How is our grasp of reasons related to motivation? Are judgments about what justifies our behavior entirely independent from the motivational processes that actually cause that behavior? Internalists think they are not: they claim that there is a necessary connection between our grasp of reasons and motivation.

The central topic of this paper will be a new approach to the debates about internalism. Proponents of internalism have traditionally drawn a sharp contrast between theoretical inquiry and practical reasoning.\(^1\) Breaking with this tradition, Philip Pettit, Michael Smith, and Tyler Burge have recently suggested that the similarities between theoretical and practical reasoning can bolster the case for internalism.\(^2\)

Focusing on the similarities between theoretical and practical reasoning provides a fruitful perspective on the internalist debates. For one thing, the new approach places the focus squarely on the core issue in the internalist debates: what are the connections between our grasp of reasons and motivation?

\(^1\) See, in particular, the standard expressivist defenses of internalism: A. J. Ayer, Language, Truth, and Logic (London: Gollancz, 1946), chap. 6; Charles L. Stevenson, “The Emotive Meaning of Moral Terms,” Mind 46 (1937): 14-31; R. M. Hare, The Language of Morals (Oxford: Clarendon Press, 1952); and Simon Blackburn, Spreading the Word (Oxford: Oxford University Press, 1984), chap. 6. Expressivists typically emphasize that theoretical inquiry is constrained by evidence and aims at representing the world while practical reasoning is constrained by desires and issues in judgments whose necessary connection to motivation jeopardizes their status as beliefs representing the world. Allan Gibbard also proposes an expressivist defense of internalism (see his Wise Choices, Apt Feelings: A Theory of Normative Judgment [Cambridge: Harvard University Press, 1990]). Gibbard, however, is concerned to emphasize some of the major similarities between theoretical inquiry and practical reasoning. In his view, theoretical inquiry is infused with normative judgments and these judgments are amenable to the same theoretical treatment as those which are the upshot of practical deliberation.


François Schroeter is a Lecturer in the Philosophy Department at the University of Melbourne.
sons – both epistemic and practical – and motivation? All too often, disputes about internalism have centered on the particular nature of moral judgments and have evaded this more fundamental question. In addition, the new approach exploits an important insight about the nature of practical reasoning: there are clearly important similarities between our recognition and assessment of reasons in the theoretical and the practical domains. Given the prominence of neo-Humean views in practical philosophy, it is easy to forget how sophisticated our reasoning about what to do typically is and how closely it parallels theoretical reasoning. Much of this paper will be devoted to exploring the similarities between theoretical and practical reasoning (sections 3 and 4). This exploration will provide the necessary background for filling in the details of the new approach to internalism and for assessing its success (sections 5 and 6).

1. A new perspective on internalism
I’ll start with some clarifications about how internalism should be understood and turn then to the new approach to internalism.

Following Stephen Darwall, metaethicists usually distinguish between two main types of internalism. (1) Judgment internalists claim that necessarily, if a person makes normative judgments (or holds normative beliefs) about what is right for her to do, she is motivated to act in conformity with these judgments (or beliefs). (2) Existence internalists, on the other hand, claim that necessarily, if an action is the right thing for an agent to do, that agent is capable of being motivated to act accordingly. In this paper, I’ll be exclusively concerned with the first of these two types of internalism, which I simply call ‘internalism’.

We can distinguish between two different types of motivation which the internalist might have in mind when he claims that normative beliefs are necessarily tied to motivation:

(i) a motivation that is “grounded” or “based” in the normative belief;
(ii) a motivation that is independent from the normative belief.

Suppose, for instance, that I judge that the right thing to do is to eat chocolate. It’s not just that it tastes good; as it happens it is also good for my health. In this case, I might have two kinds of motivation to perform the action sanctioned as right by my judgment: a motivation based on my normative belief and a brute desire to eat the chocolate. In this paper, my focus will be on motivation of type (i), which I call “mo-

---

3 See, for instance, the criticisms of internalism in David O. Brink, Moral Realism and the Foundations of Ethics (Cambridge: Cambridge University Press, 1989), 37-80; and Sigrún Svavarsdóttir, “Moral Cognitivism and Motivation,” Philosophical Review 108 (1999): 161-219. Brink and Svavarsdóttir concentrate their attention on the amoralist: they ask whether the amoralist can count as making moral judgments even if he has no moral motivation. The question of the connection between our grasp of practical reasons and motivation is left to the side in these debates.


5 Internalists are standardly interpreted as making claims about conceptual necessity, where conceptual necessity is understood as metaphysical necessity knowable a priori on the basis of reflection available to anyone who possesses the relevant concepts. In this paper, I’ll follow this standard interpretation of internalism.

6 In order to remain neutral between cognitivists and expressivists, metaethicists typically use the term ‘judgment’ rather than ‘belief’ in their explanation of internalism (a normative judgment may be a belief, but it may also be, as expressivists maintain, a pro-attitude). The use of the term ‘judgment’ may suggest that internalists are concerned with the occurrence, conscious state of making a judgment and not with the standing state of believing. This suggestion is unfortunate: if there are general motivational constraints on the deployment of normative concepts in judgments, the deployment of these concepts in beliefs should be subject to similar constraints. I’ll talk about normative beliefs in my explication of internalism and I’ll leave it to expressivists to substitute for ‘belief’ a term referring to their favorite pro-attitude.
tivation to act in conformity (or accordance) with”, or “motivation to conform to”, one’s normative beliefs. This is the type of motivation that plays the central role in Pettit, Smith and Burge’s approach to internalism. It is also the type of motivation that matters in the story I’ll sketch about the normal functional role of normative beliefs.

Traditionally, proponents of internalism have claimed that motivation is constitutive of normative beliefs. Expressivists, for instance, contend that the mental state expressed in a normative judgment is constituted at least in part by having some motivation to act as the judgment prescribes. One can be an internalist, however, while rejecting this constitutive claim. There may be necessary connections between normative beliefs and the corresponding motivation even if the motivation is not literally part of the belief itself. Thus, the internalist can simply argue that a subject who is never motivated to act according to normative considerations cannot possess the concept ‘is the right thing to do’ and thus cannot hold genuine normative beliefs. Similar claims can be found in the debates about the possession conditions for experiential concepts such as ‘experiencing red’: some theorists claim that one cannot possess such an experiential concept without being able to recognize the relevant experience from a first-person perspective. In what follows, I will approach internalism via the question of the possession conditions for normative concepts.

This approach should be acceptable to all the participants in the debates, and it is flexible enough to accommodate stronger and weaker conceptions of the internal relation between normative concepts and motivation. I’ll focus here on a weak form of internalism that is entailed by all different versions of the doctrine and suffices to demarcate it from standard externalist claims. A strong version of internalism would claim that normative beliefs must always, or in general, be accompanied by motivation. Instead, I will consider here the minimal internalist claim that, necessarily, such ties exist at least in some cases. The externalist, on the other hand, insists that a subject who is never motivated to act in accordance with his putative normative judgments can possess normative concepts and thus can have beliefs with normative content. Similarly, the point of contention in the debates over experiential concepts is whether possessing such concepts is consistent with never having had the relevant experience. As the example of experiential concepts suggests, the necessary connection between a concept and a different type of mental state (such as an experience or a motivation) can be a weak one without vitiating the interest of that necessary connection.9

My weak version of the internalist thesis is preferable to an important alternative proposal inspired by Michael Smith.10 Like most internalists, Smith rejects the strong internalist claim according to which first-person present-tense normative judgments must always be linked to the corresponding motivation. In his view, internalists should be in-

---

7 On this point, see for instance Svavarsdóttir, “Moral Cognitivism and Motivation,” 163.

8 A congenitally blind person, for instance, would not have this recognitional ability and thus would not hold genuine color-experience beliefs, for the simple reason that such a person doesn’t have genuine color-experience concepts. On this point, see for instance Michael Tye, Color, Content and Consciousness (Cambridge: MIT Press, 2000), chap. 2.

9 Most internalists believe that although there are necessary connections between normative judgments and motivation, it is possible in some circumstances to make genuine first-person present-tense normative judgments without having the corresponding motivation. My weak definition of internalism obviously allows for this possibility. I will make no attempt here to provide a precise characterization of the class of cases where, according to the internalist, normative judgments are consistent with the lack of the corresponding motivation. Such a characterization is bound to be controversial and is not necessary for the purposes of this paper.

10 See Michael Smith, The Moral Problem (Oxford: Basil Blackwell, 1994), 61. It should be emphasized that Smith’s original proposal concerns what is morally right. I am here generalizing his approach to cover what is right all told – a move Smith hasn’t objected to, at least in conversation.
Terpreted as claiming that normative beliefs are necessarily tied to motivation in those cases where the subject is being practically rational. However, this alternative proposal doesn’t succeed in capturing the internalist’s distinctive claim.

An analogy may help us to see the main problem with Smith’s proposal. Consider the following “internalist” thesis about the concept of chocolate: necessarily if one makes the judgment that chocolate is immediately available, one is motivated to polish off all available chocolate in those cases where one is being a perfectly compulsive chocoholic. Few would reject this thesis. However, the problem is that it tells us something about the nature of chocoholics and not about what it takes to possess the concept of chocolate. One simply cannot be a perfect chocoholic at a particular time without being motivated to consume whatever chocolate is available. But, on the face of it, one can possess the concept of chocolate without being a perfect chocoholic. So our definition of chocolate internalism doesn’t establish any necessary connection between possessing the concept of chocolate and the motivation to eat chocolate: one can consistently accept the “internalist” thesis while denying that there are any motivational constraints on the possession of the concept of chocolate.

Similarly, Smith’s definition of internalism seems to tell us something about the nature of practical rationality and not about the possession conditions for normative concepts. It is commonly agreed that one cannot be practically rational unless one is motivated to act in accordance with one’s beliefs about what is all things considered right. This is just part of what it means to be practically rational – or at least so it seems. The problem is that paradigm externalists who think that one can possess normative concepts while never being motivated to act in accordance with one’s normative judgments will grant Smith’s thesis: whenever a subject is being practically rational, he’ll necessarily be motivated to act in accordance with his normative judgments. But paradigm externalists will insist – against their internalist opponents – that the lack of relevant motivation and the practical irrationality that ensues doesn’t prevent a subject from possessing normative concepts. This means that Smith’s definition of internalism should be rejected: according to Smith’s proposal, even paradigm externalists count as internalists. His definition therefore cannot capture what is at stake in the debates about internalism – namely, the question whether there are motivational constraints on the possession of normative concepts and on their deployment in judgments with genuine normative content.11

Now to the new approach to internalism suggested by Pettit, Smith, and Burge. Pettit and Smith emphasize the role of normative considerations in our intellectual lives. In their view, our conversational practices make sense only if we assume that interlocutors are generally disposed to respond appropriately to normative considerations. To be “authorized” as a genuine conversational interlocutor, they claim, a subject must not only have the capacity to recognize reasons; in addition he must be generally disposed to conform his beliefs to these reasons. This is true in the theoretical domain, when we debate over what to believe about the world. If someone systematically failed to adjust his beliefs after recognizing that there are conclusive reasons against his views, conversation would break down. As soon as we realized he was impervious to reasons, it would be pointless to continue our conversation with such a person – for his views would fail to meet minimal standards of coherence. We cannot learn from such a person and he obviously will learn nothing from us if he fails to adjust his beliefs in the light of the evidence we lay before him. Hence we would no

11 For a similar criticism of Smith’s rationality condition, see Svavarsdóttir, “Moral Cognitivism and Motivation,” 163-5.
François Schroeter

longer consider him a genuine participant in our conversational practices.

In Pettit and Smith’s view, our conversations about practical matters are subject to similar requirements. When we discuss what to do, we assume that our interlocutors are generally motivated to act in accordance with the reasons they recognize. If someone systematically failed to act in accordance with what he believes to be right, conversation would break down. We would conclude that he is not seriously involved in the business of practical evaluation and we would no longer consider him a genuine participant in our discussions about what to do.

Pettit and Smith insist that the continuity between the theoretical and the practical domains is not consistent with a rejection of internalism. Externalists see no necessary connection between a subject’s grasp of practical reasons and his motivation, and hence they believe that a subject who had no motivation to act in accordance with his normative judgments could nonetheless remain competent with practical normative concepts. According to externalists, such a subject could be a genuine participant in our discussions about what to do. Still, it is hard to see how a subject who doesn’t conform his beliefs to the theoretical reasons he recognizes could count as a genuine participant in our theoretical discussions. Pettit and Smith suggest that externalists are thus committed to seeing a “huge gulf” between the role of normative considerations in the sphere of belief and in the sphere of action. Those who are impressed with the similarities between the two spheres, on the other hand, should find internalism particularly congenial.12

In a similar vein, Burge writes:13

To understand reasons, one must understand their force and application in one’s reasoning. To understand their force and application one must have some tendency normally to make them effective in forming, changing, or confirming one’s attitudes or inferences.

An instance of this sort of point is commonly associated with a view about moral reasons – the view that reasons that are associated with an obligation or with a good must, at least in normal cases and given that the person understands the reasons, be associated with some sort of motivation. [...] The point is normally applied to what are commonly called practical reasons. I think that it is embedded in the broader, less restrictive notion of a reason, and applies no more to practical reasons and practical agency than to epistemic reasons and epistemic agency. The notions of agency and practice that I am explicating are broader, and I think, more fundamental, than the standard notions of action and practical reason.

What does it take to understand the concept of a reason? Burge suggests that one of the requirements for grasping this concept is that one be motivated to follow reasons. In order to grasp the concept of a reason, one must not only be able to recognize reasons as reasons and to differentiate good from bad reasons; in addition, one must actually be moved to make one’s attitudes and actions conform to the reasons one recognizes. So, in Burge’s view, it is not enough to accept the general thesis that a person’s recognition of reasons is normally associated, in that person, with “some motive or impulse” to adjust his thoughts and actions to these reasons. In order to possess the concept of a reason, the judge himself must have such a motive, impulse, or tendency.14

14 Some readers may balk at the talk of a ‘motive’ or ‘motivation’ to believe. Speaking perhaps somewhat loosely, I will take ‘motivation to believe or act’ and ‘tendency to believe or act’ to be rough synonyms (I am following Burge on this point). ‘Motivation’ should thus be understood in the deflated, purely dispositionalist sense typically associated with the term ‘desire’ in the philosophy of action. It should be clear that no affect is required for the presence of a motivation or desire thus understood.
François Schroeter

Normative Concepts and Motivation

Burge insists that this conclusion applies to epistemic as well as to practical reasons. He sees the internalist claims found in metaethical debates as a particular application of the more fundamental point he is making about our understanding of reasons in general.

Neither Burge nor Pettit and Smith offer a fully articulated account of the analogy between our grasp of theoretical and practical reasons or of the role of this analogy for a defense of internalism. They introduce the topic of internalism en passant: in neither case is it the main focus of the paper. However their suggestions, even if sketchy, allow us to discern the broad outlines of an important new approach to the question of internalism. The new approach invites us to place the traditional metaethical debates about internalism – that is, the debates about internalism in the practical domain, or what I shall call “practical internalism” – in a broader philosophical context. When we explore the connections between normative practical judgments and motivation it is important to remind ourselves of the connections that exist between our normative epistemic judgments and our corresponding beliefs. The new approach suggests, more specifically, that there is in the theoretical domain an analog to practical internalism. According to what we might call “theoretical internalism”, there is a necessary connection between normative epistemic beliefs and the tendency to form and sustain beliefs that are sanctioned as right. Theoretical internalists thus claim that a subject who isn’t motivated to conform his beliefs to the epistemic reasons he recognizes cannot possess normative concepts like ‘is the right thing to believe’ or ‘is supported by the best reasons’. According to the new approach, theoretical and practical internalism are supported by similar considerations: as Burge puts it, internalism is a feature of our grasp of the general concept of a reason, which subsumes both theoretical and practical reasons.

What is particularly exciting about the new approach is that it attempts to integrate the traditional metaethical debates on internalism within the broader context of a general account of reasoning and concept possession. Indeed, the new approach hopes to ground practical internalism in these more general considerations. As it stands, however, the approach is too sketchy to evaluate. My aim will be to put some flesh on the bare bones of this proposal and to provide the necessary background for assessing it.

2. Strategy

One thing is uncontroversial in the debates about internalism, namely, that normative beliefs are normally tied to motivation. Typically, when a subject holds a normative belief about what is the right thing for her to do or believe, this belief is tied to a motivation to perform the action or to form or sustain the belief sanctioned as right: it is part of the normal functional role of normative beliefs that they engage our motivational capacities. What internalists and their foes disagree about is this: is it a necessary condition for the possession of normative concepts that the normal motivational role of normative beliefs be at least partially preserved?

The dispute would be easily settled if the normal functional role of normative beliefs could be entirely explained in terms of their tie to motivation. Suppose that a simple version of expressivism were true and that normative judgments functioned as expressions of one’s actual motivations. Clearly, it would then make no sense to claim that a subject could be competent with normative concepts while sincerely and systematically deploying them in the absence of the corresponding motivation. Normative concepts, according to simple expressivism, just are expressive devices for manifesting one’s conative attitudes.
François Schroeter

However, simple expressivism is a very implausible view about normative concepts and normative beliefs. The normal functional role of normative concepts and normative beliefs includes much more beside their tie to motivation. In order to determine whether competence with normative concepts requires that the normal motivational role of normative beliefs be at least partially preserved, we’ll need to explore the other aspects of the normal functional role of normative concepts and normative beliefs.

In the next two sections I’ll offer a sketch of the normal functional role of normative concepts and normative beliefs in both the theoretical and the practical domains. Of the various roles that normative concepts can play in our mental economy the most important concerns the regulation of our own attitudes and actions. The normal functional role of normative concepts, I suggest, can be explained in terms of the role these concepts play in the exercise of normative self-governance over our beliefs and actions. On the basis of a few examples, I’ll sketch a picture of normative self-governance in thought and action and of the role of normative concepts in its exercise. This sketch will make clear why, in both the theoretical and the practical domain, it is not sufficient for competence with normative concepts that the beliefs in which these concepts are deployed play their normal motivational role. My story will thus show why simple expressivism about normative judgments and normative concepts should be rejected. A subject whose putative normative beliefs were appropriately tied to motivation but who didn’t recognize any further constraints on the applicability of normative concepts would not count as competent with these concepts. Such a subject would not count as knowing what it is for an action or a belief to be right.

My sketch of the normal functional role of normative concepts and normative beliefs will provide the background for the discussion of an important asymmetry between the theoretical and the practical domains in sections 5 and 6. In these sections, I will explore what would happen to a subject’s mental economy if the normal motivational role of normative concepts were not preserved. I will suggest that a systematic lack of motivation to follow normative epistemic beliefs will have different consequences for the subject’s conceptual competence than a systematic lack of motivation to follow normative practical beliefs.

3. Normative self-governance in thought
You are approaching the end of a murder-mystery novel and you really want to know who did it. There are different sus-

---

15 This is an important and substantial task, especially since there is no standard uncontroversial model of practical reasoning or of normative reasoning in general. My sketch will remain close, I believe, to our commonsense intuitions about what we are doing when we deploy normative concepts. My aim is to articulate an account of normative concepts that is faithful to the underlying motivation of the new approach: the story I’ll tell will emphasize the similarities between theoretical and practical reasoning.

16 In order to count as possessing a particular concept, one must know what it is that the concept refers to. This constraint on concept possession is widely acknowledged, although it is controversial what sort of psychological state constitutes knowing what it is that the concept refers to. Michael Dummett is a forceful advocate of this constraint, which he understands in a verificationist way (see for instance his The Interpretation of Frege’s Philosophy [London: Duckworth, 1981], chap 3). Theorists like Gareth Evans and Christopher Peacocke have argued for a non-verificationist interpretation of the constraint (see Gareth Evans, The Varieties of Reference, ed. John McDowell [Oxford: Clarendon Press, 1982], chap. 4; and Christopher Peacocke, A Study of Concepts [Cambridge: MIT Press, 1992], 22-4, 53-9). Even a semantic externalist like Burge recognizes the need for this sort of constraint on concept possession when he speaks of the understanding required for responsible ratiocination with a particular concept. For Burge, concept possession doesn’t require knowing necessary and sufficient conditions for the applicability of concepts (this would be “full understanding” of the concept), nor does it require grasping the folk theory standardly associated with the concept in one’s linguistic community (this would be “normal understanding”). However, possessing the concept does require that one understand what it is one’s concept refers to well enough to correct one’s mistakes about the nature of the reference (see his “Intellectual Norms and Foundations of Mind,” Journal of Philosophy 83 [1986]: 697-720, especially section IV).
François Schroeter

Normative Concepts and Motivation

pects but no obvious culprit. Things are complicated and the evidence is inconclusive. To make matters worse, the main suspects all incriminate each other. Who are you to believe in the end? You carefully examine different possible scenarios, looking for whatever clues can help determine the credibility of the suspects and the likelihood of their having evil motives. At the end of your deliberation, you try to come to a conclusion about the overall plausibility of the competing scenarios. You compare the different scenarios and assess how well each of them is supported by factual and psychological evidence. This comparison may of course prove inconclusive, so that no single scenario emerges as better supported by the evidence than the others. In that case, you would remain in the dark about who the murderer is. If, on the other hand, your inquiry is successful – so that one of the scenarios emerges clearly as being supported by the best reasons – you will form a belief about who the murderer is. Your conclusion that the best reasons support the verdict that this suspect is the murderer leads you to believe that this suspect is indeed the murderer.

Your belief about the murderer is the result of a sophisticated process of reasoning in which normative concepts play a central role. This process involves a comparative assessment of the hypotheses for their overall plausibility. The reasoner evaluates how well each of the scenarios coheres with the evidence: for each of the scenarios he assesses how good the reasons are for believing that it provides an accurate representation of the facts. This assessment allows the reasoner to come to an all-out conclusion about what belief is supported by the best reasons – that is, about what belief it is right or correct to form concerning the identity of the murderer. Having reached a normative conclusion about what is the right thing to believe, the reasoner then forms the belief he takes to be correct. It may of course happen that the reasoner can’t get himself to believe what he takes to be supported by the best reasons – for instance, that this suspect is indeed the murderer. In normal cases, however, our beliefs conform to the normative opinions we form about them.

The murder-mystery example illustrates nicely the normal functional role of normative concepts. Normative concepts play a central role in a reasoning process which can be analyzed as follows. (i) The reasoner recognizes prima facie reasons for beliefs. (ii) On the basis of an assessment of the relative strength of these reasons, he comes to a normative conclusion about which of the propositions he is considering is supported by the best reasons and is thus right to believe. (iii) The reasoner forms the very belief that his normative assessment sanctions as right. This reasoning process allows the subject to exercise normative self-governance over his beliefs. In this process, the subject brings his beliefs into line with his understanding and evaluation of the reasons supporting them: the formation of his beliefs is guided by his normative assessment of reasons.\(^{17}\)

It’s important to see that part (ii) is crucial to the process of normative self-governance: if the process is to go through, the subject must recognize a particular belief as the right thing to believe. If the recognition of reasons simply led to the conclusion that while there are good reasons supporting \(P\) there are also good reasons supporting not-\(P\), the subject wouldn’t be able to form an opinion (or at least not a reasonable one) about who the murderer is. Normative self-governance requires the subject to come to a normative all-out conclusion about which proposition is supported by the best reasons (or by good enough reasons to warrant belief)

\(^{17}\) As I will suggest later in this section, this normative all-out conclusion can remain implicit. So the process of normative self-governance should not be understood as necessarily involving a sequence of three explicit, temporally distinct steps.
Francois Schroeter

and is thus right to believe. The application of normative theoretical concepts therefore plays a pivotal role in normative self-governance.

The three-part reasoning process I have sketched describes the normal functional role played by the concept ‘is the right thing to believe’ in normative self-governance. The crucial question for our purposes is: are any aspects of this role necessary for competence with that concept? The forgoing sketch of normative self-governance suggests a partial answer. In order to think about what is right to believe, a subject must be capable of performing a whole range of cognitive tasks: the subject must understand what it is for a consideration to provide justificatory support for a belief, he must be able to recognize these justificatory relations when they occur, and he must be able to weigh and assess the justificatory support afforded by competing considerations. A creature who lacked these abilities might well classify beliefs in a certain way, but that classification could not be determinately interpreted as sensitive to the all-told rightness of those beliefs, as opposed to their prima facie rightness, their pragmatic desirability, or even their immediate attractiveness. Thus competence with an all-out normative concept seems to presuppose the abilities captured in (i) and (ii): a creature who lacked the capacity to recognize and assess justificatory relations would not count as competent with the concept of ‘is the right thing to believe’. This is why simple expressivism is so implausible.\(^{18}\)

The murder-mystery example may suggest that the exercise of normative self-governance is restricted to a particular subset of our beliefs – say, our most reflective beliefs. I believe that placing such restrictions on the scope of normative self-governance will lead to an inadequate picture of the role of normative concepts in our mental lives – and to an inadequate picture of the consequences of a radical breakdown in normative self-governance. I’ll close this section by sketching two reasons that incline me to think that normative self-governance has a pervasive hold on our belief system.\(^{19}\)

In the murder-mystery case, the subject goes through an extended deliberative process in which he weighs reasons and forms an opinion about which of the scenarios is supported by the best reasons. This is the paradigm case of an exercise of normative self-governance. But normative self-governance can take place even in the absence of any explicit, conscious deliberation. We often adjust our beliefs to the reasons we recognize in an automatic way, without consciously reflecting on what is the right thing to believe. Habituation allows us to develop an implicit sensitivity to justificatory considerations. An habitué of murder-mystery novels, for instance, might find it immediately obvious who the murderer is. However, the fact that no explicit deliberation takes place doesn’t mean that, in such cases, the formation of beliefs is not guided by the subject’s recognition and appreciation of reasons and by his tacit normative assessment of what candidate belief is supported by the best reasons. The subject is not just reacting in a brute reliabilist way to changes in his environment. He is forming his belief precisely because he appreciates reasons: it’s just that he has become so good at it that he no longer needs conscious reflective effort. Nonetheless, if challenged, he would have no difficulty citing the considerations that support his belief and have led to its formation.


\(^{19}\) Let me emphasize that this is a delicate issue, whose complexities haven’t yet been thoroughly explored. I won’t therefore aim to provide a full-fledged defense of the pervasiveness of normative self-governance, but only to indicate the broad motivations for a position of this kind.
In the murder-mystery case, the subject engages in a complex inference to the best explanation which involves a difficult weighing of reasons. But reasons don’t have to be epistemically challenging for normative self-governance to take place. Consider for instance the belief that this book is a murder-mystery novel. Given the way the plot unfolds, this conclusion is evident and it would make no sense to deny it. This doesn’t mean, however, that this belief has no justification. It is just that the reasons justifying it are utterly banal and obvious. It is trivial why the book qualifies as a murder-mystery novel: its whole plot is devoted to the investigation of a murder! Now, I take it that we usually have some understanding of what we are doing when we apply our concepts: we have some awareness of the considerations, however trivial, that justify the judgments in which these concepts are deployed. Even in simple tasks like the classification of books according to literary genres, we don’t normally apply our concepts in a blind way. If we thought there were no good reasons justifying a simple classificatory judgment, we would normally not make that judgment. Moreover, in making that judgment we typically assume that there are no good defeating considerations that would outweigh the reasons we take to support our trivial beliefs. And if new evidence came along and made us realize that a belief is not supported by the best reasons, we would normally revise this belief. Thus, there seems to be a dependence, both actual and counterfactual, of our trivial classificatory beliefs on our assessment of these beliefs as being supported by the best reasons. This point, if granted, suggests that even our trivial beliefs are normally kept under our normative self-governance: they are kept in line with our recognition and assessment of reasons, however trivial these reasons may be.

If we take seriously the fact that explicit deliberation is not required for normative self-governance and that the reasons we recognize may be trivial, we should be led to think that normative self-governance is a widespread phenomenon in our mental lives. I am myself inclined to believe that normal human adults tend to keep much of their belief systems in line with their recognition and appreciation of reasons.

4. Normative self-governance in action
Like our beliefs, our actions\(^2\) are also subject to normative self-governance. We are equipped with different, evolutionarily more or less advanced systems for the regulation of behavior. The most important and most sophisticated of these systems allows us to guide our behavior according to the reasons we recognize.

Consider first the more primitive, “hot” regulatory system. You encounter a mountain lion while hiking in the Sierras. You are terrified. The sight of the predator automatically triggers in you a defense mechanism aimed at preventing you from suffering serious harm. Both your cardiovascular and your hormonal systems are aroused and you are ready to engage in one of the patterns of behavior that evolution has selected as particularly advantageous: freezing, fleeing, or fighting if escape is not possible. All your attentional resources are also mobilized to attempt to prevent harm – as we all know, it is notoriously difficult to concentrate on other tasks when gripped by fear.

The fear syndrome is a good example of the functioning of the primitive, “hot” system for the regulation of behav-

\(^2\) In the ethical and metaethical debates, actions have traditionally been conceived as intentional bodily movements. I’ll follow this tradition here and consider actions to essentially involve peripheral bodily movement. According to this demarcation between the theoretical and practical domains, beliefs do not count as actions.
What is particularly distinctive about this regulatory system is that specific stimuli set the organism in a state of arousal and preparation for the performance of pre-determined patterns of behavior. This whole process is fast, automatic, and in general hard to resist: it typically ends up in the performance of those pre-determined patterns of behavior, even when we consider different options.

Obviously, human beings have other ways to cope with their environment. Even in our dealings with danger, our behavior is not regulated by the primitive, “hot” system alone. Imagine for instance the following scenario, which provides a good example of the normal functional role of normative concepts in the practical domain. You are about to enter the last leg of your journey when trekking in a tropical park. You are looking forward to this part of your trip since it will bring you to the area of the park that is the most scenic and has the most spectacular wildlife. Unfortunately, you learn that there is recent evidence that mosquitoes carry malaria in this area of the park. What is the right thing for you to do in these circumstances? You consider your main options and their consequences. The probability is high that you will catch the disease if you enter the contaminated area of the park. On the other hand, you should be safe if you avoid that area. To do so, however, you’ll have to make an important detour that will involve extra costs and delay your return home. You assess the different reasons weighing for and against your two main options. In the end, you come to the conclusion that, all things considered, it makes no sense to expose yourself to the risk of catching the disease. You decide to avoid the contaminated area and you proceed to do so.

In this scenario, your action ends up being selected so as to avoid danger and harm. But clearly your action is not the upshot of the primitive, “hot” system. The thought of infected mosquitoes doesn’t generate in you any immediate tendency to run away, and the fear syndrome never gets activated during the process that leads to your decision. You remain calm and dispassionate throughout your deliberation. What guides our actions in scenarios of this kind is a more sophisticated, “cold” system for the regulation of behavior. The subject recognizes prima facie reasons weighing for and against different options; he then assesses the force of these reasons and comes to a normative conclusion about what is the right thing to do in the present circumstances. Finally, the subject forms the intention to do what he judges right and proceeds to execute the selected action. Thus, the sophisticated, “cold” system for the regulation of behavior allows us to exercise normative self-governance over our actions. It allows us to bring our actions into line with our recognition of reasons.²²

Normative self-governance in action and in thought share the same main structural features. Like our normative beliefs in the theoretical domain, our normative beliefs about what is the right thing to do are formed on the basis of our recognition and evaluation of reasons. Our normative be-

---

²¹ See Joseph Ledoux, The Emotional Brain: The Mysterious Underpinnings of Emotional Life (New York: Simon & Schuster, 1996), chap. 5-6, for a detailed treatment of the fear syndrome. It is worth pointing out that feelings (as for instance the feeling of being afraid) don’t play a central role in the activation and unfolding of emotional syndromes. The main function of feelings seems to be to signal to consciousness the activation of an emotional syndrome. The primitive regulatory system is called “hot” because it involves a general state of arousal of the organism, not because the presence of specific feelings is essential to its functioning.

²² In this paper, I use ‘reasoning process’ as a generic term for the process which leads from the recognition of reasons to the performance of the action which the agent takes to be supported by the best reasons. It is controversial whether the last step in this process – the transition from the normative belief to the formation of the corresponding intention and its execution – should count as genuine reasoning. In calling the process of normative self-governance a process of reasoning I mean to be speaking loosely in a way that can accommodate all the different positions on this controversial issue. In other words, my use of the term ‘reasoning’ doesn’t commit me to any specific answer to the question whether practical reasoning issues in a belief, an intention, or an action.
lies about what to do are not blind expressions of the action tendencies generated by the primitive, “hot” system. They are grounded in a rich, substantive understanding of what matters in life in general and what is important in these particular circumstances. It is this understanding which allows our normative beliefs to provide real guidance for our lives.

Having a substantive understanding of what matters is essential to knowing what it is for an action to be right and thus is essential to being competent with normative practical concepts. This point needs special emphasis, since it tends to be obscured by some of the traditional models of practical reason, such as the neo-Humean model. Consider the malaria scenario. We understand what the point of the fear syndrome is and we know what is important about the features of the environment that trigger it. We understand what can bring harm and we also understand why it is important and how important it is to avoid harm. This understanding allows us to deal in a much more flexible and sophisticated way with what is dangerous: even if the thought of catching a disease doesn’t trigger the fear syndrome, we are able to evaluate our options for their potential for harm and to come to the normative conclusion that it is right to select a course of action that doesn’t put our health at risk. Clearly, a simple cognitive system which didn’t have a substantive understanding of the importance of preventing harm would not be competent to judge what is the right thing to do in such circumstances. It might react appropriately to circumstances, but it cannot literally recognize that this is the right option unless it has the capacity to recognize and weigh competing practical reasons.

Preventing harm is not the only consideration that weighs in the deliberation about what to do in the trekking expedition. In order to reach a verdict about what to do, the subject also has to balance other factors, such as the missed excitement of a new experience and the costs of delaying the return home. Here again, the subject’s understanding of the importance of these conflicting factors plays a crucial role in his coming to a normative conclusion about what to do. We not only know what’s bad about diseases and why health matters; we also understand how important is the enjoyment of memorable nature experiences. In addition, we certainly know what’s bad about paying extra money to return home. Only if we have this background understanding of the importance of the different factors at stake do we count as making a genuine assessment of the reasons supporting the two main options.

With a sketch of the normal functional role of normative concepts in the theoretical and practical domains in hand, I now turn to the question of internalism in these two domains.

5. Lacking the motivation to follow epistemic reasons
Is the possession of normative concepts subject to motivational constraints? Our discussion of normative self-governance will allow us to see that the denial of any motivational constraint on the possession of normative epistemic concepts is faced with difficulties which find no analog in the case of normative practical concepts. Let me emphasize that this is difficult and unexplored terrain: to my knowledge, no one has yet articulated an explicit rationale for embracing theoretical internalism. My aim in this section will not be to provide a conclusive argument for theoretical internalism, but only to illustrate the type of considerations that can support such a view.

Imagine a subject who satisfies the requirements for competence with normative epistemic concepts. The subject is familiar with a whole range of platitudes defining these concepts: he acknowledges that what is right is what is sup-
supported by the best reasons, that reasons justify, and so on. In addition, the subject is generally competent at recognizing reasons supporting beliefs and at forming opinions about what is the right thing to believe. Asked to assess particular beliefs, the subject reliably identifies the considerations which are relevant to determining whether or not these beliefs are well supported by reasons. Although his performance deteriorates when the focus is on esoteric subject matters, he is typically capable of recognizing relevant reasons when the topic is of general interest and doesn’t require any special expertise. He does reasonably well, for instance, when quizzed on murder-mystery cases. Can this subject remain competent with normative concepts if the normal motivational role of his normative beliefs is not at least partly preserved? Is the subject’s competence with normative epistemic concepts consistent with a systematic lack of motivation to conform his own beliefs to his normative assessments? If the answer to these questions is no, we should accept some form of theoretical internalism. That is, a negative answer would mean that one must have some motivation to conform one’s beliefs to what one deems right in order to possess normative concepts and to hold normative beliefs.

A subject who loses the tendency to conform his beliefs to what he takes to be right is, of course, affected by a radical form of irrationality. When normative epistemic beliefs – whether explicit or implicit – stop playing their normal motivational role, normative self-governance breaks down: the subject’s beliefs simply cease to be regulated by his opinions about what is supported by the best reasons. This means that what the subject ends up believing is now dictated by those of his dispositions which are not based on his recognition and all-out assessment of reasons. If, for instance, our subject takes reasons to incriminate one suspect in the murder-mystery case, this verdict will not affect the beliefs he forms about the identity of the murderer. Instead, insofar as he makes up his mind about who the murderer is, his belief will be dictated by his immediate dislike for one of the suspects or by other factors which have nothing to do with his assessment of reasons. Often enough, what will happen in these circumstances is that our subject will form a belief that is in conflict with his recognition of reasons: although he takes reasons to support the claim that Odette is the murderer, our subject will form the belief that somebody else is the culprit. Indeed, it would be sheer luck if his non-reason-based dispositions led him to believe that Odette is the murderer. The question we need to address is whether the dissonance generated in a belief system when normative self-governance breaks down will affect the subject’s competence with normative concepts. Must we conclude that competence with normative concepts is lost when a subject’s normative beliefs stop regulating his belief system?

To answer this question, we will need to take a closer look at what happens to a subject’s belief system and conceptual competence when normative self-governance breaks down. My attention will be focused on abstract concepts – such as ‘is guilty’, ‘is a democracy’, or ‘is right’ – and on the beliefs in which these concepts are deployed. In addition, I will set aside those exceptional cases where, once normative beliefs stop playing their normal motivational role, extraordinary luck or divine intervention still ensures that the subject’s beliefs are kept in conformity with the normative opinions he has about them. My discussion will suggest that, if we set such exceptional cases aside, a subject’s competence with abstract concepts – including normative concepts – will be significantly impaired when the subject loses

\[23\] In other words, must the competent subject have abilities corresponding to all three parts of the normal functional role of normative theoretical concepts outlined in section 3?
the motivation to conform his beliefs to his normative opinions.

Consider a characteristic train of thought of a subject who has lost the motivation to conform his beliefs to his normative opinions. Our subject takes reasons to support the claim that Odette is the murderer, but his non-reason-based dispositions lead him to form the belief that Marion is the culprit. Given that Marion is the murderer and that the crime was particularly wicked, reasons support the claim that Marion is an evil person who deserves punishment. Still, our subject forms the belief that Marion is a good person and deserves no punishment: Adèle is in his view the person who deserves to be punished for the crime. And so on. Unless we assume that luck or some *ad hoc* mechanism keeps our subject’s beliefs in line with his normative opinions, we’ll have to conclude that incoherencies will spread in his cognitive system. Ultimately he will end up with a set of attitudes sufficiently dissonant that it cannot qualify as a single belief system any more. Cognitive systems can withstand a certain amount of epistemic akrasia – we are all familiar with encapsulated beliefs which conflict with our recognition of reasons and stubbornly resist revision. But once our subject is caught in pervasive incoherencies, his belief system will lose its integrity as a single belief system: it will tend to become a mere collection of beliefs.

In fact, one might wonder whether our subject’s incoherent attitudes can still count as beliefs. Looking at the fundamental contrast between the attitudes of assuming and believing can help highlight what the main worry is. It is certainly possible to make, for the sake of argument, *assumptions* which are obviously dissonant with one’s beliefs. Even if one believes that Odette is the murderer, it is unproblematic to assume, at some point in one’s deliberation, that another suspect is the culprit or that none of the suspects committed the murder. However, it is another thing to actually form the corresponding *beliefs* while being fully aware of their inconsistency. Both assuming and believing involve regarding a proposition as true, but there is a crucial difference between the two types of attitudes. As David Velleman has emphasized, believing involves aiming at regarding a proposition *as* true only if it actually *is* true. Belief has a specific aim, namely the aim of “getting right” the truth-values of the propositions regarded as true. Assuming, on the other hand, is not constrained by this particular aim: one can assume that a proposition is true without trying to “get its truth-value right”. The worry is then this: how can our subject still be in the business of believing – that is, of aiming at “getting truth-values right” – when he forms beliefs that are obviously incoherent? Shouldn’t we deny that our subject’s obviously incoherent attitudes qualify as genuine beliefs, demoting them to the status of mere assumptions or imaginings?

I won’t try to settle these questions here or to provide a precise characterization of the nature of the attitudes which our subject ends up with – for simplicity, I’ll keep calling his attitudes ‘beliefs’. Instead, I’ll focus on what is directly relevant for the question of internalism – namely, the fact that our subject’s conceptual repertoire will be seriously muddled.

Take for example abstract concepts like ‘is a culprit’, ‘is guilty’, or ‘is a democracy’. What falls into the extension of these concepts for our subject? If he finds good evidence that the sovereign power of a given country resides in its citizens, he will take reasons to support the claim that that country qualifies as a democracy. But this normative as-

---

sessment – be it an explicit judgment or an implicit belief – will have no effect on his beliefs about whether or not that country is a democracy. In these circumstances, it would be a lucky coincidence if he nonetheless came to believe that ‘is a democracy’ applies to the country in question. As we have seen, the same result applies to a concept like ‘is guilty’. If he finds evidence that Odette has committed the murder and he understands that a murder is a very serious offense, our subject will take reasons to support the claim that Odette is guilty. Still, his non-reason-based dispositions will typically lead him to form the belief that somebody else is the culprit. There will thus be a significant degree of randomness in our subject’s particular applications of concepts like ‘is guilty’ or ‘is a democracy’ when these applications are cut loose from his – implicit or explicit – assessments of reasons.

In fact, it is not just one’s capacity to apply abstract concepts correctly to particular cases that will be affected by a radical breakdown in normative self-governance. In addition, our subject will also be led to deny important general truths about the properties picked out by his concepts. Consider for instance the following questions. Should non-intelligent beings be considered guilty when they cause harm to people? Is it possible to be guilty when one hasn’t broken a duty or a requirement of some sort? Can a country count as a democracy if it has no government? Competent subjects are in a position to know which answers to these questions are best supported by reasons. They understand that only responsible agents can be guilty of crimes. Since non-intelligent beings are not responsible for their behavior, reasons will lead competent subjects to conclude that non-intelligent beings should not be considered guilty. They also understand that to be guilty is to have committed some offense or crime and that committing an offense or crime requires the violation of some duty or requirement. They’ll therefore take reasons to support the claim that guilt is impossible without the violation of some duty or requirement. Similarly, competent subjects understand that democracy is a form of government and that, for this reason, it is wrong to think that there can be a democracy where there is no government. However, since our subject’s normative opinions don’t play their normal motivational role, his beliefs won’t be affected by such – explicit or implicit – assessments of reasons. Here again, there will be a significant degree of randomness in the beliefs our subject forms or sustains concerning these and similar questions. Typically, our subject will acquire erroneous beliefs about the nature of guilt and democracy – for instance the belief that non-intelligent beings can be considered guilty and that a democracy doesn’t require the existence of a government.

These examples suggest that the radical breakdown in normative self-governance affecting our subject will seriously impair his conceptual competence, or at least his competence with abstract concepts. It is fair to assume that in order to count as competent with a particular concept a subject must have a sufficiently accurate capacity to apply a concept in judgments and that he must have a sufficiently

---

25 Recall that, in this section, my interest is focused exclusively on abstract concepts, which are the genus to which normative concepts belong. For my purposes, I don’t need to make any claim about concepts which, like ‘is red’ or ‘is loud’, are closely tied to brute recognitional capacities. I take it that it would be more controversial to claim that competence with these concepts will also be significantly impaired when normative beliefs stop playing their normal motivational role. *Prima facie* at least, the close tie of these concepts to brute recognitional capacities seems to make them less vulnerable to the consequences of radical breakdowns in normative self-governance. However, see Sellar’s “Empiricism and the Philosophy of Mind” (reprinted in his *Science, Perception, and Reality* [London: Routledge and Kegan Paul, 1963]) for the defense of a contrary view.
François Schroeter

accurate understanding of what it is the concept refers to.\textsuperscript{26} Once normative self-governance breaks down, however, our subject places in the extension of his abstract concepts objects and properties which fail to meet not only our, but also his best criteria for applying the concepts. And he forms beliefs about the nature of the properties picked out by these concepts that are not only false but also in conflict with his own recognition of reasons. The incoherencies affecting our subject’s conceptual practices will thus lead to a serious degradation of his capacity to fulfill the two central requirements for conceptual competence. The moral is that our subject’s motivational abnormalities undermine his capacity to think about abstract properties: given the incoherence of the beliefs and cognitive dispositions he associates with abstract terms, we simply cannot associate a determinate reference to his words and thoughts.

Let me emphasize that I have been assuming here that competence with abstract concepts cannot be grounded in primitive “sphex-like” dispositions to deploy concepts on the basis of brute recognitional or associative mechanisms. As many authors have argued, possessing a concept is a matter of recognizing particular deployments of the concept as correct or as supported by the best reasons. In other words, conceptual capacities are intimately bound up with the subject’s normative abilities to recognize and assess reasons.\textsuperscript{27} Of course, I cannot offer a full-fledged defense of this view here, though I believe it has a good deal of \textit{prima facie} plausibility, especially in the case of abstract concepts. More importantly, the idea of a normative infusion of our conceptual capacities seems central to the view of concept possession developed by proponents of theoretical internalism like Pettit and Burge.\textsuperscript{28} Thus it is natural to appeal to this view in order to underwrite their suggestion that there are motivational constraints on the possession of normative concepts in the theoretical domain.

I have focused on the question of our subject’s competence with abstract concepts; but what about our subject’s competence with normative concepts? The subject’s normative concepts will share the same fate as his other abstract concepts when normative self-governance breaks down. Consider what our subject takes to fall into the extension of concepts like ‘is right’, ‘is a good reason for’, or ‘is a good justification for’. Is \( P \) a good justification for \( Q \)? After evaluating the pros and cons, our subject comes to an all-out

\textsuperscript{26} On the second of these requirements, see above fn. 16. I take the first requirement to be a central feature of competence with normative concepts. In the case of theoretical concepts such as ‘phlogiston’ or ‘witch’, one might possess the concept and understand what it would take to be the referent of the concept while systematically misapplying it to features of one’s environment. However, it is implausible that one could possess the concept ‘is right’ without being able to recognize when it applies to particular cases. If the first and second requirements are not satisfied, the subject won’t have a sufficiently rich conception of what his representation is about to allow us to assign a determinate referent (even taking externalist theories of reference into account). It just won’t make sense to say he is thinking \textit{about} a particular object or property.

\textsuperscript{27} The idea that possessing a concept or grasping the meaning of a word involves a normative element is an important theme in the philosophy of Frege, Wittgenstein, and Sellars. For recent and influential developments of such an approach to concept possession, see Robert Brandom, \textit{Making it Explicit} (Cambridge: Harvard University Press, 1994), and Christopher Peacocke’s \textit{A Study of Concepts}. Brandom’s project is to reduce semantic facts to a subject’s recognition of a system of normative statuses of commitment and entitlement. Peacocke grounds the possession of a concept in the subject’s finding certain deployments of the concept “primitively compelling” – that is, obviously correct and requiring no further justification. Of course, if concept possession is indeed a matter of recognizing deployments of concepts as correct, it should be no surprise that normative self-governance is – as I have suggested in section 3 – a pervasive phenomenon in our mental lives.

\textsuperscript{28} Pettit insists that the ability to recognize and follow semantic rules, considered as norms for the correct deployment of sub-propositional elements of thought, is essential to being a thinking system (\textit{The Common Mind}, Oxford: Oxford University Press, 1993, chap. 2). Burge argues that concept possession cannot be reduced to some simple pattern of activity or functional role precisely because this reduction couldn’t capture the intellectual norms that are essential to concept possession (“Intellectual Norms and the Foundations of Mind,” see especially 713, 718-20).
François Schroeter

conclusion about what to believe in answer to this question. But since this and similar normative opinions will have no effect on his beliefs he will often enough misapply the concept ‘is a good justification for’ and other normative concepts. In addition, his assessments of reasons and his all-out conclusions about what to believe in answer to general questions about the nature of rightness and of other normative properties will also be motivationally inert. So our subject will be led to acquire false beliefs about the nature of rightness and of other normative properties. This suggests that his competence with normative concepts will be seriously impaired. The breakdown of normative self-governance will lead to a severe degradation of our subject’s capacity to correctly apply his normative concepts and of his understanding of what it is these concepts refer to. In other words, the conceptual capacities which underwrite the subject’s capacity to think about epistemic reasons and what is the right thing to believe will be seriously destabilized by his lack of motivation to conform his beliefs to his normative opinions.

Shall we then conclude that it will be impossible for our subject to remain competent with normative concepts and hence that internalism is true in the theoretical domain? A full-fledged defense of theoretical internalism would have to show that it is incoherent to suppose that a subject can possess a concept of what is right to believe while having no tendency to conform his beliefs to his applications of that concept. My discussion of the case of a subject whose normative beliefs stop playing their normal motivational role hasn’t established that result. In the first place, the argument sketched presupposes that our conceptual capacities are normatively infused, a view that could be challenged. Second, my discussion suggests that ceteris paribus – that is, once extraordinary luck and the intervention of a benevolent
demon are ruled out – the motivational inertia of normative beliefs will significantly impair the subject’s competence with abstract concepts in general and his competence with normative concepts in particular. This conclusion falls short of an unqualified vindication of internalism in two important respects: (i) the ceteris paribus clause means that internalism will have to be hedged if it is to be true in the theoretical domain, and (ii) it doesn’t follow from the fact that competence is severely impaired that it will be lost altogether. Throughout my discussion I have set aside cases where luck or a benevolent demon ensures that the subject’s beliefs are kept in line with his normative opinions when normative self-governance collapses. But these cases seem to be counterexamples to the internalist thesis: they may show that, strictly speaking, it is possible for a subject’s competence with normative concepts to remain unaltered once his normative beliefs stop playing their normal motivational role. Hence the necessity to hedge internalism to rule out such cases. More importantly, we must keep in mind that our normal standards for conceptual competence are undemanding: one doesn’t count as conceptually incompetent simply because one makes some errors in the application of concepts and forms some false beliefs about the properties they pick out. We tolerate a fair amount of confusion in the conceptual practices of competent concept wielders. Thus, in order to vindicate internalism, one needs to show that the conceptual impairment generated by a radical breakdown in normative self-governance will be severe enough to jeopardize conceptual competence. In other words, one needs to show that this conceptual impairment will be sufficiently serious that the subject will lose the capacity to think about what is right to believe altogether. I

François Schroeter

won’t attempt to provide a conclusive defense of this crucial premise of the internalist argument here, although I believe a good case can be made for it.  

I have explored the kind of considerations that would support internalism in the theoretical domain, but I haven’t provided a full-fledged argument for that thesis. For present purposes, it’s enough to see what form an argument for internalism in the theoretical domain would take, so that we can then determine whether the argument could be extended to establish internalism in the practical domain. Of course, the argument I’ve sketched may support only a hedged version of internalism. Nonetheless, this result would still be a significant one: since the counterexamples ruled out by the hedge are so outlandish, the argument would show that, for creatures who are not the beneficiaries of extraordinary luck or benevolent demons, there are indeed motivational constraints on the possession of normative epistemic concepts. What now remains to be seen is whether an analogous argument can be formulated to support practical internalism.

6. An asymmetry between the theoretical and the practical
Consider a subject who is competent with normative practi-

cal concepts. He acknowledges the central platitudes associated with these concepts: he understands that reasons justify, that the right thing to do is what it makes best sense to do, that failures to do the right thing can have serious consequences for one’s well-being, and so on. In addition, he is able to recognize reasons weighing for or against the main options available in the typical circumstances of human life. He is also able to assess the relative importance of these reasons and to come to reasonable verdicts, when particular choices need to be made, about which options are best supported by reasons. In hard cases, he sometimes remains undecided about what option is right and he occasionally fails to appreciate the real importance of some of the factors which are directly relevant to the question of what to do. Despite these limitations, he qualifies as a reliable judge of what is the right thing to do. Can this subject remain competent with normative concepts if his normative beliefs stop playing their normal motivational role in the regulation of his behavior – that is, if he loses the motivation to conform his behavior to his normative beliefs? If the answer to this question is no, we should accept some form of internalism in the practical domain.

A subject who loses this type of motivation is affected by a very radical form of irrationality: his judgments about what to do always leave him cold, and he has no disposition to adjust his behavior to conform to them. We can get an initial grip on what this aberrant condition consists in by comparing it with ordinary weakness of will. A strong-willed person is disposed to systematically act in conformity with his judgments about what is right to do. In contrast, in a weak-willed person this disposition is unreliable in cases where the subject has strong countervailing desires. The case we are imagining is like weakness of will carried to its logical extreme: since he has no motivation to act in confor-
strike us as odd – or indeed utterly crazy. However, the motivational inertia of our subject’s normative practical beliefs needn’t have the same consequences on his belief system and representational capacities as the motivational inertia of his normative epistemic beliefs.

As we have seen, the dissonance typically generated when normative self-governance breaks down in the theoretical domain will have a direct impact on a subject’s mastery of concepts. When the subject acquires erroneous and incoherent beliefs about which objects fall into the extension of ‘is guilty’ and about what it takes to be guilty, his competence with the concept ‘is guilty’ will be impaired. The reason for this is that beliefs play a constitutive role in the cognitive system which underlies a subject’s ability to represent abstract properties. When a significant segment of the subject’s belief system stops cohering with his recognition and assessment of reasons, his representational capacities will be degraded. If our subject’s belief system fails to meet a certain threshold of coherence, no determinate reference to an abstract property can be assigned to his putative concepts. In particular, we cannot assign the property picked out by our normative concepts as the reference of one of our subject’s representational states.

Notice that the difficulty in the theoretical case hinges on the output of normative self-governance – the system of beliefs the subject ends up with. Here the motivational impairment creates chaos in the belief system. As a result, the subject’s capacity to represent abstract properties in general, and normative properties in particular, is undermined. In the practical case, the output of normative self-governance is an intentional action. Would a motivational impairment in the practical domain lead to actions that would undermine the subject’s ability to refer to what is right to do?

Prima facie, this seems implausible. Actions are not as central as beliefs to our ability to represent abstract proper-

---

32 One may even doubt that our subject will still qualify as an agent. Although his behavior will be caused by his beliefs and urges, his recognition of reasons and his all-out judgments about what to do will have no impact on his behavior. If one believes that agency requires reflective control over one’s behavior, our subject won’t count as an agent in this strong sense. However, the crucial issue for our purposes is not whether our subject is an agent but whether he can think about what is right to do.

33 See Antonio R. Damasio, Descartes’ Error: Emotion, Reason, and the Human Brain (New York: G.P. Putnam, 1994) for an influential discussion of disorders involving a dissociation between cognition and motivation. Damasio’s patients suffer from a serious deficit in emotional reactivity and motivation to implement value judgments: they reason about practical issues as if they were uninvolved spectators. Although these patients can competently discuss ethical and practical issues (and perform well on Kohlberg’s tests), they typically fail to make decisions which conform to the general values they accept. I’ll come back to these cases later in this section.
ties. No one thinks, for instance, that becoming a paraplegic will necessarily involve an impairment of one’s representational capacities. In itself the fact that one can no longer perform motor tasks does not impair one’s ability to think about which actions are right. What this suggests is that actions don’t play the same role as beliefs in constituting one’s representational capacities. So we should be especially cautious about any quick generalization of the argument sketched in section 5 to the practical domain.

Still, even if the output of normative self-governance in the practical domain – action – is not crucial to our ability to represent what is right to do, the subject’s motivations to perform actions – his desires or intentions – might be essential to his ability to think about this abstract property. In that case, the parallel between the theoretical and practical cases would still go through: in both cases the subject’s motivational impairment would undermine his ability to think about the relevant normative properties. This is the issue I will be focusing on in the remainder of this section.

Let’s first consider whether the subject’s motivational impairment in the practical domain would undermine his ability to represent abstract properties in general. There is no reason to think this is the case. In particular, there is no conceptual reason why the practically impaired agent would suffer from the type of cognitive breakdown suffered by the theoretically impaired subject we considered in the previous section. When thinking about what it takes to be a democracy or which governments count as democracies his verdicts about what it makes sense to think will influence his beliefs in the normal way. Despite his motivational abnormalities, we can expect that our subject will maintain his prior set of beliefs concerning democracy and revise it reasonably in the light of new evidence. Whatever his motivation and his behavior turn out to be, it seems clear that the subject’s coherent theoretical beliefs about democracy and his capacity to identify particular instances will suffice to ground his ability to think about the relevant property. So, unlike the theoretically impaired subject whose belief system becomes radically incoherent, the practically impaired subject has no trouble thinking about abstract properties in general.

Can the practically impaired subject also succeed in thinking about the abstract property picked out by our concept of what is right to do? It is important to emphasize that the question I am asking here is framed at the level of reference. I have been assuming that there are determinate “thin” properties picked out by our normative concepts. The argument canvassed earlier for internalism in the theoretical domain suggests that a subject who lacks the motivation to conform his beliefs to his normative theoretical judgments cannot think about what is right to believe. Does a failure of motivation in the practical domain likewise undercut the subject’s ability to possess a concept that picks out the property of being the right thing to do?

Once again, there is no conceptual reason why the practically impaired subject couldn’t think about what is right to do. Our subject’s motivational defects will wreak havoc in his life, but they need not wreak havoc in his normative concepts: his defects will not necessarily undermine the system of beliefs in virtue of which his thoughts refer to what is right to do. After all, we can expect that our subject will continue to accept the standard platitudes linking reasons to

---

François Schroeter

Normative Concepts and Motivation

---

20

---

34 The assumption that normative concepts pick out properties has lost its controversial edge. Leading non-cognitivists accept that normative concepts have determinate extensions, and this is all I am assuming here. A property in the thin sense I am assuming here may involve nothing more ontologically robust than an extension in all possible worlds. Concepts, of course, are individuated more finely than by the properties they pick out. Thus there may be more than one concept that picks out the very same thin property of being the right thing to do. This feature of concepts has played a central role in the traditional debates about internalism. I’ll come back to it at the end of this section.
justification and well-being, and that he won’t revise his prior verdicts about which types of actions are right. Moreover, as we saw in section 4, normal subjects have some reflective understanding of typical right-making features, of why these features are important, and of their relative importance in grounding verdicts about standard cases. Here again, there is no reason this general theoretical understanding must be undermined by the subject’s motivational impairment. Finally, there is no incoherence in imagining that the subject will continue to assume that he is co-referring with others in his linguistic community when he uses the words ‘the right thing to do’ and to trust the judgments of those he considers particularly wise or astute. Thus, the core intellectual capacities which ground the subject’s ability to think about what is right to do needn’t be affected by his motivational abnormalities. It is natural to conclude that the motivationally impaired subject can think about what is right to do.

I should emphasize that this conclusion is not based on the assumption that our subject’s motivational defects will have no impact whatsoever on his cognitive system. Some neuroscientific evidence seems to suggest that radical motivational or emotional pathologies are associated with degraded performance in reasoning about specific practical situations. Although patients suffering from such pathologies perform normally when assessing stereotypical situations presented in verbal form, they often fail to come to the right verdicts in complex real-life circumstances. These cases suggest that, in a normal human subject, the kind of motivational impairment we have been considering might affect the subject’s ability to classify real-life actions as being right to do. Even so, this psychological evidence provides no reason to suppose that radical motivational impairments will prevent subjects from thinking about what is right to do. Although these patients are not as proficient as normal subjects in classifying real-life cases, it is uncontroversial that they manage to refer to what is right. After all, they endorse the standard platitudes, they retain their beliefs about typical cases, they demonstrate understanding of what makes particular cases right, and they continue to assume coreference with others in their community. Moreover, these patients’ motivational impairment clearly does not cause the kind of radical cognitive chaos we discussed in the theoretical case. These facts are enough to secure their ability to think about what is right to do. Thus, the neurological evidence does not support the idea that our motivationally impaired subject could not think about practical normative properties.

The upshot of our discussion is that we cannot extend to the practical domain the type of argument which, if successful, would support internalism in the theoretical domain. The major premise of that argument is that widespread dissonance between a subject’s beliefs and his normative opinions about what to believe will lead to a serious degradation of the subject’s cognitive system and in particular his capacity to think about abstract properties. We haven’t found support for the corresponding premise in the practical case: when there is a radical dissonance between the subject’s

---

35 Damasio, Descartes’ error, especially chap. 3 and 9 (see also above, fn. 33).

36 To prevent misunderstandings, it is worth noting that even if further neuroscientific evidence were to suggest that thinking about what is right to do is not compatible with a severe motivational impairment in humans, this evidence could at most establish that internalism holds for thinkers like us – i.e., thinkers with our specific psychological organization. Neuroscientific evidence could not establish internalism as a metaphysically necessary thesis: it may still be possible to imagine a slightly different psychological organization in which thinking about what is right to do is compatible with motivational impairment. So the standard philosophical thesis of internalism cannot be established on the basis of such psychological evidence. On the other hand, psychological possibility implies metaphysical possibility. So the externalist can take comfort in the neuroscientific evidence which suggests that motivational impairment does not undermine the capacity to think about what is right to do. On the status of the internalist thesis, see above fn. 5.
motivation and his normative beliefs about what to do, the subject’s capacity to think about what is right to do needn’t be seriously impaired, let alone lost altogether.

Should we conclude that competence with normative practical concepts is consistent with a lack of the corresponding motivation? Is internalism false in the practical domain? To draw this conclusion would be too hasty. I have suggested that the motivation to act in accordance with one’s normative practical beliefs is not required in order to have a concept which is co-referential with our standard concept of what is right to do. However, an internalist might insist that such a concept wouldn’t count as a genuinely normative concept: it would only qualify as a normative concept in an “inverted comma” sense. Concepts, after all, are individuated more finely than by their reference alone. Traditional internalists insist that what’s distinctive of our standard normative concepts is that they involve an element of personal endorsement which is essentially tied to motivation: the subject doesn’t qualify as competent with the standard, genuinely normative concepts unless he really cares about what is right and is therefore motivated to bring it about. Thus the traditional internalist debate hinges on a claim about concept individuation – about what is essential to possessing the standard normative concept of what is right to do as opposed to a non-normative concept which is merely co-referential with the standard one. Nothing I have said in this section helps adjudicate this question. The argument sketched in the theoretical case and rejected in the practical case hinges on reference determination, not on concept individuation.

Since the goal of this paper is not to argue against practical internalism, but simply to show that theoretical and practical internalism deserve separate treatment, I don’t need to take a position here on the traditional debate about internalism. Coming to a verdict about this debate would require an in-depth discussion of delicate issues. Most importantly, we would need to take a theoretical stand on how concepts are individuated: without some general theoretical characterization of the identity conditions of concepts, we wouldn’t know which aspects of our ordinary practice we should focus on in determining what is essential to possessing our standard normative concepts. In addition, we can also assume that we would need a more precise characterization of the psychology of the motivationally impaired subject than I have provided. I have focused on the more intellectual aspects of the impaired subject’s reasoning about what to do – aspects which appear to be sufficient to secure reference to what is right to do. But I have not attempted to decide, for instance, whether such a subject could still count as really caring about what is right. Depending on which account of concept individuation one favors, an answer to this question may well be crucial to determining whether the motivationally impaired subject succeeds in sharing our standard, genuinely normative concepts. The main virtue of the argument explored in section 5 is that it bypasses the controversial territory of concept individuation. The moral of the present section is that we shouldn’t expect any such shortcut to internalism in the practical domain.

7. In closing
Normative concepts, I have claimed, play a central role in the reasoning process which allows us to exercise normative self-governance. In this process, the reasoner (i) recognizes prima facie reasons, (ii) evaluates these reasons and comes to a normative conclusion about which belief or action is supported by the best reasons, and finally (iii) forms the belief or

---

37 See Allan Gibbard, Thinking How to Live (Cambridge: Harvard University Press, 2003), chap. 7, for a recent defense of such claims.
performs the action which she takes to be right. I have suggested that, in the theoretical domain, a systematic breakdown in the last part of this process will, *ceteris paribus*, impair the subject’s capacity to perform its first two parts. Because of this particular feature of the theoretical domain, there is reason to think that there is some necessary connection between motivation and the capacity to think about what is right to believe.

In the practical domain, on the other hand, a breakdown in the last step of the process needn’t lead to a radical degradation of the subject’s capacity to perform its first two steps. Our capacity to recognize and evaluate practical reasons and to form normative beliefs on the basis of this evaluation can be preserved even if we have lost the motivation to conform our actions to the normative opinions we form about them. Our capacity to assess what is the right thing to do is grounded in a substantive understanding of what matters in life in general and what is important in our particular circumstances. This understanding needn’t be disrupted if normative beliefs stop regulating the subject’s behavior.

I conclude that the new approach to internalism is unconvincing. Even if there are profound similarities between the role of normative concepts in the theoretical and practical domains, the issue of theoretical internalism raises questions of its own and deserves special treatment. Opponents of practical internalism thus needn’t be immediately worried if some form of theoretical internalism turns out to be true. Given the asymmetry between the theoretical and practical domains which I have highlighted, arguments supporting theoretical internalism needn’t *eo ipso* generalize to practical internalism.

I have explored considerations which, I believe, can support theoretical internalism, though I have not provided a full-fledged defense of that position. I should emphasize that the type of argument I have sketched would only support a weak version of theoretical internalism. Indeed, the argument relies on the examination of the consequences of a subject’s *systematic* lack of motivation to conform his beliefs to his normative assessments. If successful, an argument of this type would establish that a subject must be *at least sometimes* motivated to conform his beliefs to his normative assessments if he is to be competent with normative concepts. It wouldn’t establish stronger forms of internalism which claim that such ties between normative assessments and motivation must exist for the most part or in general. In view of the general shape of our discussion of the consequences of breakdowns in normative self-governance in sections 5 and 6, however, I would be surprised if proponents of internalism succeeded in vindicating such a stronger form of internalism, even in the theoretical domain.\(^\text{38}\)

---

\(^\text{38}\) Earlier versions of this paper were presented at the 2001 conference of the Australasian Association of Philosophy, at the University of Fribourg, and at the Australian National University. Thanks to these audiences for stimulating discussions. Thanks also to Philippe Chuard, Dan Cohen, David Copp, Karen Jones, Philip Pettit, Michael Smith, Dave Sobel, Daniel Stoljar, Sigrún Svavarsdóttir, Jon Tresan, and three anonymous referees for helpful comments and criticisms. Special thanks to Laura Schroeter for extensive discussion of the central issues of this paper.