

DAVID VELLEMAN

REPLIES TO DISCUSSION ON
THE POSSIBILITY OF PRACTICAL REASON

The critiques of Jonathan Dancy and Al Mele give me an opportunity to clear up some misunderstandings of my view. Nadeem Hussain's critique gives me an opportunity to revise the view in some crucial respects. I am grateful for both opportunities.

1. DANCY

Dancy asks whether the self-knowledge at which an agent aims, in my view, is knowledge of what he is currently doing or knowledge of what he is going to do. In the first instance, I think an agent aims at the former. But Dancy may find my view confusing because I also think that an agent's primary means of attaining the former is by attaining a short-term version of the latter. In my view, an agent wants it to be the case that he knows, at any given time, what he is doing at that time, but he ensures this outcome by not doing anything until he knows that he is going to do it. This knowledge is of course practical knowledge of the sort that causes what is known – causes it, in my view, by way of the agent's inclination to do what he thinks, so that his thinking it will constitute knowledge of what he is doing. But the fact remains that the agent attains contemporaneous knowledge of his actions by attaining anticipatory knowledge of them.

Dancy then objects that I frequently shift between saying, on the one hand, that an agent wants to know what he is doing and, on the other, that he wants to know why he is doing it, or to make sense of what he is doing. As Dancy later acknowledges, I raise and answer this very objection in the Introduction to the book (26–27). My view is that an agent wants to know a



description of his action that implicitly explains why he is performing it. The way I put it in my Introduction is that he wants to “grasp [his] bodily movements under concepts that set them in an explanatory context of motives and circumstances” (p. 27).

Dancy next turns to my view that a reason for doing something is a consideration in light of which, were the agent to do that thing, he would know what he was doing in this explanatory sense. Dancy’s first question about the view is how it accounts for reasons against doing something. No problem. Reasons against doing something, on my view, are considerations in light of which doing that thing would be unintelligible or hard to understand. For example, Dancy’s recognition that he does not need new shoes makes it harder rather than easier to understand why he is shopping for shoes. The description “buying shoes I don’t need” therefore incorporates a reason against buying shoes rather than a reason for buying them. If Dancy becomes puzzled as to why he is buying yet another pair of shoes, he might well express his puzzlement by asking himself “What am I doing?” “Buying shoes I don’t need” would not answer this question except in a perversely literal sense, since it offers no clue as to what would explain his doing such a thing. The question clearly aims at obtaining a description that relates the action to his motives and circumstances in a way that explains it.

Dancy next asks why the screams of a torturer’s victim do not count, on my view, as reasons for the torturer to continue, given that, upon considering his actions in light of them, he knows what he is doing – namely, causing intolerable pain. Nothing I say remotely suggests that considerations acquire the force of reasons by alerting the agent to what he is doing in the sense of what effects he is producing.

When Dancy finally considers the possibility that I am talking about knowledge that incorporates explanatory material, he goes on to argue that an agent can explain his action only in terms of reasons in favor of performing it, and he implies that I have somehow substituted this clear and plausible conception of self-knowledge for my own, unclear and implausible conception. As in the previous instance, however, I

explicitly raise and answer this objection in my book (p. 26). Before repeating the answer here, I want to forestall misunderstandings that are likely to be caused by Dancy's distinction between "normative" and "motivating" reasons.

I prefer to distinguish between justifying reasons, or reasons-for, and explanatory reasons, or reasons-why. Having Parkinson's disease can be the reason why someone trembles without in any way being a reason for him to tremble – which is to say, it can explain without justifying. What complicates this distinction is that justifying reasons must be capable of entering into explanatory reasons, since a reason for acting must have the potential to influence a rational agent, in which case it may become a part of the reason why he acts. And the psychological process by which a justifying reason influences a rational agent can also be triggered by considerations that do not in fact justify, in which case the explanatory reason for the agent's action can include a consideration to which he responded as if it were a justifying reason even though it was not. Whether or not the consideration influencing an agent in this way has the force of a justification, it is said to be the agent's reason for acting, or the reason for which he acted – expressions whose definition requires the concepts of both justifying and explanatory reasons. A reason for which someone acts is a consideration that explains his action because he responded to it as if it justified (as it may or may not have done).

Here is why I prefer to avoid the expression "motivating reason". In Davidson's philosophy of action, which still dominates the intuitions of analytic philosophers, the psychological process of being motivated by a desire and belief is said to be one and the same as the process of being influenced by a justifying reason. As I explain in the Introduction to my book and in the paper entitled "The Guise of the Good," I believe that these are two distinct processes, mistakenly conflated by Davidson. An agent can act on or out of motives without acting for any reason at all, because being motivated does not in itself entail responding to anything as a justification. I therefore prefer to reserve the term "motivating" for the one process and "reason" for the other.

Here is what I say in the Introduction to my book:

[T]here is a definition of ‘making sense’ under which it is a term of practical rationality. What makes sense for someone to do, by this definition, is whatever he has reason for doing. The statement that reasons for an action are the considerations in light of which the action would make sense can therefore sound like a tautology. But I do not mean to speak tautologically.

When I speak of “making sense,” I am borrowing the phrase from the domain of theoretical reason, where it is used to characterize phenomena as susceptible to explanation and understanding. What makes sense to someone, theoretically speaking, is what he can explain. This is what I mean when I say that reasons for doing something are considerations in light of which it would make sense. I mean that they are considerations that would provide the subject with an explanatory grasp of the behavior for which they are reasons. (26)

In other words, the “explanatory context of motives and circumstances” that is provided by a justifying reason for acting, in my view, is one that explains the action in the theoretical sense, in terms of what actuates the agent under the circumstances.

As I have just said, of course, part of what actuates an agent can be a justifying reason to which he responds in a rational manner. But in my view, that justifying reason must be a consideration citing *something else* that actuates him, antecedently to its own rational influence, because it acquires that influence precisely by rendering the action intelligible to the agent – which it cannot do solely on the grounds of this very influence. The “something else” actuating the agent, and cited in the justifying reason, is often a motive, influencing the agent through a psychological process that is pre-rational in my view. The agent then understands his behavior as shaped by two influences – his motive and this very understanding of his behavior in light of it. As I explain in the title paper of the volume, I regard action as a process of rationally regulated activity, in which the activity that is regulated by reason must be actuated by antecedent forces.

Perhaps Dancy is so confused about my conception of reasons for acting because he has ignored all but the Introduction to my book. He says:

I am conscious, in all this, of having concentrated on three pages in the introduction to a 280-page collection. My reason for concentrating on those

pages is that they contain the most explicit treatment of what it is to be a reason; the topic is also treated on pp. 197–199, but there the discussion is interwoven with other considerations and less easy to pin down.

What's odd about this admission is that the passage to which it refers, on pp. 197–198, belongs to the end of the title essay, which I disavow in the Introduction as “tortured” and “unworkable”. Stranger still is that Dancy appears to have thought of himself as criticizing a view that is included in the introductory narrative and then mentioned only incidentally, “interwoven with other considerations,” on a total of two other pages. Should not this characterization have led to the suspicion that the view he criticizes is not what the book is about?

As it happens, the book contains slightly more material on this view than Dancy appears to have recognized. The paper entitled “The Story of Rational Action” argues that the axioms of formal decision theory qualify as norms of rationality because preferences that satisfy them are guaranteed to be characterizable in integrative terms, giving the agent self-knowledge of the kind that is typically embodied in reasons, according to my view. Attention to pp. 156–169 of this paper might have saved Dancy from some of the misunderstandings on which his critique is based.

2. MELE

Al Mele appears to think that my focus on “full-blooded” action, or action “par excellence,” manifests narrow-mindedness or perhaps even elitism about kinds of behavior. Yet this focus is fairly common in the philosophy of action, and for good philosophical reasons.

The extension of a concept like “action” is not determined by necessary and sufficient conditions that philosophy can articulate. The extension of such a concept is determined by stereotypes, or paradigm cases, which specific instances approximate more or less. How closely a bit of behavior must approximate these paradigms in order to qualify as an action, or how far it can depart from them while still qualifying, are context-dependent questions, to which there is no single

answer. What is close enough in one context will be too far away in another, and the relevant variables include normative or evaluative features of the context. The range of what we're willing to count as your doing may expand or contract in response to what is at stake.

Hence there are two different questions, one of which is clearly more fundamental than the other. The first question is how to characterize the stereotypes or paradigms approximation to which determines the extension of the concept "action". The second is how to characterize the dimensions along which instances can depart from the paradigm, and the contextual variables that determine how much of a departure is too much. Given the clear priority of the former question, there should be no objection to starting the philosophy of action with the paradigm cases.

Of course, Mele's charge may be, not that I start with the paradigm cases, but that I never get beyond them. To this charge I plead guilty. But I would cite, as mitigating circumstances, the philosophical problems raised by the paradigms of action – problems that further justify focusing on those cases. The human capacity to chew gum may raise interesting questions, I suppose, but they are nothing compared with the questions raised by the capacity to deliberate between alternative futures and to bring one of them about for good reason. How can there be alternative futures, and how can we bring them about, in a world where the future is caused by the past, or by nothing at all? How can anything have the normative force of a good reason for bringing about one future rather than another?

To be frank, I suspect that Mele is not interested in these questions. As I understand him, he is interested in characterizing a particular stretch of the human world, whatever it may turn out to contain and hence whether or not it contains choices or reasons or autonomous actions worthy of those descriptions. Or, to put it another way, he thinks that "choice", "reason", and "action" are not descriptions but names, denoting whatever there is in that stretch of the human world to which they are applied.

My belief is that these terms have descriptive content in virtue of which they may not apply to anything in our world – or in any other world, for that matter, if their content is, as it sometimes seems, fundamentally incoherent. In light of this belief, I choose to focus on the paradigm cases that raise this prospect most vividly. In doing so, I am not showing a lack of interest in the full range of human actions; I am just showing an interest in those paradigm cases where the fundamental philosophical problems lie.

The primary target of Mele's criticism is my view about the constitutive aim of action, and its similarity to the constitutive aim of belief. Early in his critique, Mele cites a single parenthetical remark of mine as grounds for attributing to me the claim that constitutive aims are identical with criteria of success. I claim no such identity. What I say is that constitutive aims provide criteria of success. Unfortunately, Mele interprets me as saying that criteria of success also provide constitutive aims, and this misinterpretation undermines much of what follows.

A constitutive aim is something pursuit of which is essential to the constitution of some process or activity. Kicking a ball around on a field does not amount to playing the game of soccer unless one is trying to kick the ball into the net more times than one's opponent. Of course, one can try to lose a game of soccer, but as I explain in my book, trying to lose a game is the pursuit of a second-order goal, the goal of failing in an attempt to win. If this second-order pursuit of failing to win ever collapses into a first-order pursuit of losing *simpliciter*, then one is no longer playing the game – as one's opponent might justifiably complain – and so there is no game for one to lose.

The constitutive goal of a game – the object of the game, as we call it – determines a criterion of success at playing it. But many processes and activities have criteria of success that are not associated with constitutive aims. It is a criterion of success in the activity of contributing to a book symposium that one address the views actually expressed in the book. But one need not aim at this goal in order to engage in the activity. Indeed, this particular activity has no constitutive aim that I can think of. Hence criteria of success do not necessarily generate constitutive aims.

I will shortly come to Mele's remarks on the constitutive aim of action, but first let me deal with his remarks on the constitutive aim of belief. These remarks are introduced by a mélange of decontextualized quotations that may mislead an innocent reader as to what I actually say. Mele quotes a statement on p. 17 of my book to the effect that belief aims at the truth in the sense that it is constitutively regulated by mechanisms designed to ensure that it is true. He then quotes a statement from a footnote to another chapter: "The concept of design encompasses the work of the pseudo-designer known as evolution."

He fails to note that the book has a chapter entitled "On the Aim of Belief," from which neither of these quotations is taken and in which I clearly state that the evolutionary design of our cognitive systems is only one of many ways in which belief might acquire its aim (pp. 252–253 and note 18 on p. 253).

Mele then quotes the following passage:

If it can be a natural or scientific fact that belief aims to be true, then it can also be a natural or scientific fact that false beliefs are wrong or incorrect, which is the fact underlying the normativity generally attributed to content. The hope of naturalizing that normativity is thus a reason for being interested in how belief aims at the truth.

Mele fails to say that this passage is drawn from an explanation of why one might be interested in the aim of belief, and that it carries a footnote beginning as follows: "Naturalizing the aim of belief is not on my agenda for this paper" (p. 245, n. 3; see also p. 253, n. 18). In other words, the reason which Mele here quotes me as adducing for being interested in the aim of belief is one that I explicitly disavow as *my* reason in the paper from which this quotation is drawn.

It is true that, in the Introduction to the book, I endorse a strategy for reducing the aim of belief to naturalistic facts. It is also true that, because I regard aiming at the truth as a conceptual requirement for an attitude to qualify as a belief, I should not begin sentences with the clause "if it can be a scientific fact that belief aims to be true...". I should say, instead, "if it can be a scientific fact that some of our cognitions aim at the truth, thus qualifying as beliefs...". The fact expressed in

the latter, more precise clause is as far as I would go in treating the aim of belief as a scientific fact: it is an aim such that a token attitude's having it can be a scientific fact about that attitude token. My chapter on belief makes clear, I think, that I would never go to the extent of saying that science can discover what aim, if any, an attitude must have in order to qualify as a belief. In my view, "belief" is not a rigid designator for a natural kind of attitude, a kind whose nature remains to be discovered. Our concept of belief has a fairly rich descriptive content, and what science can discover is whether, and how, that concept is satisfied by attitudes that we have.

Mele refers to his 2001 discussion of empirical research indicating that what we call beliefs are regulated by a mechanism designed, not to ensure that they are true, but rather to prevent "costly errors". He suggests that if this research turns out to be right, then my claim about the aim of belief must be wrong. The research he cites is James Friedrich's "primary error detection and minimization" (PEDMIN) model, according to which the ordinary person's testing of hypotheses is biased by the difference between the perceived cost of mistakenly accepting an hypothesis and the perceived cost of mistakenly rejecting it. Mele neglects to mention that my chapter on belief discusses the possibility that belief-forming mechanisms are biased by design (pp. 254–255; see also n. 22 on 254):

Evolution or education may have given us dispositions to err on the side of caution in perceiving predators, to overestimate our own popularity, and so on. But my thesis is not that belief is completely shielded from mechanisms that tend to make it false; my thesis is that belief is necessarily subject to mechanisms designed to make it true.

In most cases, the latter mechanisms retain some influence, despite interference from the former. ... Unlike an optical illusion or a phantasy, a biased belief usually responds to indications of the truth, however imperfectly.

Even if Mele overlooked this passage, he might have recalled his own response to the claim that Friedrich's hypothesis militates against the truth-directedness of belief:

It might be claimed that motivation to discover the truth about *p* is PEDMIN motivation on the grounds that wanting to discover the truth about *p*

is a matter of wanting to avoid certain errors – the error of not believing that p if p is true and the error of not believing that $\sim p$ if $\sim p$ is true. ... This claim would undermine the contention that PEDMIN motivation, *as opposed to* motivation to discover the truth, is at work at a given time. The claim implies that when motivation to discover the truth is operative, it is an *instance* of PEMDIN motivation. [Mele 2001, p. 40]

I would make only one correction to this passage. The errors at issue do not consist in not believing things that are true. Failing to believe a truth is never a cognitive error (although it may be a pragmatic error, if that particular truth is of some practical importance). The errors at issue consist in believing things that are false. With that correction, the foregoing passage of Mele's explains why Friedrich's hypothesis is no threat to mine.

When Mele comes to my views on the constitutive aim of action, his discussion suffers increasingly from his mistaken but frequently reiterated belief that I identify constitutive aims with criteria of success. He offers various alternative aims that might be constitutive of action, or of decisions to act, each aim being derived from a criterion of success that might reasonably be applied to actions or decisions. None of these derivations would be supported by premises attributable to me. As one might expect, Mele finds it impossible to transform some criteria of success into plausible constitutive aims without considerable adjustment, which he accomplishes by adding clauses to exclude "Freudian" or "deviant" processes, without any explanation as to what justifies these additional clauses. Hence his examples do not even carry out the strategy of derivation that he mistakenly attributes to me. Because these examples are twice removed from any relevance to my view, I will comment only on one or two, which happen to provide some occasion for clarification.

Mele suggests at one point that we might identify the criterion of success for action as the execution of a decision, whereupon (he notes) we would have to ask ourselves what the criterion of success for a decision might be. And he considers the following possibilities:

Some theorists may claim that the criterion is "the good": a decision* is successful just in case the course of action decided on is good. Others may

assert that the criterion is “the rational”: a decision* is successful just in case the course of action decided on is rational.

Mele neglects to mention that my book contains arguments against both of these proposals – one in the paper entitled “The Guise of the Good,” the other in the title paper, “The Possibility of Practical Reason”. These arguments are prominently featured in these papers and are deserving of at least mention even if not of a response.

Mele then considers another possibility: “Suppose a philosopher, Velma, were to claim that the constitutive aim of decision* is making up one’s mind what to do on the basis of relevant states of mind from which one is not alienated and in a way that does not involve any proscribed Freudian processes.” How could decision have a constitutive aim of “making up one’s mind what to do on the basis of relevant states of mind”? To begin with, anything that qualifies as “making up one’s mind what to do” already qualifies as a decision. Identifying a constitutive aim of decision (if there were such a thing – which I doubt) would be a way of analyzing what a decision essentially is. And the analysis of what a decision is can hardly be that it is something aiming to be a decision.

What is worse, Mele’s proposed constitutive aim refers to making up one’s mind “on the basis of relevant states of mind.” The main reason for analyzing decision in terms of a constitutive aim would be to explain how a person’s attitudes acquire the relevance of reasons in light of that aim. To identify an aim that presupposes their relevance would therefore be pointless, since it would leave us still in need of some basis for explaining how and why they are relevant.

None of the candidate aims proposed by Mele is shown to solve any problems in the philosophy of action. Each is proposed solely on the grounds that it is associated with some intuitively plausible criterion of success. But as I explain in both “The Possibility of Practical Reason” and “Deciding to Decide,” my purpose in trying to identify a constitutive aim of action is to find a non-normative foundation for our norms of practical reasoning. Hypotheses about the constitutive aim of

action earn credibility, in my view, not by resting on unsupported normative intuitions, but by explaining various otherwise inexplicable aspects of agency. Supported by this inference to the best explanation, they can provide independent support for practical norms, by generating criteria of success for action.

In this connection, consider Mele's argument that action might have a constitutive aim and yet be subject to an unrelated criterion of success, such as conduciveness to eudaimonia. Mele considers how I might respond, as follows:

It is in this connection that [Velleman] presses his analogy between action* and belief. A proper understanding of what belief is reveals, one may say, a double-duty "norm of correctness" – truth. Truth, it may be claimed, is the "constitutive aim" of belief and determines what counts as a reason for believing. Suppose that claim is true. Even so, no one has shown that action* is like belief in this regard. No one has shown that action* is such that its constitutive aim determines what counts as a (good) reason for acting ... Whether philosophers of action should try to develop convincing arguments for these theses about action* is an open question.

How does Mele imagine that one might "show" that action is analogous to belief in being subject to reasons generated by its constitutive aim?

Here is how I have tried to show it: by showing *what it would explain*.¹ In "What Happens When Someone Acts," I try to show that this hypothesis would explain how events can be brought about by agents instead of other events. In "The Possibility of Practical Reason," I try to show that the hypothesis would explain how reasons for acting can depend for their influence on an agent's motivational makeup without being constrained by his contingent motives for acting – in short, how the hypothesis would resolve the debate about internal and external reasons. In "Deciding How to Decide," I try to show how the hypothesis is preferable to pragmatic considerations as an explanation for the normative force of reasons for acting. In "The Story of Rational Action," I try to show that the aim of self-knowledge, in particular, would explain how the axioms of decision theory gain their force as norms of practical reason. In "Epistemic Freedom," I show that the associated, epistemic conception of intention would explain the openness of the future

from the deliberative perspective. In “How to Share an Intention,” I show that the same conception would explain the possibility of an intention’s being shared.

None of these arguments amounts to a deductive proof that action is analogous to belief. All that they are intended to establish is that the analogy is worth exploring, because of its explanatory power. Mele complains that the analogy has not been proved, but he does so without addressing a single one of my arguments.

3. HUSSAIN

I am grateful to Nadeem Hussain for reading my book and criticizing it so incisively. His criticism gives me a welcome opportunity to rethink my view and provide some clarifications and corrections.

To begin with, Hussain is quite right to question my characterization of the “story of rational guidance” in “The Guise of the Good”. I cannot defend the passages that he quotes, in which I say that the propositional contents of reasons for acting must be intrinsically justifying, independently of the attitudes in which they are fixed. As he notes, these passages are an attempt to explain how Davidson came to associate desires with value judgments. I think that my explanation can be reformulated to avoid the problems raised by Hussain; but as he also notes, the important question is not which story of rational guidance I attribute to Davidson but which story I accept on my own behalf; and these passages cannot be salvaged for that purpose. Instead of attempting to replace them now, however, I will move on to other questions.

One of Hussain’s questions about my view is whether it conceives of an agent’s awareness of reasons as an awareness of their being reasons or rather as an awareness of that about them in virtue of which they are reasons. In my view, what makes a consideration into a reason is its relevance to an integrative conception of the action for which it is a reason – a conception that would situate the action in an explanatory

context of motives and circumstances. A reason for doing something, as I put it, is a consideration in light of which doing that thing would make sense. Hussain seems to assume that an agent's merely grasping the explanatory relevance of a consideration to an action would not constitute an adequate grasp of the reason's justifying force. And if I understand him correctly, his grounds for this assumption is that the agent's perceiving a consideration merely as explanatorily relevant would leave an "open question" as to whether he ought to do that which it would help to explain. In order to close that question, Hussain believes, the agent must grasp the consideration under the concept of a reason, since the question whether he ought to do what he has reason for doing is not genuinely open.

In my view, however, whether to do what makes sense is not an open question, either – not, at least, insofar as the subject is thinking about what to do. If a person stops assimilating his own behavior into his understanding of the world, he may then be unmoved by the greater intelligibility of some available actions over others. But in that case, he will have stopped deciding what to do – will in fact have stopped exercising agential control over his behavior. Once he stops trying to bring his behavior under descriptions that would embody integrative knowledge of what he is doing, he has stopped considering that behavior in the context of reasons, or conforming that behavior to intentions, and so he has stopped functioning as an autonomous agent. So long as he continues to function as an agent, he will aim at making sense of his behavior, with the result that explanatorily relevant considerations will not be matters of indifference to him.

Thus, there is no open question whether to be swayed by explanatorily relevant considerations – no open question for an agent, that is. There may be, in some sense, an open question whether to be an agent, whether to get into or stay in the agency game. But of course someone who is not already in the game is in no position to entertain that question, because entertaining it entails thinking about what to do, which entails trying to bring his behavior under descriptions that would

embody knowledge of what he was doing. Anyone who asks himself whether to get into the agency game *is already in the game*; and anyone who asks himself whether to stay in the game cannot answer in the negative without staying in it at least that far. Of course, such a person can leave the game in a final exercise of agency – say, by taking drugs or jumping off a bridge or just dozing off for a while. But reasons for someone to act are not reasons for him to start or continue functioning as an agent; they are reasons for him in so long and so far as he functions as such.

I therefore favor the view that Hussain assumes I must reject – namely, the view that an agent is exposed to the force of a reason when he perceives its explanatory relevance to the action for which it is a reason. One advantage of this view, I believe, is that it provides a plausible explanation for weakness of will, as I explain briefly in a footnote to the Introduction of the book (p. 28, n. 34). An agent can be influenced by considerations in their capacity as reasons without necessarily conceiving of them as such. Insofar as he does think of considerations under that description, he may be mistaken about whether they are influencing him in the relevant way, or whether he is not being influenced even more by other considerations that he has not acknowledged under that description. An agent's enumeration of the reasons for him to act may therefore be mistaken as to the considerations that are in fact weighing with him in the manner of reasons. That is why he can act for reasons while failing to do what he judges himself to have most reason for doing: because the influence of reasons upon him is not mediated by his judgment about reasons so described.

This response to Hussain's first question leads directly into his second question, which is whether I have adequately accounted for the normativity of reasons. This question has also been raised by Kieran Setiya,² and the corresponding question about my account of belief has been raised by Nishiten Shah.³ I will want to make some concessions to Hussain, Setiya, and Shah on this question, and so I should spend a moment explaining what I take the question to be.

My response to Hussain's invocation of the open-question argument was that, insofar as someone is an agent, the question whether to do what makes sense would not actually be open for him, since his agency will just consist in his behavior's being regulated by the aim of making sense. The question of normativity is whether the agent's disposition to do what makes sense should be described as a disposition to do what is justified by reasons. This disposition makes explanatory considerations influential with him, but should their influence be valorized as rational – indeed, as the very mainspring of his rationality?

This question becomes especially acute in light of my insistence (e.g., on pp. 19–20) that the constitutive aims of belief and action need not be aims of the agent's, since they may be aims of some sub-agential or sub-personal systems instead. If the processes regulating the agent's cognitions and behavior really are sub-personal, then he should be able to assess them from a detached perspective, just as he can assess the way his nasal mucosae are disposed to swell upon contact with cat dander, or the way his palms are disposed to sweat when he is in unfamiliar company. The former disposition is due to regulatory mechanisms that are constitutive of immunity, the latter to mechanisms constitutive of the fight-or-flight response. The mere fact that these mechanisms are performing according to their design does not oblige the agent to approve of them, does it? And if the processes regulating his cognitions and behavior are likewise subject to his approval or disapproval upon reflection, how can they be constitutive of his theoretical and practical rationality?

Come to think of it, though, I have reason to reject this analogy. The basis of the analogy is the possibility of a reflective assessment such as Harry Frankfurt imagines in his hierarchical theory of agency: reflecting on his allergies or his social anxieties, the agent can find them deplorable, and so he should also be able to deplore the mechanisms regulating his cognitions for truth or his behavior for intelligibility. In "What Happens When Someone Acts?",⁴ however, I argue that the agent cannot attain a perspective of fully detached reflection on

the aim constitutive of action, because it is an aim that must be operative in order for him to reflect, in the first place. At the end of the Introduction, I suggest that the resulting inalienability of our intellectual drives is what Aristotle has in mind when he says that a person appears to be identical with his *nous*, his understanding. There is no perspective in which the drive to understand is not behind the lens of introspection, even if also in front of it, because no mental processes that were not regulated by this drive would be sufficiently coherent to constitute a perspective at all:

Your understanding is ... like that point between your eyes which constitutes the visual standpoint from which you see whatever you see, even when you view that point itself in the mirror, at a distance. Just as that point is always “here,” at the origin of your visual images, even when it’s also “over there,” in the mirror; so your understanding is always “me” in your reflective thinking, even when you regard it externally, as “it.” It’s your inescapable self ... [31]

Revisiting these passages under pressure of the present criticism, I now realize that I was too hasty in saying that the constitutive aim of action can belong to the agent’s sub-personal systems rather than the agent himself. I am committed to the view that the inclination toward this aim is the agent, functionally speaking, in that behavior guided by it is in fact attributable to him, as his doing. I would still insist that the constitutive aim of action is not the agent’s end, insofar as an end is something that the agent understands himself as pursuing and thus accords a potentially justifying role in his practical reasoning. But on reflection, I do not want to say that it is alienable from the agent.

Yet the agent’s inability to withdraw from his intellectual drives does not entail that he must approve of them, and it certainly does not entail that he must approve of them as that by appeal to which considerations qualify as reasons for acting. Even if the agent is inextricably identified with these drives, what gives them rational authority as opposed to brute motivational force?

At this point, I want to borrow a page out of Shah’s book and substitute it for several pages of mine.⁵ Shah has developed

a norm-expressivist analysis of belief-attribution. To classify a cognition as a belief, he claims, is, among other things to express acceptance of alethic norms in application to it. In the course of defending this analysis (at a length to which I cannot do justice here), he points out that insofar as the human mind is designed to regulate some cognitions for truth, the norms expressed by classifying them as beliefs will be syntonically with how the mind already works. Accepting alethic norms in application to these cognitions will then amount to endorsing a part of our own nature – though it may involve a commitment to making that part of our nature more reliable and consistent.

Now, this idea of accepting norms that ratify but also regularize some part of our nature is especially congenial to my conception of practical reason. When norms are accepted consciously, they not only guide our behavior but also provide generalizations with which to understand the behavior so guided. When we adopt the posture of being “for” some things and “against” others, we thereby adopt a comprehensive description for some region of our conduct, where we then gain an additional incentive to follow suit, so as to be comprehensible under that description. The process posited by Shah is therefore a natural step along the practical short-cut to self-knowledge, a way of making sense of ourselves by making sense to ourselves. Whereas I previously suggested that the norm of truth is emergent from, or can be reduced to, the way our cognitions are regulated when they qualify as beliefs, I am now inclined to say, along with Shah, that the norm of truth is a norm that we apply to cognitions that are largely regulated for truth already but become more intelligible to us by virtue of being consciously and more thoroughly regulated under that norm.

Here I am disavowing the project, announced in the Introduction to my book, of reducing the norms governing belief to naturalistic facts about belief. One of the advantages of publishing a collection of papers is that the format accommodates inconsistencies, and one of the advantages of inconsistency is the possibility of being right half of the time. In this case, I think that I was right to say, in the final chapter,

that a naturalistic reduction of doxastic norms was not on my agenda.

But what about the practical norm that is generated, according to my view, by the constitutive aim of action? Whereas the correctness of believing the truth, and the incorrectness of believing falsehoods, are almost universally accepted as norms, the same cannot be said for the correctness of doing what makes sense, and the incorrectness of doing what does not.

No sooner is that concession uttered than I am tempted to retract it: of course the correctness of doing what makes sense is universally acknowledged. Talk of what makes sense to do is pervasive in the contexts of deliberation and advice. As I explained in my response to Dancy, however, there is a reading of the phrase “what makes sense” that is of no help to my view, and I cannot rule out the possibility that what makes sense under this reading is all that’s generally acknowledged to be correct. For if what makes sense is understood as that which is favored by the balance of reasons for acting, then the correctness of doing it is obvious but also trivial, so that it cannot be the basis on which considerations qualify as reasons, to begin with.

Mine is not the vacuous claim that reasons are considerations in light of which actions are seen to be supported by reasons; it is rather the substantive claim that they are considerations in light of which actions are intelligible, can be explained and hence understood. And this cognitive reading of “what makes sense” strikes many philosophers as quite irrelevant to correctness in action.

Here my view runs into that intellectual brick wall which is consequentialism. Unfortunately, what Aristotle had to say about means-end reasoning in the practical syllogism is easier to understand than what he had to say about the relation between *eudaimonia* and *to kalon* (“the noble”, or “the fine”), the former of which is not a result of the latter but rather supervenes on its pursuit. Philosophy thus inherited a prejudice in favor of instrumental rationality, a prejudice subsequently reinforced by Hume’s distinction between reason and passion,

and then by the formalization of decision theory, which lent means-end reasoning the aura of scientific respectability.

That the result is nothing more than a prejudice becomes clear upon an honest survey of the reasons for acting that we offer and accept in real life. To be sure, a means-end syllogism can be cobbled together for most courses of action. Cobble as you will, however, you will never produce a syllogism that expresses the force of justifications like the ones that I cited in my *Precis*:

Why are you whistling?
Because I'm happy.

Why aren't you having any wine?
Because I don't drink.

Why worry about his problems?
Because I'm his friend.

Why are you shaking your head?
Because I think you're wrong.

Why do you have her picture on your wall?
Because I admire her.

Here already?
I'm punctual.

A theory of practical reason needs to explain what these justifications have in common with one another and with instrumental justification. And the explanation cannot be that all of them can be shoe-horned into instrumental form, since not all of them can, and the ones that can are sorely pinched in the process.

The preceding, admittedly polemical remarks, and the ones to follow, are intended to sketch the philosophical outlook from which my hypothesis about practical reason originates. If you think that the norms of practical reason are already well understood, then novel hypotheses on the subject will strike you as unnecessary and perhaps even unwelcome. But I think

that the norms of practical reason are not understood at all. In my view, practical reason is a subject on which philosophy is in a backward state of ignorance. The blithe assurance with which philosophers continually assert that reasons for acting are desire-belief pairs indicates, not that we have attained a stable understanding of the subject, but rather that we are still in the grip of myth and superstition. What else can explain the unquestioning acceptance of a theory that is so obviously inadequate to the phenomena?

The rational norm of making sense, in my understanding of the phrase, can subsume and explain the norm of instrumental justification while also accounting for the other forms of justification illustrated above. Given that the norm is implicit in those forms of justification, why not entertain the hypothesis that it is explicit in our practical discourse, whenever we talk about what it makes sense to do? Why assume that such talk uses a different sense of the phrase from the one that it has in other contexts?

Thus, I do not concede that the norm of making sense is irrelevant to practical reason as we know it. All I am willing to concede is that it is irrelevant to practical reason as caricatured in philosophy. I think that we adopt this norm in order to regiment and make sense of a process by which our actions are already regulated, just as we adopt the norm of truth in order to regiment and make sense of the regulation of our beliefs. What we are making sense of, in this case, is the very process of making sense of what we do by doing what makes sense. My hypothesis, then, is that the natural process of attaining practical knowledge affirms itself, by leading to the adoption of a norm that ratifies and regularizes it as the process of practical reasoning.

NOTES

¹ See p. 158, where I characterize one of my arguments as an argument to the best explanation.

² "Explaining Action," forthcoming in *The Philosophical Review*.

³ "How Truth Governs Belief," forthcoming in *The Philosophical Review*.

⁴ See also “Identification and Identity”, in Sarah Buss and Lee Overton (eds.), *Contours of Agency*, MIT Press, 2002, pp. 91–123.

⁵ At present Shah’s book is only metaphorically such. The ideas I am borrowing were developed in the course of his dissertation research. See also note 45 of Shah’s “How Truth Governs Belief”.

Department of Philosophy
University of Michigan
435 South State Street
Ann Arbor, MI 48109-1003
USA
E-mail: velleman@umich.edu