CHRISTINE TAPPOLET is Associate Professor in the Philosophy Department at the Université de Montréal. She is the author of *Émotions et valeurs* (2000: Presses Universitaires de France) and of several articles in meta-ethics and the philosophy of emotions.

SERGIO TENENBAUM is Assistant Professor of Philosophy at the University of Toronto. He works on ethics, practical reason, and Kant.

GARY WATSON is Professor of Philosophy at the University of California at Riverside. He works primarily in moral and political philosophy, and in the theory of agency.

RALPH WEDGWOOD is a Lecturer in Philosophy at the University of Oxford and a Fellow of Merton College, Oxford. He is the author of several articles on meta-ethics, epistemology, and practical reason, and is currently planning a book that will give a unified presentation of the arguments of these articles.

Introduction

Sarah Stroud and Christine Tappolet

Imagine someone who often can't make up her mind what to do—she just keeps endlessly deliberating. Once she has decided what to do, though, she frequently ends up changing her mind, whether out of mere caprice or because she gives in to temptation. At other times, she is obstinate to the point of persisting with plans even when she has good reason to revise them. At yet other times she tends to procrastinate, leaving for later what she judges or even knows she should do now. Moreover, she often fails to do things that she believes prudence requires doing, given the desires she foresees having. Or she simply fails to do what she judges would be best and just chooses to do an inferior course of action instead. Sometimes she feels utterly depressed and is paralysed by a total lack of motivation. Things can get so bad with her that she is not even able to do what she would like to do: she just finds herself compulsively doing things she judges to be bad.

This character suffers from a number of practical failures—indecision, irresoluteness, caprice, weakness of will, obstinacy, procrastination, imprudence, akrasia, *accidie*, and compulsion—many of which have been considered to be forms of practical irrationality. One important philosophical task arising out of these all-too-common phenomena is the challenge of better understanding them, and, if possible, systematizing these different categories, which have been handed down to us by common-sense morality and psychology as well as philosophical tradition. What do these different phenomena involve, and what do they imply for our accounts of action, deliberation, and, more generally, practical rationality? What can we conclude from the nature of

these phenomena about the principles that govern practical reasoning, rational choice, and rational action?

One sort of practical failure which has been considered paradigmatic has been called alternatively 'weakness of will', 'incontinence', or 'akrasia'. Philosophical discussions have focused not so much on weakness of will as a character trait than on the sort of action that manifests it: roughly, intentional action contrary to one's better judgement, that is, contrary to the judgement that another course of action would be better. The philosophical debate about weakness of will starts from the question whether it is even *possible* freely and intentionally to act against one's better judgement, or whether on the other hand Socrates was right when he claimed that no one willingly does wrong (Plato, *Protagoras* 352a ff.).

I

Donald Davidson's seminal paper 'How is Weakness of the Will Possible?' (1970*a*) is often taken to be the starting point of the contemporary debate on weakness of will. However, it is useful first to look at R. M. Hare's discussion of weakness of will (Hare 1952), for Davidson's account is in part a reaction to Hare's denial that weakness of will is, strictly speaking, possible. Hare denies the possibility of weakness of will because he defends prescriptivism, which maintains that moral judgements like 'I ought to do *x*' entail imperatives. Hare believes that in order to guide our choices and actions, moral judgements like 'I ought to do *x*' have to be construed in such a way that to assent to such a judgement is to assent to an imperative. In full-blown cases of moral judgement, if I judge that I ought to do *x*, I address to myself the imperative 'do *x*'. And this means that if I am free to do *x*, I will do *x*.

This view amounts to a very strong version of the 'internalist' idea that there is an internal relation between moral or more generally practical judgements, on the one hand, and motivation and action, on the other. Given his commitment to strong internalism, Hare has to deny that it is possible to make a full-blown moral judgement that one ought to do *x* while failing to act accordingly. He claims, rather, that what actually happens in cases of supposed weakness of will is either (*a*) that the agent does not really judge that he ought to do *x*—he fails to realize that the general claim that one ought to do certain things applies to him, for instance, or he merely judges that convention requires such actions—or (*b*) that the agent is not free: he is

in fact physically or psychologically unable to do *x* (see Hare 1952, ch. 11; 1963, ch. 5; 1992: 1305–6).

Davidson, by contrast, wishes to deny that either (*a*) or (*b*) need follow. Not just with respect to moral judgements, but more generally practical judgements—as he shows, weakness of will is not only a matter of acting contrary to a moral judgement—Davidson aims to vindicate the common-sense idea that weakness of will as free and intentional action contrary to a full-blown practical judgement is possible. However, his account is closer to Hare than one might first think, for though he rejects prescriptivism as well as Hare's own version of internalism, he nonetheless remains faithful to the internalist idea. In particular, Davidson tries to show that weakness of will is compatible with the idea that intentional action is done in the light of what the agent judges to be good or simply better, a claim he suggests is self-evident. His aim is specifically to make weakness of will compatible with the following two principles:

(P1) If an agent wants to do *x* more than he wants to do *y* and he believes himself free to do either *x* or *y*, then he will intentionally do *x* if he does either *x* or *y* intentionally.

(P2) If an agent judges that it would be better to do *x* than to do *y*, then he wants to do *x* more than he wants to do *y*. (Davidson 1970*a*: 23)

Taken together, these two principles entail that if an agent judges that one course of action is better than another, and she judges herself to be free to do either one, then she will intentionally perform the first action if she does either one. Thus these two principles *appear* to be inconsistent with the possibility of free intentional actions contrary to one's better judgement. The problem is that both these principles, and the claim that weakness of will is possible, seem difficult to deny.

Davidson's solution to this puzzle lies in distinguishing different kinds of evaluative judgement. In particular, he distinguishes judgements as to what is better—which he calls 'unconditional' evaluative judgements, or 'evaluative judgements sans phrase'—from judgements as to what is better *all things considered*, which he considers to be merely prima facie evaluative judgements. The difference between them is that the latter type of judgement, unlike the former, is *relational*. An all-things-considered better judgement, unlike an unconditional better judgement, does not involve a commitment to the superiority of the option in question. Such relational judgements do not tell us what is better *simpliciter*, but what is better in light of some reason *r*; they

are thus rather like a judgement to the effect that a particular piece (or body) of evidence favours a particular hypothesis. Davidson proposes (1970*a*: 38) that the logical form of such relational evaluative judgements involves a 'prima facie' operator which governs the entire judgement (and thus does not permit the evaluative conclusion to be 'detached'). Such judgements, he holds, are of the general form *pf* (*x* is better than *y*, *r*), where '*r*' refers to the reason why *x* is judged better. An all-things-considered judgement is simply a relational judgement about what is better in light of all the reasons the agent considers relevant; it remains conditional in form.

With this distinction in hand, Davidson defines weakness of will as action contrary to an all-things-considered evaluative judgement, not to a judgement as to what is better *sans phrase*. His definition of akratic action is as follows:

(D) In doing *x* an agent acts incontinently if and only if: (a) the agent does *x* intentionally; (b) the agent believes there is an alternative action *y* open to him; and (c) the agent judges that, all things considered, it would be better to do *y* than to do *x*. (Davidson 1970*a*: 22)

Given that principles P1 and P2 postulate a relation between *unconditional* better judgement and intentional action, not between *all-things-considered* better judgement and intentional action, weakness of will turns out to be compatible with those principles. If the akratic agent had reached an *unconditional* judgement in favour of doing *y*, she would have done *y* rather than the akratic *x*. But she never did. She judged only that *y* was better all things considered.[1] The problem with such an agent is that she did not conclude that the action she considered to be better all things considered was also better *sans phrase*. As Davidson underlines, this is not a logical blunder: her judgements are not contradictory. But it is nonetheless a failure of rationality. The akratic agent violates what Davidson calls the 'principle of continence', that is, the rational requirement to 'perform the action judged best on the basis of all available relevant reasons' (Davidson 1970*a*: 41).

This argument secures only the possibility of intentional action contrary to a *relational* evaluative judgement—not intentional action contrary to an *unconditional* better judgement. Is that enough? It is now quite generally

---

[1] Indeed, given his conception of intentional action as action performed in light of one's unconditional evaluative judgement, it seems Davidson not only has to deny that it is possible to act against one's unconditional better judgement but, in addition, has to suppose that the akratic agent makes an unconditional evaluative judgement in favour of the akratic option. This latter implication emerges with greater clarity from Davidson 1978.

acknowledged that it is not. As Robert Audi (1979) and Michael Bratman (1979) argue, it certainly seems possible freely and intentionally to act in one way while judging that another course of action would be better *sans phrase*, or while thinking that it would be best to refrain. Accordingly, the debate since Davidson has focused on so-called 'strict akratic action' (Mele 1987) or 'last-ditch akrasia' (Pears 1982), namely free and intentional action performed contrary to a judgement held at the time of action that another course of action open to the agent would be better (period), or that there is sufficient reason against performing this action at this time.[2]

While the debate after Davidson has departed from his analysis on the important point just mentioned, it has tended to follow him in two other equally important respects. Most contemporary philosophers writing on weakness of will agree with Davidson that akratic action is not only possible, but actual; indeed, that it is quite a common feature of our lives. This means our philosophical theories ought to leave room for the possibility of weakness of will. However, most philosophers have also agreed with Davidson that akratic action, while possible, is *irrational*. They have thus sought to make room for akrasia, but *only as a species of practical irrationality*. These two general points of consensus have shaped the contemporary literature on weakness of will. However, each has also been subject to powerful challenges.

Gary Watson's influential paper 'Skepticism about Weakness of Will' (Watson 1977) threatens the first of these points of consensus. As Watson emphasizes, an important question that is pertinent to whether strict akratic action is possible is whether action contrary to one's better judgement can be free. For if it cannot, there is no clear difference between akrasia and compulsion, that is, cases in which an agent is motivated by a desire or emotion he is unable to resist. Watson asks what could explain a putatively weak agent's failure to resist a rebellious desire; he argues that what seem to be the only two possible explanations have both to be rejected. His failure cannot be explained in terms of what he chooses to do, since choice must follow better judgement. Nor can it be explained by a culpably insufficient effort to resist. Given that the action is supposedly free, we have to assume that the agent was able to control himself; thus the question of why he didn't make the requisite effort

---

[2] It is accepted that there are other places where, to borrow a useful phrase from Amélie Rorty (1980*b*), the *akratic break* can take place, such as when an agent fails to commit himself to the general value judgements from which he draws his practical conclusion. However, strict akratic action is the phenomenon that has attracted the most philosophical attention in the post-Davidson era.

immediately presents itself. According to Watson, the answer cannot be that he misjudged the amount of effort required: that would be a different fault from akrasia. Watson claims that we are entitled to conclude that the agent was unable to resist, in which case his action was unfree. As Michael Smith's chapter in the present volume attests, however, this conclusion is controversial.[3]

The second point of consensus, namely that akrasia is necessarily a failure of rationality, has also come under attack in recent years.[4] The charge has come that this belief reflects an erroneous conception of what it is to act rationally. These objectors agree that *most* action against one's better judgement is irrational. But as Robert Audi asks (Audi 1990), what if the desires which ground the agent's better judgement are highly irrational, or fail to represent his overall desires, interests, or ideals? In certain cases, it is plausible to suppose that if the agent had thought long enough, he would have come to a different practical conclusion.

Alison McIntyre (1990) also presses this point. Following Bernard Williams (1980), she claims that a consideration constitutes an internal reason for an agent if, as a result of deliberation, she would come to see it as a reason for her. Given this, an agent can be wrong about her own internal reasons.[5] Thus her better judgement might not reflect the reasons she really has, and the akratic action may in fact be the one which she has better reason to perform. McIntyre argues that an akratic action is not irrational if it is motivated by considerations that the agent would have taken to be sufficient reasons for performing that action had she properly deliberated. In such cases, the agent's sensitivity to the reasons she has outstrips her intellectual ability to see that certain considerations are reasons for her to act in a particular way. Under those circumstances, following one's better judgement might well be a mark of obstinacy rather than the rational thing to do.

As the foregoing discussion shows, akrasia has proved to be a fertile ground for reflection on our broader conceptions of intentional action and practical rationality. Indeed this can be considered one of the principal strands of the contemporary philosophical literature on akrasia and other putative

---

[3] In fact, the majority view continues to be that strict akratic actions can be distinguished from compulsion. For discussion of this issue, see Kennett 2001, ch. 6; Pugmire 1982; Mele 1987, 2002; Buss 1997; Tenenbaum 1999; and Wallace 1999b.

[4] See, in addition to the papers discussed in this paragraph, Arpaly 2000 and Ogien 2002.

[5] Of course, it is even easier to imagine that an agent could be wrong about her reasons on an 'external' construal of reasons. See Williams 1980 for the distinction between internal and external reasons.

instances of practical irrationality: exploration of the implications of these phenomena for the nature of practical reason and rationality.[6] One such issue which has been prominent in the literature concerns the implications of strict akratic action for the relation between evaluative judgements, on the one hand, and motivation and action, on the other. Debate has centred in particular on whether *internalism* should be completely abandoned or whether a weaker form of internalism compatible with the existence of akrasia can be developed. The difficulty, if one opts for the rejection of internalism (that is, for externalism), is that one seems forced to say that the relation between practical judgement and motivation is no different from that between the judgement that something is square, for instance, and motivation. That is why those who accept the possibility of strict akratic action have tended to place their hopes in the second option.[7]

Michael Smith, for instance, has argued (Smith 1994) for a weak form of moral internalism according to which an agent who judges that some course of action is right is motivated to act accordingly unless she is practically irrational (that is, unless she suffers from weakness of will or other forms of practical unreason). It is easy to generalize such a weak internalism from moral to practical judgements: the claim would be that an agent who judges that a course of action is best will be motivated to act in accordance with her judgement unless she suffers from practical unreason. Another way of formulating such a weak internalism would be to specify what a rational agent would do given her practical judgement. Thus one could claim with T. M. Scanlon that a rational agent who judges that there is compelling reason to perform a certain action normally forms the intention to do that action, her judgement then serving as sufficient explanation of both the intention and the action which is intended (Scanlon 1998: 33–4). On this view, failing to form an intention to do something while nonetheless taking oneself to have sufficient reason to do it constitutes one important form of irrationality.

We have been sketching one of the principal strands in the recent literature concerning weakness of will and other varieties of practical irrationality,

---

[6] Of the chapters in the present volume, those by Stroud, Tenenbaum, Watson, Wedgwood, MacIntosh, and de Sousa can be viewed as contributions to this broad line of inquiry having to do with general conceptions of practical rationality. Many of these look anew at practical reason through the lens of putative examples of irrationality.

[7] Velleman (1992) is an exception, for he argues that intentional action need not be directed at outcomes regarded as good. See also Stocker 1979 for the claim that one can desire something one considers to be bad.

namely that concerned with their implications for general conceptions of practical reason. A second general line of inquiry which has grown out of consideration of those same phenomena is of quite a different sort. It concerns the *explanation* of these phenomena, or the challenge of offering a plausible account of what is going on in cases of akrasia, *accidie*, or failure to act on one's resolutions (for instance). It is generally acknowledged that akrasia specifically, and practical irrationality generally, involves a coming apart of the motivational force of the agent's wants from his assessment of the objects of those wants. The course of action the agent judges to be better or more desirable turns out not to be the one he most desires (see Mele 1987, ch. 1; Pettit and Smith 1993). The question is how this gap arises.

One influential answer to this question, proposed by the later Davidson (1982), appeals to mind-partitioning. According to Davidson, action against one's better judgement involves mental causes that fail to be reasons for the mental items they cause. He claims that in order to understand how this can be so, we need to suppose that the mental cause in question is part of a semi-autonomous structure of the mind, whose boundaries are defined by the breakdown of reason-relations. However, a variety of other explanations of the gap between better judgement and motivation have been proposed that do not appeal to mind-partitioning. One such explanation, due to Ronald de Sousa, appeals to emotions (de Sousa 1987: 199–201). According to de Sousa, an emotion is responsible for the fact that a reason which the agent considers insufficient gets acted upon. He argues that emotions are perfectly tailored for this role, given their impact on attention.

More generally, the importance of attentional phenomena in the explanation of akrasia has often been noted (Bratman 1979: 156, 168; A. Rorty 1980b; Peacocke 1985; Mele 1987, ch. 6). There are, however, still other possible explanatory strategies, such as the proximity explanation, the habitual explanation, and the social explanation (A. Rorty 1980b; Mele 1987, ch. 6). As Amélie Rorty points out (A. Rorty 1980b), these strategies need not exclude each other; indeed, they often supplement each other. Alfred Mele draws on empirical studies to suggest that many strict akratic actions can be explained in terms of the perceived proximity of the rewards promised by the incontinent action, the agent's motivational level, his failure at self-control, and his attentional condition (Mele 1987: 92).

While this second, broadly explanatory, line of inquiry is in the first instance more focused on the putatively irrational phenomena themselves, it too has broader ramifications, in so far as it raises issues about the explan-

ation of action in general.[8] For instance, the question of what mental entities one has to postulate in order fully to explain someone's action is obviously pertinent not just in the context of providing an account of akrasia and other types of irrational action, but also for the project of understanding rational action. Common sense speaks readily, in this domain, of intentions, decisions, choices, preferences, and emotions, in addition to beliefs and desires, the two kinds of state that contemporary philosophers have typically taken to be primordial for explaining action. The philosophical issues which such common-sense appeals raise are, first, whether these purported mental entities survive theoretical scrutiny as states which are distinct from beliefs and desires, and, if so, what role they play in action and particularly in irrational action. The same kinds of questions apply to putative mental *faculties* such as 'the will'. Until recently—and quite ironically when one thinks of the expression 'weakness of will'—it was generally thought that there was no need, and indeed no room, in our psychology for the will.[9] That assumption has now become controversial.

However these issues concerning particular mental states and faculties are resolved, it is surely only in tandem with a detailed picture of the workings of the mind that we can begin to formulate the norms of practical rationality that apply to it. An understanding of the nuts and bolts of akrasia and other putative forms of practical irrationality ought therefore to make possible a richer understanding of human action and the norms that govern it.[10]

II

We turn now to the chapters that make up the present volume. In his 'Rational Capacities, or: How to Distinguish Recklessness, Weakness, and Compulsion', Michael Smith responds to the challenge powerfully expressed in Watson 1977 (and noted above): the need to make out a distinction between weak-willed and compelled action. As Smith notes, such a distinction seems to be required in order to legitimize our holding weak-willed but not

[8] Of the chapters in the present volume, those by Smith, Holton, Pettit, Tappolet, Tenenbaum, and Heath contribute to this broad explanatory project, either specifically with regard to (putatively) irrational actions, or with regard to action in general.

[9] As Gary Watson notes in his contribution to this volume, O'Shaughnessy was going against the then grain in writing his 1980 book *The Will*.

[10] Our thanks to Daniel Laurier for commenting on a draft of this first part of the Introduction.

compelled agents responsible for their actions. Smith offers such an account in terms of *rational capacities*. He defines each of the categories of the reckless, the weak-willed, and the compelled in terms of the agent's having, or lacking, a pertinent rational capacity. In cases of compulsion, for instance, the compelled agent lacks the rational capacity to bring her desires into line with her belief about what she ought to do, whereas in cases of weakness of will she has that capacity, although she does not exercise it. Smith then seeks to explicate rational capacities themselves via a possible-worlds analysis. He proposes that an agent has a rational capacity just in case, abstracting away from all properties except relevant aspects of her brain, she is situated in the context of a whole raft of nearby possible worlds in which she *does* form the correct belief or desire. His analysis explicitly holds of both belief and desire: for Smith, *both* can and ought to be the products of rational capacities.

Richard Holton's chapter, 'How is Strength of Will Possible?', also emphasizes, in a rather different way, the importance of a particular capacity: the capacity for strength of will, or success in persisting with one's resolutions. Holton understands weakness of will as the unreasonable abandonment of resolutions in the face of strong contrary inclinations. (He thus distinguishes weakness of will from akrasia.) At least sometimes, however, we are able to muster the opposite of weakness of will—strength of will—and stand firm in our resolutions. How is this possible? On a classical Humean approach to the explanation of action, an agent's action is determined solely by her beliefs and desires. Holton argues that this makes it difficult to account for strength of will. He proposes that we expand the conceptual repertoire used in the explanation of action by appealing to the operation of a distinct *faculty of will-power*. Our active deployment of this faculty, Holton argues, causally explains our ability to stick to our resolutions in the face of contrary inclinations. Holton adduces evidence from common sense and from psychology to elaborate on the existence and nature of this faculty. He argues that this faculty works much like a muscle: exerting it takes effort, it tires in the short term, but sustained exercise makes it stronger over the long term. Because reconsideration of resolutions is all too likely to result in revision, we exercise our will-power by refusing to reconsider our resolutions. Holton closes by making a case for the rationality of refusing to reconsider resolutions.

Philip Pettit's 'Akrasia, Collective and Individual' takes up the question of whether a *group* could properly be said to exhibit akrasia. Pettit understands akrasia as failing, even under intuitively favourable conditions, to act in the

way required by one's intentional states; he concludes that certain kinds of groups can indeed manifest akrasia in this sense. Furthermore, he suggests, such cases of collective akrasia offer important lessons for our conception of individual akrasia. The chapter begins by considering different kinds of groups and collectives, concluding that only *self-unifying cooperatives* are agents in a sufficiently robust sense to admit of akrasia. Pettit then argues that even groups concerned for their own rational unity will be subject to what he calls 'discursive dilemmas'. These arise when—with respect to a set of rationally connected issues—the votes of each individual member of a group are consistent, but these sets of votes nonetheless result in an inconsistent set of majority decisions. In such cases, the group may have trouble 'getting its act together': individuals may be loath to set aside their own votes in order to arrive at a consistent set of group decisions. But then we have collective akrasia: a failure, on the part of the group, to live up to its own standards of rationally integrated agency. Pettit describes several strategies for achieving self-control in the group context, and extends his analysis to the individual case: the different 'voices' within a person can also fail, akratically, to get their act together.

Christine Tappolet's chapter, 'Emotions and the Intelligibility of Akratic Action', takes up a somewhat neglected question in the contemporary literature on akrasia, namely the role of emotions in akratic action. Tappolet rejects the tendency to see emotions only as (non-rational) *causes* of akratic actions; she is concerned to cast them in a more positive light. Proposing that emotions be viewed as perceptions of value, she argues that as such they have the capacity not just to cause but even to render intelligible actions which are contrary to one's better judgement. She holds that non-conceptual perception of a value can make one's action intelligible even when it is opposed by one's all-things-considered judgement, and indeed even when that perception is in fact erroneous. Sometimes, on the other hand, an akratic action prompted by an emotion can be *more rational* than following one's evaluative judgement, for it may be the judgement and not the perception which is in error. By contrast, Tappolet argues, akratic actions in which no emotion is involved (cases of 'cool' akrasia) are genuinely puzzling and of dubious intelligibility. In cases of emotional akrasia we can at least point to the emotion—a product of a subsystem which is independent of the agent's higher-order cognitive faculties—in order to make sense of the agent's action. But in cases of cool akrasia no such appeal is possible: we have available only states of the same *kind* as the overall judgement contrary to which the agent acts, and which have been

judged to be insufficient. Tappolet concludes that it is not clear whether there really is such a thing as cool akratic action.

Some (but not all) *internalists* about judgement, motivation, and action would welcome this last suggestion. For—as we noted earlier—the existence of akrasia seems to make trouble for various internalist claims. The next three chapters take up this issue: all seek to rehabilitate one or another internalist approach to such matters, even in the face of weakness of will and other 'recalcitrant' phenomena. Sarah Stroud, in 'Weakness of Will and Practical Judgement', considers whether the existence of weak-willed actions puts paid to the idea that we make genuinely practical judgements. A practical judgement, she says, is one which enjoys an internal, necessary relation to subsequent action or intention, and which can serve as a sufficient explanation of such action or intention. She contrasts the idea that deliberation characteristically issues in practical judgements with what she calls a 'Humean externalist' view of practical reasoning, according to which our deliberative conclusions are merely motivationally inert judgements which must be combined with an appropriate independent desire if they are to do any work. First Stroud argues that, contrary to appearances, the possibility of akratic actions does *not* favour the Humean externalist conception over the practical-judgement view. If the latter is properly interpreted—as a constitutive norm of rational agency—then it is not threatened by the existence of akrasia. She goes on to argue that in fact the Humean externalist view is committed to a highly questionable thesis concerning weakness of will, namely that *global* akrasia is a coherent possibility. She thus suggests that a fuller examination of the implications of weakness of will actually points in favour of the practical-judgement model.

Like Stroud, Sergio Tenenbaum, in '*Accidie*, Evaluation, and Motivation', is concerned with the implications for the nature of practical reason of an anomalous phenomenon. He focuses on *accidie*, a type of apathetic or depressed state in which the unfortunate agent seems to suffer from 'loss of will': he is completely unmotivated to pursue things which he nonetheless sincerely judges to be of value. Tenenbaum aims specifically to account for *accidie* within the framework of what he calls a 'scholastic' view of practical reason, according to which to desire something is to conceive it to be good. Desires, Tenenbaum proposes, are 'appearances of the good' from particular evaluative perspectives. But not all such appearances will be incorporated into one's overall reflective conception of the good. For instance, one may take an evaluative perspective to be *conditioned*: an evaluative perspective is *conditioned* by

X if what appears to be good from that perspective can only be correctly judged to *be* good, and hence worth pursuing, if *X* obtains. Kant, for instance, thought that one's happiness cannot be considered good or worthy of pursuit unless one is virtuous. Tenenbaum uses this idea of conditioning to explain the accidic agent's malady in scholastic terms. He proposes that the accidic agent believes that *all* evaluative perspectives are subject to a condition which (he believes) does not obtain. Hence such an agent can truly be said to retain some appreciation of the value of the things he no longer finds worth pursuing.

Gary Watson's 'The Work of the Will' defends yet another type of internalism. Watson begins with the idea that *deciding* or *making up one's mind* is a primary locus of human agency. Understanding the *will* in these terms, Watson seeks to clarify the scope of its work. First, he asks whether the will has real work to do only on 'externalist' conceptions of agency: those which deny any necessary connection between the will and the good or the choiceworthy. Can internalist theories of agency support the idea of a power to make up one's mind which is distinct from normative assessment? Or do such theories simply—and wrongly—equate the will with practical judgement? Watson argues that they do not. The will can serve an executive function even on views that are not premissed on the need to make room for counter-normative agency. Secondly, Watson considers whether we ought to say that the will, and hence agency, exist not only in the practical domain, but also in the cognitive sphere. After all, we make up our minds what to believe as well as what to do. Watson argues that like 'deciding to', 'deciding that' should indeed be classified as an active phenomenon, and hence as a mode of agency. Watson in fact sees the answers to his two questions as linked. The natural line of thought supporting externalism about practical decision is at odds with the idea of doxastic activity; thus a negative verdict on the former paves the way for acceptance of the latter.

Ralph Wedgwood's chapter, 'Choosing Rationally and Choosing Correctly', is also concerned with choice and the rationality thereof. More specifically, it distinguishes, as the title suggests, between two standards for the assessment of choices. Wedgwood links the idea of choosing *rationally* to an 'internal', and that of choosing *correctly* to an 'external', 'should'. This distinction turns on whether what you 'should' do depends only on your overall state of mind, or on what your options are really like—in particular, on whether any of them really are *good things to do*. Wedgwood argues that choosing *correctly* is the basic notion, in the sense that truths about which

choices are rational are explained by truths concerning which choices are correct, rather than the other way around. Internal requirements on *rational* choice are thus derived from the ultimate practical 'aim' of arriving at *correct* choices. In order to defend this 'recognitional' view of practical reason against the 'constructivist' who sees the internal, procedural requirements on rational choice as fundamental, Wedgwood distinguishes between substantive and formal versions of the recognitional view: the former, but not the latter, offer a determinate specification of what it is for something to be a good thing to do. Wedgwood argues that the objections which constructivists have raised against the recognitional view in fact apply only to substantive versions thereof. Furthermore, he urges, objections of the very same kind can be pressed against the constructivist view itself. Wedgwood ends by elaborating on his favoured 'formal' interpretation of the recognitional view.

Duncan MacIntosh's chapter, 'Prudence and the Temporal Structure of Practical Reasons', explores an approach to rationality which was one of Wedgwood's targets: a broadly Humean, present-aim conception thereof (see e.g. Williams 1980). MacIntosh's exploration in this chapter is situated in the context of the problem of prudence. MacIntosh argues—contra, for example, Thomas Nagel (1970)—that there is no rational requirement of prudence: that it is not rationally obligatory to act in light of one's foreseen future desires as well as one's current desires. Now one might worry that if there is no rational duty of prudence—if I need not take account now of my future desires—I might be rational in acting now so as to thwart desires which I foresee having tomorrow. The acts of a rational agent could thus be absurdly incoherent over time. MacIntosh seeks to rebut this worry by showing how a Humean, present-aim approach to rationality itself generates rational constraints on the evolution of desires and hence of reasons. In fact, MacIntosh seeks to generalize the point by arguing that whatever reasons are, your future reasons need not function as reasons for you now—and that there is nothing incoherent about this. MacIntosh's ultimate aim is to establish the true temporal structure of reasons, whatever they may be. One upshot of his arguments, however, is that we ought to remove from the category of practical irrationality a phenomenon which has often been taken to be one of the leading examples thereof: imprudence.

Joseph Heath's chapter also urges that some paradigm examples of practical irrationality may have been wrongly classified as such. This conclusion emerges out of an overall strategy which Heath shares with Richard Holton: broadly speaking, to expand the repertoire of psychological states, faculties,

and intentional phenomena in terms of which we explain action. Heath argues, in 'Practical Irrationality and the Structure of Decision Theory', that our understanding and modelling of actions would be enhanced by accounting for them not just in terms of beliefs and desires as these are construed in standard decision theory, but also in terms of further factors which have not typically been incorporated into decision-theoretic explanations. Heath claims that if we do not help ourselves to these additional resources, we are liable wrongly to classify certain actions as practically irrational. He focuses on two examples: the apparently counter-preferential behaviour which many agents exhibit in a variety of games in experimental game theory, and the widespread phenomenon of yielding to temptation. Heath seeks to demonstrate that relative to an expanded set of intentional phenomena, including—crucially—deontic preferences over actions and hyperbolic temporal discount rates, the case for considering such actions as instances of practical irrationality simply vanishes. Thus, Heath suggests that the charge of practical irrationality may often just be an artefact of an unduly impoverished theory. A decision theory which incorporates more structure, Heath proposes, will better meet a standard of expressive adequacy. It will also yield a cleaner division of labour between the theory of practical rationality strictly construed, and assessments of the rationality of an agent's intentional states.

Ronald de Sousa's aims in 'Paradoxical Emotion: On *Sui Generis* Emotional Irrationality' are in a certain respect the opposite of Heath's and MacIntosh's. Whereas their arguments, if successful, would *contract* the scope of what can properly be called irrational, de Sousa's arguments, if successful, would have the effect of *broadening* the domain of application of the charge of irrationality. For de Sousa proposes in his chapter that there is a hitherto unrecognized *sui generis* framework of specifically *emotional* rationality. Attitudes and emotions can be rationally assessed within this new framework, which (de Sousa argues) cannot be reduced to either of the main existing templates for rationality, the strategic and the epistemic. However, the domain of emotional rationality, like the strategic and epistemic forms thereof, contains its own antinomies or paradoxes. De Sousa catalogues a number of emotions or attitudes which present a paradoxical aspect, in that there are good reasons both to consider them rational and to condemn them as irrational, with neither view clearly more persuasive than the other. These questionable attitudes typically have a temporal dimension: they include 'dessert last' and other principles for the temporal ordering of pleasures, and varying attitudes towards death. De Sousa highlights the exposed and seemingly arbitrary status of such assessments of

emotional rationality and urges a strongly naturalistic approach to the resolution of these antinomies. On this approach, the emotions we actually have serve as the final court of appeal, both for the assessment of other emotions and for the resolution of conflicts between strategic and epistemic rationality.

1

Rational Capacities, or: How to Distinguish Recklessness, Weakness, and Compulsion

Michael Smith

In 'Skepticism about Weakness of Will' Gary Watson invites us to consider the distinction between recklessness, weakness, and compulsion.

Suppose that a particular woman intentionally takes a drink. To provide an evaluative context, suppose she ought not to have another because she will then be unfit to fulfill some of her obligations. Preanalytically, most of us would insist on the possibility and significance of the following three descriptions of the case. (1) the reckless or self-indulgent case; (2) the weak case; and (3) the compulsive case. In (1), the woman knows what she is doing but accepts the consequences. Her choice is to get drunk or risk getting drunk. She acts in accordance with her judgement. In (2) the woman knowingly takes the drink contrary to her (conscious) better judgement;