Justifications and Reasons

JOHN GARDNER*

1. Isolating the issue

Nobody seriously denies that there is a close relationship between justifications and reasons. To claim that one has justification for doing or believing as one does is to claim, at the very least, that one has reasons for so doing or so believing. The question which arouses disagreement is merely how the reference to 'reasons' here is to be interpreted. For reasons may be either *guiding* or *explanatory*.[1] Guiding reasons are reasons which apply to one. They bear on what one ought to do or believe. One may, however, overlook or ignore these reasons. Then, even though one acts or believes exactly as the guiding reasons would have one act or believe, they are not the reasons *for* which one so acts or believes. They are not, in other words, the explanatory reasons. Explanatory reasons are logically related to

* Fellow and Tutor in Law, Brasenose College, Oxford. This paper was written during my tenure of a British Academy research leave award, and I am grateful to the Academy for the opportunity to develop this and the larger project of which it forms part. Versions of the paper were presented at seminars in Glasgow, Edinburgh, and Tel Aviv. I benefited greatly from the ensuing discussions, and would like to thank all those who contributed. For incisive written comments, thanks also to Tony Honoré, Stephen Shute, Andrew Simester, and Andrew von Hirsch.

[1] See Joseph Raz, *Practical Reason and Norms* (2nd ed., Princeton 1990), 16–20.

guiding reasons, for it is necessarily true of every explanatory reason that the person who acts on it, or holds beliefs on the basis of it, also believes it to be a guiding reason for the action or belief in question.[2] But it may or may not be the guiding reason that she believes it to be. Just as guiding reasons need not be explanatory reasons, in other words, so explanatory reasons need not be guiding reasons. Thus there arises an issue, current in epistemology as well as moral and legal philosophy, about whether justification depends on guiding reasons or on explanatory reasons. Are one's actions and beliefs justified by the reasons which actually applied to one, or by the reasons which, perhaps mistakenly, one thought applied to one and accordingly treated, in one's acting or believing, as if they were reasons which actually applied to one? Faced with this question, some have come to the view that there are two different perspectives or points of view from which one's actions or beliefs may be justified. On the one hand there is so-called 'subjective' justification, which depends on explanatory reasons, and hence bows to one's mistakes about the applicable guiding reasons when these mistakes affect what one believes or how one acts. Then there is 'objective' justification, which depends on guiding reasons, and hence extends justification even to some who had no inkling of those reasons, let alone acted or believed anything on the basis of them.[3] In the eyes of some writers, these two modes of justification simply represent irreconcilably different ways of looking at our actions and

[2] Or, strictly speaking, believes it to *disclose* or *reflect* a guiding reason. Some apparent counterexamples to this are mentioned by E.J. Bond in *Reason and Value* (Cambridge 1983), 29. They cannot be dealt with here. But I should stress that 'believes' here, and throughout this paper, carries its widest connotations. It covers everything from the firmest conviction to the merest inkling, everything from knowing to imagining, and everything from explicit awareness to latent or subconscious recognition. C.f. the remarks on 'vindication' in note 39 below.

[3] See, e.g., Alvin Goldman, *Epistemology and Cognition* (Cambridge, Mass. 1986), 73; Jonathan Kvanig, 'Subjective Justification', *Mind* 93 (1984), 71; William P. Alston, 'Concepts of Epistemic Justification', *The Monist* 68 (1985), 71. This is also, I believe, the distinction with which Paul Robinson is mainly concerned in his 'Competing Theories of Justification: Deeds vs. Reasons', this volume, 000.

beliefs. For some purposes or in some contexts we may favour one perspective, while for other purposes and in other contexts we resort to the other.

The terminology of 'subjective' and 'objective' which is used to draw this contrast is notoriously treacherous. It is particularly likely to be misleading for criminal lawyers, who also customarily use these labels to mark a number of quite different distinctions. One of these distinctions needs to be mentioned at this stage in order to move it out of the way. It is a distinction between those who attach justificatory importance to how our actions turn out, and those who decline to do so. In the lingo of some criminal law scholarship and commentary, an 'objectivist' might say that murder is harder to justify than attempted murder, because a death actually comes about; a 'subjectivist' would have to demur.[4] It is a profound and interesting disagreement. But it is not the same as the disagreement about whether justification depends on explanatory reasons or guiding reasons. Rather, it is an internecine dispute, among those who accept that justification turns in part on guiding reasons, about what kinds of guiding reasons there can be. Some hold that (a) there can in principle be no guiding reasons for or against doing what cannot be done (thus the proposition 'ought implies can') and (b) the most we can ever do is try, with no guarantee of success. It means that they insist on regarding all guiding reasons as merely reasons for or against trying, not as reasons for or against succeeding.[5] But in fact there may be reasons to succeed as well as reasons to try. Moreover the two do not automatically go hand in hand. Sometimes I have reasons to try without having reasons to succeed. Suppose that people will mistakenly take against me if I do not try to help with an unjust war effort. Then I have a perfectly obvious reason to try, but (other things being equal) no reason to

[4] See, e.g., Andrew Ashworth, 'Taking the Consequences' in Stephen Shute, John Gardner and Jeremy Horder (eds), *Action and Value in Criminal Law* (Oxford 1993), 107 at 109–10, or Antony Duff, 'Subjectivism, Objectivism and Criminal Attempts', this volume, 000 at 000–00.

[5] This appears to be the argument of W.D. Ross in *Foundations of Ethics* (Oxford 1939), 160, relied upon by Andrew Ashworth in 'Sharpening the Subjective Element in Criminal Liability', Antony Duff and Nigel Simmonds (eds.), *Philosophy and the Criminal Law* (Wiesbaden 1984), 79.

succeed. Conversely, I may have reason to succeed but no reason to try. Suppose that, since I cannot swim a stroke, it would be pointless for me to try to rescue someone from a stormy sea. Then (other things being equal) I have no reason to try to rescue them, but it scarcely means that I have no reason to rescue them.[6] If I had no reason to rescue them, after all, I would not be so horrified at the realisation that it is pointless for me to try. I could walk past without compunction. That shows why we should give short shrift to 'ought implies can': my horror as I look on helplessly reflects the fact that I ought to save this life even though I cannot. But it also shows why, if the justification of an action depends in part on the guiding reasons for performing that action, the justification of an action may sometimes be partly hostage to the action's results and sometimes not. It simply depends on what particular action the reasons in question are reasons to perform, i.e. whether they are reasons to get or avert certain results or merely reasons to try to get or avert those results. That, however, does not suggest for a moment that sometimes, in relation to some actions, one should look to explanatory reasons rather than guiding reasons to do the justificatory work. On this issue, so far as I can see, the justificatory role of explanatory reasons is neither here nor there.

In what follows, my central concern will not be with the question of what guiding reasons there are in favour of particular actions or particular kinds of actions, but with the

[6] Some who accept that I have reason to save in this case may doubt whether I have no reason to try. They picture me momentarily wavering on the cliff's edge, incapacitated by indecision, now leaning forward to jump, now pulling back. Does this not suggest the impetus of a reason to try to save? It may do: it may suggest that I believe myself, contrary to fact, to be capable of effecting a rescue, and so think there is a reason to try, and feel its pull. But my wavering may also be interepreted as the action of a man who is, in momentary defiance of logic, trying to succeed without trying, because he has reason to succeed but no reason to try. This interpretation presupposes that sound practical reasoning follows what Anthony Kenny calls 'the logic of satisfactoriness' (by which one automatically has reason to do whatever is sufficient to achieve what one has reason to achieve) and not 'the logic of satisfaction' (by which one automatically has reason to do whatever is necessary to achieve what one has reason to achieve). For argument, see Kenny 'Practical Inference', *Analysis* 23 (1966), 65.

more fundamental conceptual question of whether justification depends upon guiding reasons or explanatory reasons. The answer, irritating but unavoidable, is that it depends upon both. No action or belief is justified unless it is true *both* that there was an applicable (guiding) reason for so acting or so believing *and* that this corresponded with the (explanatory) reason why the action was performed or the belief held. It follows that the common view that there are two different perspectives on justification, a 'subjective' (explanatory reason) perspective and an 'objective' (guiding reason) perspective, must be rejected. To cite explanatory reasons as well as guiding reasons is not to provide justifications from two different points of view, nor even to provide two partial justifications, but merely to provide the two essential parts of one and the same (partial or complete) justification. Of course this is not to deny that some actions and beliefs may be justified from one point of view and not from another. A certain belief may be justified from my point of view and not from yours, or justified from the Benthamite point of view but not from the Kantian. A certain action may be justified from the point of view of a Christian but not from the point of view of a Muslim, or justified from the point of view of the army's rules of engagement but not from the point of view of the criminal law. My only proviso is that, from whatever point of view one claims justification for one's actions or beliefs, one claims justification only if one claims *both* that there were, from that point of view, reasons for one to act or believe as one did *and* that one's reasons for performing the act or holding the belief were among these reasons. Notice that this is perfectly compatible with a recognition that, within certain systematic points of view or perspectives, the word 'justification' may sometimes be appropriated to do other jobs, e.g. to refer to something falling short of, or going beyond, justification. The legal point of view, in particular, is widely noted for putting its own specialised glosses on everyday words. But English criminal law, at any rate, has not yet paid that particular compliment to the word 'justification'. So far as I can see, our judges persist in using the word 'justification' to refer mainly to legal justifications proper, i.e. to legally recognised reasons for acting which were also the relevant agent's reasons for acting in the case under consideration. This is the

essence of the famous *Dadson* doctrine.[7] But even if English criminal law were found to use the word 'justification' in some different, technical sense, that would be a matter of little concern for present purposes. Our interest is not in the legal meaning of the word 'justification'. Our interest is in the ordinary phenomenon, that of justification, which still plays a major role in the thinking of most criminal courts, and indeed in evaluative thinking at large, whatever the local lawyers and legal commentators may choose to call it.

2. Pros and cons: the basic asymmetry

What calls for justification? I already mentioned actions and beliefs, and we may add to the list a wide range of phenomena which are logically related to actions and beliefs, such as emotions, attitudes, desires, decisions, practices, and rules. But in a way, that is not an answer to the question. We still need to know whether such things *always* call for justification, or only *sometimes* do. The answer has to be an equivocation. In a loose sense, justification is always called for. That is just to say that actions, beliefs, etc. are always answerable to reason. One may always ask 'why?' But in the stricter and more important sense which concerns us here, justification is called for only when one also has some

[7] *R* v *Dadson* (1850) 4 Cox C.C. 358. Per Erle J.: 'The prosecutor not having committed a felony known to the prisoner at the time when he fired, the latter was not justified in firing at the prosecutor.' There are two conditions of justification implicit in this: (1) that the prosecutor must have been a felon and (2) that the prisoner firing upon him must have known (or, more broadly, believed) this at the time when he fired. Condition (1) specifies the necessary guiding reason, while (2) is needed to ensure that it is also the explanatory reason. It is true that believing that the prosecutor is a felon is not the same as acting because he is a felon. But the former is a necessary condition of the latter, and since the court in *Dadson* expresses the doctrine negatively, the absence of this necessary condition is all that is needed to dispose of the case. It does not mean that the knowledge condition is being elevated to the status of a sufficient condition – *pace* Paul Robinson, 'Competing Theories of Justification: Deeds vs Reasons', this volume, 000. See further *R* v *Thain* [1985] NI 457, in which the distinction between beliefs and reasons becomes pivotal.

reason *not* to act, believe, etc. as one does. The unobjectionable, in other words, is in no need of justification. In this stricter sense, justification may be either partial or complete. What one claims if one claims *partial* justification is that the prima facie reasons against one's action or belief are countered by some reasons in favour. What one claims if one claims *complete* justification is that the reasons in favour are, moreover, strong enough to prevail over the reasons against. Thus by claiming full justification one denies that the prima facie judgment against performing the action or accepting the belief should be elevated to the status of an 'all-things-considered' judgment against its performance or acceptance. All things considered, the action or belief was alright in spite of the prima facie objections to it.

Criminal lawyers should already be at ease with this distinction between prima facie and all-things-considered judgments, and should quickly be able to see its relevance to the idea of a justification. For this distinction is highly visible, as Kenneth Campbell has observed, in the familiar demarcation between criminal offences and (justificatory) criminal defences.[8] In classifying some action as criminal, the law asserts that there are prima facie reasons against its performance – indeed reasons sufficient to make its performance prima facie wrongful. In providing a justificatory defence the law nevertheless concedes that one may sometimes have sufficient reason to perform the wrongful act, all things considered. Yet the very familiarity of this point has led many criminal lawyers to underestimate its significance. They have looked upon it rather shame-faced as a kind of artificial legalistic divide. Some have thought that it can only be a matter of expository convenience whether one treats a certain issue as going to the presence of an offence or the absence of a defence.[9] Others have conceded at most an evidential significance to the demarcation, relating it only to the question of who should normally bear the burden of adducing initial

[8] See Campbell, 'Offences and Defences' in Ian Dennis (ed), *Criminal Law and Justice* (London 1987), 73. Campbell does not seem to limit the point, as I do, to *justificatory* defences.

[9] Glanville Williams, 'Offences and Defences', *Legal Studies* 2 (1982), 233 at 233–4.

evidence.[10] Things are not helped here by the fact that criminal lawyers are also accustomed to use the label 'prima facie' itself to mark an evidential classification. To say that an action was prima facie wrongful normally signals, to the legal mind, that there is some reason to believe that a wrong was committed, but that it may yet, once more evidence is presented, turn out not to be so. But in the sense which matters for an understanding of the demarcation between offences and justificatory defences, to identify a prima facie wrong is to identify an *actual* wrong, not just an apparent or putative wrong.[11] It is to claim that there were indeed legally recognised reasons against an action, not merely that there are now legally admissible reasons to believe that there were such reasons. To be sure, the reasons against the action, which are also the reasons for its criminalisation, may all have been defeated in the final analysis. It may have been alright for the defendant to act against them all things considered. But it does not mean that they dropped out of the picture. That a reason is defeated does not mean that it is undermined or cancelled. It still continues to exert its rational appeal. It may indeed be a matter of bitter regret or disappointment that, thanks to the reasons which justified one's action, one nevertheless acted against the prima facie reasons for avoiding that action. It may even be a matter of regret or disappointment to the criminal law. The law certainly need not welcome it. But by granting a defence the law concedes that any regret or disappointment must be tolerated, and that no liability can attach to the person who by her prima facie wrongful actions occasioned it. By granting a *justificatory* defence the law concedes that this is true by virtue of the fact that the defendant had, at the time of her prima facie wrongful action, sufficient reason to perform it.[12]

[10] See Michael S. Moore, *Act and Crime: the Philosophy of Action and its Implications for Criminal Law* (Oxford 1993), 179.

[11] For further discussion of this point, see John Searle, 'Prima Facie Obligations' in Joseph Raz (ed), *Practical Reasoning* (Oxford 1978), 81.

[12] Compare Paul Robinson, 'A Theory of Justification: Societal Harm as a Prerequisite for Criminal Liability', *U.C.L.A. Law Review* 23 (1975), 266 at 274: 'Justified behaviour is correct behaviour and therefore is not only tolerated but encouraged.' Robinson assumes that a justification operates to cancel or undermine the countervailing considerations rather than to defeat them. So

Whether one accepts this account of the role of justification in the law and beyond is not a matter of merely academic importance. It has far-reaching practical implications. One implication is of immediate significance to us. It stems from the fact that, as I have explained them, claims of justification cannot but exhibit one of the most striking asymmetries in all human thought and experience. This is the asymmetry between the pursuit of positive value and the avoidance of negative value, between reasons in favour and reasons against, or, as we might say in ordinary conversation, between *pros* and *cons*. The asymmetry is brought to the surface by claims of justification only because such claims implicate both reasons in favour of and reasons against the justified action or belief. If the role of justification were that of cancelling or undermining the reasons against an action or belief – if it were, in lawyers terms, a matter of 'negating an element of the offence' – then no question of the relationship between reasons for and reasons against would arise in cases of justification, since in such cases all reasons against would be 'negated', i.e. cancelled or undermined, and would not exert any countervailing force. But since a justification merely defeats the reasons against an action or belief without cancelling or undermining them, the conflict between pros and cons, and hence its asymmetrical structure, is very much at the centre of attention when claims of justification are made.

To understand the asymmetry of rational conflict properly one must begin by thinking a little about the logic of guiding reasons. The first question is: What are guiding reasons there to guide? It may seem like a silly question. Surely it goes without saying that guiding reasons for and against action are there to guide action, guiding reasons for and against belief are there to guide belief, guiding reasons for and against emotion are there to guide emotion, and so on. But there is an alternative view. Perhaps guiding reasons of all kinds

---

the law has nothing to regret. It is surprising that Robinson apparently remains attached to this position to this day in spite of his own highly effective attack on it in 'Criminal Law Defences: A Systematic Analysis', *Columbia Law Review* 82 (1982), 199 at 220: 'Where conduct is covered by an offence modification, it is not in fact a legally recognised harm … Justified conduct, on the other hand, causes a legally recognised harm or evil … It is tolerated only when, by the infliction of the intermediate harm or evil, a greater societal harm is avoided or benefit gained.'

are merely there to guide reasons. Thus reasons for or against action are there to guide the reasons for which we act, reasons for or against belief are there to guide the reasons for which we believe, etc. It is not a circular or regressive proposal so long as one remembers the distinction between guiding reasons and explanatory reasons. The suggestion that guiding reasons guide reasons is the suggestion that the real point of there being a guiding reason is that one should act, believe etc. *on the basis of that reason*, i.e. that it should also be the explanatory reason for one's so acting or believing. Thus if one merely acts or believes as the guiding reason would have one act or believe, but one's explanatory reasons are quite different, then strictly speaking one did not follow the guidance. Some have thought that guiding reasons do indeed work in this way by default, or at least that they work in this way by default if they belong to certain families, e.g. moral reasons, reasons of duty, altruistic reasons, etc.[13] If this were true then it would immediately introduce a very radical asymmetry between reasons for and reasons against action and belief. That is because, as I have pointed out elsewhere, explanatory reasons are necessarily reasons *for* action or belief rather than reasons *against*.[14] Explanatory reasons explain. They are the reasons that the agent or believer had for acting or believing as she did. To say that she also acted or believed *in spite of* certain countervailing reasons is necessarily to return to the discourse of guiding reasons, and to leave the discourse of explanatory reasons behind. For the whole point of what one says is that these reasons, since the agent or believer did not act or believe on the basis of them, do not explain her action or belief.[15] The effect of this, if we

[13] See, famously and starkly, Kant's *Groundwork of the Metaphysic of Morals* (ed. H.J. Paton, New York 1964), 65–68. For interpretation and discussion of Kant's point, see e.g. Bernard Williams, 'Persons, Character and Morality' in his *Moral Luck: Philosophical Papers 1973–1980* (Cambridge 1981), 1 at 16–19; Barbara Herman, 'Integrity and Impartiality', *The Monist* 66 (1983), 233; Samuel Scheffler, *Human Morality* (Oxford 1992), 19ff.

[14] 'Action' and 'belief', at this point and throughout this paper, must be interpreted widely enough to include inaction and disbelief.

[15] See John Gardner and Heike Jung, 'Making Sense of Mens Rea: Antony Duff's Account', *Oxford Journal of Legal Studies* 11 (1991), 589 at 569–73.

share the view that guiding reasons exist to guide explanatory reasons, is that guiding reasons against action and belief are to a large extent a debased currency. They can only exert their purchase in a derivative way. After all, to the question 'for what reason should I act?', the answer 'here's a reason not to' is, to say the least, evasive. One cannot, as a matter of logic, act for reasons against so acting (unless, of course, one mistakes them for reasons in favour). The most one can do is modify one's reasons for so acting in the light of reasons against. This endows reasons against with at most an indirect motivational role, subsidiary to that of reasons in favour. The claim is not, I should stress, that reasons against an action cannot be appreciated, considered, taken into account, given their due weight, etc. alongside reasons in favour. The point is only that they cannot, *qua* reasons against one's action, be the reasons for which one acted. It means that, on any view according to which the point of guiding reasons is to guide explanatory reasons, reasons against action and belief are bound to be poor relations which lack the independent rational force of their positive cousins.

If it were accurate, this broadly Kantian account would make it easy to explain how justifications come to take the form I claimed for them. Since on this account one follows reasons in favour of an action only if one actually acts for those reasons, there is no justificatory ground to be gained by citing a guiding reason in favour of what one did if one did not act for that reason. On the other hand, one can readily lose justificatory ground by citing a guiding reason against one's action without claiming to have acted on it, since reasons against action are merely reasons to modify the reasons for which one acted, and cannot themselves be acted on.[16] Unfortunately, however, the Kantian account quickly descends into confusion. It is not surprising that those who have espoused it, including Kant himself, have had great difficulty understanding the nature of moral, legal, and other rational conflict. For this view of theirs, by debasing the independent rational force of

[16] *Qua* reasons against. They may of course be reasons for some other action as well as reasons against this one.

negative guiding reasons, ultimately forces them to reinterpret such reasons as mere cancelling conditions on positive guiding reasons.[17] Thus it turns out that the correct view, for once, is the obvious one. Fundamentally, guiding reasons for and against action are there to guide *actions*, not to guide explanatory reasons for actions. This is the grain of truth in Paul Robinson's view that a 'deeds' account of justification should be preferred to a 'reasons' account.[18] It means that negative guiding reasons retain in full their independent rational force. Although one cannot in principle act for these reasons, one can act as they would have one act, and other things being equal that is all that guiding reasons, be they positive or negative, envisage that one should do. But does it follow that explanatory reasons do not matter at all so far as guiding reasons are concerned? Do guiding reasons wash their hands of the quality of our reasoning, so long as we do the right thing in the end? Not a bit of it. In a moment we will come to the special 'second-order' cases of reasons to act for reasons, and reasons not to do so. But even before we come to these, there are two ways in which ordinary guiding reasons for action and belief necessarily have something to say on the subject of explanatory reasons. The first point is permissive. Other things being equal, the fact that such-and-such is a reason for a given action or belief makes it *alright* for this to be the reason why one performs the action or holds the belief. The second point of contact is rather more prescriptive. For it is a basic principle of practical rationality, which is at the root of the whole pros/cons asymmetry, that one should always act for *some* undefeated reason, i.e. that at least one of the (guiding) reasons in favour of doing as one did should have been one's (explanatory) reason for doing it.

You may well be suspicious of any appeal to the 'basic principles of practical rationality.' Is this not just a device to win an argument by fiat, to make its conclusion sound incontrovertible when in fact it should be vigorously interrogated? I sincerely hope

[17] Needless to say, some regard this as a positive merit of the Kantian line of thought: see Barbara Herman, 'Obligation and Performance: A Kantian Account of Moral Conflict', in Owen Flanagan and Amélie Rorty (eds.), *Identity, Character, and Morality* (Cambridge, Mass. 1990), 311.

[18] See his 'Competing Theories of Justification: Deeds vs Reasons', this volume, 000.

not. In one way, it is true, principles of practical rationality are more fundamental than other practical principles. For they apply across the boundaries between different fields of practical reasoning. They apply equally in legal reasoning, moral reasoning, strategic reasoning, etc., and do so just by virtue of the fact that these are fields of practical reasoning. This may mean that they have a slight air of the *a priori* about them. But in fact this air is deceptive. Principles of rationality have a rather more mundane, *a postiori* basis than many ordinary moral and legal principles. For principles of rationality exist purely to guide us towards conformity with moral reasons, legal reasons, and so on, whatever those reasons may happen to be. They are sound or valid principles of rationality if it is true that, when we conform our reasoning to them, we also better conform our actions and beliefs to whatever reasons for action and belief apply to us. Thus principles of rationality stand or fall, in large measure, on their straightforwardly instrumental merits.[19] The basic principle of practical rationality that tells us always to act for some undefeated reason is no exception. For while it is true that, by default, guiding reasons are there to guide actions and beliefs rather than to guide explanatory reasons for actions and beliefs, it is also important to remember that the only way in which actions and beliefs can be guided is through the guidance of explanatory reasons. Reasons can only move us, if you like, by motivating us. And the principle that we should always act for some undefeated reason is a principle which identifies what we should be looking for, in the way of motivation, if we are to maximise our prospects of acting as reason demands. We should look for an undefeated reason to act as we do. There is no point in going further and looking for *all* the undefeated reasons to act in that way.[20] That would be ridiculous overkill. The very fact

[19] That explains why, in the work of some philosophers, the label 'principles of rationality' has been appropriated to designate the principles we should adopt if we are better to secure whatever ends we each happen to have, without touching on the intrinsic merit of those ends. See e.g. Derek Parfit, *Reasons and Persons* (Oxford 1984), 4. This also underlies the popular use of 'rational' when associated with the calculating pursuit of personal advancement.

[20] But isn't it necessary to survey all the reasons, if only in order to *work out* that the reason for which one acts is indeed undefeated? Not necessarily. On the view defended here, the fact that one

that the reason for which one acts is undefeated is enough by itself to guarantee that, all things considered, it is on the winning side in the whole overarching conflict of reasons.[21] Thus the basic principle of rationality which I cited, although modest, is as strong a principle as we need to ensure that our explanatory reasons for action reliably push us towards an action that ought, all things considered, to be performed.

These claims cut against a familiar model of rational conflict. On this model, the reasons for and the reasons against some action or belief are always up to a point mutually defeating. Some of the reasons in favour of an action or belief devote their entire motivating energy to defeating the reasons against that action or belief, and in the process become spent forces, so that the only undefeated reasons in the end are those which were not needed for the battle, the reasons which make up, so to speak, the *net advantage* which the pros had over the cons in terms of their rational force. These are therefore the only reasons left for us to act upon, if it is true that one should always act for an undefeated reason. But my claims in defence of the principle that one should always act for some undefeated reason suggest an alternative picture, according to which, if the reasons in favour of some action defeat the reasons against, then in the ordinary case it is only the reasons against which end up defeated. The reasons in favour are *all of them* undefeated, i.e.

---

acted for an undefeated reason by instinct or habit need not in any way detract from one's justification. Sometimes we identify undefeated reasons better without stopping to think. There is no basis for the view that fully deliberated action is more reliably sound than spontaneous action: often, the more we think about it, the worse our predicament gets. That is one reason (among many) why such things as emotions, passions, and raw desires are so crucial to any well-rounded life. It incidentally helps to dispel another criminal lawyer's myth: that justifications are more at home in situations where there is proper scope for clear thinking, while in situations of emergency, calling for immediate reactions, we must make do with excuses.

[21] Or perhaps we should say, to accommodate cases of incomparability, 'not on the losing side'. In some cases one acts on undefeated reasons whichever way one acts, and one is justified either way, since incomparability prevents the reasons in favour from defeating the reasons against or vice versa. Where the undefeated reasons on both sides are *protected* reasons (on which see below) these are normally known as 'dilemma' cases. For a good account of which, see Ruth Marcus, 'Moral Dilemmas and Consistency', *Journal of Philosophy* 77 (1980), 121.

they are undefeated *in gross*. The principle of rationality which demands that one always acts for some undefeated reason is therefore compatible, in the ordinary situation of an action with more pros than cons, with our acting on any one of the various reasons in favour of the action. This makes it a much more modest principle than at first sight seems to be needed to defend my promised strong conclusion that reasons in favour of an action are only justificatory if they are also the reasons for which one performed that action. And yet the modesty of this principle of rationality is also, in another way, its strength. For what this principle means is that it is quite pointless to cite, by way of justification, an undefeated reason for which one did not act, even though it would have been alright for one to act upon it if one had been minded to do so. Once one has attempted to make justificatory capital out of such a guiding but non-explanatory reason, one must still go on to identify some *other* undefeated reason which one *did* act upon in order to clinch the justification. For one must always act for some undefeated reason. But once one achieves this, the original non-explanatory reason one cited simply drops out of the picture. It adds nothing to one's case. The citation of the second reason was both necessary (it was acted upon) and sufficient (it was undefeated) to identify one's action as justified. Essentially, that is why justifications must have two matching parts: in the first place, an undefeated guiding reason, which is also, secondly, the explanatory reason for the justified action.

Notice that the reasons against one's action, which a full justification is needed to defeat, are not similarly two-pronged. It is enough that they were guiding reasons against one's action. For as I already said, there is no such thing as an explanatory reason against an action, so there can be no demand, concerning the reasons against an action, that in order to be counted against that action's performance, they must also play some explanatory role in that performance. What this means, in the everyday terms most familiar to lawyers, is that undesirable side-effects as well as undesirable intended results can count against an action, while only desirable intended results can ever count in its favour.[22] This is the basic

[22] Compare Andrew Simester, 'Intention, Recklessness, and Morality', this volume, 000. My position in this essay directly confronts Simester's view that there are two alternative bases of

all-pervading asymmetry between pros and cons, between reasons in favour and reasons against, between the pursuit of positive value and the avoidance of negative, that comprehensively carves up our lives as rational beings.

3. Fortifying the asymmetry

This basic asymmetry makes its presence felt in many ways. Most obviously and famously, it structures the distinction between people's virtues and their failings, and the related distinction between credit and blame. Ironically, this may be one of the factors which leads some to suppose that the asymmetry can have no impact on the logic of justification.[23] Their thinking goes something like this. It is one thing to say that an action is fully justified, so that blame can be eliminated. But it is a long way from there to the conclusion that the action is positively creditable. It is only when we take this further evaluative step that the claimed asymmetry between pros and cons can have a real impact. For there can be little doubt that whether an action is creditable turns primarily on the reasons for which it is performed. By making these very same explanatory reasons the key to justification, one wrongly eliminates the logical gap between justified and creditable action. One makes the earning of credit an automatic corollary of the mere elimination of blame.

---

blame, one connected with 'motive' and the other with 'reasonableness'. Reasonableness in action depends straightforwardly on the reasons for which one acts. Simester's examples do not contraindicate. In what he calls the 'Experiment Case', a great deal turns on the fact that (as Simester confesses in his note 67) the 'neutral end' is not in fact neutral, but is a reason in favour of administering the drug. On my account, *if* this reason in favour is undefeated, which on the information given by Simester it might indeed be, *then* the administration of the drug can indeed be justified. It matters not that the motive is not good, so that the action is not creditable. As I explain in the next section, the question of whether a reason for acting is good, worthy, etc. is not the same as the question of whether it is undefeated.

[23] Thus Glanville Williams, *Criminal Law: The General Part* (2nd ed., London 1961), criticising the decision in *Dadson*, above note 7, complains that 'the law of consummated crimes … governs conduct, not purity of intention.'

The mistake here lies in the failure to bear in mind that reasons for action may vary in quality as well as strength. As I have already explained, other things being equal it is alright to act for any undefeated reason in favour of one's action. But sometimes it is *more* than alright to act for a particular undefeated reason. Sometimes there is also an undefeated second-order reason to act for that particular reason, e.g. a reason to support one's friends out of affection for them, to refuse conscription out of a sense of humanitarian duty, to tell the truth because it is the truth, etc. It does not mean that the other reasons for doing what one does automatically become unacceptable, so that acting on them can provide no justification. It means only that a particular reason is privileged, so that acting for that reason lends special value to one's action, value going beyond what is needed for mere justification. Often, indeed, there are several privileged reasons, and there arises a question of which of these one should act for. At this level incomparability of value often precludes a straight answer. By acting in the same way for different privileged reasons people normally exhibit their incomparably different virtues. But whether or not one actually exhibits a virtue in the process, it is creditable to act for such privileged reasons in any case in which both they and the second-order reasons to act for them remain undefeated. This is not, I hasten to add, the only way in which one may earn credit. One's actions may be creditable because of *how* they are performed rather than why, in which case one should look to the value of skill rather than the value of virtue to supply the explanation. But either way more is required for credit than is required for mere justification. An action may still be justified even though performed without any technical proficiency and for a most banal, trivial, and unimpressive reason. But the key point still stands that, to be justified, it must nevertheless have been performed for an *undefeated* reason.[24]

[24] In some extreme cases, if one cannot act for particular reasons, or at a certain level of technical proficiency, then that strips out so much value from one's act that it would be better not to perform it at all. In such cases, the gap narrows between the point at which blame ends and the point at which credit begins.

Does it follow that second-order reasons are irrelevant to justification? If only things were that simple. In many contexts, and notably in the law, the scope of justification is pervasively determined by the operation of second-order reasons. But these are not reasons to act for reasons, of the kind which allow for creditable action. These are reasons *not* to act for reasons. Such negative second-order reasons are in operation whenever an action is required or forbidden. They are therefore the reasons at the heart of *wrongdoing*. This claim should be interpreted with care. We describe actions as 'wrong' in several senses. I already distinguished prima facie wrongness from the all-things-considered wrongness. At the same time one should be aware of a cross-cutting distinction between *advisory* wrongness and *mandatory* wrongness. Actions which are all-things-considered wrong in either the mandatory or the advisory sense can, in appropriate cases, be equally a source of blame. But the distinction between mandatory and advisory wrongness is nevertheless one of great importance. One does the wrong thing in the advisory sense if one does not do what one has reason to do. One does the wrong thing in the mandatory sense – for which we normally reserve more emphatic terms like 'wrongdoing', 'wrongdoer' and 'wrongful' – if one does not do what one has, in Joseph Raz's terms, a *protected* reason to do.[25] A protected reason differs from what I earlier labelled a 'privileged' reason. A privileged reason is a reason which one has reason to act for. A protected reason, on the other hand, is a reason for an action combined with a reason not to act for some or all of the reasons against that action. What happens when a reason is protected is that certain countervailing considerations are defeated in advance of any practical conflict. Because these reasons are pre-emptively defeated, action for these reasons does not meet the demands of the basic principle of rationality that one should always act for some undefeated reason. It may be asked how a reason can be defeated in advance of a conflict. Surely whether it is defeated depends on its relative strength when pitted against another reason, and therefore must await the conflict? This ignores the role of *rules* in practical reasoning. Rules are devices

[25] Raz, *The Authority of Law* (Oxford 1979), 17–9.

which improve our prospects of doing what reason demands by settling certain conflicts of reasons before they arise. They obviate the need for reliance on some of the raw pros and cons. It does not mean, as some have thought, that all rules are merely 'rules of thumb', or 'indicator rules', which provide a prima facie reason to believe that the action ought to be performed, without affecting whether the action ought to be performed.[26] On the contrary, real rules are capable of affecting what really ought to be done. In rough outline, they do so because some guiding reasons in favour of certain actions are less likely to be properly followed, or more likely to corrupt the following of other reasons, if one tries to follow them. They are guiding reasons which had better not become explanatory reasons. Since, by the time they are explanatory reasons, it is too late to avoid their distorting effect, they are ruled out of being explanatory in advance by the operation of the rule. Consequently, of course, the price of following the rule is sometimes that one does not act as the underlying reasons apart from the rule would have one act. That in such cases one should nevertheless follow the rule is the small price that rationality pays for avoiding the risk that, by trying to act for some of the raw underlying reasons, one will otherwise very often act against reason. That explains why there are reasons not to act for certain reasons, and why some actions are as a result not merely advisable or inadvisable, but are actually required or prohibited.

In fulfilling its primary functions, the law rarely classifies actions as merely advisable or inadvisable.[27] But it often classifies them as required or prohibited. That is particularly

[26] I.e. that they are prima facie in the lawyer's evidential sense. Some of W.D. Ross's remarks in *The Right and the Good* (Oxford 1930) may lead one to think that all sound mandatory rules are prima facie in this sense. See also J.J.C. Smart, 'An Outline of a System of Utilitarian Ethics' in J.J.C. Smart and Bernard Williams (eds.), *Utilitarianism: For and Against* (Cambridge 1973), at 42ff. On rules of thumb and indicator rules see Frederick Schauer, *Playing by the Rules* (Oxford 1991), 104ff (and especially the note at 105).

[27] By 'primary functions' I mean non-self-regulatory functions.. Among the law's functions are the governance of legislation and adjudication. In the fulfilment of these functions the law often avails itself of principles and values which provide merely advisory legal guidance to officials. But of

apparent in the criminal law. The very idea of a 'criminal offence' is the idea of an action which is in the eyes of the law not merely wrong but wrongful, i.e. which there is, in the eyes of the law, not merely a reason but a protected reason not to perform. In fact the protected reason which the law creates is, by default, *absolutely* protected. So far as the criminal law is concerned *all* reasons in favour of performing the criminalised action are defeated by virtue of the law's unquestionable and all-embracing authority. It means that one is left with no automatic access to any justificatory considerations, however powerful they might be apart from the law. What the law does, which nevertheless creates a role for some justificatory defences, is to provide us with *cancelling permissions* to perform, under certain specified conditions, the actions which it criminalises. This may seem like a rather surprising proposition. After all, it was argued above that justificatory reasons do not *cancel* but rather *defeat* the reasons against an action or belief. Now, by contrast, the claim appears to be that justifications arise in the criminal law precisely when reasons are cancelled by permissions. But the air of paradox is dispelled as soon as one realises that the law's cancelling permissions do not cancel the reasons not to perform the criminalised action, but merely cancel the second-order protective reasons not to act for certain countervailing reasons. Thus justificatory arguments which the law would otherwise disallow are specifically allowed. This means that the law does not only regulate the actions which one may perform, but also regulates the reasons for which one may perform those actions. It regulates the actions which one may perform by making some of them into criminal offences, which are prima facie wrongs. It regulates the reasons for which one may perform those actions by picking out certain reasons in favour of their performance as legally acceptable reasons. But, as ever, one benefits from the acceptable reasons in favour of one's action only if one actually acts for these reasons. For even legal rationality, with all

---

course these functions are parasitic on others. The primary functions of law are its functions in affecting people's actions other than legislative and adjudicatory actions – if you like, outward-looking rather than inward-looking functions. Here the law rarely makes do with mere advice, since it cannot normally count on automatic cooperation. On primary and secondary functions, see Raz, *The Authority of Law*, above note 25, 163ff.

of its second-order protection, is governed by the basic principle of practical rationality that one should always act for some undefeated reason. In the case of a criminal act, all reasons are defeated apart from those permitted by law. Thus to claim a justificatory defence, one must not only have, but act for, one of those permitted reasons.

The same is not true in a case to which the definition of the crime does not extend, i.e. in which 'an element of the offence is negated'. In such a case one need rely on no permission to act for specified reasons since the protected reason, which gives rise to the need for such permission, does not apply in the first place. So it does not matter why the defendant acted as he did, so long as that is how he acted. Here we see exactly what the English law student learns by reading *Deller* alongside *Dadson*: in some cases the defendant's motivation is beside the point, whereas in others it is decisive.[28] The real difficulty arises, however, when we try to draw the line in practice between the two types of case, between *Dadson* cases and *Deller* cases. Where we are dealing with raw pros and cons, it is easy to distinguish a justificatory argument from an argument to effect that there is nothing to justify. A justificatory argument is an argument which points to reasons in favour of the action performed rather than to the absence of reasons against it. But when protected reasons enter the equation, this test is no longer adequate to draw the distinction. That is because some of the reasons in favour of an action are already taken into account in the structure of the protected reason. The rule has been shaped with an eye to many of the pros and cons of action in accordance with it. It means that sometimes, when one cites a reason in favour of one's action, one cites a consideration which should bear on the interpretation of the rule one is accused of violating rather than a consideration which bears on whether one is justified in violating it. Accordingly, the mere fact that one points to a reason in favour of one's action does not mean, in this context, that one asserts a justification as opposed to denying the application of the law to the case. Only if one

[28] *R* v *Deller* (1952) 36 Cr. App. R. 184; *R* v *Dadson*, above note 7 See the excellent discussion in J.C. Smith and Brian Hogan, *Criminal Law* (7th ed., London 1992), 33–5.

asserts a justification, however, does it matter whether one actually acted for the reason in favour that one cites. It follows that in law one may sometimes benefit from a reason in favour of one's action which was not a reason for which one acted, even though in other cases one must have acted for the reason before one may benefit in law from its application. Unfortunately, there is no general test for telling the two kinds of case apart. It is a question of law, on which different legal systems may of course arrive at different answers, whether a given argument is to be regarded as justificatory rather than as bearing on the scope of the offence. In fact, for all I know, some legal systems may not bother to recognise justifications at all, so confident are they of the moral finesse of their offence definitions. This would tend to lumber them with a regrettable inflexibility, or force them to undesirable vagueness, or both. I will return to these possible moral objections below. But logically it is a perfectly possible solution. It has the pay-off that the legal system in question may take an interest in the reasons in favour of the action which the defendant performed without caring whether the defendant acted for those reasons. All I wish to stress is that this does not turn it into a legal system which endorses an 'objective' rather than a 'subjective' theory of justification, or into a legal system which allows justification to depend upon guiding reasons irrespective of explanatory reasons. It makes it into a legal system which, strictly speaking, does not care about justification at all.

4. The priority of justification over excuse

It is widely thought that excuses are more 'subjective' than justifications.[29] In one sense of 'subjective', as we will see, this is perfectly true. But it is not true if we are using the labels 'subjective' and 'objective' to mark the contrast between explanatory and guiding reasons. Over a wide range of cases, excuses, just like justifications, depend on the union of

[29] Kent Greenawalt, 'The Perplexing Borders of Justification and Excuse', *Columbia Law Review* 84 (1984), 1897 at 1915–18.

explanatory and guiding reasons. Whenever excuses depend on the union of explanatory and guiding reasons, moreover, they do so precisely because justifications depend on the union of explanatory and guiding reasons. The structure of excuse derives, in other words, from the structure of justification, and thus shares in its combination of subjective (explanatory) and objective (guiding) rationality.

Some theorists have associated excuses with character traits.[30] They are mistaken if they think that every excuse is concerned with character. Many excuses are of a technical nature. They relate to levels of skill rather than degrees of virtue. Their gist is that the person claiming them does not possess the skills needed to do better, and should not be expected to possess those skills. Whether one should be expected to possess certain skills, or skills to a certain degree, depends, to some extent, on one's form of life. A doctor who tries to excuse her blundering treatment by claiming lack of diagnostic skill should not get far, whereas an amateur first-aider may be able to extinguish her blame, under similar conditions, by making exactly the same argument. But such excuses, even though they are of great legal importance, will not concern us here. Our concern will be with that major group of excuses which do indeed relate to character evaluation. These include excuses very familiar to criminal lawyers, such as excuses based on provocation and duress. Their gist is similar to that of technical excuses. It is that the person claiming them does not possess the virtues needed to do better, and should not be expected to possess those virtues. Again, which virtues one should be expected to possess, and to what extent, depends largely on one's form of life. A police officer is expected to exhibit more fortitude and courage than an ordinary member of the public, a friend is expected to be more considerate and attentive than a stranger, etc. What exactly does this mean? Essentially, it means that where there is a conflict of reasons, some people are expected to act for some reasons, whereas others are expected to act for other, often incompatible and

[30] See the useful bibliographical note in Michael Moore, 'Choice Character, and Excuse', *Social Philosophy and Policy* 7 (1990), 29 at 40–1. A subtle and sympathetic reconstruction of this approach to excuses is offered by Bob Sullivan in 'Making Excuses', this volume, 000.

incomparable, reasons. But obviously the need to claim an excuse from one's action arises only if one fails to establish a full justification. A fully justified action needs no excuse. So the point cannot be that those who act with excuse act for undefeated reasons, i.e. that it is alright for them to act for those reasons. That would yield a full justification for their actions. The point must be that there is something suspect about the reasons for which they act. And indeed there is. They are not valid reasons. They are what the person acting upon them takes to be valid reasons, *and justifiably so*. Thus the structure of excuse derives from the structure of justification. To excuse an action is not, of course, to justify that action. Rather, one justifies one's belief that the action is justified.[31]

This explanation of non-technical excuses has to be modified and extended somewhat to accommodate unjustified actions upon justified emotions, attitudes, passions, desires, decisions, etc. as well as unjustified actions upon justified beliefs. Provocation, as Jeremy Horder has explained, involves unjustified action out of justified anger.[32] Duress, or a certain central kind of duress, can be similarly analysed as involving unjustified action out of justified fear. But these are, in a sense, derivative cases. Emotions like anger and fear are mediating forces between beliefs and actions. They enhance or constrain the motivating force of certain motivating beliefs. Their justification therefore turns in part on the justification of the beliefs which partly constitute them. Of course, there is still a justificatory gap: an emotion is not fully justified merely by the justification of its cognitive component. But justified emotion (and in similar vein justified attitude or desire or

[31] C.f. Suzanne Uniake, *Permissible Killing: The Self-Defence Justification of Homicide* (Cambridge 1994), 15–25.

[32] *Provocation and Responsibility* (Oxford 1992), 158ff. This explains why women who have killed after prolonged domestic violence and are denied provocation defences on the ground that their reaction was not immediate do not much like the response that they should instead claim 'diminished responsibility'. They want it to be acknowledged that their anger was (at least partly) justified, and that *this* is why their admittedly unjustified action is (at least partly) excused. 'Diminished responsibility' is a claim which suggests irrationality all the way down (meaning that it is not strictly an excuse: see note 35 below). It thus reduces the moral status of those who claim it.

decision) nevertheless entails justified belief. Thus the most basic or rudimentary case of non-technical excuse remains that of unjustified action upon justified belief. One must therefore consider what is needed to make a belief justified. It is of course one of the great problems of epistemology, and we cannot do justice to it here. Suffice it to say that the general account of justification applicable to action is also broadly applicable to belief. One must have an undefeated reason for one's belief, and that must moreover be the reason why one holds the belief. This explains the nature of epistemic faults, such as prejudice, gullibility and superstition. One cannot understand these faults unless one appreciates that a belief is justified, not only by the reasons there are for holding it, but also by the process of reasoning by which it came to be held, i.e. not only by guiding reasons but also by explanatory reasons. The same facts also explain why a requirement of *reasonableness* has traditionally been imposed upon excuses in the criminal law. It is not enough that one made a mistake as to justification, if it was not a reasonable mistake, it is not enough that one was angry to the point of losing self-control, if one's anger was not reasonable, etc.[33] By 'reasonable' here is meant, in my view, much the same as 'justified'. There must have been an undefeated reason for one's belief, emotion, etc. which also explains why one held the belief or experienced the emotion, etc. The fact that sometimes this element of reasonableness is dispensed with in the law does not show a drift towards a more purely 'subjective' account of excuses, i.e. one depending on explanatory reasons without regard for guiding reasons. It shows, rather, that some excuse-like arguments, in common with some justification-like arguments, may actually serve to negate an element of the offence rather than to excuse or justify its commission. Some mistakes, as the courts put it, may simply serve to negate the mens rea for the particular crime; and if, as may be, the mens rea required is, e.g. knowledge, then of course the reasonableness of one's mistake is neither

[33] See, among many examples, *Albert* v. *Lavin* [1981] 1 All ER 628 (reasonable belief required for mistaken self-defence), *R* v *Graham* [1982] 1 All ER 801 ('fear for good cause' required for duress), *Phillips* v *R* [1969] 2 AC 130 9 (loss of self-control must be reasonable for provocation).

here nor there.[34] The extent to which legal systems will tolerate such arguments depends on many contingencies about them, including the extent to which and way in which they implement the demand for mens rea. But this has nothing to do with excuses, in which an element of reasonableness, at some level, is conceptually necessitated whether the crime is one of full subjective mens rea or one of no mens rea at all.

Requirements of 'reasonableness' in criminal excuses also sometimes go beyond what the logic of excuses requires, and in that case they normally serve another role. They serve to orientate the law towards general application to people living many different forms of life, rather than tailoring it to suit the expected virtues of a certain kind of person leading a certain kind of life. The debate about the extent to which the reasonable person should be 'individualised' to the characteristics of the defendant in the definition of criminal excuses is partly a debate about the extent to which the criminal law should aspire to this kind of generality. Should the 'reasonable person' in provocation become the 'reasonable police officer' when the defendant is a police officer? Should the 'reasonable person' in cases of drunken mistake become the 'reasonable drunkard'? Once again there is no universal theoretical solution to this problem. Within broad limits, legal systems may quite properly vary in their willingness to individualise excuses and the general principles, if any, upon which they do so. But legal systems cannot, consistent with the logic of (non-technical) excuses, vary in the importance they attach to the combination of guiding and explanatory reasons in the excusatory scheme of things. Thus they cannot altogether eliminate the

[34] That is the logic of the decisions in *R* v *Williams* [1987] 3 All ER 411 and *Beckford* v *R* [1987] 3 All ER 425, which should therefore not be understood as authorities on the mental element in excuses, let alone as authorities on the mental element in defences in general. At most they show that self-defence is no longer regarded as a genuine defence in English criminal law. Instead absence of self-defence is regarded as an implied element of every offence, and one to which an implied element of mens rea is automatically attached. It is an absurd rule, and one which should be overturned, but the reliance on *DPP* v *Morgan* [1975] 2 All ER 347 in both of the cases makes it the only viable interpretation of what they stand for. For excellent discussion, see Andrew Simester, 'Mistakes in Defence', *Oxford Journal of Legal Studies* 12 (1992), 295.

essential 'objective' dimension of excusatory claims. They cannot ignore the important point that excuses rely on reason, not on the absence of it. That is, they rely on the ability of the person who claims to be excused to believe and feel as reason demands, and because reason demands it. Those people who cannot meet this condition do not need to bother making excuses. Such people are not responsible for their actions, and are free from blame as well as being improper targets for criminal liability, irrespective of both justification and excuse.[35] Justification and excuse both belong to the realm of responsible agency, and that is precisely because both depend, to put it crudely, on the ability to live within reason.

The logical relationship between justification and (non-technical) excuse helps to explain the so-called 'quasi-justificatory drift' of many familiar excuses.[36] In English law this is compounded by the law's cautious insistence on having a belt as well as braces: in general no excuse is accepted into the criminal law which is not also a partial justification,

[35] Notice that the effect of this claim is to deny the status of excuses to, e.g. insanity, infancy, and 'diminished responsibility'. I do not shrink from this pay-off. Excuses have a built-in precariousness. 'Don't make excuses' is sometimes a legitimate stricture. For many actions are inexcusable in some situations thanks to the fact that everyone is expected to have the virtues and skills necessary to perform them in those situations. But even in connection with these actions in these situations we should not blame the very young or (often enough) those suffering from serious mental illness. Notice that this does not commit me to the view that very seriously mentally ill people should get away with everything they do. I share Anthony Kenny's view that most mental illness is selective in its impairment of rationality, and should only preclude blame by extinguishing responsibility where relevantly operative. See Kenny, *Freewill and Responsibility* (London 1978), 80–4. Even so, that is not enough to make an argument based on mental illness into an excusatory argument, since whether mental illness affects blame depends not on the nature of the action but on the relevance of the illness to its performance. Illustrated crudely: even if cannibalism were inexcusable, some mad cannibals would not be to blame for eating people. The blameless mad cannibals would be those mad cannibals, roughly speaking, who were cannibals because mad.

[36] I borrow the label from Simon Gardner, 'Instrumentalism and Necessity', *Oxford Journal of Legal Studies* 6 (1986), 431 at 433.

and no justification is accepted which is not also a partial excuse.[37] The drift of the excuse is not so much quasi-justificatory as truly justificatory. But neither of these facts should obscure the crucial conceptual distinction between justification and excuse. Nor should one be distracted by the paradoxical sound of the claim that an action which is justifiably believed to be justified is excused rather than justified. It only goes to show that, as between the concepts of justification and excuse, justification is the more fundamental. The same proposition also brings out the true sense in which excuses may be regarded as more 'subjective' than justifications. For by their nature excuses take the world as the defendant justifiably sees it rather than as it is. They look to what the defendant believes to applicable reasons for action, so long as she does so on the basis of genuinely applicable reasons for belief. Justifications, meanwhile, look directly to the genuinely applicable reasons for action, without stopping to look for applicable reasons for belief. But in this whole contrast the talk of 'reasons' is talk of guiding reasons. It leaves on one side the fact that, in both justification and excuse, explanatory reasons also play a key role, and that, in this sense and to this extent, each is just as subjective as the other.

5. Institutional objections

Many of the arguments which lawyers give for dividing up justification and excuse along different lines from these are of a broadly institutional character. They are based, not on general considerations of rationality and value, but on views about the limitations which are imposed upon the logic of the criminal law by its legitimate social functions and roles. Such arguments often lead to the conclusion that, while the general spirit of justification and

[37] Hence the Law Commission's difficulties in *Legislating the Criminal Code: Offences Against the Person and General Principles* (London 1993), 63–4. The Commission finds excusatory as well as justificatory strands in the cases on 'necessity' and jumps to the conclusion that there must be two defences rather than one. Not so: one defence, compound rationale.

excuse as I have explained them may remain broadly visible and important in the law, the distinction is not one to which the law can or should give exact doctrinal expression. Translation into the legal context means dispensing with some of the finer points. I will briefly mention two arguments which may be thought to point in this direction.

a. The rule of law. One alluring line of argument goes something like this. The question of whether an action is justified is, on any view, a question of whether it ought all things considered to be performed. If they are to have a role in the criminal law, justifications must therefore serve as guidance to potential offenders as to what, in law, they ought to be doing. Justification doctrines must belong to the law's 'conduct rules'.[38] But if it is to be possible to rely upon the law's justificatory doctrines in one's reasoning about what to do, it must be permissible in law to do as the justificatory doctrines require *for the very reason that they are part of the law*. Accordingly, the law cannot consistently demand, in the guidance it gives to potential offenders, that the action be performed instead for some *other* reason. It means that the law must introduce a schism between guiding and explanatory reasons in its institutional adaptation of the justificatory framework. The relevant guiding reasons are, of course, those which the law mentions as justificatory: the fact that one is under attack and self-defence is called for, the fact that crimes are being committed and need to be prevented, and so forth. But the law cannot demand that these also be the explanatory reasons for one's act of self-defence or crime-prevention. The best it can demand by way of explanatory reason is that one acted thus because that is how the law would have one

[38] The contrast between 'conduct rules' and 'adjudication rules' which I introduce here is drawn from Paul Robinson, 'Rules of Conduct and Principles of Adjudication', *University of Chicago Law Review* 57 (1990), 729. Similar contrasts have been drawn by Peter Alldridge in 'Rules for Courts and Rules for Citizens', *Oxford Journal of Legal Studies* 10 (1990), 487 and by Meir Dan-Cohen, 'Decision Rules and Conduct Rules: On Acoustic Separation in Criminal Law', *Harvard Law Review* 97 (1984), 625. All three authors make interesting remarks on where justificatory defences might fit into such a scheme, and how rule of law requirements might apply to them. I should stress that none of the three is seduced by the argument being outlined here, although none of them is entirely immune to its charms either. Dan-Cohen makes the best job of resisting it.

act. Thus even if there is a pros/cons asymmetry which lies at the heart of rationality, it cannot in principle be directly replicated in the institutional context of the criminal law. When justifications are being framed, the law can only state the guiding reasons, and leave the explanatory reasons to look after themselves. Explanatory reasons, however, can readily be counted on the excusatory axis instead. For excuses do not, on any view, bear on what all things considered one ought to be doing. Legal excuses therefore do not belong to the law's 'conduct rules' which are there to guide potential offenders, but rather to the 'adjudication rules' which are there to guide judges in dealing with potential offenders. They can thus remain sensitive to considerations which cannot in principle be the subject of guidance directed to potential offenders. It means that, while justifications become more 'objective' in their adaptation to the law's demands, excuses end up taking all the more 'subjective' elements under their wing.

This argument, which I have made as good as I can make it, nevertheless harbours many errors. At its heart lies the idea that justifications, as I analyse them, cannot play a role in the criminal law without violating the rule of law, which requires that the law's conduct rules should be capable of guiding those who are subject to them. But this complaint betrays a false assumption about the sense in which legal justifications are there to provide guidance to potential offenders in the first place. When the law grants a justification, as I explained above, it provides a *cancelling permission* to act for certain reasons which would otherwise be automatically defeated by the prohibition. But a permission to do something is, by itself, no reason to do it. Thus the law does not provide any reasons for one to do what the law holds to be justified. It simply allows that one may have such reasons and act on them. To say that justificatory rules belong to the 'conduct rules' of the law, and must serve to 'guide' potential offenders, is to give the impression that the law gives people positive reasons to do what the justificatory rules allow. But the law does no such thing. Thus the idea of someone who tries to follow a justificatory rule of criminal law, in the sense of acting because of the rule, is the idea of someone who mistakes it for something quite different from what it is. Anyone who sees a justificatory legal rule for

what it really is will know that it cannot in itself motivate action in accordance with it since it gives no reasons for action but only cancels the law's otherwise pre-emptive reasons for not acting on certain independent reasons for the justified action, the latter being the reasons which account for its being justified. So there is no point in a law which attempts to turn its justificatory rules into rules which can be followed directly by potential offenders. In the process all that the law does is to defeat its own object, which is that people should act for the reasons which the law permits them to act for, and not for other reasons which the legal prohibition preempts. That being so, perhaps justificatory rules of law are not best labelled as 'conduct rules' at all. Of course it is true that sometimes the law *combines* them with what are more perspicuously thought of as 'conduct rules.' Sometimes a police officer may be required by law to arrest using reasonable force, as well as being justified in using such force, which would otherwise be criminal, in effecting a legally required arrest. Such cases may be thought of as self-referential, since they look within the law rather than beyond it for the justifying reasons. But again they create no schism between explanatory and guiding reasons which could conceivably affect the conditions of legal justification. For the legally recognised and protected guiding reason for the arrest must also be, in such cases, among the reasons why the arrest was made if it is to count as fully justified in law. In other words, a police officer who does not act upon the legal rule requiring arrest in such a situation also does not benefit from it.[39]

[39] Opponents will no doubt respond by pointing out that, in English law at least, arrests may be made lawful *either* by the fact that they were made on reasonable suspicion of the commission of an offence *or* by the fact that, although the arrest was not made on reasonable suspicion, it turned out that a real offence was committed: Police and Criminal Evidence Act 1984, s24. The second option is a special case of justification, normally known as *vindication*. It is the case in which one is justified in taking a chance that one will turn out to be right. It is not a counterexample to my account of justification nor to my remarks on arrest, since the law still requires, in the vindication cases in s24, that the action be an *arrest*, which is *by its nature* an action performed for certain legal reasons. See J.C. Smith, *Justification and Excuse in the Criminal Law* (London 1989), 33–4. Tony Honoré has suggested to me, however, that there is a more general issue here. Cases of vindication, he believes, point to the need for me to make certain modifications to my analysis of justification.

These remarks show that a stark distinction between 'conduct rules' and 'adjudication rules' is inadequate to capture the complex inner logic of the criminal law. But they also show, for our purposes more importantly, that the rule of law does not militate against a role in the criminal law for justifications as I analyse them. On the contrary, it tends to count in favour of allowing them a significant role. I said earlier that legal systems may differ in the way in which they deal with particular rational conflicts in the structure of criminal liability. Some may incorporate into the very definition of the offence the same considerations which others treat as going purely to justification. But I also added that legal systems which try to follow the former route exclusively are apt to suffer from moral shortcomings, either being excessively rigid or excessively vague. Why do these dangers come of failure to recognise justificatory defences? The answer relates, predictably, to the demands of the rule of law.[40] In a legal system which adheres conscientiously to the rule of law, offence definitions will be so far as possible clear, accessible and certain in their application, so that they can be used for guidance by potential offenders as well as by courts and officials.[41] This means that actions which will fall outside the law must be largely decided upon in advance, and closure rules provided for any unexpected cases. But defences in general, and justifications in particular, are largely exempt from these tough demands which the rule of law places upon the definition and drafting of offences. I have

---

In these cases, one only has a hunch that undefeated guiding reasons exist, and one acts on this hunch. This, in Honoré's words, represents 'a tertium quid between acting for the guiding reason and acting as the guiding reason requires but not for it.' I am not convinced. In vindication cases one does not act because of one's hunch, but because, if I may put it this way, of what it is one hunches. What one hunches is the guiding reason. It is the fact that one acts for this reason, always assuming it is undefeated, that justifies one's taking the chance that one's hunch is right. That one hunches the reason rather than, e.g. knowing of it, does not diminish its explanatory role.

[40] The answer is suggested in George Fletcher, 'The Nature of Justification' in Stephen Shute, John Gardner and Jeremy Horder (eds.), *Action and Value in Criminal Law*, above note 4, 175.

[41] I have explored some aspects of these demands, and in particular the 'so far as possible' proviso, in 'Rationality and the Rule of Law in Offences Againat the Person', *Cambridge Law Journal* 53 (1994), 502.

just explained why. Legal justifications are not there to be directly followed by potential offenders. They merely permit one to follow reasons which would otherwise have been pre-emptively defeated. So there is no need for them to aspire to standards which apply to legal rules when they *are* there to be directly followed. It means that justifications can introduce an element of flexibility into the law which often cannot be accommodated, compatibly with the exacting demands of the rule of law, in the very definition of the offence. A legal system which tries to do without justifications in its criminal law is thus likely either to violate the rule of law by allowing its offence definitions to remain vague enough to accommodate judicial deliberations in novel and difficult cases, or else it is likely to conform to the rule of law in its offence definitions but at the high price that the novel and difficult cases will be decided without adequate scope for judicial deliberation. Justificatory defences provide a way out of the dilemma. It is not true that the same can be achieved by granting purely excusatory defences. For, as defenders of the 'conduct rules'/'adjudication rules' view of the justification/excuse distinction will be the first to point out, only actions which are all-things-considered wrong need to be excused. What a sophisticated system of criminal law has to have space to do is to grant that some actions covered by a legal prohibition but not properly taken into account or accommodated by the formulators of that prohibition are not all-things-considered wrong in the eyes of the law. By and large that can only be achieved, in conformity with the rule of law, by the continuing judicial development of genuinely justificatory defences.[42]

b. The harm principle. A related but distinct objection to the full legal implementation of the justification/excuse contrast as I have explained it points to the limited moral resources of the criminal law. My analysis depended at more than one point on the invocation of what may be called 'perfectionist' categories, such as those of virtue and skill. I also relied upon the importance of basic principles of rationality. But it may be objected that it is not a proper function of the criminal law in a modern society to reflect judgments of virtue and

[42] Contrast Simon Gardner, 'Instrumentalism and Necessity', above note 36, at 436–7.

skill, nor even to enforce norms of rationality. The criminal law in a well-ordered modern society is restricted to the task of harm-prevention. The famous 'harm principle', first defended by J.S. Mill, sets the law's legitimate agenda, and constrains the law's perfectionist ambitions. Some have thought that this prevents legal defences, and in particular legal justifications, from carrying all the baggage of their moral counterparts. Notably, the demand that one acted for the justificatory reason to which the law gave recognition strikes some people as being quite out of place in a legal régime directed exclusively towards harm-prevention. In such a legal régime, acceptable action is action which prevents, or on some versions is expected to prevent, more harm than it brings about. So long as this condition is met there is no further harm-prevention advantage to be gained out of legal inquiry into whether the defendant did or did not have the prevention or avoidance of harm close to her heart at the time when she acted. That did not in any way affect the amount of harm she did or was expected to do. Nor will it affect the amount of harm done by others who, for whatever reason, follow her lead in performing analogous actions in similar situations. Thus her reasons for doing as she does are not the law's proper concern.[43]

Again, errors abound in this argument. The most striking is a far-fetched misinterpretation of the harm principle. The harm principle is a principle which exists to protect people from having to surrender their worthwhile pursuits and ways of life merely because those pursuits and ways of life are morally imperfect. Now, owing to the diversity of ultimate moral values, every valuable pursuit and way of life is morally imperfect. To possess the virtues and skills needed for one way of life one must forego the virtues and skills needed for another. One effect, particularly in a complex and highly mobile modern society in which people are widely exposed to strangers, is that intolerance is widespread. We all find it hard to appreciate and respect fully the different virtues and skills exhibited by those many people that we encounter daily whose ways of life and pursuits are so

[43] This seems to be the key to understanding Robinson's position in 'A Theory of Justification', above note 12, at 273 and 292.

markedly different from our own. We correspondingly inflate the importance of their limitations, and are continually tempted by the path of suppression. The harm principle, when conscientiously followed, provides some protection against the institutionalisation of such temptations in cases in which certain ways of life and pursuits are unpopular with those who hold, whether through democratic or undemocratic channels, the power of suppression. The harm principle, thus defended, is the principle that harmless immoralities should not be officially prohibited or punished, and that harmful immoralities should not be officially prohibited or punished disproportionately to the harm they do.[44] But it is a long way from this to the proposition that official prohibitions and punishments should be tailored solely to the single overarching aim of harm prevention. What is lost in the transition to this proposition is the important point that, within the boundaries set by the harm principle, the law may be tailored to reflect countless considerations which have nothing much to do with harm. Compatibly with the harm principle a legal system may limit its prohibition and punishment of harmful activities to those which are also, e.g., dishonest or malicious or inconsiderate. To put it another way, it does not follow from the premiss that the law should not institutionalise intolerance of any harmless immoralities that the law should not institutionalise the tolerance of some harmful ones. On the contrary, there are many harmful immoralities which should arguably be tolerated by law for the sake of other values, including perfectionist values. This is a matter of particular salience where the scope of criminal law defences is concerned. For, as I explained earlier, when the law grants a defence it tolerates, perhaps regretfully, a prima facie wrong. Assuming that the law is otherwise in conformity with the harm principle, this makes the granting of defences into an act of tolerance rather than an act of intolerance. Thus the harm principle has nothing significant to add beyond what it already contributed to the construction of the offence. Of course that is not to say that the law should ignore

[44] I have based my remarks on Joseph Raz, 'Autonomy, Toleration and the Harm Principle' in Ruth Gavison (ed), *Issues in Contemporary Legal Philosophy* (Oxford 1987).

considerations of harmfulness in constructing defences. On the contrary, other things being equal it should always give such considerations whatever rational weight they have, be it great or small. Nor is it to say that there cannot sometimes be positive moral principles requiring that particular harms no longer be officially tolerated. There is certainly no general principle of toleration which would allow the authorities to wash their hands of every problem and dispense with legal regulation altogether. The point is only that intolerance, not tolerance, is the problem which the harm principle, in particular, exists to counter. It is therefore not a general constraint on the legal perfectionism, but a constraint on perfectionist considerations invoked by themselves as if they provided a sufficient ground for official prohibition and punishment.

So, even if it is true that my analysis of justifications and excuses makes them necessarily sensitive to perfectionist considerations, that presents no automatic obstacle to their legal implementation. But, so far as justifications are concerned, I am not even sure that the sensitivity to perfectionist considerations is inevitable. My main argument for the combined subjective/objective (or explanatory/guiding) account of justifications was based on the fundamental principle of rationality that one should always act for some undefeated reason. That principle, I explained, has a broadly instrumental grounding: by acting for an undefeated reason one is more likely to do what one ought, all things considered, to do. That claim, it seems to me, is no less applicable where (let us suppose) what one ought to do, all things considered, is minimise the harm one does. And the claim is therefore no less applicable in some strange legal context to which a reductive harm-prevention objective has (myopically) been applied. Assuming that e.g. actions in self-defensive situations are generally harm-preventing, one will prevent harm more reliably by acting out of self-defence than by acting in such situations out of, e.g., spite or fear of the legal consequences. One should thus look for and react to the self-defensive aspects of one's situation, not to some other aspects. That is no less true in subsequent cases in which others follow one's example than it is in one's own case. Thus the fictitious, narrow-minded legal system we are considering does well, even by its own excessively

parsimonious harm-prevention standards, to reflect the importance of explanatory reasons as well as guiding reasons in the law relating to self-defence. Justificatory defences as I analyse them are accordingly no less at home in a legal system intent on harm prevention and nothing else than they are in a legal system, like our own, with a less monomaniacal outlook.