be faultless, and hence beyond reproach. But suppose Stewart were to replace 'reproach himself or herself for acting' with 'be sad that he acted'. This would correct the

[43] Above, 181.
[44] Stewart, 'The Role of Reasonableness in Self-Defence', *Canadian Journal of Law and Jurisprudence* 14 (2003), 317 at 322-3.

misinterpretation. But it wouldn't affect the more important challenge in Stewart's second sentence. For Stewart's main quarrel with me is this: Whether or not there is any difference in warranted self-reproach, there is no difference in warranted self-*respect* as between D1 who defends himself against an actual attack and D2 who similarly defends himself against what he reasonably imagines to be a similar attack.

I can see now why my remarks were taken to suggest that the excused D2 should respect himself less than the justified D1. But that is not what I actually said. What I actually said, and certainly what I meant, is that a self-respecting D2 would *want* to have been justified rather than excused. He would wish that his situation had been D1's situation, i.e. that he had not been wrong-footed by reason. That is because, as a rational being, he cannot but aspire to conform maximally to reason, and his self-respect is bound up, not with the conformity itself, but with the aspiration. His self-respect is threatened by his lowering his sights as a rational agent - by his being no less content to make an excuse than to claim a justification, or no less content to deny his responsibility for his actions than to do either of those things. It does not follow that he compromises his self-respect by the mere fact that he actually *is* excused rather than justified, or actually *does* lack responsibility rather than being excused or justified – so long as he is not indifferent to that being the case, so long as he does not approach with equanimity the question of which argument he is going to offer in his defence.

Ronnie Mackay and Barry Mitchell take exception to the claim, emphasized in 'Provocation and Pluralism' and in 'The Mark of Responsibility', that a denial of responsibility is, all else being equal, the least appetizing of all defences for a self-respecting person to offer. They focus on the negative attitude to mental illness which they find implicit in this view and in the examples that were used to illustrate and corroborate it:

> Small wonder that defendants are reluctant to plead insanity when we encounter the type of stigma which is being promulgated here. In essence what Gardner and Macklem are telling us is that a diminished responsibility verdict, like insanity, is one to which stigma is attached. This not only perpetuates an unfortunate attitude towards the mentally disordered but also relegates the interest in avoiding a murder conviction as being less important than what is referred to as a defendant's 'interest in being accorded their status as fully-fledged human beings'.[45]

The final salvo in this passage (after the 'but also') ignores the 'all else being equal' proviso which is an important part of the claim to which Mackay and Mitchell are objecting. The claim is not that the interest in avoiding a murder conviction is less important than the interest in being treated as a fully responsible person. The claim is only that the latter is also an interest that the law (and lawyers) should protect. The more severe the penalties he faces, the more a defendant is under rational pressure to find a way to secure his own acquittal. The point may come at which using a demeaning argument to get off the hook is reasonable. If so, it is the law that is being unreasonable. Its penalties should not be so severe that a defendant has no reasonable alternative but to demean himself in order to avoid incurring them.

But does a defendant really demean himself by relying on his mental illness for exoneration? I think that he does. No self-respecting person, in my view, would wish this line of defence upon himself. Mackay and Mitchell do not actually say that this view is mistaken. They seem to be more concerned about the way in which it might be used against mentally ill people, or the way that mentally ill people might feel about it. If so I share their

[45] Mackay and Mitchell, 'Provoking Diminished Repsonibility: Two Pleas Merging Into One?', *Criminal Law Review* [2003], 745 at 757. The quotation at the end of this passage is from Gardner and Macklem, 'Compassion without Respect? Nine Fallacies in *R* v *Smith*', *Criminal Law Review* [2001], 623, a paper that is not included in this volume.

concern. Mentally ill people have often been persecuted, neglected, patronized, and treated as objects of mirth. Their woes have often been compounded by quack treatments, pointless incarcerations, and brutal 'care' regimes. They have often fallen victim to bizarre superstitions and prejudices. But one should not conclude from the fact that mentally ill people have been on the receiving end of so much baseness and stupidity that their mental illness should be regarded with equanimity. Mental illness is not like homosexuality or left-handedness, unobjectionable traits that do not need any remedy. Mental illness really is a kind of illness and illnesses by definition call for treatments and cures. So long as it can be done without mistreating anyone, or committing other wrongs, the world would be a better place with all illnesses eradicated. Mental illnesses, in particular, make the following case for their own eradication. One cannot live a distinctively human life without a full range of rational faculties (cognitive, affective, deliberative, conative) in decent working order. Mental illnesses are illnesses that consist in the malfunction, partial or complete, of one or more of these rational faculties. In more severe cases they restrict, or even sometimes prevent, participation in a distinctively human life. And the ability to participate in a distinctively human life is one of the conditions for being a fully-fledged human being.

This suggestion rings alarm bells because it sounds like (and has often been taken for) an invitation to treat mentally ill people as if they were not people, to treat them like wild animals or even like plant life. But it is no such thing. We shouldn't want a wolf or a palm tree to be able to hold down a job or give a rationally intelligible account of itself, but we should want a mentally ill person to be able to do these things. Why? Because a mentally ill person meets the *other* conditions for being a human being (notably the genetic and physiognomic conditions) and this makes the aspirations and expectations of a successful human life applicable to her. We should therefore want to see her illness cured, and failing that, its symptoms alleviated. We should aspire

that she lives the best human life that is possible for her, with the maximum possible participation in distinctively human value. And we should all want her to have rights and resources that will protect her from further injustices and inhumanities, including those that have characterized some past misguided models of treatment and cure for mental illness. We should want a mentally ill person, in short, to be as fully human as possible.

Being responsible for one's own wrongs is a distinctively human capacity and one that is central to all distinctively human lives. It is the capacity that makes self-respect possible, for self-respect is the attitude of one who can sincerely say, of anything she did wrong, that she was justified in doing it, or, failing that, excused. We should all wish mentally ill people to have the opportunity to lead self-respecting lives. 'All' here includes mentally ill people themselves. They should prefer to have justifications for their actions, or failing that excuses, and should prefer not to have to fall back on their illnesses to furnish them with a defence. Alas, some mentally ill people do have to fall back on their illnesses to furnish them with a defence on at least some occasions. Thanks to mental illness, their actions sometimes defy rational explanation and must be put down to pathology. To deny that this is regrettable, it seems to me, is to claim that mentally ill people have nothing wrong with them. Is this what Mackay and Mitchell are claiming? I hope not.

8. Responsibility, relations, and relativities

Antony Duff picks up another theme from 'The Mark of Responsibility'. I argued there that basic responsibility, unlike some other kinds of responsibility, is non-relational. One is not responsible to anyone in particular. Duff objects:

> An Old Bailey judge or a stranger on the bus to whom I sought to answer for my unfeeling behaviour towards my aunt would naturally reply that I did not have to answer for such conduct to them. The

more significant point is that they have no right to demand that I account to them for myself or my conduct in this respect: if they tried to call me to account for my unkindness towards my aunt, I could properly reply that it was none of their business (unless they were suitably connected to my aunt). ... [F]or many moral wrongs, and for other non-moral matters, I am responsible only to some specifiable people, not to all.[46]

I hope that my postscript on accountability, which was added to 'The Mark of Responsibility' after Duff wrote this passage, may now help to clear up the apparent disagreement between us. Being basically responsible, on the view I defended, is not a normative position that one occupies. It is not, for example, a duty or a power or a cluster of duties and powers. Rather, it is a capacity and a propensity to answer for oneself. It may figure in the justification of various duties and powers, including a duty on me to answer to some people as opposed to others, or a power in some people as opposed to others to impose a duty on me to answer to them. Such duties and powers may also be referred to as my responsibility (or, more fashionably, as my accountability). But they are not my *basic* responsibility because they are *based* on my basic responsibility: my basic responsibility, my capacity and propensity to answer for myself, forms part of the case for my being responsible (=accountable) to someone in particular.

It is true, as I explained in 'The Mark of Responsibility' that those who are basically responsible always have a reason to explain themselves. Their propensity, in other words, is a rational one. But a reason is not a duty. One can owe a duty to somebody, but one cannot owe a reason to somebody. Of course one can have a reason to do something with or to someone. One can have, for example, a reason to give a gift to someone in particular. In that sense a reason can be relational. Yet it cannot

[46] Duff, 'Answering for Crime', *Proceedings of the Aristotelian Society* 106 (2006), 85 at 88–9.

be relational in the further sense in which a duty can be relational, viz. in having another person, a rightholder, to whom it is owed. And it is this further relationality of duties, their oweability to rightholders, that Duff invokes in the passage quoted above. This shows that he is talking about a slightly different topic from the one I was talking about in the passage he criticizes. He is talking about the next topic along. After we have established that a given agent is basically responsible, capable of giving an account of herself, we may raise the further question of who, if anyone, is entitled (=has a right) to call her to account. The position I took in 'The Mark of Responsibility' does not commit me to giving 'anyone and everyone' as the answer to this follow-up question. I agree with Duff that one needs to have some *locus standi* to be entitled to call another to account.

There is something misleading, however, about Duff's use of the phrase 'none of your business' to convey the absence of the required *locus standi*. In 'Complicity and Causality' I aligned myself with the agent-neutralist view according to which wrongdoing by anyone is fundamentally everyone's concern. We all have one and the same reason (the fact that the wrong is a wrong) to avoid the commission of any wrong by anyone. But it is a different matter to ask which of us has a *duty* or a *right* to avoid the commission of a particular wrong by another. The answer depends mainly on what would make for (as I called it) 'the efficient use of rational energy'. It depends mainly on how well a particular person is placed to secure the avoidance of the wrong in question, thereby conforming to the reason that he has, in common with everyone else, to avoid its commission. What does this have to do with the relationality of responsibility? The fact that everyone has a reason to avoid the wrongdoer's committing a wrong does not by itself entail that everyone has a reason to get the wrongdoer, having committed the wrong, to account for herself. Nevertheless the agent-neutralist line of thought can readily be extended. Just as everyone's wrongdoing is everyone's concern, so everyone's responsibility is everyone's

concern. We each have reason to see to it that people in general answer for their wrongs, and one way in which I can see to it that people answer for their wrongs is to have them answer to me. That is one way, of course, but is it the optimal way? Not always. What we face here is once again mainly a question of the efficient use of rational energy. It may not be my place (my role) to extract justifications and excuses. It may be the law's place, or the place of the person who was wronged, or the place of the wrongdoer's friends and family, etc. But where this is so, it is mainly (not only, but mainly) because and to the extent that this person with *locus standi* is the one who is best-placed to do the extracting, i.e. who will do the best job of conforming to the reason that we all have in common to see to it that the wrongdoer answers for her wrongs.

I did not mount a full defence of this agent-neutralist view in 'Complicity and Causality'. I only paused to indicate how, once the efficiency principle (as I will call it) is factored in, the agent-neutralist view is compatible with a great deal of superficial agent-relativity in everyday moral experience. From the fact that all wrongdoing and all responsibility is everyone's concern it does not follow that one should be even-handed in one's efforts to avoid every wrong or to hold everyone responsible. Often one should begin by trying to put one's own house in order, for here one may perhaps be less prone to counterproductivity, self-defeatingness, overzealousness, and similar errors. I suspect that I did not make this point clearly enough in 'Complicity and Causality'. For I left Tatjana Hörnle with this worry:

> [T]he amount of intrusion is limited only by the fact that it might sometimes be inefficient. If one takes this stance seriously, everybody would constantly be trying to be as efficient as possible in morally improving others.[47]

[47] Hörnle, 'Commentary', above note 14, 148.

Actually, I said 'limited mainly' rather than 'limited only'. But leaving this quibble on one side, is it true that, on my view, 'everybody [should] constantly be trying to be as efficient as possible in morally improving others'? No. Indeed this is exactly the worry that I was trying to defuse by invoking the efficiency principle in the first place. That *ceteris paribus* everyone should minimize wrongdoing by everyone does not mean that *ceteris paribus* everyone should *try* to minimize wrongdoing by everyone. For much of this effort would be inefficient, i.e. would not succeed in reducing wrongdoing and might even increase it. Hörnle is making the common mistake of confusing what one should do with what one should try to do, forgetting that often it is the trying that is the problem.

To be fair, Hörnle has deeper worries about my agent-neutralism, and perhaps Duff shares them. Hörnle worries, in a neo-Kantian vein, that my agent-neutralism 'violates ... human dignity' by licensing 'trade-offs' of one person against another. I do not share this worry. I cannot speak for all agent-neutralists. It may be that some (e.g. classical utilitarians) leave little space for anything resembling human dignity. But if so I clearly part company with them. As 'The Wrongness of Rape' and 'The Functions and Justifications of Criminal Law and Punishment' both make clear, I share Kant's view that human beings have a value beyond price (i.e. a value beyond their use-value) that we might reasonably call their dignity. My difficulty is in seeing why this dignity should be valued agent-relatively, not agent-neutrally, and hence should not be traded-off between persons just like any other value. Fundamentally the wrongdoing and responsibility of those who join death squads in Iraq, of those who traffic people in Albania, of those who attempt internet scams in Nigeria, and of those who preach pseudo-Biblical resentments in the USA are no less my concern than are my own wrongdoing and responsibility. That is precisely because these wrongdoers, just like me, have a value beyond price; they belong, just like me, to the Kingdom of Ends. I have no reason

to regard their moral fates (which are bound up with their wrongdoing and their responsibility) as any less important than my own. I only have reason to make extensive allowances in what I do for my own relative impotence in improving the moral fates of others, and in particular for the risk that I will make those moral fates even worse in my bungled efforts to improve them. If I should avoid trading off one person's dignity against another's – if I should adopt superficial agent-relative conceits of the 'none-of-my business' variety, and thereby pass the buck for trade-offs onto somebody else - then that (I repeat) is mainly because and to the extent that I am not best-placed to do the trade-offs well.

9. Crimes and punishments

There are various applications of the efficiency principle in 'The Functions and Justifications of Criminal Law and Punishment' and also in 'Crime: in Proportion and in Perspective.' A central theme of the former paper is that a certain reason can contribute to the justification of criminal punishment even though it is not the case that a particular official (or indeed any official) should punish anyone for that reason. Normally this is because acting for that reason is not a good way for the official concerned to do what the reason would have him do. Highlighted in the latter paper, meanwhile, is the localized agent-relativity of the state's duties of humanity and justice, an agent-relativity which is said to have a deeper agent-neutral basis. Again, the efficiency principle plays the dominant role in explaining this.

These papers have not attracted as much critical attention in print as have the others in this volume. Perhaps they are more banal. It does not follow that their claims are more widely endorsed. In discussion, many of my colleagues and students have doubted my view that punishment calls for (or is even open to) a pluralistic, cumulative justification. This is thought by many to be an incoherent view, or at least a view that condemns the

practice of punishment itself to incoherence. Several people have reminded me of H.L.A. Hart's remark that

> what is most needed is *not* the simple admission that instead of a single value or aim (Deterrence, Retribution, Reform or any other) a plurality of different values and aims should be given as a conjunctive answer to some *single* question concerning the justification of punishment. What is needed is the realization that different principles (each of which may in a sense be a called a 'justification') are relevant at different points in any morally acceptable account of punishment.[48]

Hart himself thought that deterrence (or more generally the prevention of wrongdoing) is the 'general justifying aim' of criminal punishment. He thought that guilt (=culpability, fault) is relevant only to the distribution of criminal punishment, coming into play only once it had been determined that criminal punishment in general is justified. Importantly, and unlike some who have cited Hart's remark to me, Hart agreed that it would be possible to hold a view according to which 'giving guilty people what they deserve' and 'preventing future wrongs' figure side-by-side as criminal punishment's twin general justifying aims. His resistance to doing so did not come of the thought that one could not reasonably endorse 'conjunctive' general justifying aims for criminal punishment. His objection was simply a moral objection to the view according to which a person's guilt is a reason in favour of punishing her. He thought that a person's innocence serves as a (powerful, usually decisive) reason for *not* punishing her, and hence that the importance of considering questions of guilt in connection with an anticipated punishment could only be negative: to limit and regulate the use of criminal punishment, granted that such punishment, as a general practice, is justified on other grounds.

[48] Hart, 'Prolegomenon to the Principles of Punishment' in his *Punishment and Responsibility* (1968), 3.

Personally, I think that Hart's philosophical judgment here is sound, but his moral judgment is flawed. The principle of desert (as I will call it) cuts both ways. The fact that someone is innocent is a reason against punishing her; the fact that she is guilty is also a reason in favour of punishing her. So giving people what they deserve figures among the general justifying aims of punishment. Yet the justificatory importance of the principle of desert is not symmetrical. It justifies the non-punishment of the innocent much more readily than it justifies the punishment of the guilty. That is because, in general, it is much easier to justify not punishing than it is to justify punishing. Indeed it is easier to justify that many guilty people go unpunished than that one innocent person is punished. And why, you may ask, is that? Primarily, it is because punishment involves the intentional infliction of disadvantage or suffering upon the person who is punished. The disadvantage or suffering is not a side-effect (as it is with the payment of reparative damages, the protective detention of the seriously mentally ill, the pre-trial custody of a suspect, and so on), but part of the punisher's plan. This morally repugnant feature – the wishing of evil upon another – makes punishment extremely hard to justify.

Because of the argumentative force of the other side of the desert principle (i.e. against punishing the innocent) guilt is usually a necessary condition of justified punishment, and when it is present it is also (without further ado) a reason in favour of exacting the punishment for the justification of which it is a necessary condition. So as well as neutralizing a major objection, it forms the first part of the positive case for punishing. But even though it is usually a necessary condition of justified punishment, guilt is rarely a sufficient condition. More than nominal punishment is almost never justifiable on the strength of guilt alone. There must be further benefits, which may either be consequences of honouring the principle of desert or independent benefits, before there is an adequate case for punishing any but the very guiltiest. That, at any rate, is my sense

of how we are forced towards a generally pluralistic approach to the justification of punishment, and thus (inter alia) to the justification of criminal punishment.

This pluralistic approach is advocated in 'The Functions and Justifications of Criminal Law and Punishment' and echoed in 'Crime: in Proportion and in Perspective'. All the same, looking back on the latter paper, I now think that it starts off on a misleading note. The 'displacement function' of criminal law – its ability to interpose itself coolly between wrongdoers and those they have wronged - is presented as a key part of the case for criminal law's continuing existence. This was taken by some of my readers to entail that the displacement function is also a key part of the case for punishment. This interpretation was understandable in view of some of my remarks and examples, but it was not what I intended. I do not regard every argument for punishment as an argument for criminal law, or for that matter *vice versa*. In writing about the displacement function I assumed that punishment of the wrongdoer, if justified at all, falls to be justified by an accumulation of other considerations, including the wrongdoer's guilt. What the paper was supposed to offer was an simple explanation of why, *assuming that the guilty wrongdoer should indeed be punished by someone*, the criminal law (as opposed to victims or their families or their sympathizers) should be the one to exact the punishment. I then tried to investigate what some of the consequences of this explanation would be for the criminal law itself, including its future legitimacy.

Clearly, in framing this investigation, I should have paid more attention to Hart's warning quoted above. As well as there being 'a plurality of different values and aims' that are needed to provide an adequate answer to the general question 'why punish?', it is also true that different considerations 'are relevant at different points' in the inquiry. Notably, some considerations are relevant mainly to the question 'who gets to punish?' rather than the prior question 'why punish?' Of course one can run the two questions together by asking 'why punish?' with a particular

punisher (such as the criminal law) already in mind. But it is better to separate the two questions. Already, too much academic writing about the criminal law casually runs together questions about the justification of criminal law and questions about the justification of punishment. Such writing often harbours an excessively statist view of punishment and/or a too narrowly punitive view of the purposes of the criminal law. In real life, for better or worse, most punishments are exacted, not by the criminal law, but by parents, spouses and ex-spouses, friends and ex-friends, business partners and ex-business-partners, etc.[49] At the same time the criminal law should not be seen as a merely, or even a mainly, punitive institution. Rather, as I explained in 'In Defence of Defences', the criminal law

> is primarily a vehicle for the public identification of wrongdoing ... and for responsible agents, whose wrongs have been thus identified, to *answer for* their wrongs by offering justifications and excuses for having committed them. By calling this latter function 'primary' I do not mean to suggest that it is socially more important. I mean that the proper execution of the other functions depends upon it. Criminal law can be a proper vehicle for ... punishment only because it is a vehicle for responsible agents to answer for their wrongs.[50]

This remains my view. It explains the emphasis, in many of the foregoing essays, on questions about responsibility, justification, and excuse. I am sorry if 'Crime: in Proportion and in Perspective' accidentally gave succour to a rival view according to which the purposes of the criminal law are primarily punitive, and according to which whatever justifies the criminal law also justifies (*pro tanto*) the practice of punishment.

[49] Hart relegates such punishments to the margins in ibid, 5, favouring (without argument) the view that official state punishment is the 'central case'.
[50] Above, 107.