

# Normativity and the Will

*Selected Papers on Moral Psychology  
and Practical Reason*

R. JAY WALLACE

CLARENDON PRESS • OXFORD

# Contents

Introduction	1
I. REASON, DESIRE, AND THE WILL	
1. How to Argue about Practical Reason	15
2. Three Conceptions of Rational Agency	43
3. Explanation, Deliberation, and Reasons	63
4. Normativity and the Will	71
5. Normativity, Commitment, and Instrumental Reason	82
II. RESPONSIBILITY, IDENTIFICATION, AND EMOTION	
6. Reason and Responsibility	123
7. Moral Responsibility and the Practical Point of View	144
8. Addiction as Defect of the Will: Some Philosophical Reflections	165
9. Caring, Reflexivity, and the Structure of Volition	190
10. <i>Ressentiment</i> , Value, and Self-Vindication: Making Sense of Nietzsche's Slave Revolt	212
III. MORALITY AND OTHER NORMATIVE DOMAINS	
11. Virtue, Reason, and Principle	241
12. Scanlon's Contractualism	263
13. The Rightness of Acts and the Goodness of Lives	300
14. Moral Reasons and Moral Fetishes: Rationalists and Anti-Rationalists on Moral Motivation	322
<i>Index</i>	343

# Introduction

Moral philosophy has turned increasingly to topics in moral psychology and the theory of normativity in recent years. But there are very different ways of approaching both of these clusters of issues. Some philosophers treat moral psychology as a largely empirical domain, dedicated to the description and explanation of human thought, emotion, and behavior, through methods that are broadly continuous with those of the sciences. The moral psychologist, on this conception, tries to get clear about what people are like as a matter of fact, ignoring for these purposes normative questions about how people ought to behave or what it would be valuable for them to do. On the other side, normativity is sometimes taken to constitute an autonomous intellectual realm, one that can be studied largely in abstraction from questions about human psychology. Normative considerations define ideals for human thought and action, and it is natural to suppose that our conception of the ideal should not be held hostage to messy facts about what human beings actually think and do.

There is no doubt something importantly right about the distinction between fact and value on which these approaches rely. It is one thing to ask what people are like, quite another to consider how they ought to behave. While acknowledging the distinction between these questions, however, I myself do not believe that they can effectively be addressed in isolation from each other. Normativity in the domain of practice is fundamentally about reasons for action, the considerations that count for and against actions in the perspectives of deliberation and advice. But reasons can be normative in this sense only if they are considerations that agents are able to acknowledge and to comply with, insofar as they are rational and are otherwise deliberating correctly. To the extent this is the case, the study of normativity in practice must attend to the psychological capacities that undergird normative response, and that make it possible for normative reasons to figure properly in the deliberations of the agents to whom they apply. Conversely, human motivational psychology distinctively involves capacities to respond to considerations whose normative significance for action the agent acknowledges, as well as motivations and emotions that can interfere with these forms of rational response. These reciprocal connections between normativity and motivation raise a series of large and difficult questions for philosophy, centering on the interpretation of our capacities for rational agency, the nature and conditions of normativity in general,

and the possibilities for motivated departures from our own judgments about what we have reason to do.<sup>1</sup>

The present volume collects fourteen papers on these central questions in moral psychology and the theory of practical reason. All of the papers reflect my commitment to the general idea that normativity and moral psychology are best pursued together. They might be thought of as advertisements for this idea, attempts to explore the interpenetration of the normative and the psychological in a series of debates that lie at the heart of moral philosophy.

Substantively the essays are united in their allegiance to three broad claims:

- (a) Rationalism in ethical theory, which holds that moral considerations are reasons for action.
- (b) Realism in the theory of normativity, the thesis that there are facts of the matter about what we have reason to do that are prior to and independent of our normative convictions.
- (c) An anti-Humean approach to motivational psychology, which denies that desires have a substantial role to play in explanations of rational action.

The essays that have been selected pursue these central philosophical issues from a variety of perspectives. I have organized them into three parts, to emphasize thematic continuities between individual papers; a brief account of each part follows.

## 1. REASON, DESIRE, AND THE WILL

This part addresses general issues about the relation between normative considerations and motivation. It collects five papers on these issues.

Chapter 1, ‘How to Argue about Practical Reason’, was written as a survey of contemporary approaches to practical reason. A main focus here is the relation between normative reasons for action and the dispositions and desires of the agents to whom they apply. Many philosophers, taking their inspiration from a perhaps inaccurate reading of Hume, have held that normative reasons for action must be grounded in the antecedent desires and dispositions of the agents to whom they apply. (This is what Bernard Williams has called the internal reasons model,<sup>2</sup> or ‘internalism’, as I shall refer to it.) Chapter 1 traces the intuitive appeal of this approach to the ideas that normative reasons must be capable of being acted on in deliberation, and that intentional action in turn involves states of

<sup>1</sup> For more on the interpenetration of normative and psychological issues in these domains, see my ‘Moral Psychology’, in Frank Jackson and Michael Smith (eds.), *The Oxford Handbook of Contemporary Philosophy* (Oxford: Oxford University Press, 2005), and ‘Practical Reason’, in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <http://plato.stanford.edu/entries/practical-reason/>.

<sup>2</sup> See Bernard Williams, ‘Internal and External Reasons’, as reprinted in his *Moral Luck* (Cambridge, England: Cambridge University Press, 1981), 101–13.

desire. But the paper goes on to show why these considerations do not in fact support the internalist approach to normative reasons, identifying the more specific question that properly divides Humeans and their opponents about the conditions of normative reasons. This question concerns the explanation of motivation: internalists maintain that deliberation can give rise to a new motivation only if it begins in some sense from desires that are already to hand (in accordance with what I call the principle of ‘desire-out, desire-in’), while externalists deny that practical reflection must accord with this principle.

Chapter 2, ‘Three Conceptions of Rational Agency’, is a later exploration of the basic idea that deliberation on our normative reasons must be capable of giving rise to corresponding motivations to action. The paper begins by noting that rational agents are not merely motivated in accordance with their normative reasons, but guided in their deliberation by their reflection on those reasons. The question is, what must be true about rational deliberation if it is to satisfy this ‘guidance condition’? I identify and assess three frameworks for answering this question. Internalists hold that normative reasons are grounded in the antecedent desires of the agent, and they appeal to these desires to make sense of the capacity of deliberation to generate new motivations and actions (in accordance with the principle of ‘desire-out, desire-in’). I argue, however, that this approach does not really do justice to the guidance condition, insofar as it leaves no room for genuinely normative thought to figure in deliberation. A second approach, which I call ‘meta-internalism’, does better in this respect, tracing deliberated revisions in our motivating attitudes to the operation of abstract or second-order dispositions that are partly constitutive of our standing as (rational) agents. But this approach comes to grief over cases of irrationality, in which we act in ways that conflict with our own judgments about our normative reasons. A third alternative, ‘volitionalism’ as I call it, rejects the empiricism about motivation that is implicit in the other two approaches, postulating motivating attitudes with respect to which the agent is fundamentally active. I offer a tentative defense of this approach, arguing that volitionalism can account for the guiding role of normativity in the deliberative reflections of rational agents, while leaving the right kind of space for cases in which we freely defect from our own normative views in action.

Chapter 3, ‘Explanation, Deliberation, and Reasons’, offers a slightly different perspective on the role of normative considerations in deliberation and action. The paper is a critical response to Jonathan Dancy’s contention—developed in his book *Practical Reality*<sup>3</sup>—that the considerations we cite to explain motivation and action are not psychological states or facts, but rather normative considerations, as they struck the agent at the time of action. Dancy is correct to stress the role of normative reasons in relation to rational agency, but I argue that we can do justice to their significance without denying that explanations of action are a kind of

<sup>3</sup> Jonathan Dancy, *Practical Reality* (Oxford: Oxford University Press, 2000).

psychological explanation. To this end, I distinguish between the prospective standpoint of practical deliberation and advice, within which normative considerations are front and center, and the retrospective standpoint of explanation. From the latter standpoint, we consider an action that has already been performed, and ask why the agent carried it out. I suggest that the general form of an answer to this explanatory question will cite the agent's normative beliefs or convictions; these are psychological considerations, but ones that precisely capture or constitute the agent's normative point of view. (The discussions in Chapters 1 and 2 might be construed as addressing the question, what *further* psychological conditions must be satisfied in order for normative convictions of this kind to guide reflection and generate corresponding motivations to action?)

Chapters 4 and 5 continue the focus on the general connection between motivation and normativity, but they approach the connection from the other direction. Here the question is not about the role of normative considerations in guiding deliberation, but about the normative significance and implications of motivation itself. Chapter 4, 'Normativity and the Will', discusses the constructivist approach to the sources of normativity that has recently been developed by Christine Korsgaard. The basic constructivist idea is that normative principles are not prior to and independent of the will, but somehow constituted by it; but how should we understand the metaphor of construction that is central to this approach? I offer an interpretation of Korsgaard's constructivist program, and contrast it with a realist approach to normativity. According to the interpretation I propose, constructivists hold that a commitment to comply with principles of practical reason is built into every act of will, and that this element of commitment accounts for the idea that such principles are binding or normative for the agent. I argue, however, that a conception of commitment adequate to this theoretical task is elusive, and suggest that we do better to think about the will within the framework of a realist conception of normativity.

Chapter 5, 'Normativity, Commitment, and Instrumental Reason', continues to explore the interplay of the normative and the psychological in volition. The paper begins where Chapter 4 leaves off, with Korsgaard's claim that volition is to be understood in terms of principles that are essentially normative. Korsgaard contends that our activity as agents can be made sense of only if we take normative principles to be implicit in each act of willing. I show that the argument from activity fails, appealing to cases of *akrasia* to support the conclusion that one can be committed actively to achieving some end or goal, without believing that the end or goal would be valuable or justified. The remainder of the paper considers the principle of instrumental reason, which specifies that one should take the means that are necessary relative to one's ends. The challenge for an account of this principle is to explain why it applies even to acts of will that are not normative in Korsgaard's sense—why, that is, even *akratic* agents are subject to a kind of internal irrationality if they fail to take means that they know to be necessary for the attainment of their *akratic* ends.

My response to this challenge builds on the idea that intention or volitional commitment involves an element of belief: the belief, namely, that it is possible that one will attain the end one has set for oneself. If (as I contend) this idea is plausible, we can see why a minimally reflective agent would be subject to a kind of incoherence in belief if they failed to take the necessary means to their chosen ends. The instrumental principle thus derives from basic requirements of theoretical rationality, together with a plausible assumption about the nature of volition or intention. In a new postscript to this chapter I develop this cognitivist account of the instrumental principle further, defending it against some objections and alternatives that have recently been proposed. I emphasize in particular the role of the instrumental principle as a source of rational pressure that we feel and respond to when we recognize that a given means is necessary if we are to achieve the ends we have set ourselves to pursue.

## 2. RESPONSIBILITY, IDENTIFICATION, AND EMOTION

This part too collects five papers. Here the general focus is on issues of responsibility and identification, especially as they intersect with questions about the norms that apply to our actions and the rational and volitional capacities that enable compliance with such norms.

Chapter 6, 'Reason and Responsibility', takes as its starting point the general approach to moral accountability defended in my book *Responsibility and the Moral Sentiments*.<sup>4</sup> This approach holds that accountable agency is to be understood not in terms of freedom of will, but rather in terms of general rational powers or capacities, specifically those that enable us to grasp and respond to moral principles. But what if an agent who possesses these general powers of reflective self-control should lack compelling reason to do what morality prescribes? I argue that such agents would not be fully responsible for their failure to comply with moral principles in that case. The rational powers approach takes responsible agency to be grounded in our general capacities for critical reflection and self-determination. But if a given agent has no compelling reason to do the right thing, then even the most conscientious application of their general rational powers would not bring them to comply with the requirements of morality. I then address the question of whether people in general have good reason to comply with moral demands. I suggest that even if moral reasons are in some sense inescapable, that alone would not secure their normative grip on all agents. There

<sup>4</sup> R. Jay Wallace, *Responsibility and the Moral Sentiments* (Cambridge, Mass.: Harvard University Press, 1994).

is a further dimension of practical reason to consider, having to do with the relation between moral requirements and the demands of a good or a meaningful human life. If these two sources of norms cannot be reconciled with each other, then practical reason will be divided within itself; the result of such a division in practical reason would be that we are not fully to blame when we sacrifice morality for the sake of ends that are personally of great human significance.

Chapter 7, ‘Moral Responsibility and the Practical Point of View’, looks at the powers or capacities that underwrite moral responsibility, on the approach I favor, and in general make it possible for agents to comply with the normative requirements. I begin by suggesting that these capacities should be understood as including an active power of self-determination, in line with the volitionalist conception of rational agency sketched in Chapter 2. But this volitionalist picture seems potentially problematic. It will perhaps be surprising to readers of *Responsibility and the Moral Sentiments*, which argued that responsible agency does not require the kind of metaphysical freedom that incompatibilists have traditionally insisted on. Doesn’t the postulation of an active power of self-determination make responsibility hostage to questions of freedom of the will, in ways that the book attempted to resist? Furthermore, whether or not it is compatible with the argument in *Responsibility and the Moral Sentiments*, the volitionalist conception of the will might appear independently implausible. It is reminiscent, for instance, of the problematic theory of agent-causation, according to which agents intervene in the causal order of nature from a position indeterminately external to it.

Chapter 7 addresses these worries. To this end, it invokes the distinction between the practical standpoint of deliberation and the theoretical standpoint of explanation drawn in Chapter 3, arguing that the capacity for volitional self-determination—like the other powers of reflective self-control—should be understood in relation to the distinctively practical point of view. Once we are clear about this, I contend, concerns about the mysteries of agent-causation can be disarmed. The theory of agent-causation goes wrong in supposing that the volitional capacities with which agents are equipped constitute a framework for explaining actions by reference to the agent who performed them; but this is not the right way to understand volitionalism. The paper goes on to consider whether the volitional picture involves a commitment to the kind of freedom that would be threatened by determinism. I argue, first, that the association of our volitional powers with the practical point of view does not automatically insulate them from any possible conflict with deterministic approaches to explanation. But I contend, second, that the only real threat from this direction would be posed by a distinctively psychological version of determinism, such as we have no real reason to take seriously.

This general way of conceptualizing the will involves a distinction between two classes of desire: the volitional forms of intention, choice, and decision that are themselves paradigms of agency, and states of inclination, longing, and attraction with respect to which we are merely passive. If we accept this distinction, questions naturally arise about the role of the second class of desires—given desires, as



we might call them—in relation to agency and intentional action. These questions are addressed in Chapters 8 and 9. Chapter 8, ‘Addiction as Defect of the Will’, looks at the kinds of desires involved in addiction, and asks how we should think about their influence on our capacities for deliberated agency. The paper makes the case that the basic features of addiction can best be accommodated within the framework of volitionalism, which rejects the assumption that intentional action is a simple causal function of the agent’s beliefs and given desires, operating in accordance with a ‘hydraulic’ conception of the mind. On the volitionalist alternative I favor, states of given desire involve, *inter alia*, the direction of one’s attention onto possibilities for action, as attractive or potentially valuable along some dimension (e.g. as an opportunity for pleasure or visceral satisfaction). Addiction renders one susceptible to given desires of this kind that are both resilient and strong, but it is not clear how these notions of resilience and strength are to be interpreted. I reject the causal understanding of them that is latent in the hydraulic conception, defending in its place a phenomenological interpretation of the phenomena of strength and resilience of given desire. This phenomenological account is then deployed to explain the ability of addictive desires to interfere with the good functioning of our deliberative capacities. In particular, I show that addictive desires may affect the volitional as well as the cognitive side of rational agency, and trace some implications of such ‘defects of the will’ for questions of responsibility in this domain.

Chapter 9, ‘Caring, Reflexivity, and the Structure of Volition’, turns to a different phenomenon, that of identification. The pioneering work of Harry Frankfurt has drawn attention to many important complexities of human agency, including above all the possibilities for desire and intentional action from which the agent is estranged. Frankfurt himself favors an approach to the twin phenomena of identification and estrangement that makes use of the notions of reflexivity and of a hierarchy of desire. Chapter 9 is an extended response to Frankfurt’s approach. I argue that the idea of a hierarchy of desire, taken literally, distorts more than it illuminates the phenomena with which Frankfurt is concerned. To improve on it, we need to move away from the noncognitivism about given desire that seems implicit in much of Frankfurt’s work (and that has, I contend, contributed to its appeal and influence). In its place, I recommend the quasi-perceptual conception of given desire sketched in Chapter 8, and show how this conception leads to an improved understanding of identification and estrangement. To identify with a given desire is to affirm through reflection the normative content that the desire presents, in ways that would remain stable if subjected to further critical scrutiny. With this account in place, I turn next to the notions of caring and reflexivity that have figured prominently in Frankfurt’s more recent work. Among other things, the paper argues that there is a distinctive context of eudaimonistic reflection—already anticipated in the argument of Chapter 6—in which we deliberate reflexively on the things that we care about, reflecting on their contribution to the goodness of our own lives. I suggest, however, that we cannot capture the potential critical

force of this kind of reflection unless we depart from the noncognitivist assumptions about caring that Frankfurt evidently favors.

Questions about the relation between desire, emotion, and value are pursued within a very different context in Chapter 10, '*Ressentiment, Value, and Self-Vindication: Making Sense of Nietzsche's Slave Revolt*'. As its title suggests, this paper offers a sustained interpretation of Nietzsche's idea that modern moral consciousness has its origin in a slave revolt. This episode or process, as Nietzsche describes it, involves strong feelings of *ressentiment* that build to monstrous proportions in the psyches of the powerless masses, eventually giving birth to a new table of values—a democratic, leveling, universalistic conception of morality that challenges the older aristocratic values of good and bad. The fundamental question that is raised by this account concerns the relation it posits between *ressentiment* and the values to which that emotion is said to give birth. Most commentators interpret this causal nexus in strategic terms, assuming that the new moral values are adopted by the slavish masses as part of a plan to achieve revenge against the masters who have oppressed them. I argue that this way of thinking about the slave revolt is deeply problematic, and propose in its place an expressive interpretation of the slave revolt. The *ressentiment* of the masses gives them an emotional orientation to the social world that does not fundamentally make sense, so long as they accept the aristocratic values of good and bad. *Ressentiment* involves profound hostility to the very people who are, in terms of the aristocratic value scheme, paradigms of the good, and this combination of attitudes is essentially unstable and conflicted. According to the expressive account, the adoption of new values by the slavish serves to resolve this psychic tension, and the slave revolt can in this way be construed as the expression of slavish *ressentiment*.

The paper aims in the first instance to reconstruct Nietzsche's position, but it is not possible to pursue this goal without addressing systematic issues of independent interest in moral psychology and the theory of value. Reflection on Nietzsche's account helps us to understand how unconscious emotions can distort and corrupt evaluative reflection, leading to forms of moral thought that amount to ideology and false consciousness. Important in these processes, I suggest, is the widespread need people have to understand themselves and their world in ways that provide a kind of vindication of their position within it.

### 3. MORALITY AND OTHER NORMATIVE DOMAINS

This part collects four papers in which morality is in the foreground. Issues that are addressed include the structure and normative significance of morality, the relation between moral and other reasons, and the distinctive sources of moral motivation.

Chapter 11, 'Virtue, Reason, and Principle', grapples with John McDowell's account of the nature of moral reasons and moral reasoning. In a series of challenging and influential papers McDowell has drawn on both Wittgenstein and the Aristotelian tradition to defend an interpretation of morality as a domain of considerations that make rational claims on agents. Central to McDowell's approach is his denial that the deliberations of the virtuous agent can be reconstructed in terms of general principles that capture the rational requirements to which the agent responds. Grasp of the reasons provided by morality requires habituation into a comprehensive form of life, one whose requirements resist discursive formulation. Virtue, on this account of it, involves a responsiveness to reasons that are not fully intelligible to those who are not themselves virtuous already.

I offer an interpretation of this central idea, proposing that McDowell's conception of practical reason be understood on analogy with the phenomenon of connoisseurship. The virtuous agent, like the connoisseur, has a refined, quasi-perceptual capacity to make reasoned discriminations of a kind that can be justified on a case-by-case basis, but that resist capture in a set of general principles or norms applicable across the board. I argue that this represents a legitimate model of rational discrimination and response, but raise some questions about its applicability to morality as a normative domain. The connoisseurship model seems most plausible if we think of virtue as involving a comprehensive conception of how to live, something that would require a corresponding capacity for responding to a range of normative considerations too diverse and complex to permit perspicuous representation through general discursive principles. At least since the modern period, however, there has been a tendency to think of morality as a distinctive subdomain within the broader normative landscape, involving reasons that are binding on and accessible to people who accept a plurality of comprehensive conceptions of the good. This modern conception of the moral remains attractive, and I suggest that it goes together with the idea that moral requirements admit of discursive formulation in terms of principles accessible equally to all the members of pluralistic communities.

On the Aristotelian conception implicit in McDowell's approach, morality does not constitute a unified subdomain within the landscape of reasons. Virtuous agents have a special, habituated capacity for responding to the normative considerations that bear on their choices, but there is nothing presumptively moral about this capacity, nor is there any nonsuperficial way of carving up the reasons to which the virtuous respond into moral and nonmoral classes. This deeply pluralistic understanding of normativity contrasts with the conception implicit in such modern theories as utilitarianism and Kantianism, which take morality to collect a unified set of reasons for action. Among the most impressive and influential recent unifying theories of this kind is the contractualism of T. M. Scanlon, which forms the subject of Chapter 12 ('Scanlon's Contractualism'). According to Scanlon, the unity implicit in morality should be understood in terms of the

notion of justification to others; moral considerations are genuine reasons for action, and what these reasons have in common is their connection to principles for the general regulation of behavior that permit us to justify our actions to those potentially affected by them. Actions are morally wrong, for instance, if they are prohibited by principles of this kind that nobody could reasonably reject, and their being wrong in this way is something that constitutes a strong reason against them.

My discussion of Scanlon's theory focuses on four different sets of issues. I begin with the conception of rational agency that forms the background to Scanlon's account. This conception—a version of what I call 'meta-internalism' in Chapter 2—holds that intentions, like beliefs, are attitudes that are intrinsically sensitive to our normative judgments. I argue that this approach exaggerates the continuities between theoretical and practical reason, in a way that does not do full justice to the kinds of irrationality that seem possible in the practical sphere.<sup>5</sup> It also, I contend, leads to a distorted interpretation of our responsibility for the kinds of wayward desires and emotions discussed in Chapter 9, which persist in the face of one's reflective rejection of their contents. Scanlon treats such attitudes as fully attributable to us, insofar as they are open to assessment in terms of reasons; but in cases in which wayward desires and emotions resist our best efforts at reflective scrutiny and control, it seems to me they do not constitute a ground for moral blame. A further element in Scanlon's comprehensive theory is his 'buck-passing' account of the value, which holds that goodness is not to be understood as a substantive property that grounds reasons for action. Though broadly sympathetic to this way of thinking about the relation between the normative and the evaluative, I raise some questions about the specific version of the buck-passing account that Scanlon seems to favor.

The core of Scanlon's contractualism is his suggestion that the contractualist formula accounts for the reason-giving force of what he calls the morality of right and wrong, in a way that supports the idea that morality in this sense is a unified normative domain. Scanlon himself traces moral reasons to the value of the distinctive kind of relationship with other people that compliance with moral principles makes possible. I develop this suggestion by situating the appeal to what Scanlon calls mutual recognition within the context of the kind of eudaimonistic reflection identified in Chapters 6 and 9. Attention to the valuable forms of human relationship made possible by compliance with moral principles helps us to see how compliance with such principles can make our own lives better. But it is not clear that this line of thought alone vindicates the basic thought that the morality of right and wrong is a unified normative realm. Scanlon himself accepts a pluralistic account of the broader domain of reasons, and against this

<sup>5</sup> Compare the discussion of theoretical and practical irrationality in Chapter 5, section 2 of this volume.

background the question can be pressed of whether morality is just a convenient way of collecting a variety of reasons that exhibit no interesting substantive unity. I propose some tentative answers to this question on behalf of contractualism, which seems to me the most promising unifying story about morality on the contemporary scene.

Among the most important recent defenders of a thoroughgoing pluralism about the normative is Joseph Raz. According to what he calls the ‘classical view’, there is no context-independent way of categorizing reasons as distinctively moral in nature. Normative reasons are grounded in values, and there is a diversity in the realm of the normative that reflects the deep variety of ways in which the actions open to us can be valuable. It is a consequence of this picture, as Raz develops it, that there is no interesting global contrast to be drawn between morality and self-interest, a position that seems to undermine the challenge to morality that is posed by one traditional form of skepticism about its normative significance. I take Raz’s view as my starting point in Chapter 13, ‘The Rightness of Acts and the Goodness of Lives’. This paper develops further the idea—presented in Chapters 6, 9, and 12—that there is a distinctive perspective of eudaimonistic practical reflection, and that the normative importance of morality may be threatened if moral considerations cannot be justified from this point of view. I offer an interpretation of the terms in which eudaimonistic reflection is framed, distinguishing it from the standpoint of narrowly egoistic concern about our own interests or well-being. Against Raz, I contend that the eudaimonistic perspective collects a significant set of normative considerations, and show that an appeal to such considerations lies behind the important challenge to morality presented in the work of Bernard Williams. To respond adequately to this challenge, it needs to be shown that compliance with moral requirements can make a direct contribution to the goodness or meaning of the agent’s own life. I argue that part of the appeal of traditional unifying approaches to morality, such as utilitarianism and Scanlon’s contractualism, is that they help us to understand how acting morally can contribute along this important dimension of normative assessment.

Chapter 14, ‘Moral Reasons and Moral Fetishes’, explores the interplay of the normative and the psychological as it bears on the interpretation of distinctively moral motivation. Rationalists hold that morality is a set of unconditional (or external) reasons for action, considerations that have normative significance for agents in a way that is not dependent on their subjective interests and desires. Proponents of this approach have held that its denial would render the concern to act rightly a kind of fetish. I take up this charge, with the aim of getting clear about what the fetishism objection comes to, and assessing its force against anti-rationalist accounts of morality. A form of motivation may be said to be fetishistic, I suggest, when it cannot be understood as a response that is merited by its proper object. I develop this interpretation, contrasting it with Michael Smith’s contention that the concern to act rightly is fetishistic if it is understood *de dicto* rather than *de re*. I argue that the objection, on the account of it I favor, poses a significant challenge

to anti-rationalists. To parry it, the anti-rationalist must either defend a global nihilism about normative reasons, or abandon the distinctively anti-rationalist account of moral motivation. Normative and psychological issues are thus revealed, once again, to be inextricably intertwined, insofar as a plausible and attractive conception of moral motivation rests on normative assumptions about the character and conditions of moral reasons.

# 1

## How to Argue about Practical Reason

What are the comparative roles of reason and the passions in explaining human motivation and behavior? Accounts of practical reason divide on this central question, with proponents of different views falling into rationalist and Humean camps. By ‘rationalist’ accounts of practical reason, I mean accounts which make the characteristically Kantian claim that pure reason can be practical in its issue. To reject this view is to take the Humean position that reasoning or ratiocination is not by itself capable of giving rise to a motivation to act. This alternative position is most famously expressed in Hume’s polemical assertion that ‘reason is, and ought only to be the slave of the passions’ in influencing the will or moving us to act.<sup>1</sup>

To fix terms, let us say that a process of thought is an instance of reasoning or ratiocination just in case it is governed by the principles or norms of rationality. To say that a principle or norm *governs* a process of thought is in turn to make an explanatory claim: it is to say, not just that the process of thought is in accordance with the rational principle or norm, but that the process of thought occurs *because* the person believes it to be in accordance with the principle or norm. Thus, if we say that it is a principle or norm of rationality that people should not hold inconsistent sets of beliefs, this does not just mean that they revise those sets of beliefs which they know to be inconsistent, but that they revise them *because* they know those sets of beliefs to be inconsistent.

In these terms, the dispute between the Humean and the rationalist is a dispute about the capacity of rational principles or norms to contribute to explanations of motivation. The rationalist holds that such rational principles have a *primary* role to play in the explanation of motivation, that the psychological processes which originally give rise to motivation can be processes which are governed—in the sense I have specified—by the principles or norms of reason. By contrast, the

Originally published in *Mind*, 99 (July 1990), 355–85. Copyright © Oxford University Press 1990. Reprinted by permission of Oxford University Press. I have had helpful comments on predecessors of this paper from Simon Blackburn, John Collins, Samuel Freeman, Gilbert Harman, Sally Haslanger, Katharina Kaiser, Wolfgang Mann, and an audience at the University of Pennsylvania. I owe a special debt to Michael Smith, with whom I have had the benefit of many stimulating discussions about practical reason. Work on this paper was partially supported by a grant from the Research Council of the University of Pennsylvania.

<sup>1</sup> See Hume, *A Treatise of Human Nature*, ed. Selby-Bigge, 2nd edn. (Oxford: Clarendon Press, 1978), 415.

Humean maintains that such rational principles never have a primary role to play in the explanation of motivation or the fixing of our ends. Rather, the explanatory contribution they make is exclusively secondary, accounting for the extension of motivational influence (for example, along means/end lines), but never explaining the original formation of motivation.

The significance of this dispute about practical reason lies in its connection with central issues in moral philosophy, concerning the nature and scope of moral requirements. It is a common thought that moral requirements, if they are to provide reasons for action, must be capable of guiding behavior, leading those who are aware of the requirements to be motivated in accordance with them. (This is, presumably, one thing that is meant by the multiply ambiguous word 'internalism'.) If we combine this internalist position with the Humean picture of practical reason, however, it seems to follow that moral requirements can only provide an agent with reason to act if they are appropriately related to the agent's antecedent desires; for all motivation and behavior, on the Humean view, must be explained by reference to the agent's given, prior desires. The resulting account represents moral behavior as dependent on the agent's existing dispositions, in a way that could restrict the scope of moral reasons (since the required prior desires may not be universally distributed). A Kantian approach to practical reason, by contrast, suggests a different picture of the psychological bases of moral behavior, and a straightforward development of the idea that moral requirements are universal or inescapable in their scope. It does this by opening up the possibility that there are processes of pure reasoning or ratiocination which can explain moral behavior by themselves, and which could equally lead all agents, regardless of their background desires, to be motivated to act on moral requirements.

All of this should be familiar to students of the history of philosophical ethics since Hume. Indeed, so familiar has it become that one might reasonably be skeptical whether there is anything to be contributed on the topic which has not already been said. In the event, however, such skepticism has not deterred contemporary philosophers from entering the fray, and recent years have seen a flurry of philosophical discussions purporting to defend or refute the different approaches to practical reason. My aim in this essay is to sort through these recent discussions, with an eye to reaching a clear assessment of the current state of argument between the opposing camps. I hope to show that recent work has in fact helped to advance the old debate between Humeans and rationalists. For one thing, it has become increasingly clear that the appeal of the Humean position is linked to the teleological character of intentional action, consideration of which suggests an *a priori* argument for the Humean claim that action must be explained ultimately in terms of desires. Rationalists have generally not paid sufficient attention to these teleological considerations and the *a priori* arguments they suggest, and this has made their pronouncements about the possibility of pure practical reason vulnerable. Or so I will suggest.



My own view is that the rationalist position can, in the end, be sustained against the challenge of these Humean arguments. To see why, however, it will be necessary to get clear about what is really at stake in the debate about practical reason. A further aim of my discussion will accordingly be to sharpen our understanding of the issue that divides Humeans and rationalists. Here, I think it will be helpful to turn to the somewhat less recent work of Thomas Nagel, which contains important suggestions about how the issue dividing Humeans and rationalists should be conceived. This is, or ought to be, familiar territory—Nagel's work has hardly wanted for readers. But the issues are complicated, and Nagel himself has neither explained nor developed his proposals adequately. Hence, despite the influence his work has enjoyed, its significance for the debate about practical reason remains rather poorly understood. A clearer account of what is at stake in the debate, which draws on and develops Nagel's suggestions, should help to reveal the inadequacy of recent arguments for the Humean view, and lead to an improved understanding of how we should be thinking, and arguing, about practical reason.

## 1.

Rationalist accounts of practical reason claim that principles or norms of reason can play a primary role in the explanation of action and motivation. In an incisive recent discussion, Christine Korsgaard has distinguished two kinds of skepticism about this rationalist position.<sup>2</sup> *Content* skepticism, as she describes it, is doubt about whether specific principles or norms of rationality are sufficient, by themselves, to guide practical reflection and to explain motivation and action. *Motivational* skepticism, by contrast, is not directed at specific proposals about the content of rational principles and norms. Rather, it purports to offer general grounds for doubting whether there could be such a thing as pure practical reason, grounds which are antecedent to consideration of rationalist proposals about the norms or principles of reason, and which turn on the alleged incapacity of reason to give rise to motivation.

Korsgaard herself rejects motivational skepticism. She says: 'motivational considerations do not provide any reason, in advance of specific proposals, for skepticism about practical reason'.<sup>3</sup> In support of this conclusion, Korsgaard suggests that motivational skepticism typically rests on a misinterpretation of the internalist requirement on practical reasons.<sup>4</sup> Humeans, she contends, often construe internalism as the requirement that rational considerations (or reasons) necessarily succeed in motivating us. So construed, internalism would lead fairly

<sup>2</sup> Christine M. Korsgaard, 'Skepticism about Practical Reason', *Journal of Philosophy*, 83 (1986), 5–25.

<sup>3</sup> *Ibid.* 25.

<sup>4</sup> *Ibid.* 15.

directly to motivational skepticism, since it is doubtful that there are any rational considerations which *necessarily* succeed in motivating us, independently of our desires. Korsgaard argues, however, that this is a misinterpretation of internalism, which does not require that reasons necessarily succeed in motivating us, but only that they 'succeed in motivating us insofar as we are rational'.<sup>5</sup> This, however, is a fairly trivial condition on reasons, which does not place any *a priori* constraints on the possible norms of practical rationality. Whether a candidate principle could motivate rational agents depends on what it is to be rational, and that in turn depends on what the norms or principles of reason are. Hence, Korsgaard concludes, 'motivational skepticism must always be based on content skepticism';<sup>6</sup> it has no independent force.

This conclusion, if correct, would already be a kind of victory for the rationalist approach. Motivational skepticism, as Korsgaard describes it, purports to offer a general argument against the very possibility of a principle or norm of pure practical reason. But if there is no such argument, then it is already possible that pure reason might be practical in its issue, and this, in its weakest form, *is* the rationalist position. A more substantial victory for the rationalist, however, would require specification of the norms or principles of practical reason. Here, as Korsgaard notes, the internalist requirement entails that rationalist proposals about the norms or principles of practical reason will have psychological implications, telling us something about what it would be like to be rational.<sup>7</sup> Some proponents of the rationalist approach, notably including Thomas Nagel, have seen in these psychological implications of internalism a fertile source of arguments in favor of specific rationalist proposals. Focusing on the case of prudence, Nagel argues that this class of motivations can better be explained in terms of principles or norms of reason than on the Humean assumption that motivation always has desire at its source.<sup>8</sup> His strategy is to show that the psychological implications of rationalist accounts, for the motivation of ideally rational agents, are more plausible than those of alternative, Humean accounts.

If we are to conceive the debate in this way, however, it must indeed be the case that there is no reason for questioning the very possibility of a rationalist explanation of motivation. On this point, Korsgaard's own argument seems to me too swift. She takes it that internalism will be seen as innocuous, for the rationalist, once it is correctly interpreted as a thesis about the motivations of the fully rational agent.

<sup>5</sup> 'Skepticism about Practical Reason', 15.

<sup>6</sup> *Ibid.* 6.

<sup>7</sup> *Ibid.* 23–5.

<sup>8</sup> Thomas Nagel, *The Possibility of Altruism* (Princeton: Princeton University Press, 1978), chs. vi–viii. Nagel's discussion of prudence is guardedly endorsed by Philippa Foot, in 'Reasons for Action and Desires', as reprinted in her *Virtues and Vices* (Oxford: Blackwell, 1978), 148–56; and there is a similar account of the rationality of prudence in Martin Hollis, *The Cunning of Reason* (Cambridge, England: Cambridge University Press, 1987), ch. 6. Derek Parfit offers some powerful objections to accounts of this sort in his discussion of the self-interest theory of practical rationality, in *Reasons and Persons* (Oxford: Oxford University Press, 1984), pt 2. See also Richard Kraut, 'The Rationality of Prudence', *Philosophical Review*, 81 (1972), 351–9; and Janet Broughton, 'The Possibility of Prudence', *Philosophical Studies*, 43 (1983), 253–66.

But rationalism is not simply a stipulative claim, about the motivations that the fully rational agent will happen to have; it also makes an explanatory claim, to the effect that the rational agent's motivations can be explained in terms of norms or principles of correct reasoning. As I show in what follows, there is something about the teleological character of motivation that has seemed to rule out the possibility of such explanations in principle, and so to provide an *a priori* argument for motivational skepticism about the rationalist approach. It is this argument, and not simply a misunderstanding of the internalist requirement, that accounts for the persistence of Humean skepticism about the very possibility of pure practical reason. Only once this argument is understood, and conclusively laid to rest, can we proceed to assess the plausibility of specific rationalist proposals about the content of the principles or norms of reason.

## 2.

The *a priori* argument I want to consider may be called the teleological argument, because it takes as its starting point the essentially teleological character of both motivation and intentional action. In saying that these are teleological phenomena, I mean that the person who acts intentionally, or who is motivated so to act, is in a goal-directed state. Any psychological explanation of these phenomena must account for the fact that to act intentionally, or to be motivated so to act, is necessarily to be in a goal-directed state. The teleological argument aims to show that conformity to rational principles cannot alone account for this fact.

The argument—which has been given its clearest and most vigorous statement in a recent discussion by Michael Smith<sup>9</sup>—proceeds as follows. To be in a goal-directed state, it is claimed, is to be in a distinctive kind of psychological condition. Specifically, it is to be in a state whose content is not meant to match or represent the way things are in the world, but which is such that the world is to be made to match or fit the content of the state.<sup>10</sup> This reflects itself in the fact that people who are in a goal-directed state will not, in general, give up the goal, upon learning that it has not been realized in the world, but will instead take steps to change the world so that the goal can be realized. The question arises, however, as

<sup>9</sup> See Michael Smith, 'The Humean Theory of Motivation', *Mind*, 96 (1987), 36–61. Smith's article makes the significant contribution (which I have followed) of formulating the argument in terms of claims about the 'direction of fit' of psychological states, vis-à-vis the world. I suspect that similar considerations, less cogently expressed, have historically tended to move proponents of the Humean approach. They also seem to lie behind the common thesis that intention entails desire; for a recent discussion of this idea, see Robert Audi, 'Intending, Intentional Action, and Desire', in Joel Marks (ed.), *The Ways of Desire* (Chicago: Precedent Publishing, 1986), 17–38.

<sup>10</sup> On the idea that psychological states may be differentiated according to their 'direction of fit' with the world, see e.g. G. E. M. Anscombe, *Intention* (Ithaca: Cornell University Press, 1963), sects. 36, 40; Bernard Williams, 'Consistency and Realism', as reprinted in his *Problems of the Self*

to what it is to be in a psychological state which has this peculiar direction of fit vis-à-vis the world; and a plausible answer to this question must hold that, whatever else is involved in being in a goal-directed state, it cannot simply be a matter of having a certain belief or set of beliefs. Beliefs are precisely those psychological states which aim to match or represent the world, and their direction of fit is therefore just the converse of that which characterizes goal-directed states. Some other kind of state must thus be present, whenever one acts intentionally or is motivated so to act. Moreover, it is plausible to suppose that these further goal-directed states will characteristically be constituted by desires; for there is a general conception of desires according to which they are the psychological states one is in whenever one is in a state such that the world must be made to fit the content of the state (rather than vice-versa).<sup>11</sup>

This argument, which looks fairly strong as far as it goes, establishes that beliefs alone cannot account for motivations to action, but that desires must also be present whenever an agent is so motivated. The Humean, however, wishes to draw the stronger—or at any rate, different—conclusion that we cannot account adequately for motivation and intentional action solely in terms of the following of rational principles or norms, and to reach this conclusion on the basis of the teleological argument it is necessary to make some further assumptions.<sup>12</sup> The most important such assumption is the following: that rational principles will only be capable of contributing to the explanation of motivation to the extent that desires are not implicated in motivation. This assumption seems to be part of a broader picture of rationality, according to which reasoning and ratiocination are associated exclusively with the cognitive side of human psychology—that is, with beliefs and relations among one's beliefs—and contrasted with such nonrational

(Cambridge, England: Cambridge University Press, 1973), 187–206, at 203–5; John Searle, *Intentionality* (Cambridge, England: Cambridge University Press, 1983), 7–9; and Richard Wollheim, *The Thread of Life* (Cambridge, Mass.: Harvard University Press, 1984), 52–3. Talk about 'direction of fit' applies literally to propositions; its application to psychological states may seem metaphorical or otherwise problematic. For attempts to explain and to defend talk about the 'direction of fit' of propositional attitudes, see Andrew Woodfield, 'Desire, Intentional Content and Teleological Explanation', *Proceedings of the Aristotelian Society*, 82 (1981–2), 69–88, at 82–6; and Smith, 'The Humean Theory of Motivation', sect. 6.

<sup>11</sup> The conception in question is dispositional (but nonbehaviorist). See e.g. the dispositional conceptions of desire sketched by Richard B. Brandt and Jaegwon Kim, in 'Wants as Explanations of Actions', *Journal of Philosophy*, 60 (1963), 253–66; and by William P. Alston, in 'Motives and Motivation', in Paul Edwards (ed.), *The Encyclopedia of Philosophy* (New York: Macmillan and Free Press, 1967). Apparent counter-examples to the claim that goal-directed states are always realized by desires—such as hopes and wishes—are plausibly understood as involving an element of desiring; on this point, see Wayne Davis, 'Two Senses of Desire', in Joel Marks (ed.), *The Ways of Desire*, 63–82, at 64. Some philosophers, such as Brandt and Kim, find it more felicitous to use the term 'want' to refer to the general, dispositional conception of desire, reserving the term 'desire' for appetitive states which have a distinctive phenomenology; but nothing significant hangs on this terminological issue.

<sup>12</sup> This gets obscured in Smith's discussion, because he is content for the most part to represent the Humean view as the claim that explanatory reasons for action are partly constituted by desires. This formulation, however, does not bring out adequately the central point at issue between the Humean and the rationalist, which is the extent to which rational processes of thought—those which are governed by rational principles or norms—can contribute to the explanation of motivation.

states as desires and emotions (what Hume refers to collectively as ‘the passions’). Hume himself endorses some version of this picture in the *Treatise*.<sup>13</sup> He distinguishes between the passions, as ‘original existences’, and those states of the understanding which admit of truth and falsity, suggesting that the latter states are alone the province of rationality. But this picture, when combined with the teleological argument, appears to make the Humean conclusion irresistible: if desires are necessarily present on occasions of motivation, and if desires are not themselves rational psychological states, what scope can there possibly be for the rational explanation of motivation and intentional action?

This, in outline, is the teleological argument for the Humean position. If sound, it would constitute an *a priori* argument against the very possibility of a rationalist account of practical reason, for it would show that rational principles or norms cannot contribute to the primary explanation of motivations, in the way the rationalist supposes. Motivation, being a teleological phenomenon, requires the presence of desire, and desires are simply beyond the range of explanation in terms of rational principles or norms.<sup>14</sup>

### 3.

If we are to leave open the possibility of a rationalist account of practical reason, some response to the teleological argument will have to be made. In fact I think there are two strategies that might be followed in responding to the teleological argument. One strategy would be to question the conception of desire that figures in the argument. In particular, we might challenge the Humean assumption that to be in a teleological or goal-directed state is necessarily to be in a state of desire. Though some have followed this strategy, it does not seem a very promising line to take, because the conception of desire underlying the teleological argument is independently more plausible than the alternative accounts that have been proposed.<sup>15</sup> For this reason, I think that a different strategy will have to be pursued.

<sup>13</sup> Bk. II, pt. III, sect. iii (p. 415). I assume here that when Hume describes the passions as original existences, he is not necessarily denying that they have propositional content, but only denying that their content is ‘representational’—such as aims to fit the way the world is. For discussion of this and other possible interpretations, see Mark Platts, ‘Hume and Morality as a Matter of Fact’, *Mind* 97 (1988), 189–204, sects. 6–9.

<sup>14</sup> This teleological aspect of action-explanations is entirely left out of account in the argument which Richard Warner gives for the coherence of a rationalist account of motivation, in *Freedom, Enjoyment and Happiness* (Ithaca: Cornell University Press, 1987), 42–5. Warner describes a thought experiment in which we are to imagine a creature which takes thoughts as inputs and produces behavior as output; the coherence of the description is then taken to show that a rationalist account is at least possible. But the description is coherent only if we interpret the creature’s behavior as mere bodily movement rather than as intentional action. The point of the teleological argument is that it is *not* coherent to suppose that intentional action could take place in the absence of a state of desire.

<sup>15</sup> The alternative accounts treat desires—or at any rate, ‘genuine’ desires, as opposed to mere formal desires—as states that have a distinctive phenomenology: see e.g. Mark Platts, *Ways of Meaning* (London: Routledge and Kegan Paul, 1979), ch. 10; and Don Locke, ‘Beliefs, Desires and

The immediate conclusion of the teleological argument, it should be recalled, is that a person cannot simply be in a state of belief, on an occasion when he acts intentionally or is motivated to act intentionally; some state of wanting or desiring is additionally required. As I said earlier, however, this conclusion by itself is not yet damaging to the rationalist, or supportive of a Humean account of practical reason. To derive a Humean moral from the teleological argument, it is necessary to make the further assumption that the presence of desire precludes the rational explanation of motivation; that because the desires involved in motivation are themselves nonrational states, there is no scope for distinctively rational principles to enter into the explanation of motivation. The second and more promising strategy for challenging the Humean is to question this crucial assumption.

To understand the problem with this assumption, however, it will be necessary to provide a clearer account of what the debate between the Humean and the rationalist is all about. Now I suspect—though I am not certain about this—that the key to a proper understanding of the debate is to be found in chapter V of Thomas Nagel's book, *The Possibility of Altruism*. I say I am unsure about this, because Nagel's account of the debate turns on distinctions and concepts that are not adequately explained; certainly, his discussion has not prevented his readers from continuing to misunderstand the issue that divides Humeans and rationalists. In what follows I shall develop an interpretation of that issue which is at least broadly inspired by Nagel's discussion, with the aim of making more perspicuous than Nagel himself succeeded in doing the flaw in the Humean appropriation of the teleological argument.

Nagel starts by accepting the immediate conclusion of the teleological argument, and the associated conception of desires as states that realize one's having of an aim or goal. He says, for instance, that '*whatever* may be the motivation for someone's intentional pursuit of a goal, it becomes in virtue of his pursuit *ipso facto* appropriate to ascribe to him a desire for the goal'.<sup>16</sup> Desires, then, must always be present on occasions of motivation and intentional action. But a further issue arises, concerning the explanatory role of desires in accounting for motivation, and the resolution of this issue is absolutely crucial to the Humean interpretation of the teleological argument. Nagel raises this further issue by drawing a distinction between two broad categories of desires, which he calls 'motivated' and 'unmotivated'.<sup>17</sup> He explains the distinction in the following terms. Unmotivated desires are desires which simply assail us or come over us (Nagel cites as examples the appetites and certain emotions); whereas motivated desires are '*arrived at* by

Reasons for Action', *American Philosophical Quarterly*, 19 (1982), 241–9. Smith effectively criticizes the idea that all genuine desires are states with a distinctive phenomenology, in 'The Humean Theory of Motivation', sect. 5. See also Alston, 'Motives and Motivation', 402–3.

<sup>16</sup> *The Possibility of Altruism*, 29.

<sup>17</sup> *Ibid.* Various historical precedents for this distinction have been cited. Nagel himself now claims that it has affinities with the Kantian distinction between inclination and interest; see his *The View from Nowhere* (New York: Oxford University Press, 1986), 151 n. 3. N. J. H. Dent suggests a

decision and after deliberation'.<sup>18</sup> Desires of both types are presumably susceptible of explanation—we can, for instance, explain the onset of an unmotivated desire to eat by citing the physiological factors associated with a lack of food. But only motivated desires admit of what Nagel calls 'rational or motivational explanation'.<sup>19</sup>

Now this account of the difference between motivated and unmotivated desires is not very helpful. It suggests that motivated desires will always be formed after a prior episode of deliberation, and that all unmotivated desires are like bouts of animal lust in 'assailing' the agent unfortunate enough to be in their grip; but I do not think this can have been the kind of distinction that Nagel was trying to draw. For one thing, Nagel's distinction looked like offering a comprehensive typology of desires; but a great many of the desires that we ascribe to people are neither states that simply assail the person who has them, nor states that the person goes into following an episode of deliberation. Consider, for instance, the important class of long-term or dispositional desires that are formed as a result of moral education, and help to constitute a person's overall character: these can hardly be said to 'assail' the person who has them, like lust or thirst, but at the same time they are not arrived at by decision and after deliberation, either.

More promising, in my view, is Nagel's suggestion that what marks the difference between motivated and unmotivated desires is the kind of explanation to which each is susceptible; in particular, the idea that motivated desires are distinctive in admitting of what he calls 'rational or motivational explanation'. So far, however, this idea is merely a suggestion, for Nagel's own discussion does nothing to clarify the important notion of rational or motivational explanation. Moreover, without a precise explanation of this notion, the significance of the distinction between motivated and unmotivated desires is apt to remain obscure—a point which I shall have occasion to illustrate, later in my discussion. Let me now offer my own account of the notion of rational or motivational explanation, and say in terms of it what is significant about the distinction between motivated and unmotivated desires.<sup>20</sup>

The basic idea, I would suggest, is that when a person has a motivated desire, it will always be possible to explain that desire in a way that shows it to be *rationalized* by other propositional attitudes that the person has. That is, psychological

different parallel between Nagel's distinction and the distinction in Aristotle and Aquinas between 'deliberated appetites' and 'sense appetites', in *The Moral Psychology of the Virtues* (Cambridge, England: Cambridge University Press, 1984), ch. 4. I am myself doubtful how close Nagel's distinction really is to either of these antecedents—the category of motivated desires is much more encompassing than the earlier counterparts, at least on the interpretation of it I shall go on to offer.

<sup>18</sup> *The Possibility of Altruism*, 29.

<sup>19</sup> *Ibid.*

<sup>20</sup> There is, in fact, a variety of ways of distinguishing between different kinds of desires, many of which are close to being coextensive with the distinction between motivated and unmotivated desires, as I shall construe it. Consider, e.g., Stuart Hampshire's distinction between desires which are, and desires which are not, mediated by descriptions or conceptions, in *Freedom of the Individual*, expanded edn. (Princeton: Princeton University Press, 1975), ch. 2; Stephen Schiffer's distinction

explanation of motivated desires is not restricted to causal claims, about the states or conditions that trigger the onset of the desire. Rather, motivated desires also (and *necessarily*) admit of a different kind of psychological explanation, in which the propositional content of the desire is shown to be rationalized or justified by the content of other of the person's attitudes.<sup>21</sup> Of course, it is possible for one propositional attitude to be rationalized by other attitudes of the agent's, without the rationalizing attitudes *explaining* the formation of the state that is rationalized; a rationalizing explanation requires, more strongly, that the person should be in the rationalizing state *because* he has certain other attitudes that rationalize that state. But it is plausible to think that motivated desires can be explained in this distinctive, rationalizing way. Thus—to take Nagel's own simple example—if Wotan's desire to shop for groceries is a motivated desire, then there will be an explanation of it that reveals it to be rationalized by other propositional attitudes that Wotan has: for example, a desire to eat something, a belief that there is nothing at home to eat, plus various other beliefs of Wotan's (about grocery stores, shopping, etc.). It is because Wotan has these rationalizing attitudes that he forms the motivated desire to shop; but at the same time those rationalizing attitudes provide *reasons* for Wotan's motivated desire.<sup>22</sup>

To put the point this way, however, is potentially misleading. Rationalizing explanations, as I have introduced the notion, explain propositional attitudes in terms of other attitudes whose content rationalizes the state which is to be explained, where rationalization is construed as the provision of reasons or justifications. Strictly speaking, however, the *contents* of desires never provide reasons or justifications for other propositional attitudes; nor do they themselves appear directly susceptible of rationalizing explanation. For example, the propositional content of Wotan's desire to shop for groceries—that is, 'that he (Wotan) shop for groceries'—simply seems to have the wrong form either to justify, or be justified by, other propositions; read literally, it is not even in the indicative mood.<sup>23</sup>

between reason-providing and reason-following desires, in 'A Paradox of Desire', *American Philosophical Quarterly*, 13 (1976), 195–203; and the distinction Wayne Davis draws between volitive and appetitive desires, in 'Two Senses of Desire'. The near coextensiveness of these various distinctions may make it tempting to suppose that they are all, at bottom, different ways of marking the same basic difference between kinds of desires; but I suspect that this is not in fact the case. At any rate, my discussion in the text is meant only to give a sharper interpretation of a single distinction which is especially important for the debate about practical reason, that between what Nagel has called motivated and unmotivated desires.

<sup>21</sup> I do not mean to suggest that rationalizing explanations may not also be cause-giving explanations. For a useful recent discussion of this point, and of the nature of rationalizing explanation more generally, see Philip Pettit, 'Broad-Minded Explanation and Psychology', in Pettit and John McDowell (eds.), *Subject, Thought, and Context* (Oxford: Clarendon Press, 1986), 17–58.

<sup>22</sup> There is an obvious analogy here with the case of beliefs. Many of a person's beliefs are held for reasons, where the reasons are other propositional attitudes which both rationalize the belief and explain the holding of it.

<sup>23</sup> Michael Smith made me see this point. It raises a potential problem for all accounts which hold that some desires are held for reasons, such as those of Stephen Schiffer, in 'A Paradox of Desire', and Wayne Davis, in 'Two Senses of Desire'.



It would be a mistake to conclude from this, however, that desires cannot enter into rationalizing explanations at all. To see how they can, we need only note that desires are characteristically associated with evaluative beliefs. Thus, if Wotan wants to shop for groceries, it will—in the normal case, at least<sup>24</sup>—be legitimate to ascribe to him an evaluative belief, to the effect that shopping for groceries is (*prima facie*) desirable.<sup>25</sup> The content of this evaluative belief, however, can straightforwardly enter into relations of rationalization and justification with other contents of propositional attitudes. For example, we might deploy the schema of the practical syllogism to explain the belief that shopping for groceries is (*prima facie*) desirable in terms of the following, further propositional attitudes: the evaluative belief that eating is (*prima facie*) desirable; and the belief that in order to eat it is necessary to go shopping for groceries.<sup>26</sup> The content of these propositional attitudes justifies or provides a reason for the conclusion that shopping for groceries is (*prima facie*) desirable. And, a person might have reached this conclusion *because* that person holds these further, rationalizing beliefs.

Evaluative beliefs can in this way straightforwardly enter into relations of rationalizing explanation. If this much can be admitted, however, then we have a way of drawing the distinction between motivated and unmotivated desires in terms of the notion of rationalizing explanation. We may say, to start with, that motivated desires are desires whose associated evaluative beliefs admit of a rationalizing explanation. Furthermore, when the evaluative belief associated with a desire admits of a rationalizing explanation in this way, the factors which justify and support the belief can equally be said to justify and support the desire associated with the belief; so that the reasons for the belief may be considered reasons for acquiring the desire as well. To give a rationalizing *explanation* of the desire in terms of these reasons, it need only be supposed in addition that the desire has

<sup>24</sup> The abnormal cases are those in which one desires something which one does not value at all, not even instrumentally. Such cases are discussed by Harry Frankfurt, in 'Freedom of the Will and the Concept of a Person', as reprinted in his *The Importance of What We Care About* (Cambridge, England: Cambridge University Press, 1988), 11–25; and Gary Watson, 'Free Agency', as reprinted in Watson (ed.), *Free Will* (Oxford: Oxford University Press, 1982), 96–110. See also Frankfurt's more recent discussion 'Identification and Wholeheartedness', as reprinted in *The Importance of What We Care About*, 159–76. For present purposes it is enough to acknowledge that such desires exist, and to observe that they cannot, strictly speaking, provide material for practical reasoning or deliberation. (Reasoning or deliberation involving such desires requires the *pretence* that their objects satisfy some evaluative predicate or other.)

<sup>25</sup> The predicate 'is (*prima facie*) desirable' is meant only as an example; it should be construed as ranging over both instrumental and intrinsic kinds of desirability. Similar points might be made, with appropriate modifications, in terms of other specific evaluative predicates; or in terms of 'ought'-judgments, or judgments about reasons for action.

<sup>26</sup> That these apparently simple syllogistic explanations mask the complexity of even instrumental reasoning is a point made by Jaakko Hintikka, in 'Practical vs. Theoretical Reason—An Ambiguous Legacy', in Stephan Körner (ed.), *Practical Reason* (Oxford: Blackwell, 1974), 83–102; see also David Wiggins, 'Deliberation and Practical Reason', in Amélie Oksenberg Rorty (ed.), *Essays on Aristotle's Ethics* (Berkeley and Los Angeles: University of California Press, 1981), 221–40. The point does not seem to me to detract from the usefulness of the practical syllogism, as a framework for certain kinds of rationalizing explanations.

been formed *because* the agent has endorsed the evaluative judgment and the reasons that directly support it. This would often seem to be a plausible assumption to make.

A distinction can thus be drawn between two broad classes of desires—the motivated and the unmotivated—in terms of whether the evaluative beliefs associated with the desires admit of a distinctively rationalizing explanation. Moreover, it should be clear that this distinction has some relevance to the Humean appropriation of the teleological argument. The Humean concludes, from the fact that desires are always present on occasions of motivation, that motivation cannot be explained rationally, assuming that the presence of desires necessarily limits the scope for explanation in terms of rational principles or norms. But Nagel's distinction, at least on the interpretation of it I have offered, calls this assumption into question. Many of the desires that figure in motivation are themselves motivated propositional attitudes—in my terms, states which admit of a further rationalizing explanation, via their associated evaluative beliefs. Their presence on occasions of motivation therefore does not necessarily preclude the purely rational explanation of motivation. For all that the teleological argument shows, the rationalizing explanation of motivated desires may sometimes rely solely on the agent's beliefs, together with rational principles or norms, and in such cases it would be possible to give a purely rational explanation of both motivation and the desires necessarily implicated in it.

#### 4.

It will be well to pause over the distinction between motivated and unmotivated desires, to get clearer about its implications for the debate between the Humean and the rationalist. On the interpretation I have urged, a motivated desire is one whose associated evaluative belief admits of a rationalizing explanation, where the desire is formed *because* the agent has arrived at the evaluative belief. A crucial assumption here is that the rational explanation for an evaluative belief may account for the formation of the motivated desire as well, so that the reasons which explain the belief will equally be reasons for the motivated desire. To say this, it seems, is to admit that it is an independent principle or norm of rationality that one should desire in accordance with one's evaluative beliefs, where this means that one should desire those ends and activities one takes to be desirable, to the extent one takes them to be desirable. Only on this assumption are we entitled to maintain that the rational explanations and justifications of evaluative beliefs may extend as well to the motivated desires associated with them.

This would seem to be a plausible minimal assumption to make about the content of the principles or norms of practical reason, since we do in fact try to adjust our desires to our evaluative beliefs, and take ourselves to be subject to rational criticism when we fail. Officially, of course, Hume himself would deny that this is

a legitimate assumption, since he is on record as holding that rational criticism is restricted to the theoretical sphere of beliefs and the relations between them. But as Nagel and Korsgaard have persuasively suggested, Hume's own discussion of practical reason appears to take for granted some irreducible principles of practical rationality, such as the principle that one should adjust one's desires to one's evaluative beliefs.<sup>27</sup> Thus, Hume notes that people who cease to believe that a certain action or object is a means to some valued end will immediately desist from desiring that action or object.<sup>28</sup> But how are we to understand this tendency? Behind it, there seems to lie the assumption (or something equivalent to it) that it is rational to adjust one's desires at least to one's beliefs about what is instrumentally valuable; plus the assumption that people are, in this respect, characteristically rational. Hume himself thus seems poorly placed to reject out of hand the principle that rational agents adjust their desires to their evaluative beliefs.

This has important implications for the debate about practical reason. It shows that if we are to make sense of the notion of a motivated desire—a desire, that is, which is explicable in terms of the agent's reasons—then we must deny the most extreme Humean claim that there are no irreducible principles or norms of practical reason, only principles or norms of theoretical rationality. To deny this claim, however, is not yet to settle the issue between the Humean and the rationalist. That issue concerns the possibility of explaining motivation and intentional action in purely rational terms. But for all that has been established so far, it might be that the irreducible principles of practical reason are exclusively principles of instrumental or derivative rationality, accounting for the rational extension of motivational influence, but not capable of explaining the original formation of motivation or the original fixing of an agent's ends.

Rationalists, of course, deny this. They maintain that the principles or norms of practical reason are such that reasoning in accordance with them can explain, not just the extension of given motivations, but the original formation of motivation and the fixing of one's ends. If I am right, explanations of this sort would have to be capable of yielding explanations of an agent's motivated desires, since, by the teleological argument, desires are invariably present on occasions of motivation. What would a distinctively rationalist explanation of motivated desires look like?

For purposes of illustration, we may consider two possibilities. The rationalizing explanation of desires, I have suggested, will proceed via the rationalizing explanation of their associated evaluative beliefs. One way we might provide these explanations is to deploy the schema of the practical syllogism, relating specific evaluative conclusions to more general evaluative premises, by way of intervening factual beliefs, until we arrive at basic principles about what is intrinsically valuable. The rationalist might then attempt to establish that these basic evaluative principles can themselves be given a rational justification, one which shows them

<sup>27</sup> Nagel, *The Possibility of Altruism*, 33–4; Korsgaard, 'Skepticism about Practical Reason', sect. III.

<sup>28</sup> Hume, *A Treatise of Human Nature*, 416.

to be valid independently of considerations about the desires that particular agents may happen to have.

Consider, for example, Nagel's argument in part III of *The Possibility of Altruism*.<sup>29</sup> We might reconstruct Nagel's rationalist position, in accordance with the model I have just sketched, in the following terms. Suppose someone has a desire to perform a specific action of type *r*, and an associated evaluative belief that doing that action is (*prima facie*) desirable (where *r* = an act of relieving someone's pain). On one reading of Nagel's account, the rationalizing explanation of this evaluative belief may relate it to a basic evaluative principle that the agent accepts, to the effect that *any* action which would relieve someone's pain is (*prima facie*) desirable. This basic evaluative principle may then in turn be given a rational justification, on the reading of Nagel's argument now under consideration, in so far as rejection of the principle would commit one to the constitutively irrational doctrine of solipsism.<sup>30</sup> But if such a justification can successfully be carried through, then the rational considerations which figure in it could be invoked to explain, not just the basic evaluative principles and the specific evaluations derived from them, but also an agent's desires in accordance with these (basic and derived) evaluations. For we are assuming that it is an independent principle of rationality that one should desire in accordance with one's evaluative beliefs.

This is not, however, the only way that a rationalist explanation of motivated desires might be developed. A second possibility is that the rationalizing explanation of specific evaluative beliefs may depart from the schema of the practical syllogism; so that we are sometimes able to explain a specific evaluation rationally by relating it *directly* to further beliefs that the agent holds, together with background principles of rational reflection. To return to the example of Nagel's argument, this model suggests a different way of explaining the specific evaluative belief that it is (*prima facie*) desirable to perform a particular action of type *r* (where *r* = an act of relieving someone's pain). In particular, we might rationally explain the belief by relating it, simply, to the agent's belief that the action is of type *r*; for on a second interpretation of Nagel's argument, background principles of rationality license a *direct* inference from the belief that an act would relieve someone's pain to the evaluative conclusion that the act is (*prima facie*) desirable.<sup>31</sup> Furthermore, this

<sup>29</sup> What follows is based only very loosely on Nagel's complicated discussion, but I hope the parallels will be apparent. It should be stressed, too, that I am not trying to argue for Nagel's position, but only trying to see what a rationalist position might look like, by considering how Nagel's claims might variously be formulated.

<sup>30</sup> Nicholas Sturgeon seems to interpret Nagel this way, as trying to provide a *justification* for a basic evaluative principle in favour of altruistic motivation, in 'Altruism, Solipsism, and the Objectivity of Reasons', *Philosophical Review*, 83 (1974), 374–402 esp. 375, 393–4. Aristotle's ethical theory is sometimes interpreted as an attempt to provide a similar rational justification for such basic evaluative principles (fixing the ends of action). See e.g. T. H. Irwin, 'Aristotle on Reason, Desire, and Virtue', *Journal of Philosophy*, 72 (1975), 567–78, esp. 574–6; and Norman O. Dahl, *Practical Reason, Aristotle, and Weakness of the Will* (Minneapolis: University of Minnesota Press, 1984), pt. 1.

<sup>31</sup> An interpretation on these lines seems the more proper way of reading Nagel's argument. The burden of the argument is not to provide a rational justification for some basic evaluative premise, such as figures in the first model of the rationalist position which I sketched. Rather, it is to provide an

rationalizing explanation of the evaluative belief may carry over to the desire which is associated with it, since (again) we are assuming the rational requirement that one should desire in accordance with one's evaluative beliefs.

We have, then, two models for explaining motivated desires in accordance with the rationalist claim that pure reasoning can be practical in its issue. What, on the other side, would a Humean position look like? Here, the simplest approach would be to deploy the schema of the practical syllogism to explain specific evaluative beliefs. That is, we should suppose that, when specific evaluative beliefs admit of a rationalizing explanation, the explanation offered will relate the beliefs to more general evaluative premises, by way of factual beliefs, until basic principles are reached about what is intrinsically desirable. So far, the Humean approach would follow the first model for a rationalist account which I sketched above. Unlike the rationalist, however, the Humean must deny that the basic evaluative principles with which syllogistic explanations terminate can be given an independent rational justification. Rather, the Humean will suppose that these basic evaluative principles are fixed by the agent's intrinsic desires, desires which cannot themselves be given a further, rationalizing explanation.<sup>32</sup> Only if we make this assumption will it be the case—as the Humean claims—that practical reason is restricted to accounting for the extension of motivational influence from fixed, antecedent ends.

To put the issue in these terms is to admit that, even on a Humean account, rational explanations of motivation can be given exclusively in terms of the agent's (evaluative) beliefs. For on the kind of Humean position I have sketched, the rational explanation of motivation terminates with citation of the agent's basic principles or beliefs about what is intrinsically valuable. It is true that, on the Humean account, these basic beliefs will in turn be fixed by the agent's intrinsic desires; but those desires do not *rationalize* or *rationaly* explain the basic evaluative beliefs. To think that they do is to suppose that it is a basic principle of practical rationality that one should adjust one's evaluative beliefs to one's (intrinsic) desires. But in fact rationality does not require that we adjust our evaluative beliefs to our desires in this way (it can be perfectly rational to hold that one's intrinsic desires sometimes aim at objects or activities which are not valuable at all). Still,

*interpretation* of the inference patterns characteristic of moral reflection, showing those inference patterns to be rational by displaying their connections with broader patterns of inference that are paradigmatically rational. See Nagel's remarks about the 'method of interpretation', in *The Possibility of Altruism*, 4, 18–23.

<sup>32</sup> The most straightforward way to develop this claim would be to say (with Hobbes) that a person's basic values are, simply, those things which the person desires for their own sakes. A more plausible proposal would allow room for discrepancies between a person's actual, first-order desires and the person's values, treating values (for instance) as some function of the agent's second-order desires (cf. Frankfurt, 'Freedom of the Will and the Concept of a Person' and 'Identification and Wholeheartedness').

Note too that it will often be possible to provide *non*-rationalizing explanations of an agent's intrinsic desires. For instance, in some cases they might be taken to result from a process of Aristotelian habituation or training.

even if intrinsic desires do not rationally explain an agent's basic evaluative beliefs, on the Humean view, there is a looser but more important sense in which they determine the starting points for practical reasoning and deliberation. So long as basic evaluations are fixed by an agent's intrinsic desires, rational criticism of the agent's ends will not be a possibility, and practical reason will be restricted to accounting for the extension of motivational influence from given, antecedent ends.<sup>33</sup>

This discussion may be summarized by saying that Humeans are committed to a distinctive thesis about the form taken by rationalizing explanations of desires. In particular, they are committed to the view that rationalizing explanations of desires must terminate, at some point, with the citation of a basic evaluative belief of the agent's which cannot itself be justified or explained in rational terms. An agent's particular evaluative beliefs and motivated desires may be explained, in the first instance, by being related to basic evaluative principles that the agent holds; but on the Humean view these basic principles are always fixed or determined by the agent's intrinsic desires, and so beyond range of rational justification or explanation. This thesis about the form taken by rationalizing explanations of desires might be called the 'desire-out, desire-in' principle, since it maintains that processes of thought which give rise to a desire (as 'output') can always be traced back to a further desire (as 'input'), one which fixes the basic evaluative principles from which the rational explanation of motivation begins. An adequate defence of the Humean position must provide a reason for accepting this distinctive thesis about the form taken by rationalizing explanations of desires.

Once the issue is seen in this way, however, it becomes apparent that the teleological argument by itself has no direct bearing on the dispute between the Humean and the rationalist. That argument establishes only that desires must be present on occasions of motivation. It leaves it an open question whether the present desires are themselves motivated or unmotivated; and, still more significantly, it says nothing at all about the form that must be taken by rationalizing explanations of motivated desires. An *a priori* argument for the Humean account would have to be a defense of the desire-out, desire-in principle, and the teleological argument sketched earlier does not by itself provide a defense of that crucial principle.

I would suggest, then, that the significance of the distinction between motivated and unmotivated desires is that it sharpens our conception of the debate

<sup>33</sup> This point seems to be neglected by Don Locke, in 'Beliefs, Desires and Reasons for Action'. Locke argues against the Humean approach by insisting that rational explanations of motivation can be given exclusively in terms of an agent's evaluative beliefs (in particular, he suggests explanations in terms of what he calls 'sufficient reason' beliefs: see 'Beliefs, Desires and Reasons for Action', 246–7; also Locke's 'Reasons, Wants, and Causes', *American Philosophical Quarterly*, 11 (1974), 169–79, at 170–2). But Locke admits that these sufficient reason beliefs may simply 'derive from' the agent's desires ('Beliefs, Desires and Reasons for Action', 247). In the terms I have proposed, however, this is not an admission which it is open to an anti-Humean to make, since the distinctive Humean claim is precisely the claim that the evaluative beliefs in terms of which we explain desires are themselves 'derived from' or fixed by the agent's desires.

between the Humean and the rationalist. Interpreting this distinction as I have done, we see that the real burden on the Humean is to defend a claim about the rationalizing explanation of desires, the claim I have called the desire-out, desire-in principle. It is because the teleological argument by itself lends no support to this crucial principle that it fails to settle the issue between the Humean and the rationalist.

## 5.

Of course, the failure of one attempted *a priori* argument for the Humean position does not rule out other strategies that might be pursued in support of the Humean approach. In this section I wish to consider, briefly, three further arguments that purport to offer general support for a broadly Humean account. The interpretation of the debate I have developed in the preceding sections should help us to see why none of these further arguments succeed.

An assumption commonly made about rationalist accounts is that they end up identifying desires and beliefs, saying that there are certain beliefs which are desires, or which serve as desires, or which are necessarily connected with desires.<sup>34</sup> But the idea that desires might be identified with beliefs in this way has recently come under attack, on grounds that it is incompatible with the accounts of belief and desire provided by decision theory.<sup>35</sup> The idea, roughly, is that on decision-theoretic accounts, beliefs and desires should evolve in different ways, when new information is added to an existing set of attitudes, thus precluding the identification of desires with beliefs. If this is right,<sup>36</sup> it should provide a general reason for anyone who accepts decision theory to resist the identification of desires with beliefs.<sup>37</sup> It is doubtful, however, whether this result has any direct

<sup>34</sup> See e.g. Warner, *Freedom, Enjoyment and Happiness*, ch. I, where it is suggested that, on a rationalist view, certain thoughts may 'serve as' desires; and Philip Pettit, 'Humeans, Anti-Humeans, and Motivation', *Mind*, 96 (1987), 530–3, who proposes a version of rationalism on which the presence of desires is sometimes entailed by the presence of beliefs.

<sup>35</sup> See David Lewis, 'Desire as Belief', *Mind*, 97 (1988), 323–32, where the argument is presented using Bayesian decision theory; and John Collins, 'Belief, Desire, and Revision', *Mind*, 97 (1988), 333–42, who reaches a similar conclusion using nonquantitative decision theory.

<sup>36</sup> For some doubts, see Huw Price, 'Defending Desire-as-Belief', *Mind*, 98 (1989), 119–27. A better objection, found in very recent work by John Collins, challenges the argument's premise that rational belief-revision is always by conditionalization (or its nonquantitative analogue). Collins now thinks that there are independent reasons for holding that belief revision follows two, distinct methods: conditionalization, which is appropriate for genuine updating of belief; and the merely *hypothetical* revision implicit in the decision theoretic definition of expected value, which has quite different formal properties. The theorems proved in the papers by Collins and Lewis (see n. 35, above) give us further reasons for acknowledging these two distinct methods of belief revision; but they do not rule out the possibility that beliefs and intrinsic desires are necessarily connected. Collins's new position is developed in his paper 'Updating and Supposing', read at the A.A.P. Conference in Canberra in July 1989.

<sup>37</sup> Stephen Darwall has argued from decision theory to the opposite conclusion that the Humean approach must be wrong, in *Impartial Reason* (Ithaca: Cornell University Press, 1983), ch. 6. He maintains that we can only make sense of the decision-theoretic requirement of transitivity of

bearing on the debate between the Humean and the rationalist, as I have reconstructed it. For on neither of the models of a rationalist account which I have sketched is the rationalist committed to identifying desires and beliefs, or to holding that they are necessarily connected. What the rationalist does maintain is that certain kinds of desires can be given a rational explanation, in terms of an agent's beliefs (together with principles or norms of practical rationality). This would establish, perhaps, a *rational* connection between certain beliefs and certain desires. But to conclude, on this basis, that the beliefs and desires in question are identical, or necessarily connected, is to confuse the rational connections of explanation and justification with laws of psychological necessity. People are irrational, much of the time, and for this reason alone no rationalist account should identify beliefs with desires, or hold that they are necessarily connected.<sup>38</sup>

Still, the rationalist must establish the claim that there is at least a rational connection between beliefs and desires, and attempts to make out such a connection might appear to run into a different kind of problem. To see this, observe that the two models for a rationalist account which I sketched above both apparently license an inference from *factual* premises to an *evaluative* conclusion. On the first model, this occurs at the point at which a rational justification is provided for basic evaluative principles; on the second model, it occurs in the allegedly rational inference from specific factual beliefs to specific evaluative conclusions. Noticing this feature of rationalist accounts, Mark Platts has recently suggested that the force of Hume's own argument against the rationalist rests on his famous strictures against deducing 'ought'-conclusions from 'is'-premises.<sup>39</sup> For the ban on such inferences, if it could be sustained, and extended to all inferences from factual premises to evaluative conclusions, would indeed seem to rule out the rationalist explanation of motivated desires.

As an interpretation of Hume, however, this seems to me to get things the wrong way around: coming long after the Hume's discussion of the influencing motives of the will,<sup>40</sup> and lacking any independent support, the is-ought considerations appear simply to reflect Hume's basic anti-rationalist convictions, not to

individual preferences on the assumption that an agent's preferences are criticizable in terms of reasons. Even if this is correct, however, it does not yet establish that the terms for criticism of desires conform to a rationalist rather than a Humean pattern: the basic values which lie behind an agent's preferences, and render them commensurable, may themselves merely reflect the agent's intrinsic desires.

<sup>38</sup> This is the moral of Korsgaard's discussion of internalism, in 'Skepticism about Practical Reason'. See also Nagel, *The Possibility of Altruism*, 20–2, 65–7; Michael Smith, 'On Humeans, Anti-Humeans, and Motivation: A Reply to Pettit', *Mind*, 97 (1988), 589–95, at 591–2; and Michael Smith, *The Moral Problem* (Oxford: Blackwell, 1994), ch. 3.

<sup>39</sup> Platts, 'Hume and Morality as a Matter of Fact', 201–3. It should be stressed that Platts offers this as an interpretation of Hume, not as an argument he himself endorses. In discussing Platts's suggestion, I shall put aside the question of whether Hume's intent in the is-ought passage was really to propose a strict ban on deriving 'oughts' from 'is'-premises.

<sup>40</sup> The is-ought passage is in bk. III, pt. I, sect. iii of the *Treatise*, a full fifty pages after the discussion of the influencing motives of the will, which appears in bk. II, pt. III, sect. iii.



provide an argument for them. The point, moreover, is a general one. *Any* argument against rationalist accounts which merely invokes the alleged gap between is and ought (or between facts and values) seems bound to fail, because Humean strictures against deriving 'oughts' from 'is'-premises are not independent of Humean accounts of practical reason. An argument for accepting the ban on deriving 'oughts' from 'is'-premises would already *itself* be an argument for rejecting the rationalist approach to practical reason. But then the Humean needs to show us what that argument is; it is no use simply invoking the is-ought strictures, as if they had the status of an independent and established conclusion in the context of the debate between the Humean and the rationalist.

A third general argument against the rationalist approach has recently been suggested by Michael Smith.<sup>41</sup> Like the argument I presented in section 2, above, this argument turns on the teleological character of reason explanations. The apparent aim of Smith's new argument, however, is to show, not just that desires must be present on occasions of motivation, but that the explanation of motivated desires must conform to the principle of desire-out, desire-in. As we have seen, this is precisely what a general argument for the Humean account must establish.

Smith's own presentation of his position is succinct enough to be quoted virtually in its entirety:

A motivated desire is a desire had for a reason; that is, a desire the having of which furthers some goal that the agent has. The agent's having this goal *is*, in turn, *inter alia*, the state that constitutes the motivating reason that he has for having the desire. . . . But if the state that motivates the desire is itself a reason, and the having of this reason is itself constituted by his having a goal, then, given that the having of a goal is a state with which the world must fit rather than *vice versa* . . . , so it follows . . . that the state that motivates the desire must itself be a desire. Thus, the Humean will say, the idea that there may be a state that motivates a desire, but which is not itself a desire, is simply implausible.<sup>42</sup>

The structure of this argument is extremely straightforward. A motivated desire, Smith notes, is one that is explicable in terms of reasons. But reason explanations are essentially teleological, attributing a goal to the person who has the reason; and to have a goal is already to be in a state of desire. Of course that further desire may itself be motivated by a reason, but simple iteration of Smith's teleological argument suffices to show that the chain of explanations must eventually terminate in an unmotivated desire. Hence it is not an open question whether explanations of desires themselves always terminate in a desire. On the contrary, teleological considerations concerning the nature of reason explanations suffice to establish the Humean principle of desire-out, desire-in.

<sup>41</sup> See Smith, 'The Humean Theory of Motivation', 58–60; the argument is repeated in Smith's book *The Moral Problem*, ch. 4. A different argument, which like Smith's assumes that the states which rationally explain a desire must themselves be motives, may be found in Joel Marks, 'The Difference between Motivation and Desire', in Marks (ed.), *The Ways of Desire*, 133–48, at 136–42; my remarks about Smith's argument tell equally against the one that Marks offers.

<sup>42</sup> 'The Humean Theory of Motivation', 59.

So Smith argues; but his argument seeks to prove too much. It turns on the claim that reason explanations are necessarily teleological, in the sense that the psychological states which constitute one's reasons are always goal-directed states. But if this were true, then to have a reason for *believing* something would equally be to have some desire which enters into the explanation of the belief. This is clearly not the case, however, since rationalizing explanations of beliefs are ordinarily given exclusively in terms of further beliefs that the agent holds (together with principles or norms of theoretical rationality). Thus, when Smith confronts the rationalist with what is supposed to be a dilemma—that is, a choice between denying that to have a reason is necessarily to have a goal, and denying that desires are the states that realize one's having a goal<sup>43</sup>—the appropriate response is to grasp the first horn. Far from its being, as Smith suggests, a conceptual truth, the principle that 'to have a reason is to have a goal' is simply *false*, on the interpretation of it that is relevant to the dispute between the Humean and the rationalist. What the Humean needs to establish is something at once less general and more difficult: not that to have a reason which explains a propositional attitude is always to have a goal, but that it is to have a goal whenever the attitude to be explained is itself a goal-directed state. That is the desire-out, desire-in principle, and there is nothing in the considerations Smith adduces that would support this crucial claim.

In fairness, however, it should be mentioned that Smith's apparent misinterpretation of the rationalist position is one that is directly encouraged by Nagel's own presentation of that position. As I explained in section 3, above, Nagel attaches great significance to the distinction between motivated and unmotivated desires, without making it adequately clear what it is for a desire to be motivated. Taking the terminology of 'motivated' and 'unmotivated' states quite literally, Smith seems to have assumed that a motivated desire is one that is formed for something like an *ulterior* motive.<sup>44</sup> That is, he interprets motivated desires as states which are precisely *not* rationalized by other of the agent's propositional attitudes, but rather formed under pressure of some further aim or goal that the agent has (as in a case of wishful thinking). This is what leads him to suppose that explanation of motivated desires has to postulate some further, goal-directed state of desire. Once we are quite clear about what is at stake in the debate between the Humean and the rationalist, however, it also becomes clear that teleological considerations are not going to determine the outcome of the debate.

## 6.

To this point, I have offered an interpretation of the debate between the Humean and the rationalist, and in light of this account I have tried to show why teleological

<sup>43</sup> 'The Humean Theory of Motivation', 59–60.

<sup>44</sup> On this point, see Smith's more recent paper 'Reason and Desire', *Proceedings of the Aristotelian Society*, 88 (1987–8), 243–56, at 251–2.

and other considerations do not provide general grounds for preferring the Humean to the rationalist account. We are, I would submit, now in a much stronger position to accept Korsgaard's contention that motivational skepticism about the rationalist position cannot be sustained. Still further support for this conclusion may be provided by considering one additional and extremely influential discussion of practical reason, that found in Bernard Williams's paper 'Internal and External Reasons'.<sup>45</sup> Williams's discussion has widely been interpreted as an attempt to raise a quite general problem for rationalist accounts of practical reason, and while there is now some evidence that Williams himself does not share this interpretation,<sup>46</sup> it will be useful to take up his argument on the assumption that it does in fact aim to support the Humean account.

Williams's argument starts from what might be called an internalist view of practical reasons. That is, Williams assumes that a person's reasons for action must be deliberately accessible to the person, in the sense that it must be possible for the person to become motivated by her reasons for action as a result of purely rational reflection.<sup>47</sup> Given this assumption, Williams goes on to distinguish two positions about the conditions under which one may have a given practical reason. On the first view—in Williams's slightly confusing terminology,<sup>48</sup> the 'internal reasons theory'—people may only have a given reason if they could come to be motivated to act on the reason by deliberating *from* some desire in their 'subjective motivational set'. On an 'external reasons theory', by contrast, one's reasons are not required to stand in this kind of deliberative relation to one's antecedent desires.

Williams correctly takes the decision between these two positions to depend crucially on the question of the possibilities for rational explanation of motivations—the question, that is, that divides Humean and rationalist accounts of practical

<sup>45</sup> Reprinted in his *Moral Luck* (Cambridge, England: Cambridge University Press, 1981), 101–13.

<sup>46</sup> See his book *Ethics and the Limits of Philosophy* (Cambridge, Mass.: Harvard University Press, 1985), 223–4 n. 19, where Williams attributes to Kant the denial 'that there can be an absolutely "external" reason for action, one that does not speak to any motivation the agent already has'. The clear implication here is that the internal reasons theory which Williams has defended is something to which *both* Humeans and (Kantian) rationalists are committed; so that the difference between Humeans and rationalists would become a difference about the content of the dispositions which are to be included in the subjective motivational sets of rational agents. (See also Korsgaard, 'Skepticism about Practical Reason', sect. vi, where a similar reading of the internal reasons model is offered as part of an argument against Williams's position in 'Internal and External Reasons'.) If this is Williams's present view, however, it is difficult to reconcile with the text of 'Internal and External Reasons', which connects the internal reasons theory much more closely with the Humean approach to practical reason.

<sup>47</sup> 'Internal and External Reasons', 108–9. Actually Williams attributes this assumption to the 'external reasons' theorist, rather than endorsing it *in propria persona*; but it seems clear from the course of his argument that he takes the assumption to describe a genuine condition on practical reasons for action.

<sup>48</sup> The terminology is confusing because both Williams's 'internal reasons theory' and his 'external reasons theory' are, in a more conventional sense, *internalist* accounts, postulating a necessary connection between agents' reasons for action and their motivational capacities.

reason. Given the general internalist assumption, the external reasons theorist must hold that agents can acquire the motivation to act on their reasons as a result of rational reflection. At the same time what is distinctive about this view is precisely the denial that one's reasons need be restricted by one's prior desires, as considerations one could come to be motivated on by deliberation *from* those prior desires. The external reasons theorist thus needs to defend a rationalist account of practical reason, to show that rational reflection can give rise to new motivations without taking the form of deliberation *from* prior desires in the agent's subjective motivational set. Once these requirements are clearly set out, however, Williams apparently finds the external reasons theory easy to dismiss. He says, simply: 'I see no reason to suppose that these conditions could possibly be met.'<sup>49</sup>

The real problem, however, is to see how, at this level of generality, any such statement about the prospects for an external reasons account could possibly be defended. Williams may be taking for granted here some version of the teleological argument, correctly assuming that the explanation of motivation requires the postulation of a desire on the part of the agent who is motivated, and inferring from this that practical deliberation must be deliberation *from* the desires in one's prior motivational set. But the inference is unsound, as we have seen: the desire implicated in motivation, by the teleological argument, may itself be a motivated desire, in which case it will constitute not the starting point for practical deliberation but its conclusion. General skepticism about the possibility of pure practical reason would have to be based on a defence of the desire-out, desire-in principle, in rationalizing explanations of desires. But Williams, in line with most proponents of the Humean view, provides no grounds for this crucial Humean principle.

He does, it is true, at one point allege that there is a circularity in what he presumably takes to be the most plausible version of an external reasons theory.<sup>50</sup> This is an account according to which agents can come to be motivated in a certain way simply by coming to believe that they have reason to act in that way, where coming to believe such a thing is not necessarily a matter of deliberating *from* the motives in one's prior subjective motivational set. About this proposal, Williams demands to know in what the content of such a belief could possibly consist. Answering this question himself, he suggests that the content of the belief must consist in 'the proposition, or something that entails the proposition, that if [the agent] deliberated rationally, he would be motivated to act appropriately'.<sup>51</sup> But this answer, Williams observes, merely takes for granted the possibility of pure practical reason, without showing us how reasoning which does not start from the agent's prior desires can generate by itself a new motivation.

To see why this is a weak objection, consider the following analogy with a case of theoretical reasoning, or reasons for belief. Presumably, everyone would agree

<sup>49</sup> 'Internal and External Reasons', 109.

<sup>50</sup> Ibid. 109–10. Cf. Brad Hooker, 'Williams' Argument against External Reasons', *Analysis*, 47 (1987), 42–4, for a similar criticism of this part of Williams's discussion.

<sup>51</sup> 'Internal and External Reasons', 109.

that an agent could come to draw a new theoretical conclusion by coming to believe that there is reason to draw that conclusion. But suppose we now ask the question that is the analogue of Williams's question in the practical case: namely, in what might the content of the belief possibly consist? Here, it appears, there are any number of answers that might be given, such as that the new conclusion is a logical consequence of other beliefs of the agent's which she is not prepared to give up. To follow Williams's treatment of the practical case, this propositional content may indeed *entail* (in conjunction with some minimal assumptions about theoretical rationality) that if the agent were to reason correctly, she would draw the new theoretical conclusion. But this does nothing to show that the original answer was circular, or question-begging, or otherwise uninformative.

Perhaps Williams was misled here by an ambiguity in his formulation of the rationalist position. It is, he says, the view that an agent can acquire a new motivation as a result of coming to believe that there is reason for him to act in a certain way. But the existential proposition that gives the content of this belief may be read in two different ways: either as the claim that there is some such reason for action or other, where the agent does not necessarily know what that reason is; or as the claim that there is a particular practical reason which the agent grasps and understands.<sup>52</sup> Naturally the first interpretation of the proposition would leave us unsatisfied about the possibility of a purely rational explanation of motivation and motivated desire, if it were indeed all the rationalist had to say about the matter. But as the analogy with belief shows, the interpretation to opt for would anyway be the second. That is, the rationalist should say that pure practical reason is possible, because agents can acquire both new motivations, and the motivated desires implicated in such motivations, by coming to grasp and understand the particular reasons that they have for acting in certain ways. Of course, for this to be a satisfying account we will need to be convinced that the motivated desires involved here really are explicable solely in terms of the agent's new beliefs, plus principles or norms of rationality; otherwise what is represented as 'coming to grasp and understand the particular reason for action that one has' will not be a genuine case of *pure* practical reflection. But again, Williams has said nothing that would rule out this form of explanation in principle.<sup>53</sup>

## 7.

My topic in this essay has been a question about the explanation of motivation and action, the question, namely, whether we can explain motivation and action

<sup>52</sup> Williams himself, in a different context, notices precisely this ambiguity in existential claims about reasons; see *ibid.* 107.

<sup>53</sup> For further discussions of Williams's account, see Rachel Cohon, 'Are External Reasons Impossible?', *Ethics*, 96 (1985–6), 545–56; John McDowell, 'Might There Be External Reasons?', in J. E. J. Altham and Ross Harrison (eds.), *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams* (Cambridge, England: Cambridge University Press, 1995), 68–85; and Martin Hollis, *The Cunning of Reason*, ch. 6.

in distinctively rational terms. To provide explanations of this kind—particularly explanations of moral motivation and behaviour—has been a characteristic aspiration of Kantian approaches to moral philosophy, and I have urged that there is no reason to think that this aspiration would be impossible to satisfy. Some recent work, however, has suggested a different way of developing a Kantian (or at any rate, anti-Humean) approach to practical reason, one which is not committed to the possibility of explaining motivation and action in purely rational terms. This possibility requires brief consideration.

Even if, as I have argued, rational explanations of action could in principle be carried through, it might be that they would add nothing significant to the formulation of a Kantian position in ethics. This, in effect, has been suggested by Michael Smith.<sup>54</sup> Smith distinguishes between two ways of understanding the Kantian claim that moral requirements are, or are based on, norms or requirements of practical reason. ‘Belief-rationalism’, as he calls it, takes the norms or requirements of practical reason to be capable of explaining motivation and motivated desire, in line with my depiction of the rationalist position in this paper. What Smith calls ‘desire-rationalism’, by contrast, holds that explanations of motivation and desire conform to the Humean principle of desire-out, desire-in. Unlike the Humean, however, Smith’s desire-rationalist maintains that certain basic desires are intrinsically rational, in the sense that all rational agents must have those basic desires. On this view, the norms or requirements of reason tell us that certain intrinsic desires are rationally required, without being able to explain the formation of such desires.<sup>55</sup>

Having introduced these two rationalist positions, Smith proceeds to question the significance of the differences between them, and hence to raise a doubt about the importance of the aspiration to explain motivation in rational terms. He notes that both positions will be able to identify the same sorts of behavior as irrational: where the belief-rationalist says (for instance) that the amoral person is failing to reason correctly, in accordance with norms or principles of rationality, the desire-rationalist will say that the person lacks a basic desire which is intrinsically rational.<sup>56</sup> The same conclusion holds for rational behavior, on Smith’s view. On the belief-rationalist account, the rational agent will have specific moral desires which can be explained rationally in terms of principles or norms of reason. But Smith argues that any agent who has such specific moral desires could equally be credited with a basic dispositional desire with moral content, which the desire-rationalist may characterize as intrinsically rational.<sup>57</sup>

This seems to me correct, as far as it goes, but Smith is wrong to conclude from it that the differences between his two forms of rationalism are insignificant.<sup>58</sup>

<sup>54</sup> See Smith, ‘Reason and Desire’, sect. 4.

<sup>55</sup> Smith suggests that the critical present aim theory discussed by Derek Parfit might be a version of desire-rationalism, in so far as it claims that certain desires are rationally required, certain others intrinsically irrational; see Parfit, *Reasons and Persons*, pt. 2 (esp. sect. 46).

<sup>56</sup> Smith, ‘Reason and Desire’, 253–4.

<sup>57</sup> Smith, ‘*The Moral Problem*’, Oxford University D. Phil. (1989), ch. 8, sect. 4.

<sup>58</sup> Still less does it seem correct to say that belief-rationalism ‘collapses into’ desire-rationalism, as Smith suggests, *ibid.*

Take the case of the fully rational agent. It is true that on both views, such an agent may be credited with the same standing or dispositional desires. But these desires will have very different explanatory roles in the two kinds of account. For the desire-rationalist, they fix the starting points for the rational explanation of any specific moral desires that the rational agent may have, since the desire-rationalist only admits explanations of desires that satisfy the Humean desire-out, desire-in principle. On a belief-rationalist account, by contrast, it is possible to explain specific intrinsically moral desires in purely rational terms; and these explanations may equally account for the dispositional or standing desires that we ascribe to rational agents in virtue of their specific moral desires.<sup>59</sup>

Thus, Smith's two kinds of rationalist position continue to differ in their implications for the rational explanation of motivation. Nor is this difference an insignificant one. Its importance is clearest in cases where a formerly amoral agent comes to acquire the desires and motivations characteristic of virtue. On both of the rationalist positions Smith describes, this transition will mark a change from irrationality to rationality. But if desire-rationalism is correct, the transition cannot itself be explained in rational terms; rather it must be described as something akin to a conversion—the acquisition of a new motivation which is constitutively rational, but which does not admit of a rational explanation. Belief-rationalism, on the other hand, would enable us to see the transition from amorality to virtue as a *reasoned* change in the agent's motivations, the result of a process of reasoning or reflection in accordance with the norms or principles of practical rationality. Thus belief-rationalism promises to offer terms for a rational discussion with the amoralist, terms which—in principle, at least—the amoralist might be persuaded by, leading to a reasoned adjustment of or addition to the amoralist's motivations.<sup>60</sup> There is no such possibility on the view that Smith calls desire-rationalism.

I conclude that the aspiration to explain motivation and behavior in purely rational terms makes a distinctive contribution to the Kantian approach to ethics. Once this is granted, however, a further question arises, concerning the extent to which other characteristically Kantian or anti-Humean claims about practical reason can be made by a theory which renounces the aspiration to provide purely rational explanations of motivations. This question is posed, albeit in very different ways, by the recent work of John Rawls, T. M. Scanlon, and John McDowell.<sup>61</sup> These philosophers reject the Humean claim that practical reason is

<sup>59</sup> I take it this is the possibility Nagel is suggesting, when he offers the following comparison with the beliefs involved in deductive inference (in *The Possibility of Altruism*, 31): 'If someone draws conclusions in accordance with a principle of logic such as *modus ponens*, it is appropriate to ascribe to him the belief that the principle is true; but that belief is explained by the *same* thing which explains his inferences in accordance with the principle.'

<sup>60</sup> It need not be supposed that such reasoned discourse with the amoralist would be an easy thing to bring off in practice; only that it would be possible.

<sup>61</sup> See John Rawls, 'Justice as Fairness: Political Not Metaphysical', *Philosophy and Public Affairs*, 14 (1985), 223–51; T. M. Scanlon, 'Contractualism and Utilitarianism', in Amartya Sen and Bernard Williams (eds.), *Utilitarianism and Beyond* (Cambridge, England: Cambridge University Press, 1982), 103–28; and John McDowell, 'Virtue and Reason', *The Monist*, 62 (1979), 331–50.

exclusively instrumental in its function, restricted to identifying means to the satisfaction of individual ends; and they reject equally the modern development of the instrumentalist approach, according to which practical reason is directed toward maximizing the joint satisfaction of sets of desires or preferences. Instead, Rawls, Scanlon, and McDowell all contend that there are noninstrumental patterns of practical reasoning in terms of which we can criticize actions and social institutions morally, and explain distinctively moral kinds of motivation and concern. And yet, none of these philosophers seems to suppose that the morally distinctive patterns of practical reasoning are equally accessible to all agents, regardless of those agents' antecedent desires.<sup>62</sup>

If this is right, it suggests that there may be considerable work for a Kantian or anti-Humean<sup>63</sup> approach to ethics, even if it renounces or suspends the aim of explaining motivation and action in purely rational terms. The work would consist in showing that there are patterns of reasoning *from* an agent's desires—to use Bernard Williams's expression<sup>64</sup>—which are not forms of instrumental or maximizing reasoning. This seems to me a fertile area for further research and discussion: questions arise, for instance, concerning the explanatory *role* that desire plays in forms of practical reasoning that are neither instrumental nor maximizing, but that nevertheless, in some sense, begin from an agent's desires. But however these questions are answered, the possibility that reasoning from one's desires need not take a maximizing or instrumental form is an important one. It does not undermine the aspiration to provide purely rational explanations of motivation, or show it to be misguided. But it suggests that rather less may hang on the satisfaction of the aspiration than some Kantians have supposed.

## 8.

I have argued that Humeans have so far failed to give us good reasons for accepting motivational scepticism about the rationalist approach. To support this conclusion, it has been necessary to get clearer about what is at stake in the debate

<sup>62</sup> On this point, see Rawls's remarks about the role of consensus in political justification, in 'Justice as Fairness: Political Not Metaphysical', 229, 246–7; Scanlon's suggestion that the desire to be able to justify one's actions may be the basis of moral motivation, in 'Contractualism and Utilitarianism', 116–18; and McDowell's denial that those who lack the desires characteristic of virtue need be irrational, in 'Are Moral Requirements Hypothetical Imperatives?', *Aristotelian Society, suppl. vol.* 52 (1978), 13–29, at 13, 24.

<sup>63</sup> Rawls's and Scanlon's accounts are broadly Kantian, whereas McDowell's is more directly inspired by his reading of Aristotle.

<sup>64</sup> See 'Internal and External Reasons', 104–5, where Williams too contends that reasoning from the elements in one's subjective motivational set need not take an instrumental or maximizing form. I do not mean to suggest that Rawls, Scanlon, and McDowell would all agree with Williams's way of putting the point (see McDowell, 'Might There Be External Reasons?', for some objections). Whether disagreements on this point are significant, however, will depend on how the notion of reasoning *from* one's desires is developed.



between Humeans and rationalists. Developing a suggestion of Thomas Nagel's, I have proposed that the issue turns on the form taken by rationalizing explanations of desires and motivations; in particular, the question is whether such explanations conform to the Humean principle of desire-out, desire-in. Once the question is correctly seen in this way, however, it becomes clear that recent arguments, purporting to offer general support for the Humean approach, do not succeed.

How are we to proceed beyond this point? One approach that suggests itself is to turn to phenomenology to settle the issue. The idea would be that, once we see that there is nothing to rule out or determine either kind of account on general or *a priori* grounds, we may appeal to the evidence of the moral life to see what kinds of explanations of motivation are actually accepted in practice. Evidence of this sort, however, seems to me of extremely limited value in the context of the debate between the rationalist and the Humean. The main reason for this is that the conception of desire with which the Humean operates is a dispositional rather than a phenomenological conception, a conception, that is, on which one may be in a state of desire without being aware that one is in such a state. But if desires are not phenomenological states, it is unclear how phenomenological or experiential evidence could possibly settle the issue of whether desires always serve as the ultimate source of our motivations.

Indeed, so long as the issue is joined in experiential terms alone, it is open to the Humean to confront the rationalist with a dilemma. The rationalist will want to say that certain motivations and motivated desires may be explained in terms of beliefs, because in moral experience it seems (for instance) as if the desire to perform a given action may be both justified and explained by the agent's belief that the action would be, say, helpful, or just. If this is an ordinary belief, however, then it should be equally open to an amoral agent to have the belief; and yet the belief that an action would be helpful or just will not give rise, in the amoralist, to the motivated desire to perform the action.<sup>65</sup> Pointing to this difference, the Humean will say that it can best be explained by supposing that the virtuous agent has, while the amoral agent lacks, a dispositional desire to perform actions that are helpful or just, where this dispositional desire is the real source of the virtuous person's motivations. Phenomenological evidence, in other words, seems rather to support the Humean approach than to confute it. The rationalist may try to avoid this outcome, perhaps, by denying that the beliefs which seem to explain motivations and motivated desires are ordinary beliefs, available equally to those who are

<sup>65</sup> Richard Warner offers an experiential argument against the Humean which allows for this possibility (namely, that the very same beliefs which motivate may sometimes fail to serve as motives), in *Freedom, Enjoyment and Happiness*, 46–51. His anti-Humean position is different from the one I am considering in the text in holding, not that beliefs explain desires, but that they sometimes 'serve as' desires (see sect. 5, above, for a discussion of this proposal). I am supposing, however, that an experiential argument similar to Warner's could be mounted in support of the conclusion that beliefs may sometimes explain desires; and that both arguments would be subject to the same objection.

not motivated by them.<sup>66</sup> But this conception of ‘extraordinary’ beliefs seems questionably coherent; and in any case, its invocation is completely *ad hoc* so long as the debate is being conducted on phenomenological or experiential grounds alone.

What the rationalist needs to establish is that it is *irrational* to have certain sorts of beliefs, but lack the corresponding desires and motivations. To show this, it is necessary to go beyond phenomenology, by specifying the relevant principles or norms of reason, or by offering a conception of rationality in terms of which the inferences in question can be shown to be rationally required. The real burden of the rationalist position is to find a way of defending such a conception of rationality or of the specific principles or norms of reason. One strategy, adverted to in section 1, above, would proceed by showing that a rationalist account of the principles or norms of reason yields a more plausible ideal of rational attainment than alternative, Humean accounts. Or, rationalists might try to establish that their conception of practical rationality develops or extends certain aspects of theoretical rationality (such as impersonality, consistency, or universality) that may be taken as paradigmatic.

Much work remains to be done, before we can definitively judge the outcome of these strategies (which are themselves rather poorly understood). At this point, I would only venture to say that the success of the rationalist strategies is likely to be affected by the increasing sophistication of accounts which deny that pure reason can be practical in its issue (or which at least remain noncommittal on this question). Thus, to the extent that such accounts can be freed from an instrumental or maximizing conception of practical reason (see section 7, above), it will become correspondingly difficult to show that rationalist accounts provide a uniquely plausible ideal of rational attainment. Difficult, but not impossible: further progress with this debate is only going to be made once we give up the thought that there must be something wrong in principle with the attempt to explain motivation in rational terms, and begin assessing particular rationalist and Humean proposals on their merits.

<sup>66</sup> See John McDowell, ‘Are Moral Requirements Hypothetical Imperatives?’, 16–17, for this suggestion. The extraordinary character of the beliefs McDowell here takes to be explanatory of moral behaviour is revealed by his admission that ascription of the beliefs is independent of ‘ordinary’ tests for mastery of the language in which the content of the beliefs is expressed (Ibid. 22). Doubt about the coherence of this notion is suggested by Sabina Lovibond in *Realism and Imagination in Ethics* (Minneapolis: University of Minnesota Press, 1983), sect. 12. (I take it that in ‘Virtue and Reason’, referred to in sect. 7, above, McDowell offers further, nonexperiential support for his anti-Humean conception of practical reason.)