How can we motivate ourselves to do what we think we ought? How can we deliberate about personal values and priorities? Bennett Helm argues that standard philosophical answers to these questions presuppose a sharp distinction between cognition and conation that undermines an adequate understanding of values and their connection to motivation and deliberation. Rejecting this distinction, Helm argues that emotions are fundamental to any account of value and motivation, and he develops a detailed alternative theory both of emotions, desires, and evaluative judgments and of their rational interconnections. The result is an innovative theory of practical rationality and of how we can control not only what we do but also what we value and who we are as persons.

BENNETT HELM is Assistant Professor of Philosophy at Franklin & Marshall College, Pennsylvania. He has published articles in Hume Studies, American Philosophical Quarterly, Pacific Philosophical Quarterly, Southern Journal of Philosophy and Noûs.
Emotional reason
Emotional reason
Deliberation, motivation,
and the nature of value

Bennett W. Helm
Franklin & Marshall College, Pennsylvania
To Karen
Contents

Acknowledgements  
1 Two problems of practical reason  

PART I  FELT EVALUATIONS  
2 Emotions and the cognitive–conative divide  
3 Constituting import  
4 Varieties of import: cares, values, and preferences  

PART II  PRACTICAL REASON  
5 Single evaluative perspective  
6 Rational control: freedom of the will and the heart  
7 Deliberation about value  
8 Persons, friendship, and moral value  

Select bibliography  
Index
Acknowledgements

I began thinking systematically about the topic of this book in 1997 during Simon Blackburn's National Endowment for the Humanities Summer Seminar, "Objectivity and Emotion in Practical Reason." Thanks to Simon and the other participants in that seminar, especially Justin D'Arms and Daniel Jacobson, for their thoughtful criticisms. Thanks also to Yaroslava Babych and Aleksandra Markovic who, under the aegis of the Franklin & Marshall College Hackman Scholar's Program, provided helpful discussion of the issues raised in §4.3. An initial draft of this book was written with the financial support of a National Endowment for the Humanities Summer Stipend (1998) and an American Council of Learned Societies Fellowship (1998–99); I am grateful to these institutions as well. Finally, thanks to Michael Murray for providing comments on an early draft of the complete manuscript.

Parts of this book have appeared in embryonic form elsewhere. The discussion of freedom of the heart in §6.4 derives from my "Freedom of the Heart," *Pacific Philosophical Quarterly*, 77 (1996), pp. 71–87; the example of Betty in §5.2 and its moral are based on my "Integration and Fragmentation of the Self," *Southern Journal of Philosophy*, 34 (1996), pp. 43–63; and a simplified version of chapter 7 was published as "Emotional Reason: How to Deliberate about Value," *American Philosophical Quarterly*, 37 (2000), pp. 1–22. I am grateful to the editors and publishers of these journals for their permission to use material from these essays.
Two problems of practical reason

1.1 Rational Animals

The idea that a person is a rational animal, as a formula for understanding what is distinctive about persons, has been enormously influential in attempts to understand ourselves as human agents. Of course, insofar as we recognize other animals like dogs and cats as agents, we thereby implicitly understand them as rational in a certain way, for psychological explanation is essentially explanation in terms of reasons. What make us different from mere animals, however, are the distinctive abilities to reason we exhibit. Thus, roughly, our reasons are at least potentially articulate and informed by linguistic concepts, and it is because of our abilities to articulate, clarify, and criticize these reasons that we can self-consciously choose what to believe, do, and value, and why. The possibility of such choice brings with it the possibility for distinctive kinds of freedom: freedom not only to act but also to choose our visions of the good, visions which partly define the kind of persons we are. Moreover, such articulateness and freedom make intelligible our being responsible not only for what we do, but also (and more importantly) for who we are.

In this brief description, I have already laid out some important features of the kind of reason that defines us as persons: it is linguistically informed, articulate, self-conscious, critical, and reflexive – i.e., about not only the world but also ourselves. Merely to identify these features is not, of course, to have a complete account of the nature of human reason, let alone of what it is to be a person, for many puzzles remain. Two such puzzles will be my primary concern. The first, the motivational problem, is a puzzle about the connection between our choosing something as the outcome of deliberation and our being motivated to pursue it. The other, the deliberative problem, concerns the nature of our deliberation about personal values, deliberation which can seem paradoxical insofar as
Emotional reason

what values we decide upon both is a matter of individual choice (and so is something we freely invent) and is simultaneously accountable to standards of reason (and so is something we can discover). Solving these problems is a necessary step in understanding the kind of reason that defines us as persons. The trouble is, these problems have proved extraordinarily recalcitrant.

As I shall argue in this chapter, such recalcitrance has its source in common philosophical ways of understanding these problems, ways that are rooted in commonly accepted conceptions of rationality (as either epistemic or instrumental) and the nature of the mind quite generally (as divisible into the cognitive and the conative). In §1.2, I shall articulate these common conceptions, arguing in §§1.3–1.4 roughly and in outline how they result in an understanding of the motivational and deliberative problems that blocks any satisfactory solution. This motivates my central claim that if we are to solve these problems we must rethink this common understanding of the mind, rejecting the sharp divide between cognition and conation and the correlative conception of rationality as either epistemic or instrumental. Instead, I shall offer an alternative conception of reason that accords desire and emotion a more central place, an alternative I sketch in §1.5 by way of a preview of the rest of the book. As I shall ultimately argue, this alternative enables me to formulate the motivational and deliberative problems in a way that makes them more tractable and so provides a better understanding of human practical reason as well as of persons.

Insofar as my aim is to understand persons and our mental lives, however, it presupposes a general conception of the nature of the mind and psychological explanation, which I shall now make explicit. Rationality, Donald Davidson rightly claims, is the constitutive ideal of the mental.\(^1\) Part of the point of this claim is to identify a form of explanation distinctive of mental phenomena. Thus, whereas explanation in the physical sciences works by locating physical phenomena within a broader context of other physical objects and events related by laws, explanation of mental phenomena proceeds by locating these phenomena within a broader context of other mental states or events (including intentional action) in such a way as to reveal their rationality. So whereas physical explanation reveals the explanandum as to be expected given the

Two problems of practical reason

antecedent conditions, psychological explanation reveals the explanandum as what rationally ought to happen.²

The claim that rationality is the constitutive ideal of the mental, however, does not concern merely the nature of psychological explanation, for the possibility of explanation in terms of rationality is a condition of the intelligibility of the mental as such. Again, compare this to analogous claims about the physical. We could not make sense of something as physical unless we are potentially willing to locate it within a broader pattern of law-governedness in terms of which it can be explained. Thus, numbers (or, as Davidson has argued, mental phenomena) are not governed by strict laws and so are not physical. If some apparently physical object were consistently to behave in ways that manifestly violate the physical laws as we understand them, our first inclination would be to dismiss it as not real: as an illusion or hallucination. Of course, our first inclination may turn out to be wrong, and so the object may force us to revise, even substantially, our conception of the relevant laws, but this only reinforces the point: the intelligibility of something as physical presupposes that it fits within a pattern of law-governedness, and it is this fact that guides revisions in our understanding of the relevant laws. Hence, law-governedness is the constitutive ideal of the physical in that the possibility of explanation in terms of laws is a condition of the intelligibility of the physical as such. Likewise, mental phenomena are intelligible only as located within a broader pattern of rationality in terms of which they can be explained; rationality in this way is constitutive of the mental as such. This needs further clarification.

It would be clearly false to claim that every particular mental state must be identifiable as rational in order to be a mental state at all: irrational beliefs, desires, emotions, etc. are all too common. Rather, rationality is the constitutive ideal of the mental in the sense that a creature is intelligible as an agent, as having various mental capacities, only if its exercise of those capacities is for the most part rational. This "looseness" in the constitutive ideal of the mental therefore makes possible isolated instances of irrationality, but only if they are isolated: too much irrationality destroys the essential background that makes agency possible. Thus, it is possible for a person irrationally to believe falsehoods even when the truth is staring her in the face; yet, as the number of false beliefs and failures to be responsive to manifest truth

Emotional reason

increases, it becomes increasingly difficult to make sense of her as understanding the content of what she professes to believe and so as having any beliefs at all about the relevant domain. (Witness, for example, my lack of understanding of or even belief about esoteric areas of theoretical physics.) Too many false beliefs, too many failures of responsiveness to the truth, spread widely enough through the person’s system of beliefs therefore erodes the background of epistemic rationality against which her having the capacity for belief at all is intelligible. Likewise, it is possible for someone irrationally to fail to take obvious means to his ends; yet too many such failures, widespread throughout his practical activity, erodes the background of instrumental rationality against which his having the capacity for desire at all is intelligible.

In short, a creature’s generally being rational is a condition of the possibility of its being a subject of mental states and so having various mental capacities, though this allows for isolated instances of irrationality. To articulate various kinds of rationality more clearly, such as epistemic or instrumental rationality, is therefore to understand better what it is to be an agent and so to have the relevant mental capacities. It is in this sense that rationality is the constitutive ideal of the mental. Moreover, we can identify and distinguish different kinds of agents by identifying and distinguishing differences in the kinds of rationality proper to them. Thus, to understand the kind of rationality distinctive of persons is nothing other than to understand the nature of personhood itself: a person just is a distinctive kind of rational animal.

1.2 THE COGNITIVE–CONATIVE DIVIDE

Ordinary philosophical understandings of intentional mental states divide them into cognitions and conations. The fundamental difference here is typically captured in terms of the notion of a direction of fit as a way of thinking about the conditions of success of these intentional mental states. Thus, our cognitions, such as belief and judgment, have a mind-to-world direction of fit in the sense that when there is a conflict between what we cognize to be the case and what really is the case, it is our cognitions that ought to change in order to conform to the world in order for these cognitions to be successful — to be true. This is because cognitive states are about how things are and so are to be assessed in

---

Two problems of practical reason

terms of standards of truth; the world, as that which is true or not, is consequently the source of these standards. By contrast, our conations, such as desire, have a world-to-mind direction of fit, for when a conation is unfulfilled, it is the world that ought to be changed in order to conform to our minds in order for our conations to be successful – to be satisfied. Thus, conation is the source of the standards by which we assess the state of the world as satisfactory or not. In each case, it might be thought, the direction of fit provides motivation for change in cases of mismatch: to change one's mind in cases of cognitions, and to change the world in cases of conations. Such motivation for change is rationally required by the kinds of intentionality these states have; consequently, that one generally is motivated in these ways is a condition of the possibility of one's having the capacities for cognition and conation.

Clearly, it seems, these two directions of fit are mutually exclusive. For if the world is the source of the relevant standards for assessing the fitness of the connection between mind and world, as is the case with mind-to-world direction of fit, then the world must have a kind of rational priority over our cognitions; as such, the world cannot simultaneously be liable to correction by our minds, on pain of giving up this priority and so the idea that cognitions are at issue. The converse is true for world-to-mind direction of fit. Consequently, cognitive and conative states seem to be fundamentally different kinds of mental states. Moreover, these two directions of fit seem exhaustive, for the notion of a direction of fit is the notion of the source of a normative standard, and there seems to be no alternative source of such standards than either the mind or the world. Consequently, any intentional mental state must be either a cognition or a conation but not both.

This understanding of cognition and conation as fundamentally different both reinforces and is reinforced by a common conception of reason according to which reason is in general a matter of figuring out how to achieve the requisite correspondence between mind and world, given the appropriate direction of fit; differences in kinds of reason therefore correspond to differences in the requisite direction of fit. Thus, in theoretical reason we try to figure out how things are by articulating and applying the norms of epistemic rationality so as to get our thoughts to correspond to the world with the sort of mind-to-world direction of fit characteristic of cognition. Likewise, in practical reason we try to figure out how to get the world to correspond to our thoughts by articulating and applying the norms of instrumental rationality; by acting in the ways
Emotional reason

prescribed by practical reason, we thereby impose on the world the sort of world-to-mind direction of fit characteristic of conation. As with the distinction in kinds of direction of fit, this distinction of reason into theoretical and practical is typically understood to be mutually exclusive and exhaustive.

It might be thought, however, that emotions are exceptions to this claim that all contentful mental states must be either cognitions or conations, with either one or the other direction of fit. Consider even a simple emotion like fear. Fear is a response to one's situation being of a certain kind: something is dangerous. As such, the rational appropriateness of fear depends on the situation really being that way: fear is in part a cognitive state, having a content with mind-to-world direction of fit. Yet fear is also evaluative and motivating. The content of fear is not merely that something is about to be destroyed or harmed, but that this is a bad thing, and as a result one is motivated to do something to prevent or mitigate that harm. In this way, fear is a conative state, with world-to-mind direction of fit. So in a sense fear involves both directions of fit.

Nonetheless, according to philosophical orthodoxy, emotions can be easily accommodated within the theory if we understand these directions of fit as issuing from conceptually separable components of the emotions. Thus, emotions should be understood as compound states and so as consisting of both cognitive and conative states somehow bundled together in a unitary package. This is the idea behind cognitivist theories of emotions, which understand emotions to be reducible to beliefs (as the prototypical cognitive state), desires (as the prototypical conative state), and perhaps some bodily sensation. (So as to avoid confusion of cognitivism about the emotions with cognitivism about values, I shall refer to the former as "emotional cognitivism.") Fear, therefore, can be understood as a belief that danger looms, the desire to avoid that danger, and, perhaps, the sensation of a knot in one's stomach.\footnote{Not all theories of emotions can comfortably be understood as cognitivist in the sense of reducing emotions to beliefs and desires. Robert Roberts and Patricia Greenspan each reject the idea that fear, for example, involves the belief that a danger looms, arguing that we can make better sense of the conflict within a person who is both afraid and believes that there really is no danger if we understand the fear to involve merely the appearance of danger. Nonetheless, the fundamental strategy of understanding emotions in terms of distinct cognitive and conative states remains. (For details, see Roberts' "What an Emotion Is: A Sketch," Philosophical Review, 97 (1988), pp. 183–209; and see Greenspan's Emotions and Reasons: An Inquiry into Emotional Justification (New York: Routledge, 1988).) More} Moreover, this basic reductivist strategy can be applied to all our complex mental states
Two problems of practical reason

and attitudes, including caring, valuing, loving, trusting, and respecting. Hence, the mind is understood to be fundamentally composed of three kinds of mental state: cognition and conation, as well as non-intentional qualia.

Both this conception of rationality as either epistemic or instrumental (but not both) and the conception of contentful mental states as either cognitions or conations (but not both) seem inevitable. The result is that cognition comes to be seen as intrinsically more rational than conation in two ways. First, we can understand only our cognitions as fully rational in that they are always subject to correction because of the norms of theoretical rationality in light of the mind-to-world direction of fit they display. Conations, with their world-to-mind direction of fit, cannot be rationally assessed in this way, and so are intrinsically less rational than our cognitions. Of course, our conations are "correctable" insofar as we can subject them to norms of consistency and coherence, especially in light of instrumental rationality, but there is nothing outside conation to which they must be answerable, on pain of giving up the idea that these are conations and not cognitions. Second, insofar as reasoning is a matter of figuring out how to achieve the relevant correspondence between mind and world and so issues in a practical or theoretical judgment, our ability to reason is a cognitive rather than a conative ability. For, as a result of deliberation, we come to articulate in judgment our best understanding of how things are or of what to do, an understanding that, because it can be explicitly confronted with the world through the application of epistemic and instrumental reason, has a claim to objectivity. By contrast, our emotions and non-instrumental desires are subjective and less rational: to the extent that our emotions and desires are merely affective, they do not reflect at all how things objectively are and so cannot be held to rational standards except those of consistency and coherence.5

This account of cognition and conation as mutually exclusive and exhaustive categories of intentional mental states, together with the correlative account of epistemic and instrumental rationality, I shall call the "cognitive-conative divide." Of course, the cognitive-conative divide is somewhat of a caricature; nonetheless, the basic view is extremely seductive and is deeply entrenched in philosophical thinking about the

5 Of course, to the extent that our emotions have a cognitive element and so are about the world, they are answerable to judgment, which always rationally trumps them in cases of conflict.
Emotional reason

mind quite generally and moral psychology in particular. As I shall now (in §§1.3—1.4) suggest more than argue, this cognitive—conative divide underlies our ordinary understanding of the motivational and deliberative problems and makes the solutions to these problems intractable. This makes way for my central claim, sketched in §1.5, that the real solution is to reject the cognitive—conative divide and thereby reconceive the motivational and deliberative problems in a way that makes their solutions more readily apparent.

1.3 THE MOTIVATIONAL PROBLEM

As I indicated in §1.1, the motivational problem is a problem about the connection between deliberative choice and motivation. On the face of it, it seems there must be a tight connection between the decisions we make as the result of deliberation about what to do and what we are motivated to do, for it is by making such decisions that we exercise control over our motivations. Indeed, it can seem, somewhat naively, that to make such a decision just is to be motivated to act accordingly. The difficulty arises, however, when we are confronted with phenomena like those of weakness of the will and listlessness, which seem to undermine any straightforward conceptual connection between our considered evaluative judgments and motivation.

To see this, consider the following standard example. You can look at the chocolate cake, decide that, all things considered, it is best not to have a second piece right now, but nonetheless take and eat it, even while saying to yourself, “I shouldn’t be doing this,” and trying not to. Here you evaluate more cake as bad, struggle not to eat any, but end up doing so anyway; you suffer from weakness of the will. In extreme cases, you can make this evaluation that having more cake is bad but not feel at all motivated to act in accordance with it and so not struggle at all to overcome your akratic appetite; such complete failures of motivation are cases of listlessness. On the face of it, the risk of weakness or listlessness is ever present, and this is what makes it so philosophically troubling. For such a risk seems completely to sever any conceptual connection between evaluation and motivation, between practical reason and action. The naïve view of evaluative judgment as itself essentially motivating must simply be wrong, for the connection between evaluation and motivation seems to be, even when it is in place, only contingent and fortuitous. Consequently, the motivational problem emerges as the problem
of understanding how, in the face of the possibility of phenomena like weakness of the will and listlessness, it is possible to have control over what we do.

In response to the motivational problem, a common strategy is simply to accept the division between our practical, evaluative judgments and motivation, evidenced in these cases of weakness of the will and listlessness, as a persistent feature of our psychological heritage: there is an *inevitable motivational gap* between practical judgment and desire, a gap that must somehow be bridged by something other than further deliberation and judgment if such deliberation is to have an effect on our actions.\(^6\) The motivational problem is thereby conceived as the problem of how it is possible to bridge this inevitable motivational gap in cases in which we are not listless or weak willed, and it can be solved, therefore, only by postulating some contingent mechanism outside deliberation that connects the two.

Indeed, such a strategy seems to be forced on us by the assumption of the cognitive–conative divide, for our practical judgments, it seems, cannot themselves be conations, or listlessness – our making practical judgments but being wholly unmoved by them – would be impossible.\(^7\) Yet if our practical judgments are cognitions, then, according to the cognitive–conative divide, they must be conceptually distinct from particular conations: cognitions and conations are mutually exclusive. Consequently, it might seem, one can always have the cognition without having any of the relevant conations motivating one to act accordingly, and one can always have the conation without making any corresponding practical judgment. To be committed to some content having one direction of fit in no way requires or presupposes that one is committed to a similar content with the opposite direction of fit. Cognitions and conations are, in this sense, *distinct existences*, and it is precisely for this reason that the motivational gap between them seems inevitable.\(^8\)

\(^6\) Michael Smith, for example, adopts this strategy; see his *The Moral Problem* (Oxford: Blackwell, 1994). I shall discuss Smith, and motivational externalism more generally, in more detail in §6.2.1.

\(^7\) Actually, this strategy is not forced on us: some motivational internalists who accept the cognitive–conative divide do not adopt it. I shall discuss motivational internalism in more detail in §6.2.2.

\(^8\) Notice that in understanding cognition and conation to be distinct existences, those who accept the cognitive–conative divide need not understand our capacities for cognition and conation to be conceptually independent. As Carl Hempel has argued, we must at least accept an epistemic interdependence: we cannot know what cognitions a creature...
Emotional reason

I shall now suggest that this account of the motivational problem, in light of the assumption of the cognitive—conative divide, makes its solution impossible. Part of the problem lies in understanding the relevant kind of control in question. Consider, by way of comparison, the way in which I can exercise control over a remote-controlled car. When I control the car, I do so by exploiting the causal connections between the positions of the various switches and dials on the control box and what happens in the car. Thus, to make the car go straight, I must continually make readjustments to the dials on the control box so as to overcome extraneous causal influences on the car: a bumpy road, a strong cross-wind, etc. Such an exploitation of causal connections is the nature of my control over the car, and in the absence of these causal connections I lose control entirely.

By contrast, the exploitation of such causal connections cannot be the whole story about how we can exercise control over our motivation. For, in controlling my motivation, I exploit not merely contingent causal connections between my practical judgments and my desires but also rational connections between the contents of my judgments and the objects of my desires: I exploit my appreciation of the reasons my judgments provide for having and acting on these desires. Thus, faced with the possibility of a second piece of chocolate cake, I judge that I should not have one now because, although eating it would be pleasurable, my concern to eat well outweighs the momentary pleasure I would get here and now from eating it. It is my appreciation of this reason that can motivate and ought to motivate me to resist the cake or even no longer to find it tempting. Hence, the relevant kind of control we can have over our motivations is rational control — control by virtue of an appreciation of reasons — and should not be assimilated to the kind of causal control we can have over ordinary physical objects.9

The trouble is, such distinctively rational control is incompatible with the assumption of the cognitive—conative divide. For this assumption, by

has without also knowing what conations it has, and vice versa. (See his "Aspects of Scientific Explanation," in Aspects of Scientific Explanation and Other Essays in the Philosophy of Science [New York: The Free Press, 1965], pp. 331–496, especially §10.3.2.) Moreover, it is consistent with the cognitive—conative divide to take this interdependence of the capacities to be ontological, a condition of the possibility of having the capacities at all. That is why I put the claim of conceptual distinctness at issue in the cognitive—conative divide in terms of any two particular instances of cognition and conation.

forcing us to conceive of cognition and conation as distinct existences, separated by an inevitable motivational gap, requires that we somehow bridge the gap. But what bridges the gap cannot be either a cognition or a conation, for these merely stand on opposite sides of the gap. This means that what bridges the gap cannot be intentional, for intentional states must be, given the assumption, either cognitive or conative. Yet, if it is not intentional, then it is mysterious how it could be properly responsive to the reasons we have for acting, and any coincidence between our reasons and our motivation would then only occur by chance or good fortune. At best, therefore, what bridges the gap seems to present merely the illusion of control and not genuinely rational control. That is not so much a solution to the motivational problem as capitulation in the face of it.

As I shall suggest in §1.5, the real solution is to reject the assumption of the cognitive–conative divide and so to reconceive the relation between practical judgment and desire: what is needed is a conception of desires as themselves susceptible to the reasons for action we come to accept in judgment and so as more than mere conations. Of course, what I have offered here is far from a complete argument against this assumption, and I shall, in chapter 6, discuss a variety of specific positions, both internalist and externalist, on how control over our motivations is possible given the assumption of the cognitive–conative divide. Nonetheless, we already have one reason motivating a reconsideration of this assumption. Another stems from the deliberative problem.

1.4 THE DELIBERATIVE PROBLEM

As I suggested in §1.1, part of what is definitive about persons is our ability to deliberate about and so choose our visions of the good, our values, which partly define the kind of persons we are; the possibility of such reasoning, therefore, enables us to have a say in, and so be responsible for, who we are. The questions at the heart of the deliberative problem are: how is such deliberation possible and what is the source of the norms governing both how we deliberate and the conclusions we reach?

In asking these questions about how deliberation about value is possible, I have in mind primarily personal values rather than moral values. That is because my concern is with how practical reasoning
Emotional reason

enters into our freedom to define ourselves and so our responsibility for who we are. For personal values, unlike moral values, are values that are relative to the individual in the sense that what personal values it is right for me to hold may well differ from those it is right for you to hold, and so personal values are what make a distinctive contribution to my identity as a person. Henceforth, when I speak of values, I shall intend personal values unless explicitly noted otherwise. (Nonetheless, I shall also be concerned with how we can reason about those cares and concerns that are not so "deep" as personal values insofar as they do not enter into one's conception of the kind of person it is worth being.)

In general, questions about what is worth pursuing or who it is worth being cannot be answered by appealing to instrumental reasoning. For instrumental reasoning is simply a matter of figuring out how to attain a given end; as a result, the value of the means we discover by instrumental reasoning derives from the antecedent value of that end, which is simply presupposed. However, because deliberation about value, especially at the deepest level, is concerned precisely with the value of these "ultimate" ends, instrumental reasoning alone is not sufficient as an account of such deliberation. What, then, can we add to instrumental reasoning to fill in this gap and so provide an account of how deliberation about value, with its requisite normative force, is possible?\(^{10}\)

What makes this question so difficult to answer is that our understanding of personal value is pulled in seemingly opposed directions of objectivity and subjectivity. On the one hand, a person's values are at least in part up to her, and are in this sense subjective: she can have a say in creating or inventing the kind of person it is worth her being. To be able to invent ourselves in this way is to have a kind of freedom that is distinctly human: a freedom not merely to control our actions but more

\(^{10}\) Recent interpreters of Aristotle, such as John Cooper (Reason and Human Good in Aristotle [Cambridge, MA: Harvard University Press, 1975]), Miles Burnyeat (“Aristotle on Learning to Be Good,” in Amélie Rorty, ed., Essays on Aristotle’s Ethics [Berkeley: University of California Press, 1980], pp. 69–92), David Wiggins (“Deliberation and Practical Reason,” Proceedings of the Aristotelian Society, 76 [1975–76], pp. 29–51), and John McDowell (“The Role of Eudaimonia in Aristotle’s Ethics,” The Proceedings of the African Classical Associations, 15 [1980]) have thought that he distinguishes instrumental reasoning from constitutive reasoning, which concerns how more completely to articulate vaguely specified ends such as that of a good vacation or a good life (eudaimonia). Merely to speak of constitutive reasoning, however, is so far merely to provide a label for a kind of reasoning rather than a clear account of how such reasoning can proceed or of its normative foundations. Thus, “How is constitutive reasoning about eudaimonia possible?” is just another way of formulating the problem.
Two problems of practical reason

fundamentally to govern ourselves; call this freedom "autonomy". Thus it is inconsistent with our autonomy that the source of the norms at issue in the kind of person it is worth our being be wholly external to us; rather, to be autonomous is for the source of those norms to be at least partially within one's understanding of who one is to be. In being autonomous, therefore, we can choose certain values, and by making these choices we determine our reasons.

On the other hand, there seems also to be an element of objectivity in what values a person holds in that she can deliberate about them correctly or incorrectly. Deliberation is a matter of choosing for reasons, thereby making possible the articulation of why one course of life is better than another, so that it is not intellectually arbitrary which values we choose. Hence through deliberation we can discover the values things really have and so the kind of person it is worth our being, potentially overcoming delusions or misunderstandings about ourselves. The possibility of such discovery means that there are rational constraints on which values we can autonomously choose. The claim here is not merely that there are some values it is (perhaps morally) impermissible to choose, such as the value of bigotry; nor is it that the constraint is merely that of internal consistency. If those were the only constraints, we could view the choice of personal values as one of selecting arbitrarily from among the many internally consistent, morally permissible values. Yet our choice of personal values and so our understanding of the kind of person it is worth each of us being is not arbitrary in this way. For we think it is possible to have better or worse reasons for such choices, reasons we might discover only through much effort and soul-searching.

The demand for discovery and objectivity is a demand to rule out this kind of arbitrariness. This means that the relevant reasons must govern our choices not merely in the sense that we happen to feel compelled to choose in accordance with these reasons (because, perhaps, they merely seem to justify our choices), but that these reasons succeed in justifying those choices, for it is only this success that can ground the relevant kind of objectivity.

These pulls toward autonomous invention and rational discovery seem, on the face of it, to be equally essential to our understanding of ourselves as persons. On the one hand, to give up on the pull toward invention is to give up on our autonomy, our ability to have a say in the kind of person it is worth our being and thereby to take responsibility for who we are. Yet this ability both to call our lives into question and,
having done so, to determine what life it is worth leading is fundamental to distinguishing ourselves as persons from the animals. There are, surely, other differences between persons and animals, such as our linguistic ability and our self-consciousness, but autonomy is fundamental insofar as it is because we are autonomous that we can understand how these other differences contribute to a distinction in kind and not merely in degree. On the other hand, to give up on the pull towards discovery is to concede that our choices of values are arbitrary, lacking any rational justification. Yet an arbitrarily chosen life lacks the kind of grounding that is precisely the point of our deliberation about value: to attain through the discovery of such grounds a depth of meaning in our lives that would be wholly lacking in a life chosen arbitrarily. Consequently, to give up on either invention or discovery is to give up on a dimension of ourselves that diminishes our status as persons.

The problem is that such talk of rational discovery seems to leave no room for autonomous invention, and vice versa. How can we make sense of the possibility of getting our values (objectively) right or wrong when we are the ones (subjectively) determining the standards of correctness? This difficulty, which I shall call the apparent paradox of simultaneous autonomous invention and rational discovery, seems to undermine our best attempts at getting clearer on the kind of deliberation at issue here. This apparent paradox therefore seems to force us to give up on one or the other of these dimensions of our understanding of persons, both of which seem essential. The deliberative problem, then, emerges as the problem of understanding how deliberation about value is possible while both finding a way out of this apparent paradox and maintaining a reasonably robust conception of ourselves as persons.¹¹

I shall now suggest more than argue that the deliberative problem is unsolvable if we make the assumption of the cognitive–conative divide.¹² This assumption forces us to conceive our deliberation about values in one of two ways, depending on whether the task is understood as a cognitive or a conative matter. First, according to cognitivist theories of value, there are facts about the values things really have that are

¹¹ In setting up the deliberative problem in this way, I have been strongly influenced by Charles Taylor; see especially his Human Agency and Language: Philosophical Papers 1 (Cambridge University Press, 1985).

¹² Once again, my aim here is only to motivate a careful re-examination of the cognitive–conative divide; clearer arguments will be provided in chapter 7 in light of a detailed examination of the literature.
Two problems of practical reason

ontologically prior to what we think these values to be. Consequently, the task in deliberation is a cognitive task: one of getting our evaluative beliefs accurately to mirror the world in light of evidence that justifies those beliefs epistemically. Except insofar as they can serve as a source of evidence, emotion and desire are therefore simply irrelevant to the task of deliberation. Second, according to non-cognitivist theories of value, ontologically prior to conation there is no fact about what values things have; rather these values are intelligible as existing in the world only as projected there by our evaluative sensibilities, paradigmatically our emotions and desires. Given this ontological priority, the non-cognitivist claims, we cannot appeal to these projected values in order to explain or justify our sensibilities. The task of deliberation, therefore, is to get our evaluative sensibilities somehow into good shape, for example in light of non-epistemic demands of consistency.

Cognitivist theories may seem clearly able to make good sense of the idea that we discover value, for discovery is fundamentally a cognitive affair, a matter of getting our minds to conform to the world. This might be thought a strength of the view when it comes to moral value, the primary target of cognitivism, for we do not think moral values are simply up to each of us to choose individually. It is, however, problematic when we consider personal value, for it is precisely because of its construal of discovery that cognitivism cannot make sense of the idea that we invent value, that we can have a say in defining the kind of person it is worth being. If discovery were understood as the confrontation of an antecedent fact of the matter concerning what values we should adopt, there would simply be no room for our input; rather we would simply have to conform our evaluative attitudes to these facts on pain of irrationality. To be a cognitivist about personal values, therefore, is to give up on the conception of ourselves as autonomous, as able to have a say in the kind of life it is worth our living, and, as I argued above, this is to give up too much. Cognitivism therefore fails as a solution to the deliberative problem.


Emotional reason

Of course, a cognitivist might try to make room for the idea of invention in one of two ways. First, a cognitivist might attempt to accommodate this idea by noting that which values we discover depends on the particular route we take in deliberation. Since we can control the course of our deliberation, we can therefore control which values we discover thereby and in this way have a say in defining the kind of person it is worth our being. However, this tack leaves us with too thin a notion of discovery, of objectivity. For it conceives of deliberation as a matter of coming across only some values out of the total range of values there are, such that our input to deliberation is merely a matter of selecting some subset of these values as properly ours. Yet this selection is itself merely arbitrary: we can offer no reason for adopting the values thereby selected over other values other than that we happened across these first. The point of the requirement of objectivity, however, is precisely to rule out such arbitrariness so as to be able to attain a depth of meaning in our lives. The first tack therefore fails.

A second way in which a cognitivist might try to make room for invention is this. Insofar as we can act so as to change the world, we can change in particular those aspects of the world that have a bearing on what values there are so that, when we go to discover these values, we discover different values than we otherwise would have. In this way, it might be thought, we can have a say in what these values are and thereby accommodate the idea of invention. However, this accommodation either is an illusion or gives up too much by way of objectivity. For, in acting so as to change the values things have, we either act for a reason or not. If we act for a reason, such a reason — to change the values things have — must itself be a value and so (on the cognitivist conception) must be a potential object of discovery over which we have no control; thus, the idea that we can by so acting invent value must be only an illusion. If we act for no reason, then we reveal our choice of which values to invent to be merely arbitrary, and we have thereby lost the relevant notion of objectivity. This second tack therefore fails as well.

If cognitivism fails to solve the deliberative problem because of its emphasis on discovery at the expense of invention, it may seem that we are forced to non-cognitivism in order to make sense of that invention. After all, insofar as, according to non-cognitivism, values are projected by our evaluative sensibilities, we are the ones who determine what values things have for us. How, then, can non-cognitivism handle the idea that our evaluative sensibilities can go right or wrong, that there are
rational standards at issue here? We cannot understand our values to serve as the relevant standard, for that would be to understand our sensibilities as answerable to the world, as on the cognitivist account. Rather, non-cognitivism understands our sensibilities to be ontologically prior to the values things have, and that is what the metaphor of projection is meant to convey. Consequently, to make sense of such a standard, non-cognitivism can appeal only to the sensibilities we actually have and to standards of rationality internal to these sensibilities themselves: standards of coherence and, perhaps, the stability of these sensibilities in the face of pressures to change them. This would mean that deliberation can go right or wrong in the sense that it attains or fails to attain a pattern of sensibilities that is stable and coherent; when it goes right, a non-cognitivist can say, we have discovered the values things have for us.

This account of discovery, however, is too weak to ground the relevant kind of objectivity our values can have. For presumably there will be many stable and coherent patterns of sensibilities we could have, and so many systems of values that can be understood as correct for us to adopt. On what basis, then, is a person to choose from among these systems of values the one that defines the kind of person it is worth his being? To this question the non-cognitivist can have no answer: each is as fully "right" as any other, and you simply have to pick arbitrarily among them. Again, such arbitrariness is at odds with the robust notion of objectivity and discovery at issue in our conception of ourselves as persons. Put another way, the objection is this: in response to the question of why you value something, it is insufficient to answer that this value figures in a stable and internally coherent system of sensibilities. For such a formal requirement on our sensibilities does not address the question, which is after the content of the value: what is so valuable about it? The demand for objectivity is the demand that the value something has itself serve as a justification for our evaluative sensibilities, and this is precisely what the non-cognitivist rules out by the appeal to conation. Non-cognitivism can therefore offer only a thin and seemingly inadequate surrogate for discovery, and this pushes us back to cognitivism as an account of the element of discovery.

In short, if we accept the cognitive-conative divide, we must understand the discovery of value in cognitive terms as a matter of matching our thoughts to a mind-independent world, and we must understand the invention of value in conative terms as a matter of projecting our
Emotional reason

evaluative attitudes onto the world. Autonomous invention and rational discovery are therefore mutually exclusive and exhaustive: we must either accept the priority of value over our thoughts (and so be cognitivists) or accept the priority of our evaluative sensibilities over value (and so be non-cognitivists), but not both. Consequently, the deliberative problem is thereby conceived as a genuine paradox of simultaneous invention and discovery, thus forcing us to choose (as cognitivists and non-cognitivists do) discovery instead of invention or vice versa, thereby giving up part of what is fundamental to personhood.

It should be clear that these are not by any means knock-down arguments against cognitivism or non-cognitivism. More sophisticated versions of these doctrines purport to answer the objections I have just raised, and I shall consider these views in more detail in chapter 7. My aim here, however, is merely to outline the difficulties for each view and to suggest the following diagnosis: what forces the choice between a robust notion of invention and a robust notion of discovery is the assumption of the cognitive—conative divide on which these two doctrines depend; we can avoid having to make the choice, and so the difficulties that choice presents, by rejecting that divide in the first place.

This can be seen in light of the following objection. It may seem that I am illegitimately casting the deliberative problem in terms of a contrast between invention and discovery. Really, one might say, the issue ought to be cast in terms of autonomy and rational constraint, and here it might seem that there is no conflict: our freedom to determine the kind of person it is worth being is not at all undermined by the presence of reasons supporting one view of the matter over another. Indeed, following Descartes' account in the 4th Meditation, we might say that we are most free precisely when our decisions are made in the light of substantive reasons rather than arbitrarily. Consequently, it might seem, the deliberative problem is merely the result of my miscasting of the issues.

So far, I think this view of autonomy and reason is exactly right, though I do not think it renders the deliberative problem moot. The difficulty arises when we notice that the relevant kind of autonomy at issue in the deliberative problem is in effect freedom to determine what our reasons are. To have a say in the kind of person it is worth my being is not merely to defer to pre-existing reasons and shape my life accordingly; it is, rather, to shape the reasons and so the worthiness of this kind of life themselves. So the problem becomes one of how, given this
Two problems of practical reason

conception of autonomy, we can be under substantive rational constraint in determining what our most basic reasons are. Indeed, given the cognitive–conative divide, this can only be paradoxical in precisely the way the image of simultaneous invention and discovery suggests: either we alone determine what our reasons are by virtue of conation, in which case we are not subject to any substantive rational constraint while doing so, or, if we are subject to such constraint, then it is cognition that is so constrained and so its object, as ontologically prior to that cognition, is not something we have a say in determining. Consequently, one who accepts the cognitive–conative divide must cast the issue in terms of a contrast between invention and discovery: either these reasons must be potential objects of cognition and, because of its mind-to-world direction of fit, discovery, or they must be potential objects of conation and, because of its world-to-mind direction of fit, invention, but not both. Once again, the apparent paradox at the heart of the deliberative problem has become a genuine paradox given this acceptance of the cognitive–conative divide.

If we reject the cognitive–conative divide, however, we can thereby reject the implicit assumption that either the values things have or our evaluations must be ontologically prior to the other. Thus, we can say that the kind of person it is worth our being is subjective and so potentially an object of autonomous invention insofar as there is no fact independent of the evaluative perspectives we actually adopt as to what this kind of person is. This is a denial of the ontological priority of value over our sensibilities, though it does not amount (as it would if we assumed the cognitive–conative divide) to an affirmation of the reverse priority. Likewise, we can say that the kind of person it is worth our being is nonetheless objective and so potentially an object of discovery insofar as we cannot make out what that kind of person is without appealing to the values things actually have. Once again, this is the denial of the ontological priority of our sensibilities over the values things have, and it is a denial wholly consistent with the above conception of subjectivity and invention.15

This understanding of simultaneous invention and discovery of value

is, therefore, circular, but that does not mean that the circularity is vicious. (Indeed, what would make the circularity seem vicious is the assumption of the cognitive–conative divide – that there must be ontological priority one way or the other.) The trouble is that to make such an account convincing we need to provide a clear understanding of our mental states and their connection to value that is an alternative to those permitted by the cognitive–conative divide (i.e., cognitivism and non-cognitivism), for it is only in light of such an account that the charge of vicious circularity can be rebutted. To a large extent, that is my aim in this book.16

I. 5 REJECTING THE DIVIDE: A PREVIEW

I have suggested that the assumption of the cognitive–conative divide and of the corresponding conception of rationality as either epistemic or instrumental underlies our conceptions of the motivational and deliberative problems and so has resulted in an unsatisfactory understanding of ourselves and how we can reason practically. Thus, on the one hand, the split between cognition and conation results in an understanding of the motivational problem as that of how to bridge the inevitable gap between cognition and conation. As I argued, this understanding of the problem seems to force us to see practical reason as only contingently and fortuitously connected with motivation, thereby undermining any resulting account of how we can rationally control what we do. On the other hand, this split results in an understanding of the deliberative problem in terms of a genuine paradox of simultaneous invention and discovery, a paradox which forces the choice between cognitivism and non-cognitivism, between an emphasis on the cognitive and the idea of rational discovery and an emphasis on the conative and the idea of autonomous invention. No matter which of these options we choose, it seems, we are led to give up an important dimension of our understanding of ourselves: as autonomous and so responsible for the kind of persons we are, or as able non-arbitrarily to reason about who to be and so to discover what has meaning in our lives.

In spite of their entrenchment in philosophical understandings of the mind, these conceptions of cognition, conation, and rationality are not

16 Indeed, part of my complaint about McDowell's, Wiggins', and Taylor's theories is that they fail to present their views in light of an explicit account of what our mental states would look like once we reject the cognitive–conative divide.