A “quietist” reading of Wittgenstein is one that attempts to do justice to his conviction that it cannot be the job of philosophy “to reduce anything to anything, or to explain anything”—more pithily, that philosophy “leaves everything as it is.”¹ A central difficulty facing such a reading of Wittgenstein’s remarks on rule-following is his tendency to characterize putative instances of rule-following in ways that seem pointedly to omit something central to our ordinary, pre-philosophical understanding of the phenomenon. Rule-followers are variously described by Wittgenstein as just reacting as they were trained (§198), as proceeding in whatever way comes naturally to them (§185), as obeying a rule blindly (§219), and as acting without reasons (§211). What these characterizations fail to register, what they might even seem outright to deny, is the commonsense view that rule-following is an activity that engages the agent’s understanding. On that view, to follow a rule is to put one’s understanding, one’s comprehension, of the rule into practice: a rule-follower, as we say, acts on or in light of her understanding of the rule. Rule-following thus manifests what philosophers of mind sometimes call sapience.² And so the question arises: if philosophy must leave our ordinary ways of conceptualizing the phenomena of human life as they are, why does Wittgenstein opt for characterizations of rule-following that do not acknowledge a central component of our ordinary conception?

² I believe the modern philosophical use of this term originates with Feigl. See for example Feigl and Meehl 1974.
Of course, in the context of a philosophical work driven by constructive or revisionary ambitions, we can readily imagine a point to Wittgenstein’s “thin” (as I will call them) characterizations of rule-following behavior. For a skeptical account of rule-following, one which denies that genuine rule-following ever does or could take place, the thin characterizations could serve as a deflationary assessment of what’s really going on in those cases we mistakenly view as involving a person’s following a rule. And for a reductive account of rule-following, which would aim to explain that phenomenon in “naturalistic” (i.e., non-intentional, non-mentalistic) terms, the characterizations might be taken to provide materials for such an explanation, perhaps by gesturing toward a form of behaviorism or dispositionalism. Skeptical and reductive accounts of rule-following have occasionally been credited to Wittgenstein. But the whole point of a quietest reading of Wittgenstein is to make sense of the text without ascribing to him such conspicuously substantive philosophical theses.

This is to cast our question as a puzzle about motivation: what leads Wittgenstein to describe rule-following in ways that do not explicitly portray it as a form of sapient activity? If the role of understanding is a central element in our ordinary conception of what goes on in rule-following, and if it is not philosophy’s place either to cast doubt on or to naturalistically reconstruct our ordinary conceptions, then why take pains to avoid acknowledging this element? But the problem for a quietest interpretation can be made more serious. For it can easily seem that the thin characterizations of rule-following behavior do not merely leave unexpressed, but in fact are incompatible with the thought that rule-following manifests understanding. Surely, one might suppose, to act with understanding is not just to do whatever comes naturally or to react however one was trained, still less to act blindly or
without reason. These are characterizations we reserve for behavior whose determinants are
the brute forces of habituation and innate disposition, and actions performed in the light of
understanding are precisely not that.

This apparent conflict between Wittgenstein’s characterizations and our ordinary
conception has been recently been pressed by Thomas Nagel in his thought-provoking book,
*The Last Word*. Nagel grants that it would be preferable to find a reading of Wittgenstein
showing him to live up to his official repudiation of a view of philosophy as tasked with
producing reductive accounts of the phenomena of human life, or failing that, skeptical
denials of their reality. What stands in the way of this interpretive goal, for Nagel, is just
those characterizations we have been discussing, which Nagel calls Wittgenstein’s “facial
descriptions of our practices”. These descriptions “suggest that the final and correct
conception of what I am doing when I add, for example, is that I am simply producing
responses which are natural to me, which I cannot help giving in the circumstances (including
the circumstances of my having been taught in a certain way)” (Nagel 1996: 48). So viewed,
says Nagel, the practices “lose their meaning”; they appear as mere “impotent rituals” (Nagel
1996: 51, 53). The problem is that we cannot square this conception of our behavior with
how our performances look to us from the “inside”, a vantage point from which they are seen
as flowing from the thoughts and beliefs that constitute our understanding of, our insight into,
what the relevant rules dictate. To suppose, for example, that when I do arithmetic I am
“simply producing responses that are natural to me…would be to get outside my arithmetical
thoughts in a way that would be inconsistent with them” (Nagel 1996: 48-49). Adopting the
picture of rule-following behavior that Wittgenstein seems to encourage upon us—in which
the source of our behavior is our natural inclinations to go on in certain ways—we are unable
to conceive rule-following as the sapient activity that, when are actually engaged in doing it, we cannot help but take it to be.

There is no doubt that Nagel’s worry has a *prima facie* force. (In calling it “Nagel’s worry”, I don’t mean to suggest that it is idiosyncratic to him. I think the discomfort with Wittgenstein’s characterizations that Nagel articulates is a common reaction to the text.) At the same time, whatever the tension may be between saying, e.g., “In doing X, S just did what came naturally,” and saying “In doing X, S acted on her understanding of a rule,” it obviously doesn’t attain the status of a logical or analytic contradiction. If we want to pinpoint the precise nature of the incompatibility, then, we need to do some work unpacking the relevant concepts. This at least opens the possibility that Wittgenstein’s characterizations are in the end incompatible, not with the very idea that we follow rules, but only with a misconception of what that involves, one that may tempt us but which ought to be ferreted out from our thinking. And indeed, that is what I will try to show is the case.

In so doing, I follow a familiar template for quietest readings of Wittgenstein. Typically, such readings acknowledge that Wittgenstein is denying *something* with respect to the phenomenon at issue. But they hold that what is being denied is not the reality of the phenomenon per se. What is being denied is rather an imposition on our ordinary conception of that phenomenon, a confused theory or picture that distorts our understanding—in other words, a piece of philosophy, in Wittgenstein’s pejorative sense of that term.

But although the strategy is familiar, the applications of it in the recent work on the rule-following remarks do not provide a successful response to the worry we have been discussing. Indeed, they bypass this worry completely. This state of affairs is owed, I believe, to the

---

3 A contradiction is analytic, let us say, if it is immediately evident from our competence in the meanings of the relevant expressions.
enormous influence of Kripke’s *Wittgenstein on Rules and Private Language*. What Kripke cast as the main issue raised by Wittgenstein’s rule-following remarks is rather different from the one we have been discussing, and Kripke’s reading set the terms for the subsequent secondary literature. The result is that the issue that we have identified here has largely gone missing.

In fact, the situation is more complicated than this: Kripke’s exposition of his “skeptical paradox” intertwines two distinct trains of thought, one of which concludes in the, as it were, “official” skeptical thesis, and the other of which engages an issue close to the one that will be our concern. A subsidiary aim of this paper is thus to elucidate an aspect of Kripke’s discussion that, perhaps owing to Kripke’s own lack of clarity on the matter, previous commentaries have been unable to get into focus.

Although I will begin in the next section by saying something about why those quietest readings that were developed in reaction to Kripke’s official story fail to address the worry that is our current topic, my primary aim is simply to address the worry, and that is what the bulk of the paper will be given over to doing. Let me register a few caveats in this regard before proceeding. First, I shall follow Nagel in taking talk of “doing what comes naturally” as the principal representative of the various “thin” characterizations of rule-following behavior cited above. We need to make some such choice in order to give focus to the discussion, and for various reasons that will emerge, this is a good one. Second, quietest readings that adhere to the template just described inevitably face the question of where, and how, to draw the line between our ‘ordinary’ conception of the target phenomenon and the allegedly confused philosophical design that is put upon it. There is one point on this score whose plausibility we have already noted and that I will henceforth assume without argument:
anything recognizable as an activity of rule-following must involve the agent’s acting upon her understanding of a rule. An interpretation that buys quietism at the cost of taking Wittgenstein to deny that rule-following is a manifestation of sapience just blurs the worthwhile distinction between a quietest and skeptical interpretation.

1. Kripke and McDowell

On Kripke’s official story, the primary question raised by Wittgenstein’s rule-following remarks is not “What is it to act upon one’s understanding of a rule?” but “What is it to so much as have an understanding of a rule?” According to Kripke, Wittgenstein’s answer to this question is the dark claim that there is nothing that it is to have an understanding of a rule—that there is simply no such mental state as ‘understanding a rule’ (or, correlative, ‘grasping the meaning of a word’). Any candidate for such a state would have to stand in a “normative” relation to our behavior, in the sense that it would have to embody a view about which courses of action count as being in accord with the rule or meaning and which courses of action count as being in conflict with it. But, says Kripke, the tendency of Wittgenstein’s remarks is that no mental state could possibly have these “normative” implications. This skeptical view is said to be most explicitly stated in §201, where Wittgenstein speaks of “our paradox” that “every course of action could be determined by a rule, because every course of action could be made out to accord with the rule”, and adds that if this were so, “there would be neither accord nor conflict here.” Kripke reads this passage as summarizing a complex set of arguments contained in sections §§138-242, the upshot of these arguments being that nothing that might enter our minds in the course of an attempt to understand (“make out”) a rule could have any determinate implications for what counts as being in accord or in conflict.
with the rule. Hence nothing could add up to our understanding the rule in one way rather than another (Kripke 1982: chapter two).

Commentators concerned to present a quietist alternative to Kripke’s skeptical interpretation are thus led to offer a different reading of §201 and comparable passages. Assuming they follow the template described above, the thrust of their readings will be that Wittgenstein in these passages is exposing the “paradoxical” character, not of the very idea of an understanding of a rule, but rather of a confused conception of what that idea requires. The most well-known interpretation of this sort is due to John McDowell. According to McDowell, Wittgenstein’s target, in its most general form, is a conception of the mind as a realm of items that do not intrinsically stand in relations of accord or conflict with anything else. The contents of a mind are in this respect conceived like physical objects. A physical object—say, a piece of steel affixed to a post and bearing the inscription “KEEP RIGHT”—does not inherently, in its very nature as the physical object it is, have a normative bearing on our behavior. We may take it to have that bearing on our behavior, but what then gives it that bearing for us will be precisely our so taking it. By the same token, if items in the mind intrinsically lack normative significance, they can acquire such a significance only by our interpreting them as having that significance. And it is just here that McDowell’s Wittgenstein finds trouble. For interpretation is a mental activity. Given the governing conception of the mind, any item that might enter our minds in the course of our interpreting something will itself, pending interpretation, lack normative significance. And what could our interpreting our interpretation amount to, on this conception, but the occurring to us of some further item that now stands in need of interpretation? We are locked into an infinite regress.

---

4 My discussion of McDowell draws primarily on McDowell 1998. For a detailed reading along similar lines, see Finkelstein 2000.
On McDowell’s reading it is this threat of regress to which Wittgenstein is drawing attention in his discussion of “our paradox” in §201. However, the cure Wittgenstein recommends is not to give up the very idea that we understand things but to abandon the ‘sign-post’ conception of the mind’s contents, according to which mental items are intrinsically normatively inert. Whatever the attractions of that conception may be—and McDowell faults Wittgenstein for not saying enough on this score—it is a mistake we ought to exorcise from our thinking. Having done so, the question to which Kripke despairs of giving a positive answer will no longer seem pressing: “There seemed to be problems about the normative reach of meaning, but since they depended on a thesis we have no reason to accept, they stand revealed as illusory” (McDowell 1998: 274).

It is not my purpose here to query McDowell’s correction of Kripke, or to deny that the conception of the mind McDowell identifies figures as a nemesis in Wittgenstein’s rule-following remarks. The point I want to make here is simply that McDowell’s reading is of no help in confronting our worry about those remarks. Our worry, recall, is that several of Wittgenstein’s characterizations of putative rule-following behavior, of which “doing what comes naturally” may be taken as emblematic, do not square with the idea that rule-following involves acting on one’s understanding of the rule.

Suppose we were to say to someone, like Nagel, who sees a conflict between this idea and Wittgenstein’s thin characterizations: “Look, Wittgenstein does not mean to reject the very idea that rule-following involves acting on one’s understanding of the rule. The characterizations that bother you are incompatible only with a particular conception of the nature of understanding. This is a conception according to which understanding a rule, or indeed being in any mental state that has ‘normative reach’, is a matter of having in one’s
mind an item that in and of itself lacks such reach, and so can possess it only courtesy of an interpretation. That you mistake a rejection of this conception for a rejection of the very idea that rule-following involves acting on an understanding just shows how in thrall you are to the mistaken conception. Free yourself of it, and your disquiet with the thin characterizations of rule-following behavior will disappear.”

Surely our interlocutor can be forgiven if he finds this line unconvincing. We can readily imagine the response: “By all means, let us allow that the mind is populated by items with intrinsic normative significance. Let us allow in particular that a person’s understanding of a rule or meaning does not just stand there like a sign-post, awaiting interpretation, but is inherently such as to have a particular normative bearing on our behavior. Let us even allow that a good part of Wittgenstein’s remarks is given over to urging this view upon us. My problem is that I fail to see how making any of these allowances should affect my perception that in characterizing a rule-follower as just doing whatever comes naturally, acting blindly, and so forth, Wittgenstein is denying the sapient character of that activity. What bothers me about those characterizations is that they appear to overlook that rule-following involves a certain transition from understanding to action: namely, the transition embodied in a person’s acting upon her understanding. Of course, I don’t believe that my seeing a problem here is due to my being unknowingly captive to a philosophical misconception. But if that is indeed what is going on, the culprit must surely be a misconception of what is involved in this transition from understanding to action, not a misconception (such as the ‘sign-post’ conception of the mind’s contents) of what states of understanding are in and of themselves. Again, I’m happy to credit a wholesomely quietest attitude to Wittgenstein on the latter front. But it would then be a way of putting my objection to Wittgenstein that there seems little
accomplished by our protecting from philosophical confusion the idea of a state of understanding, if we then turn right around and characterize our behavior in ways that preclude the thought we ever put our understanding of rules into practice, by acting upon or in light of it.”

That this response is well-taken indicates that we need to look elsewhere for a quietest answer to Nagel’s worry. We need a different candidate for the role of the philosophical misconception than McDowell’s ‘sign-post’ conception of the mind’s contents.

Where do we look for an alternative? Here is one clue. The question that vexes Kripke—“What is involved in understanding a rule (or grasping a meaning)?”—is a question belonging to the philosophy of mind. It is a question about the nature and status of a certain kind of mental state. The question, “What is involved in following a rule?”, by contrast, is most naturally understood as a question for the philosophy of action. Philosophy of action is not, of course, divorced from the philosophy of mind, but its focus is questions about the transition from mental states to action rather than questions about mental states as such. What troubles us in Wittgenstein is precisely his apparent denial of one such transition, namely, that which holds between understanding and behavior when one acts upon one’s understanding of a rule. And so whereas Kripke and McDowell draw on contemporary ideas in the philosophy of mind in their attempt to think through for themselves the issues Wittgenstein raises (as Wittgenstein famously encourages his reader to do in the preface to the *Investigations*), we may do well to turn to contemporary philosophy of action. In particular, to the concept of rational action, of acting for a reason.5

---

5 Throughout, I use “rationality” and its various forms in their more general sense according to which any action done for a reason counts as rational and any being that can appreciate and be moved by reasons counts as rational. In this sense, “rational” contrasts with “non-rational”, not “irrational".
2. Rule-following and rational action

Start with the difference between being in accord with a rule and following a rule (cf. Wittgenstein 1960: 13). Obviously these relationships are not the same. If I jot a sequence of numbers down on a page, what I write is no doubt in accord with any number of algebraic rules for the development of a series. But of course it doesn’t follow that I am following all of these rules. I’ve never given a thought to the vast majority of them; indeed, many would require such complex formulae for their expression that I couldn’t understand them if I tried. And even if I were aware that what I was writing was in accord with a particular rule—say, with the rule for producing the Fibonacci sequence—that by itself would not imply that I was following that rule, for it would not imply that I was acting on my understanding of that rule. I may be writing what I do to some other end entirely—perhaps I’m recording business expenses, or just listing my favorite integers—and happen to notice the correspondence.

What more is required if a performance is to count as a case of acting on one’s understanding of a rule? One element obviously missing in the case just described is any explanatory connection between my understanding of the rule for the Fibonacci sequence and my writing a series of numbers in accord with that rule. This suggests the following condition: when an agent follows a rule, as opposed to just producing behavior that happens to accord with the rule, her understanding of the rule helps explain what she goes on to do.

This answer immediately raises a further question. What kind of explanatory connection obtains between one’s understanding of a rule and one’s action when one acts upon one’s understanding? How does the former explain the latter? Again, an obvious initial answer suggests itself: a person’s understanding of a rule explains what she does when she acts upon it in virtue of her understanding of the rule being her reason (or at least part of her reason) for
doing what she does. This is to represent acting on an understanding of a rule as a species of rational action, of acting for a reason. And it is to represent explanations of what people do in terms of the rules they follow as belonging to the familiar genre of (as it is sometimes called) \textit{rational-psychological} explanation, of explaining what people do by giving their reasons for doing it.

Again, this answer is illuminating only to the extent that we have a clear view of what distinguishes the kind of explanation at issue from others. The first point to register in this regard is that when we speak of a person’s reason for doing something, what is in question is not merely a reason in the undemanding sense in which any explanans counts as a reason for what it explains. If last night’s heavy wind explains why the tree in my yard fell over, then last night’s heavy wind is the reason the tree in my yard fell over. But it was not the tree’s reason for falling over. The tree had no reasons. Nor does the mere fact that the explanandum of a given explanation involves a person ensure that the explanation will involve appeal to reasons in the sense we are concerned with here. My clumsiness and distraction may be the reason I trip down the stairs, but they are not my reason for tripping down the stairs. Contrast a case in which I trip down the stairs in the belief that this will attract your sympathy. Here my belief is not merely \textit{the} reason I trip. It is \textit{my} reason for tripping—it is the reason I act \textit{for, on or in light of}, the reason upon which my action is \textit{based}.

That the explanation of my tripping in the latter case involved an ascription to me of a belief is not accidental; a little reflection suggests that all rational-psychological explanations of actions involve ascriptions of contentful attitudes, states or occurrences to the agent. But although it may be a necessary condition, the involvement of a contentful state or occurrence in the production of a given bodily movement is clearly not sufficient for the mover to thereby
have done something for a reason. Suppose I am suddenly struck by the thought that I left a burner on in my kitchen this morning, and this so disquiets me that I lose my balance and trip down the stairs. Here my believing that I left a burner on is a crucial part of the reason why I trip down the stairs, but it is no part of my reason for tripping—in tripping, I did not act on or for this reason (or, in this case, any other reason).

What further conditions must be met if a person’s doing something is to count as a case of her doing something for a reason? Here is one way to approach this question. Suppose I ask what reasons our mutual acquaintance Rachel has for moving to Detroit. As it stands, the question is ambiguous: I may wish to know why someone in Rachel’s situation should move to Detroit, or I may wish to know why Rachel actually is moving to Detroit. Call these respectively the normative and explanatory senses of the question. If the question is intended in the normative sense—as asking why Rachel should move to Detroit—its answer calls for you to state considerations that, in your view, count for (or against) Rachel’s moving to Detroit. Here a reason is simply a circumstance, a fact, a way the world is; and what makes a particular circumstance or fact into a reason for something is that in some respect or other it counts in favor of or supports it. Let us call this notion of a reason—according to which a reason is a circumstance or consideration that counts in favor of or supports doing something—the normative concept of a reason.

Now clearly the normative concept of a reason is not identical to the concept of a reason according to which a person’s reasons explain why she does what she does. If my question were intended in the explanatory sense—as asking why Rachel is in fact moving to Detroit—it will obviously not do for you just to retail your own view of the considerations that speak

---

6 Many philosophers of action have noted a distinction in this vicinity. See for example (and there are many others) Scanlon 1998, Dancy 2000, and Smith 1994.
for or against that course of action. These considerations, however plausible or decisive you may rightly take them be, might have nothing to do with why Rachel is actually making the move. It does not follow, however, that the normative and explanatory concepts of reason are unrelated. On the contrary, reflection on how answers to the normative and explanatory questions are interwoven in everyday discourse on human action strongly suggests the following view: that offering an answer to the explanatory question, while it does not commit the answerer to any particular view about what counts in favor of what, does involve ascribing to the agent a conception of what considerations support the action at issue. On this view, an agent’s reason for doing X, in the sense of what explains her doing X, is a function of what she takes to be a reason for doing X, in the sense of what she takes to count in favor of doing X. More fully, a rational-psychological explanation of an agent’s doing X involves these three components: an ascription to the agent of certain beliefs, desires, or other propositional attitudes, an ascription of a further element of the agent’s taking her beliefs or desires to be reasons (in the normative sense) for doing X, and finally, the claim that her doing X is explained by these ascriptions. 

It’s true that philosophers of action of a “Humean” stripe sometimes deny the view that a person’s acting for a reason involves her taking something to be a reason in the normative sense. I shall not defend the view here beyond noting that its denial prevents one from taking

---

7 The statement of these three conditions in the text is not intended to constitute a non-circular specification of necessary and sufficient conditions for rational action. Indeed, counter-examples can easily be constructed, in the spirit of Davidson’s famous rock-climber example, that show that the conditions are not sufficient. Or to put it another way, they are sufficient only if we question-beggingly append “in the manner of a rational-psychological explanation” to the statement of the third condition. But the claim of necessity is all that is relevant for present purposes.

8 Often denials of the view stem from misunderstandings of what it requires. See for example Velleman 2000, in which Velleman assumes that a person’s seeing herself as having a reason in the normative sense for a course of action implies that she sees the course of action as serving some value or good. Thus a view of rational action like the outlined here is seen to portray human beings as saints, always acting for the sake of the good (as they see it). But the assumption is simply a mistake: one can take a consideration to count in favor of an action
everyday discourse about reasons at face-value. Taken at face value, everyday discourse about reasons everywhere betrays the assumption that answering the explanatory question about reasons involves taking up the agent’s perspective on the corresponding normative question. We take it, for example, that our account of why Rachel should move to Detroit might potentially identify precisely what explains why Rachel did move to Detroit. And we hope, and generally assume, that we can explain our own actions simply by answering the relevant normative question, by saying what reasons (in the normative sense) there were for that action. It is an evident commitment of our ordinary discourse on action that an agent’s perception of what counts as a reason for what figures centrally in explanations of why she acts as she does.

In sum, if we take ordinary discourse about reasons at face value and accept the view of acting for a reason it appears to entail, and if we grant as well that acting on an understanding of a rule or meaning fits the mold of acting for a reason, we arrive at the following thesis about the relationship between understanding and behavior: when a person follows a rule, she does what she does because she takes her understanding of the rule to be a reason (in the normative sense) for doing that.

Another example may help to bring all this abstraction down to earth. Suppose you are following the rule add two and at a certain juncture write down “68”. According to the account we have arrived at, the following are all implications of that supposition:

1. In writing down “68” at this juncture, you are acting on your understanding of the
rule add two.
2. Your understanding of the rule is a reason (in the explanatory sense) for your writing down “68” at this juncture.
3. You take your understanding of the rule to be a reason (in the normative sense) for your writing down “68” at this juncture.

Perhaps it is worth making explicit that none of this implies that your understanding of the rule is your whole reason for doing what you do. You may, for example, take your understanding of the rule to give you reason to write “68” only because you take yourself to have reason for following the rule to begin with, and you may take this to be so only because you take yourself to have some further end that would be served by following the rule. These will then be further reasons for your action—here as everywhere, rational-psychological explanation, the cataloguing of reasons, is a complex and open-ended enterprise.⁹

A final caveat. The question of what precisely we are to identify as the agent’s reason, in the explanatory sense, for doing X, is not answered unambiguously by ordinary practice: sometimes we identify the reason upon which a person acts with what she believes or desires and sometimes with her believing or desiring it. Similarly, sometimes we identify a rule-follower’s reason with the rule (that she understands), and sometimes with her understanding of the rule. We are especially prone to talk in the second way if we take the person to misunderstand the rule she is trying to follow: if we think she is wrong that the rule for a

---

⁹ But it is also worth noting that, so far as the present view of rational action is concerned, we don’t have to suppose you have such further reasons for your rule-following behavior to be intelligible as a rational action. When one is in a situation in which a given rule applies to one’s behavior—so that what one goes on to do will count as correct or incorrect in light of the rule—one may well feel a prima facie obligation to do what one takes the rule to require, even if one does not have in view any further point or purpose that would be served by doing so. The mere fact that the rule applies to one’s behavior in the situation (for whatever reason) may be enough for one to take oneself to have a reason to conform to one’s understanding of the rule. This disposition—to accede a primitive reason-giving force to rules—is, I think, quite common. (Contra the old saw, our tendency is to feel that rules are made to be followed, not broken.) No doubt we can imagine various sociological or evolutionary-biological explanations for its existence. But the availability of such explanations has no bearing on the current point, which is about the agent’s view of her reasons for acting. The salient point is that in viewing rule-following as a species of rational action, and in understanding rational action in the manner suggested in the text, we are not thereby committed to the perhaps overly rationalistic claim that whenever a person follows a rule, she is bringing to bear some conception of the point or justification for following the rule. To paraphrase Wittgenstein, a person may act for reasons even if those reasons soon give out (cf. §211).
given arithmetical series dictates writing “68” at a certain juncture, then we cannot sensibly say “Her reason for writing ‘68’ at that juncture was that it accorded with the rule for the series,” although we can say, “Her reason for writing ‘68’ at that juncture was that it accorded with her understanding of the rule for the series.” These two ways of specifying explanatory reasons seem to suggest two radically different conceptions of what such reasons are—on the one conception, worldly states of affairs, and on the other, psychological states of the agent—and the question of which conception is correct has in very recent years become a central topic in the philosophy of action. My own view is that this question is misguided: that ordinary talk tolerates both ways of specifying a person’s reasons is harmless, and does not call for some principled disambiguation on the part of the philosopher. But we do not need to get into this vexed issue here. I shall continue to employ both ways of speaking as occasion warrants, and to exploit the convenient act/object ambiguity of words such as “understanding” to avoid committing to one or the other. As we’ll see, however, the line of thought we will trace out in this paper, and with which we will ultimately part company, has the effect of emphasizing an ‘inner’ (i.e., psychological) aspect to talk of understanding.

3. A preliminary result

Our task is to understand the basis for the impression that Wittgenstein’s ways of characterizing putative instances of rule-following behavior do not square with our ordinary idea of what rule-following involves. How do the foregoing considerations, assuming we accept them, contribute to this goal?

Well, at least one thing that is immediately clear is that it would be a gross mistake to try to leverage Wittgenstein’s characterizations into an account of what it is to follow a rule, of

---

10 See, for example, Dancy 2000.
what rule-following \emph{consists in}. For example, suppose we hold that what makes it the case that a person, in producing behavior in accord with a given rule, is \emph{following} that rule is simply that this behavior is what comes naturally to that person in those circumstances. This proposal could seem even remotely plausible only in advance of our thinking through the implications of the idea that rule-following is a species of rational action. Clearly, conceiving a person’s behavior as a case of her doing what comes naturally cannot by itself fund the idea that she is acting out of a perception of what she has reason to do. Even the most committed naturalist would be hard-pressed to deny that idea must involve something more, or something different.

Talk of a creature’s doing what comes naturally has an obvious affinity with talk of a creature’s doing what it is disposed to do, and this suggests we ought to be able to reformulate the current objection to apply to the dispositionalist view of meaning and understanding that has been a central topic in the literature initiated by Kripke. And indeed we can. This objection will be different from Kripke’s own (at least his official objection—see section 4). Kripke’s objection, in accordance with his focus on the question discussed in section 1, concerns the inability, as he sees it, of a dispositionalist account of states of understanding and meaning to provide for the normative bearing of such states on our behavior. The gist of the objection is that, while one’s behaving in a certain way may be inconsistent with the hypothesis that one possesses a certain disposition, there is no sense in saying that one’s behavior fails to accord with that disposition, or that it is incorrect in light of that disposition.

The current objection to dispositionalism, by contrast, concerns a different “normative” relationship between understanding and behavior than that of mere accord; it concerns the relationship that obtains between behavior and understanding when an agent acts on her
understanding. (One way to keep in mind the difference between these relationships is to remember the simple point that a person’s behavior can accord with her understanding of a certain rule even if she is not acting on that understanding.) The objection is that the resources to which dispositionalism restricts itself do not come into contact with what we have seen to be a central component of that relationship: that the person takes her understanding to give her a reason for what she goes on to do. The core conviction of dispositionalism is that all we need to make sense of the thought that a person understands a rule is to register the fact that the relevant behavior is a manifestation of a certain physically specified disposition. But while indicating that a bit of behavior is the manifestation of such a disposition tells us something about the causes of the behavior, it does not bring anything into view that would portray the behavior as prompted by a perception of something’s counting as reason for that behavior.

It is important for placing this objection, as with Kripke’s own, to see that it does not concern the extensional adequacy of the target accounts. The claim is not that we can be sure ahead of time that no specification of behavioral dispositions, however complicated, will yield necessary and sufficient conditions for a creature’s possessing an understanding of a given rule. It’s obvious that, at least as a matter of empirical fact, those of us who grasp rules have all sorts of dispositions not possessed by creatures who don’t grasp those rules. If the dispositionalist can find a rigorous and non-circular way of specifying these dispositions (according to Kripke, a doubtful prospect), perhaps she will arrive at naturalistic necessary and sufficient conditions for, say, meaning plus by “plus” or understanding the rule add two. But that result, if achieved, would not allay the present worry. What is at issue here is the concept of acting on an understanding of a rule; the suggestion is that an essential kernel of
this concept—namely, the idea of acting out of a perception of what counts as a reason for so acting—resists assimilation to the conceptual materials to which the accounts under consideration restrict themselves. These accounts stand or fall with our being able to hold onto the thought that people act on their understanding of rules even as we strip our conception of human action down to the bare notion that human beings react in various situations in ways that come naturally to them, in ways they are disposed to do as a result of training or inborn proclivities. And the objection is that this is something we cannot do. There is nothing in the restricted conception to give life to the thought that a person producing behavior in accord with a rule is doing so because what the rule as she understands it dictates is taken by her as a reason for that behavior. Thus there is nothing in the restricted conception to give life to the thought that she is following a rule.

4. The ‘guiding’ conception of understanding

We may put the upshot of the preceding section as follows: that a given stretch of behavior counts as a person’s following a rule cannot consist solely in its typifying a way of proceeding that comes naturally to that person. Now, this conclusion is consistent with either of two further diagnoses of where the proposal goes wrong. The first is that viewing an agent as following a rule and viewing an agent as just doing what comes naturally are incompatible, mutually exclusive ways of conceiving behavior. Behavior that is an instance of the former cannot be an instance of the latter. This assumption, of course, funds Nagel’s objection to Wittgenstein. But nothing said in the previous sections rules out a different possibility: that characterizing a bit of behavior as a natural reaction to the circumstances is simply not enough to bring into view that the agent is following a rule. To accept this diagnosis of the proposal’s
error is to allow that an agent’s just doing what comes naturally can count as her following a rule; the thought is rather that doing what comes naturally cannot be all that is involved in following a rule, and so cannot be what following a rule consists in. Until we can rule out this possibility, we do not have a clear ground for our discomfort with Wittgenstein’s thin characterizations.

Nonetheless, we seem to be on the right track. The problem we found with the simpleminded accounts of rule-following considered in the preceding section is their failure to provide for the idea that following a rule is a form of acting for a reason, which, we have seen, involves the rule-follower’s taking her understanding of the rule to be a reason (in the normative sense) for what she goes on to do. There seems nothing to prevent us from putting this idea as follows: when one follows a rule and hence acts on one’s understanding of the rule, the connection between understanding and action is effected by a perception on one’s part of a normative (reason-involving) connection between them. This way of talking in turn encourages us to think of what is at stake here as a kind of mental process. And now it is very tempting to suppose that we have found something whose presence is not merely unprovided for in a description of the agent as, say, just reacting to the situation in whatever way she is naturally inclined to do, but whose presence is positively incompatible with that description.

It turns out to be a matter of some subtlety to develop this intuition in a satisfying way. Certainly it is not enough attach the label “mental process” without any further gloss: the connotation of that phrase is too plastic, too indefinite, for introducing it to do the requisite work on its own.\(^\text{11}\) The idea must be rather that the considerations we have retailed in the

\(^{11}\) Wittgenstein himself is often read as denying that acting on an understanding is a mental process and that understanding is a mental state. But nothing Wittgenstein says commits him to the implausibly strong view that there is no sense of the terms “mental state” and “mental process” with which they might reasonably be applied to understanding and activities that engage it; indeed, to see him as in the business of making such blanket
previous sections suggest a particular model of a mental process at work when one follows a rule, and that nothing of that sort could obtain in cases where a person just does what comes naturally. Nor will it do to opt for perhaps the most obvious candidate in this regard and suggest that the process be conceived as one of thinking, in the sense of active deliberation or reflection. That may seem a tempting idea in the present context, for viewing a person as doing what comes naturally to her does seem to contrast with viewing her course of action as shaped by a conscious process of deliberation, reflection or the like. The problem is that it’s obvious that many cases of rule-following, including those cases Wittgenstein focuses upon, simply do not involve any such process. It would seriously over-intellectualize matters to insist that, in following a simple rule like add two, a competent adult must engage in anything akin to deliberation or reflection. If the source of our disquiet with Wittgenstein’s characterizations were their denying the presence of an active deliberative process, the culprit would not be the characterizations themselves but the over-intellectualized picture of rule-following to which we are apparently party.  

But there is another possibility. The source of the disquiet may lie, not in the clumsy thought that acting on an understanding always involves a process of deliberation per se, but rather in the subtle thought that it essentially involves a process of guidance. There are a variety of ways of conceiving this process that suggest themselves. When you follow a rule, you refer to or call upon your understanding of the rule. You are instructed or directed by your understanding. Your understanding shows or tells you what to do. These formulations

pronouncements is to drastically misconceive his approach. It is essential in interpreting a remark like, “Try not to think of understanding as a ‘mental process’ at all.—For that is the expression which confuses you” (§154), to ask to whom Wittgenstein is offering this advice and why. No one who thinks through these questions will find it plausible to see Wittgenstein as flatly giving out that acting on an understanding is not a mental process.

portray acting on an understanding as involving a kind of inner consultation, in which your understanding advises you what you are to do next and you act on that advice.

Note that this idea—call it the ‘guiding’ conception of understanding—is distinct from, and does not entail, the discarded idea that acting on an understanding always involves deliberation or reflection.\(^\text{13}\) It is certainly possible to act upon guidance, direction or instruction without actively thinking about it. The clearest example of this we might draw from ordinary life is a case in which, as we say, one’s “mind is elsewhere.” Suppose you are driving someone to a location unfamiliar to you, which necessitates her occasionally directing you to turn this way or that. Simultaneously you are engrossed with her in a conversation on some other topic. You grasp her directions and put them into action, but that happens subliminally: your attention is wholly focused on the conversation. A proponent of the ‘guiding’ conception might appeal to this sort of situation as an analogy for what goes on when one is guided by one’s understanding of a simple rule like \textit{add two}.

The ‘guiding’ conception of understanding seems otherwise well suited to our purposes. On the one hand, it provides an appealing elucidation of the idea of the distinctive mental process involved in rule-following. The idea of such a process was suggested to us by our reflection on the rational character of rule-following, in particular on the mediating role played by your taking your understanding as a reason for going on in a certain way. Here this mediating stage of taking-as-a-reason is concretized as a moment of communication: you look to your understanding, and it advises you how to proceed. That acting on an understanding of a rule is thus represented as involving an uptake on your part seems to well capture our master thought that rule-following is a manifestation of sapience. And on the other hand, viewing a

\(^{13}\) In light of the distinction drawn in section I between our topic and Kripke’s official one, a more perspicuous label for the idea would be “the ‘guiding’ conception of \textit{acting on} an understanding”. But this is just too much of a mouthful.
person as being guided in her performance obviously does not comport well with viewing that person as just doing what comes naturally. Indeed, on reflection, this contrast emerges as a significant part of the connotation of the phrase “doing what comes naturally”. To do what comes naturally is precisely not to proceed on the basis of directions, instructions or any other form of guidance. When you do what comes naturally, you don’t consult anyone or anything on what to do; you just do it. And the conflict is no less acute with respect to the rest of Wittgenstein’s thin characterizations. Someone who proceeds with the benefit of active guidance or instruction is not aptly described as just reacting as she was trained, still less as acting blindly or without reasons. We have thus found exactly what we’ve been seeking: a conception of rule-following that 1) seems congenial in light of various considerations to which we have been led in unpacking the thought that rule-following is a form of rational action, and 2) wears on its sleeve its incompatibility with Wittgenstein’s thin characterizations of rule-following.

Both the idea that acting on or with understanding involves a mental process that is not present when a creature just goes on in the way that comes naturally to it, and the companion idea that this intellectual activity is to be conceived as a sort of inner guidance, are present in Kripke. Indeed, it is striking that, in spite of his ostensible focus on the question of how states of understanding can so much as accord with behavior, the ‘guiding’ conception plays a central role at many points in the exposition of his skeptical paradox. In his first, “intuitive” formulation of the paradox, the normativity condition is cast as the demand that the state in which your meaning what you do by “plus” consists must be such as to direct you in your application of the expression in novel situations: “The ‘directions’...that determine what I should do in each instance, must somehow be ‘contained’ in any candidate for the fact as to
what I meant” (Kripke 1982: 11). Kripke represents the demand as common sense: “Normally, when we consider a mathematical rule such as addition, we think of ourselves as guided in our application of it in each new instance” (Kripke 1982: 17). Or similarly: “Even now as I write, I feel confident that there is something in my mind—the meaning I attach to the 'plus' sign—that instructs me what I ought to do in all future cases” (Kripke 1982: 21). And the requirement emerges time and again in the later discussions of specific attempts to solve the puzzle. The dispositional account is said to fail “as a candidate for a ‘fact’ that determines what I mean [because] it fails to satisfy the basic condition on such a candidate…that it should tell me what I ought to do in each new instance” (Kripke 1982: 24). And the view that the meaning of an expression is determined by its association with a mental image cannot succeed because no image could “tell me how I am to apply a given rule in a new case” (Kripke 1982: 43).

These formulations suggest a rather different source for Kripke’s skeptical conclusion than the puzzle about the possibility of accord that, as we noted, he finds encapsulated in §201. This alternative source would be the discovery that what the ‘guiding’ conception of understanding envisions cannot in fact obtain, that our understanding cannot intelligibly be regarded as guiding our behavior. If Kripke were to take himself to have discovered this, but were nonetheless to remain committed to the proposition that what the ‘guiding’ conception envisions is what acting on an understanding would have to be, he would be forced to conclude that there is no such thing as acting on an understanding. I think this train of thought is indeed plausibly ascribed to him. For the moment, what I want to emphasize is that this is a route toward skepticism additional to the one discussed in section 1. Even someone who, unlike Kripke, embraced the idea that items in the mind have intrinsic normative
significance, and so saw no problem with the very idea of a state of the mind with which our behavior can accord, might still be led, on the basis of a worry about the ‘guiding’ conception, to conclude that acting on an understanding of a rule, hence rule-following, is impossible.\textsuperscript{14}

5. The incoherence of the ‘guiding’ conception

5.1. The regress objection. However tempting the ‘guiding’ conception of understanding may seem, it does not survive scrutiny. Consider one of Kripke’s favored formulations: that my understanding “tells me how I am to apply a given rule in a new case”. There is a simple point to be made here: if my understanding’s telling me how to apply the rule is to be of any help, I must understand what my understanding tells me, and then put that understanding into action. A successful telling requires two contributions: one from the teller and one from (if I may coin an ugly expression) the tellee. The contribution of the latter is to understand what she is being told, and, if appropriate, to act on it. Consider our earlier example, in which your passenger tells you how to get to your destination. Clearly, nothing she tells you will succeed in directing you to the destination unless two further circumstances obtain: a) you understand what she tells you, and b) you act on what you understand. The moral is perfectly general. Thus if we hold that acting on an understanding of a rule is effected by your understanding “telling you how to apply the rule in a new case,” we represent you as understanding what your understanding tells you, and then acting on that understanding. What is involved in your acting on this further, second-order understanding? According to the conception of acting on an understanding now in question, it will involve your second-order understanding telling you

\textsuperscript{14} It is not a coincidence, I think, that there is no mention in McDowell or Finkelstein of Kripke’s many references to the “guidance”, “directions” and so forth that “we normally think of” our understanding of rules as providing. One would have to suppress this aspect of Kripke’s discussion to believe that getting Kripke to give up the ‘sign-post’ conception could by itself serve as an adequate course of treatment for his skeptical doubts.
what to do—in this case, what it tells you is how to apply your first-order understanding. But the same point must now be reapplied: nothing your second-order understanding endeavors to tell you will be of help unless you understand what you are told and act on what you understand. There is now a third-order understanding in play, and the same point applies yet again. If we allow the regress to begin, we lose the idea that people act upon their understanding altogether. At each stage of the regress, we come to an item that is supposed to tell us how to act on the previous item, but is unable to do so unless a further item tells us how to act on it.

The point is well-expressed by Wittgenstein in his treatment of the idea that intuition is required to develop an algebraic series. Wittgenstein’s response to this idea is a rhetorical question: “If intuition is an inner voice—how do I know how I am to obey it?” (§213) An inner voice can enable your correct development of the series only if you understand what the voice is telling you and then act on that understanding. And the conception of acting on an understanding that motivates the initial appeal to the inner voice now motivates positing a second inner voice—one that tells me how to understand the first inner voice. The predicament we are now in is obvious.

Is this problem specific to the “telling” formulation of the ‘guiding’ conception? The answer is no. The problem would be localized in this way if it depended upon a feature distinctive to the concept of telling, for example on the fact that telling, in the most literal sense, is a linguistic performance addressed by one speaker of a language to another speaker of that language. Obviously, telling so construed is not a plausible candidate for what transpires between a person and her understanding. But the objection does not depend upon this feature of telling. It depends on something that is common to all of the phenomena in
terms of which we have depicted the ‘guiding’ conception—telling, showing, guiding, instructing, directing, etc. The essential claim of the ‘guiding’ conception is that the role of your understanding, when you follow a rule, is to convey or impart to you how you are to go on. Speaking is not the only means of conveying or imparting something; an appropriately situated stop sign, to take one example, conveys to a driver where she is to stop, and it does without saying a word, indeed, without doing anything at all. To generate the current difficulty, it doesn’t matter which method we envision your understanding using to convey to you how to proceed—whether we envision it as speaking, gesturing, signaling, just standing there and confronting you (like a stop sign), or some hitherto unimagined possibility. The problem applies at a more general level: no matter how your understanding tries to convey something to you, its succeeding in doing so depends upon your grasping—comprehending, understanding—what is conveyed. Successful communication requires an uptake on the part of the recipient; otherwise, nothing is communicated. Indeed, as we noted in the previous section, the fact that the ‘guiding’ conception envisions an uptake on the part of the rule-follower is what makes it seem an appealing elucidation of the sapient character of rule-following. But as soon as we think through this implication of the ‘guiding’ conception, we see that it makes a regress inevitable. If your uptake, your understanding, is to bear on your subsequent performance, you must act on it. According to the ‘guiding’ conception, acting on an understanding involves the understanding’s conveying to you how to go on. Regress follows.

It is important to see that the target of this objection is a particular philosophical conception of understanding; it is not directed to just any employment of the words “guidance”, “instruction” and so forth in characterizing one’s relationship to one’s
understanding. For one thing, these words have multiple senses, and not all of these senses are apt for framing the ‘guiding’ conception. Consider, for example, the word “guidance” itself. We sometimes speak of one item’s guiding another and mean nothing more than that the first item influences the second, with no implication that this influence was mediated by any uptake, any comprehension or awareness, on the part of the item being guided. Thus we speak, for example, of guiding a ship into harbor. But the relation between agent and understanding posited in the ‘guiding’ conception of understanding is not one in which a person is an uncomprehending instrument of his understanding, as a ship is of its navigator, but rather one in which there is something presented or conveyed to the agent, something that the agent grasps or appreciates. What is at stake is not the guidance of brute causal influence, but what we might call informed guidance, guidance by way of being told or shown how to proceed. The idea that our understanding guides us in this way was the upshot of our attempt to substantiate our (i.e., Nagel’s) worry about Wittgenstein’s thin characterizations. It is this idea we are now seeing cannot be sustained.

There is a further point to be made here. We may sometimes find ourselves employing “guidance”, “instruction” and similar locutions in an attempt to articulate the phenomenology of particular instances of rule-following, and mean precisely that our understanding of that rule conveys to us how we are to go on. We may be inclined to speak this way, for example, in a situation where there is some question about how to apply a given rule: in such a situation, we may ask ourselves what our understanding of the rule “says” to us about how to proceed, and we may even then think of ourselves as subsequently acting on our understanding of what our understanding says to us. It’s crucial to see that nothing in the current argument legislates against this way of describing a particular case of acting on an

---

15 I thank David Finkelstein for encouraging me to be clearer on the issue discussed in this paragraph.
understanding. The current objection is to a proposed philosophical elucidation of the concept of acting on an understanding, in particular, to the claim that acting on an understanding consists in one’s understanding guiding one’s performance. This purports to be an account of what the mental process of acting upon an understanding is, and so it implies, unlike the sort of isolated remark just mentioned, that every instance of acting on an understanding involves the process it delineates. We need the constitutive claim to motivate our intuition that no instance of acting on an understanding, and so no instance of following a rule, can be an instance of a person’s doing what comes naturally to her. But the fully general claim also guarantees the regress, for each of the iterating moments of acting on an understanding must be conceived in its terms.

It is illuminating to contrast the regress now in play with that which we saw McDowell to take to be visited upon the ‘sign-post’ conception of the mind’s contents. On that conception, items in the mind, like physical objects in the world, do not intrinsically possess intentional content, hence do not intrinsically stand in relations of accord and conflict with one’s actions. A mental item has a particular content only if we interpret it as having that content. The regress materializes when we realize that, given the ‘sign-post’ conception, whatever further mental items we marshal in the course of interpreting the original mental item will themselves be intrinsically contentless. And so until they are themselves interpreted, they are unable to fix the interpretation of the original item. Since the same problem arises at each stage, we never succeed in endowing the original item with content.

The regress that is our topic, by contrast, does not depend upon the ‘sign-post’ conception. It is generated by the view that acting on an understanding consists in one’s understanding guiding one’s action. To reiterate: given what this view is supposed to
accomplish (given the motivation I have traced out for it in the earlier sections of this paper),
the notion of guidance appropriate to it is that of informed guidance, guidance by way of
being told, shown or otherwise instructed how to proceed. But one can be thus guided only if
one understands what one is told or shown and acts on that understanding. According to the
governing view, acting on this further understanding is itself a matter of being guided by it.
The train of thought then iterates ad infinitum. It simply has no bearing on this regress
whether the original state of understanding is conceived as having content only courtesy of an
interpretation or is conceived as inherently contentful. Whatever view we take on the
constitution of a state of understanding, the ‘guiding’ conception construes acting on that state
as a matter of being guided or directed by it. And it is that latter assumption, coupled with
completely general considerations about what is involved in being guided or directed by
something, that engenders our regress.

At the same time, there is a sense in which the ‘sign-post’ conception and the ‘guiding’
conception can end up in the same place: they can suggest identical models or pictures of the
mental items that constitute our understanding. In the effort to make concrete our often
elusive conceptions and assumptions about mental phenomena, we will be inclined to picture
mental items and processes on analogy to items and processes in the physical world. (Or so,
at any rate Wittgenstein supposes.) And in the case of the two conceptions at stake, the
attempts at concretization can easily converge. In both cases, we may be led to think of our
understanding of something as consisting of the presence in our mind of, say, a sign, an
image, a set of written instructions, or a speaker or voice. The reason for the convergence is
that these items seem to satisfy the requirements of both conceptions: they have content only
courtesy of their use or interpretation (as the ‘sign-post’ conception requires) and they guide
our actions (as the ‘guiding’ conception requires).

The conceptions themselves, however, remain distinct. Indeed, as we have seen, they are not even conceptions of the same thing: the ‘sign-post’ conception is a view about what constitutes a state of understanding’s possession of content, and the ‘guiding’ conception is a view of what it is to act on an understanding. This raises the question, for any given appearance of these models in the course of the *Investigations*, of which conception lies behind it. Although I shall not defend this interpretive claim here, my own view is that most references to these models—of understanding as image, voice, etc.—are more aptly taken as manifestations of the ‘guiding’ conception than of the ‘sign-post’ conception. The salient question to ask in any given case is whether the train of thought motivating the introduction of the model (a train of thought expressed or implicit in the comments of one of Wittgenstein’s “interlocutors”) would be neutralized by acceding to McDowell’s suggestion that items in the mind can have their content intrinsically. The answer, I think, will generally be no. We can generally imagine the interlocutor responding thusly: “But how does the understanding, even if conceived as intrinsically contentful, bear on our subsequent actions? How does it factor into the production of our actions in the way it must for our performance to exhibit sapience? It was with an eye to this question that I was led to picture the understanding as an image (or voice, or whatever). For so pictured, the understanding looks like something that could show or tell us what to do, and that is what it seems to me we need to make sense of the role of understanding in the production of action.”

5.2. Logical compulsion. I will not pursue this interpretive issue further. Let us return instead to the objection to the ‘guiding’ conception, for there is a further wrinkle here that we
must consider. The wrinkle is this: when we are in the grip of the ‘guiding’ conception, we may be inclined to suppose that our understanding is different from other items that guide or instruct us in just the way needed to block the regress.

To elaborate: the source of the regress lies in the fact that there are two contributions to a scene of informed guidance: one from the object providing the guidance and one from the person receiving it. Simply being confronted by a stop sign as you drive along does not ensure that you will stop. The stop sign will guide you in this regard only if you understand it as showing you where to stop and then act on that understanding. Obviously, the sign itself cannot force you to do this. In that sense, there is a gulf between your confronting the sign and your stopping. The gulf is filled by your understanding what the sign shows you and acting on that understanding.

But suppose we could make out the following thought: the mental item that you consult when you act on your understanding differs from stop signs, car passengers and other physical objects that guide us in that no such gulf arises. We might put the thought this way: the mental item provides for its own understanding. The proper understanding of what the item is saying or showing is built into the item itself. Hence this understanding is not something you must bring contribute to the transaction: it is part of the package that the item presents you. Here we can situate Wittgenstein’s invocation of “a picture, or something like a picture, that forces its application on us,” and the resultant imagined distinction between “psychological” and “logical” compulsion (§140). As things stand, you cannot help but understand a stop sign as telling you to stop. Even so, the compulsion at work here is only ‘psychological’: it is a contingent feature of your psychological makeup. Had you been trained differently, you might now be compelled to understand the stop sign as telling you something else entirely.
What we are now envisioning, by contrast, is a compulsion of a very different character. In the case of the mental item that constitutes your understanding, the question of how you are disposed to understand what the item presents to you does not so much as arise, for the understanding of what it presents to you is already logically \textit{contained} in what it presents to you. We can no more isolate something about which the question of understanding might arise than we can isolate, say, a triangle from its angles.

Here, it might seem, we have a key for ending the regress. The regress is fueled by the need at each stage to account, in a way that squares with the ‘guiding’ conception, for your acting on what your understanding shows or tells you. But if there is a logical unity between what your initial state of understanding shows you and the understanding of what it shows you, then, it might seem, there is no further, distinct state of understanding with respect to which the question of how you are to act on it arises. The regress cannot get started. There is no call to look into the basis or nature of your contribution to the scene of guidance, for there is no such contribution. Your understanding takes care of everything. It \textit{makes} itself understood.\footnote{One way we might try to further elaborate the idea of logical compulsion implicitly plays upon the two senses of “guidance” distinguished earlier. We may, without quite realizing that is what we are doing, conceive the informed guidance allegedly provided by one’s understanding on analogy to the guidance of brute causal influence, the sort of guidance at work when a pilot guides a ship into harbor or, to take the example Wittgenstein discusses in this context, when the hammers in a player piano are guided by the perforations in the paper roll (See §157 and Wittgenstein 1960: 118). Such physical mechanisms of guidance do not depend, of course, on the comprehension of the item being guided: the piano knows nothing of its transactions with the paper roll. Conflating these two notions of guidance, we may be led to a picture of the guidance provided by one’s understanding as the workings of a special mental or psychological mechanism. This mechanism partakes of the self-sufficiency or self-standingness of an ordinary physical mechanism, in particular in its lack of dependence upon the uptake of the guided subject. But since the mechanism exists in the “queer” medium of the mind, its functioning can be understood even so as a matter of its conveying to the subject how she is to proceed. I take this move to be part of what is at stake in Wittgenstein’s lengthy (and insufficiently discussed) examination of the “process” or “experience” of guidance involved in reading text aloud and related activities like writing from dictation or copying a doodle (§§156-178). Reading aloud and the other activities discussed in this section seem well-placed to straddle the two types of guidance we’ve been discussing. On the one hand, all of the activities considered suggest the idea of a mechanism. They are all aptly described as “mechanical” procedures, both in the sense that they do not require comprehension (Wittgenstein is “not counting the understanding of what is read as part of ‘reading’”, §156), and in the sense that the transition from input to
Consider the following remarks from Wittgenstein’s interlocutor:

‘There is a gulf between an order and its execution. It has to be filled by the act of understanding.’
‘Only in the act of understanding is it meant that we are to do THIS. The order—why, that is nothing but sounds, ink-marks.’ (§431)

One might suppose the ‘gulf’ in question here is generated by the fact that words, either spoken or written, are not intrinsically contentful. When I say, “Add two,” the noises I utter, considered in and of themselves, have no special association with the written marks, “2, 4, 6, 8, 10”. It is only in virtue of the meaning or content of what I say—something to which one can be privy only if one understands me—that what I say counts as an order to write, “2, 4, 6, 8, 10,” and so on. So interpreted, the remarks seem perfectly true, and one may well wonder why Wittgenstein finds them meriting scrutiny. The answer is that the interlocutor’s remarks are intended to gesture at a different contrast: that between an item whose power to guide a subject in a particular way is secured simply by its being present to the subject and an item whose power to guide a subject in a particular way is not so secured. On this reading, what is seen to constitute a gulf between order and execution is the fact that the order’s guiding you in the right way is contingent upon your understanding it that way. What could fill this gulf?

Not anything whose power to guide you is itself contingent on how you understand it;

output seems susceptible to characterization in terms of relatively simple algorithms. It may seem a short leap from here to the thought that our performances of these tasks are literally the product of mechanisms. Wittgenstein encourages this move by various means, for example, by imagining a case in which human beings are used as “reading-machines” (§157). But now on the other hand, if there are indeed mechanisms responsible for these performances, they would seem to have to operate at the level of the subject’s sapience, broadly construed. For we are aware of, conscious of, the transitions involved in these performances—for example, the transition from text to speech—and this suggests that we must be conscious of the mechanisms that effect them: “We imagine a feeling enables us to perceive as it were a connecting mechanism between the look of the word and the sound that we utter. For when I speak of the experiences of being influenced, of causal connection, of being guided, that is really meant to imply that I as it were feel the movement of the lever which connects seeing the letters with speaking” (§170). Thus we are led to the idea of a mental mechanism, one which has the autonomy, the independence from our uptake, of a physical mechanism, but which is nonetheless configured out of material available to our awareness. This might be imagined to provide a model for the process envisioned by the ‘guiding’ conception of understanding.

Wittgenstein’s treatment of this turn of thought warrants extended attention, but I do not have the space to discuss it further here.
precisely the same gulf would remain between such an item and the execution of the order. To fill the gulf would require rather an item that provides for its own understanding, and so whose power to guide you in that way is not beholden to contingencies of your understanding. That is just the idea of logical compulsion.

But however tempting this idea may be—and according to Wittgenstein it belongs among a set of ideas or pictures of the nature of mental processes that we find well nigh irresistible—it cannot be sustained.\(^\text{17}\) There can be no exception to the principle that an item’s capacity to guide us, in the “informed” sense, is contingent upon on, among other things, what guidance we take the item to offer. The reason is not that we know a priori that mental items cannot have powers that physical items do not possess. The problem, rather, is that there is simply no making sense of the idea of an item that ‘logically’ forces us to understand it in a particular way. It is perhaps enough just to be fully self-conscious about what it is we are expecting from the envisioned mental items to realize that it cannot be forthcoming. A person will not be guided in a particular way unless she is prepared to be guided in that way, and there is no getting around the fact that her being so prepared is a feature of her—it is something she must bring to the table, not something that the guiding object can itself provide. No matter what we place before a person, no matter how rich or extensive the

---

\(^\text{17}\) What do I mean by saying that this idea belongs among a set of ideas Wittgenstein thinks we find irresistible? It is a familiar point of interpretation that Wittgenstein is concerned throughout the *Investigations* and cognate works to battle the perception that mental phenomena are “queer” or “occult.” (For pertinent uses of these terms, see respectively §195 and Wittgenstein 1960: 5.) On Wittgenstein’s diagnosis, this perception arises at a particular stage in the thinking of someone who is captive to confused ideas about the phenomena in question. At this stage, one realizes that the envisioned mental phenomena would have to be radically different from other phenomena with which we are familiar—hence the sense of queerness—but one has not undertaken the examination that shows that one’s ideas of the phenomena are simply incoherent. “Incoherent” is not an overstatement, and for that reason perhaps “ideas” is too generous: the ‘ideas’ Wittgenstein takes to underlie the perception of queerness are nothing more than attempts to fuse together, using various ill-defined images and analogies as solder, assumptions that are literally contradictory. It is thus enough to be fully explicit about the content of these ‘ideas’ to realize that they cannot be sustained—they are like those deep-sea creatures that disintegrate when brought to the surface. But until we achieve full explicitness about them, they may exert a powerful grip on our thinking. Logical compulsion is one such idea.
materials we present, there can be no guarantee that the person will take those materials to offer the guidance we intend, or indeed to offer any guidance at all. How could it be otherwise? In the end, the person either understands—or she does not. Of course, the mental item guiding the subject’s performance is not conceived to be something that we onlookers exhibit to her; it is something she has in her mind. But nothing can circumvent the dependence of guidance on a person’s uptake; thus, merely stipulating that the guiding materials are mental rather than physical cannot help.

It is true that we have difficulty conceiving how the understanding of some kinds of objects might differ from our own. While the shape and color of a stop sign may be arbitrary, the ways in which we take items such as color samples and images seem to us to be based on relationships of similarity and resemblance for which there are no real alternatives. Philosophers are fond of pointing out that everything is similar to everything else in some respect or other, but we may find it hard to take seriously the possibility of creatures who find it natural, say, to take a bit of green paper as a sample of Goodman’s grue rather than of green. And that is a perfectly reasonable attitude. Beings whose standards of salience and similarity were such that they understood pictures and samples in ways that differed radically from our own would strike us as extremely, perhaps incomprehensibly, peculiar. We get into trouble, however, if we allow this attitude to generate the sense that the guidance yielded us by color samples, images, or diagrams somehow arises wholly out of the existence of the relevant similarities. Even in these cases, there can be no guidance unless we understand the items in such a way as to yield that guidance, and the items themselves cannot force us to do that. Our inability to imagine alternative understandings shows only how strong (if it pleases us to put it this way) the ‘psychological compulsion’ is here; it cannot vindicate the confused idea of
‘logical compulsion’.

The regress cannot be halted once we have allowed it to begin. As soon as we picture the agent’s taking her understanding as a reason for going on in a certain way as constituted by her being guided by her understanding, we are committed to picturing the subject as having to understand her understanding. We will then have to ask what is involved in her acting on that further state of understanding; given the assumption that acting on an understanding involves guidance, regress will ensue. The idea of logical compulsion tries to gloss over this difficulty by having it both ways. It supposes both that understanding operates by guiding performance, and that no further question of how one acts on one’s understanding of one’s understanding will thereby arise. To give explicit voice to this idea is to see it for a delusion.

5.3. Kripke on irreducibility. I believe it is a recognition of the bankruptcy of this idea that prompts Kripke to his brusque dismissal of the proposal that mental states and processes are *sui generis* (Kripke 1982: 51ff). His treatment of this proposal is perhaps the most poorly regarded part of his presentation of his skeptical argument, widely viewed as consisting in nothing more than an unexplained refusal to countenance the possibility of the irreducibility of the mental.\(^{18}\) Crispin Wright’s reading is representative. He takes Kripke to assert without argument that contentful mental states, left unreduced, have unacceptably mysterious features. According to Wright, this move is vulnerable to the reply that it is our ordinary concepts of contentful mental states that ascribe these features to them, and that we need to be given some substantive reason for supposing that our ordinary concepts are not in order (Wright 1984: 775-777).

But to opt for this reading, I want to suggest, is to miss that Kripke’s discussion at this

juncture is controlled by his subscription to the ‘guiding’ conception of understanding. Acting on an understanding, say of a rule, cannot for Kripke be a matter of acting “blindly”; we are genuinely following a rule only if we are “guided in our application of it to each new instance.” Kripke recognizes both that this conception is threatened by a regress, and that we can get nowhere with the alleged idea of logical compulsion intended to neutralize this threat. But given his conviction that a mere “blind” response cannot count as an action that manifests understanding, he remains committed to the conditional premise that, if there were such a thing as acting on an understanding, it would have to conform to the model of the ‘guiding’ conception. In light of this, Kripke can make nothing of the appeal to irreducibility but an attempt to protect the ‘guiding’ conception from the apparently devastating objections he has raised, and to do so just by insisting that the objections are improper because understanding is sui generis and so not to be subject to examination. With the irreducibility appeal understood in this way, the charge of mystery mongering is apt. When a compelling doubt has been raised about the very intelligibility of an idea, it is no defense to just style the idea as sui generis. If we simply cannot make sense of the idea that one’s understanding guides one’s performances, an assertion of irreducibility is beside the point.

I suspect one reason commentators have missed the importance of the ‘guiding’ conception for Kripke’s dialectic (in addition to the fact that, as we have noted, his own exposition obscures it) is that they take the conception for granted; it controls their thinking without their recognizing it as a substantive idea. Paul Boghossian writes of Kripke’s dismissal of the sui generis proposal: “[The] objection to the anti-reductionist suggestion is that it is utterly mysterious how there could be a finite state, realized in a finite mind, that nevertheless contains information about the correct applicability of a sign in literally no end
of distinct situations. But...this amounts merely to insisting that we find the idea of a
contentful state problematic, without adducing any independent reason why we should”
(Boghossian 1989: 542). The idea that a person’s understanding of how to use a sign is such
in virtue of being a state that “contains information about the correct applicability of the sign”
is a corollary of the idea that the state’s role in the production of action is to provide the
person with that information. Conceiving states of understanding as providing us with
information is of a piece with conceiving them as speaking to us, or showing us something, or
guiding us. To equate this idea with the idea of a contentful state, as Boghossian does in the
quote, is to tacitly assume the ‘guiding’ conception of understanding. That conception, as we
have seen and as I believe Kripke sees, is unsustainable. It may be that the difference
between Kripke’s “semantic irrealism” (as Boghossian calls it) and Boghossian’s version of
non-reductionism is just a matter of Kripke’s greater self-consciousness about his embrace of
the ‘guiding’ conception, and his recognition of the crippling problems besetting that
conception.

5.4. The spuriousness of the ‘guiding’ conception. At the same time, however, Kripke’s
position is deeply puzzling. Surely the natural thought, once we have acknowledged the
problems with the ‘guiding’ conception described above, is that the conception should be
discarded—that it is an incoherent and hence failed attempt to elucidate its target
phenomenon. But that is not Kripke’s attitude. He wants to retain the ‘guiding’ conception as
the proper account of what it would be to act on one’s understanding even after it is
established that the conception entails an impossible regress. Apparently, by Kripke’s lights,
what the objections to the ‘guiding’ conception establish is only that the conception is (as we
might put it) *unsustainable*—that it imposes conditions that nothing that happens in the world could possibly meet. It does not follow, and indeed we should not conclude, that the conception is *spurious*—that it is a misreckoning of the conditions that must be met if a person is to act on her understanding. Kripke doesn’t draw this distinction explicitly. But it is essential to his (unofficial) case for the irrealist conclusion: we get irrealism because, although the conception is unsustainable, it is not spurious—it captures what acting on an understanding, were it to occur, would have to be.

How do we decide which conclusion to draw from the preceding critique of the ‘guiding’ conception: that the conception is simply a mistake, or that it is, although unsustainable, not spurious? The first point to register in thinking about this question is that the onus is decidedly upon Kripke’s side of this dispute. There are two obvious and related reasons for this. First, Kripke’s stance has an extremely unpalatable consequence: irrealism. It is, to say the least, difficult to accept that we never follow rules or act upon our understanding. Since we avoid this consequence if we take the ‘guiding’ conception to be no less spurious than it is unsustainable, this is a powerful prima facie consideration for favoring that view over Kripke’s. Second, even independently of its consequences, Kripke’s severing of the question of spuriousness from the question of sustainability is an unusual and highly dubious maneuver. Suppose a philosopher offers a philosophical account of some familiar, integral, apparently ubiquitous phenomenon of our lives. Suppose it is then demonstrated that the account implies that it is logically impossible for the phenomenon to obtain—say, because it would engender an impossible infinite regress. Ordinarily, it goes without saying that the appropriate response here is to withdraw the account. Indeed, such a demonstration is

---

19 Nor is the issue one of mere unpalatability. It is doubtful that we *could* accept the conclusion that we don’t follow rules or act on our understanding. For doubts about Kripke’s own attempt to live with this conclusion (i.e., his “skeptical solution”), see McDowell 1984, Goldfarb 1985, and Stroud 1990.
ordinarily seen as the best sort of evidence that the account is failed attempt at elucidation or analysis—faced with such a demonstration, we simply conclude that the account is incoherent, confused. We would surely not be inclined to acquiesce unquestioningly if the account’s proponent, when confronted with the demonstration, were to respond: “This result merely increases the interest of my work. My account shows that our belief in the existence of the phenomenon in question was all along an illusion.”

This point is reinforced by reflecting on the particular nature of the objection to the ‘guiding’ conception we have registered. As we have noted, the conception purports to be an account of the nature of acting on an understanding. It purports to answer the question, “What is it to act on an understanding of a rule? What does acting on an understanding consist in?” But one way of putting the gist of our objection to the ‘guiding’ conception is that as soon as we attempt to think through the terms in which the conception is couched, we see that it doesn’t really answer this question. It merely begs it. The conception maintains that acting on an understanding is a matter of being guided by one’s understanding. That seems an illuminating thought only in advance of the realization that being guided by something, in the pertinent sense, involves acting on one’s understanding of what that conveys to one. That is what generated the regress. And that is what also justifies concluding that the conception provides no account or explanation of acting on an understanding at all. It offers at best a pseudo-explanation, one whose surface appeal lies in its tacitly presupposing precisely what it sets out to explain.

So the problem with the conception is not merely that it imposes a set of requirements that nothing in the world could possibly meet. (Such is how, for example, an error theorist like Mackie regards our ordinary conception of value.) The problem is that the terms in
which the conception is couched are circular, and so that the conception is, in a real sense, empty. I do not claim it is impossible for a ‘conception’ of this sort to ever find a way into our thinking and in so doing infect at least some range of our thoughts. But it would surely be remarkable if such an evidently central element in our everyday thought about the actions of ourselves and of other people should turn out to be empty in this way. Surely it is more plausible to source the circularity to a philosophical imposition upon our everyday thought.

These observations perhaps do not prove that Kripke’s answer to our question is erroneous. But it is at least clear that we need to be given strong positive reasons for opting for that answer over the alternative. We need to be told why we should assent to Kripke’s view, given that the view both has a disturbing and incredible consequence and depends upon a dialectical maneuver (the separation of the issue of spuriousness from the issue of sustainability) that is, especially in this case, highly suspect.

Whatever such an argument might involve, it is not to be found in Kripke. In Kripke’s exposition of his skeptical puzzle, the ‘guiding’ conception appears more or less from out of the blue. Little effort is made to justify it; it is simply presented as intuitively compelling. When it becomes clear that nothing could possibly conform to the dictates of the conception, one might be forgiven for expecting Kripke to reassess our initial intuitive attraction to the conception: after all, intuition is, epistemically speaking, a rather frail reed. But no such reassessment is forthcoming.²⁰

Here, by contrast, we have tried to provide a more substantive motivation for the

---

²⁰ One explanation for Kripke’s failure to undertake this reassessment may be that, as I noted at the outset, he never clearly distinguishes issues about the ‘guiding’ conception from his official skeptical puzzle about the normative reach of states of meaning and understanding. My ascription to him of an appreciation of the regress objection to the ‘guiding’ conception involves some idealization: it is a matter of extracting and clarifying a strand of his discussion that he himself does not clearly earmark. This may help explain why the need for a substantive defense of commitment to the conditional correctness of the ‘guiding’ conception never comes into focus.
'guiding' conception. We have tried to show, as methodically as possible, how one may be led to that conception via a series of steps that begins with the uncontroversial thought that rule-following is a manifestation of sapience. But it misconceives the intellectual exercise in which we have been thus engaged to suppose that it yields anything like the sort of argument that we now see Kripke’s view to require. Our aim was to provide a basis for Nagel’s intuition that Wittgenstein’s thin characterizations of rule-following are incompatible with an acknowledgment of the sapient character of that activity. None of the steps of the resultant train of thought were deductive or anything close to it; at each step, the claim was merely that the considerations on the table suggested the next step—and this always with one eye toward where we wanted to end up, namely, with a conception of acting on an understanding that could underwrite Nagel’s intuition. When it turns out that the upshot of this exercise is a conception that is incoherent, the proper conclusion is that our attempt to underwrite Nagel’s intuition has failed, not that we have established a radical new form of skepticism.

6. The appeal to practice

Where, then, did the train of thought go wrong? Certainly not at the beginning—we do not want to give up the idea that rule-following is an exercise of sapience. Nor, I think, should we be eager to abandon the linkage of sapience to rationality, or the set of claims about rational action (i.e., acting for a reason) laid out in section 2. But there is a better candidate for the location of where we went off the rails. In section 3, we considered the view that that the fact that one is following a rule consists in the fact that in acting in accord with the rule, one is doing what comes naturally to one. We saw that this view (along with its close cousin, dispositionalism) is unable to accommodate the considerations about rational action discussed
in section 2: we cannot hold onto the thought that you are acting on an understanding, and so acting for a reason, if the sole fact we have in view about your behavior is that you are, say, reacting to my order, “Add two,” in the way that comes naturally to you. We noted at the beginning of section 4 that this objection to the view was consistent with two further diagnoses of where it went wrong. The first possible diagnosis finds an inconsistency in simultaneously viewing an agent as acting on an understanding and doing what comes naturally. According to the second diagnosis, by contrast, acting on an understanding can involve just doing what comes naturally, say, in light of one’s training. It is just that doing what comes naturally cannot be all that there is to acting on an understanding, and so cannot be what acting on an understanding consists in. At the time, we endorsed the first diagnosis, framing our perception of an inconsistency between the two characterizations of behavior in terms of the idea that acting on an understanding involves a distinctive mental process, a mental process that is not in play when a creature does what comes naturally to it. This thought ultimately led to the ‘guiding’ conception. The failure of that conception thus suggests that we opted for the wrong diagnosis.

Now, the second diagnosis itself leaves open a wide range of possibilities. One might take it as simply calling for a more sophisticated naturalistic account of understanding than those broached in section 2. I will not discuss the prospects for such an account here.  

21 Here is one possibility. Dispositionalism, in the form Kripke considers it, depends upon the logical behaviorist idea that a person’s body of propositional attitudes and other mental states can be reduced to an assemblage of independently specifiable dispositions. This idea was long ago superseded in mainstream philosophy of mind by functionalism, which views a person’s body of mental states as essentially a system, composed of items defined by interrelated causal roles, hence specifiable only collectively and then only in light of the person’s overall set of possible perceptual inputs and behavioral outputs. This is a superior proposal for a number of long-recognized reasons, and one might hope it will solve our current problem as well. One might hope to accommodate acting on an understanding in our naturalistic picture of the mind by holding that when a creature does what comes naturally in such cases, the physical dispositions thereby activated are necessarily part of a complex, interrelated set of physical dispositions.

However, there are good reasons for thinking that, as a solution to the specific dilemma raised here, the
limited space that remains, I want rather to briefly look at Wittgenstein’s own and very
different way of developing (what I am calling) the second diagnosis. His basic thought here
is encapsulated in the following remark: “What, in a complicated surrounding, we call
‘following a rule’ we should certainly not call that if it stood in isolation” (Wittgenstein 1978:
§VI-34). Abstracted from the complicated surroundings in which it takes place, your doing
what comes naturally in response to my order, “Add two,” is not readily conceived as your
following a rule. But seen in the context of these surroundings, which is how we see it in
everyday life, no such difficulty arises. With that context in view, we can simply take in
stride that you react to my order in the way that comes naturally to you in light of your
training. Registering that fact will have no tendency, in and of itself, to undermine our sense
of the sapient character of what you do. Your doing what comes naturally, in these
circumstances, given these surroundings, is something we all recognize as a case of following
a rule.

For Wittgenstein, the relevant “complicated surroundings” encompass a great deal
indeed. They include no less than the “whole hurly-burly” of human activity, in all its
variegation, richness and complexity:

How could human behavior be described? Surely only by sketching the actions of
a variety of humans, as they are all mixed up together. What determines our
judgment, our concepts and reactions, is not what one man is doing now, an
individual action, but the whole hurly-burly of human actions, the background
against which we see any action (Wittgenstein 1967: §567).
In discussing rule-following and related phenomena, Wittgenstein is especially prone to describe this background of human activity in terms of talk of “practices”, “customs” and “institutions”. He does not invest these terms with any special or technical meaning. Indeed, it would subvert his strategy to do so, for he wants it to be a wholly uncontroversial claim that each of us participate in numerous customs, practices and institutions. We find a relatively straightforward illustration of what Wittgenstein aims to accomplish with his appeal to customs and practices in the following much-discussed passage:

“But how can a rule show me what I have to do at this point? Whatever I do is, on some interpretation, in accord with the rule.”—That is not what we ought to say, but rather: any interpretation still hangs in the air along with what it interprets, and cannot give it any support. Interpretations by themselves do not determine meaning.

Let me ask this: what has the expression of a rule—say a sign-post—got to do with my actions? What sort of connection is there here?—Well, perhaps this one: I have been trained to react to this sign in a particular way, and now I do so react to it.

“But that is only to give a causal connection; to tell how it has come about that we now go by the sign-post; not what this going-by-the-sign really consists in.” On the contrary; I have further indicated that a person goes by a sign-post only in so far as there exists a regular use of sign-posts, a custom (§198).

The first paragraph raises the regress problem for a version of the ‘guiding’ conception, and extracts the following moral: we must reject the idea that the relationship between the expression of a rule and what I do in response to it is mediated by a mental item (an “interpretation”) guiding my performance (“showing me what I have to do”). In the wake of this rejection, Wittgenstein proposes an alternative account of the relationship: I was trained to react in a certain way when confronted with the expression of a rule, and that is how I now react. But the interlocutor is baffled by this suggestion. The point of the ‘guiding’ conception is to underwrite the thought that what is going on here is precisely a case of going by a sign—

---

22 This passage is discussed by McDowell in McDowell 1984. Following his practice, I have put quotes around the first sentence of the final paragraph to indicate the interlocutor’s voice. My treatment of this passage is indebted to McDowell’s in other ways as well.
which is to say, a case of following a rule—rather than a brute, non-sapient reaction. The contrasting picture of the relationship between expression and action drawn by Wittgenstein, in which I am portrayed as just reacting as I was trained, seems to the interlocutor to fail on this score. But Wittgenstein takes this reply to miss his point (which, in fairness to the interlocutor, has not at this stage in the passage been made very explicit). When we train a child to react in a certain way to, say, a walk sign, we introduce the child to a going practice in the use of those signs. This is a practice of going by or following the sign: of walking when the sign says “walk” and of not walking when it says “don’t walk”. To recognize that a person has been initiated into our practice in the use of walk signs is thereby to recognize that what that person does now when confronted by such a sign counts as a case of her going by the sign, and so of her acting on an understanding of a rule, and so of her acting for a reason. In referencing our training, Wittgenstein seeks to remind us that we have been initiated into such practices, practices of following rules and as such, practices whose exemplifications in our actions thus manifest our sapience.

We seemed to need the idea that our performance is guided by our understanding because the alternative—that we just act, without guidance—seemed to us incompatible with the idea that we act on our understanding, and thus for a reason. I take Wittgenstein’s invocation of custom and practice in these and related passages to be a strategy for getting us to question this perception. Practices and customs are usefully pressed to this service because they have two salient features: 1) once a person has been successfully trained into a given practice, it will often just come naturally to her to do whatever conforms to the practice on the appropriate occasions—it will just be a matter of her reacting as she was trained; and 2) many practices are rule-following practices, practices of following or going by particular rules,
signs and the like. These facts about practices are facts that everyone knows; they are truisms. But taken together, they entail that it is a mistake to think that doing what comes naturally and following a rule cannot comport. For most adults, participation in the practice of walking only when the sign says “walk” does not require any deliberation, reflection or internal guidance; one just reacts in the pertinent situation as one was trained. But those who participate in this practice are following certain rules; these just are practices of following rules. Thus reminding ourselves that people who follow rules are in so doing enacting customs and practices ought to undercut the intuition expressed by Nagel. We might put Wittgenstein’s suggested therapy this way: if you find yourself doubting that any given tract of behavior can satisfy both the descriptions, “following a rule” and “doing what comes naturally”, then remind yourself that the behavior may well satisfy a third description: “instantiating a practice”.

I do not pretend that it is anything but extremely difficult to get a clear view of the character and content of the kind of move at which I have just gestured. The hardest thing to grasp is how slight the move really is, and yet how even so it can accomplish its intended purpose. This paper is already very long, and I cannot satisfactorily address all the issues that arise here. I will close by very briefly mentioning two of them.

First, there is little chance that we will find the appeal to custom even remotely convincing if we construe it as an attempt at a self-standing, knockdown rejoinder to, say, the intuition expressed by Nagel. It is crucial to see the appeal as supplemental to the other considerations, not least of which are those I have tried to embody in the dialectic discussed here, in which the intuition is traced to a conception that proves incoherent. That argument is the primary case for denying the conception; the appeal to custom is one strategy for
becoming comfortable with that denial.

Second, if Wittgenstein’s appeal has the character I have just suggested, it ought not to imply any strong externalist claims about the necessity of a sapient creature’s participation in communal practices. If the aim is just to correct our impression that following a rule cannot be a matter of doing what comes naturally by pointing out that actions that exemplify practices are readily conceived as having both features, there is no need, and no basis, for insisting that all instances of rule-following must exemplify practices (an insistence those who adopt a “community interpretation” of Wittgenstein tend to ascribe to him). This is all to the good, from the perspective of our search for a quietist reading of Wittgenstein’s remarks: Wittgenstein read quietistically is not in the business of offering sweeping claims about metaphysically necessary conditions for a creature to be in a position to follow rules. But a fuller treatment of this topic would then have to address the fact that Wittgenstein does sometimes appear to be gesturing in this direction. Such might seem the case with the final sentence of the passage just quoted. And all the more so in a passage like the following:

“But it is just the queer thing about intention, about the mental process, that the existence of a custom, of a technique, is not necessary to it. That, for example, it is imaginable that two people should play chess in a world in which otherwise no games existed; and even that they should begin a game of chess—and then be interrupted.”

But isn’t chess defined by its rules? And how are these rules present in the mind of the person who is intending to play chess? (§205)

Obviously much of this passage is congenial to the present interpretation: it confirms that one of Wittgenstein’s primary stalking-horses is a picture of rule-following as involving a queer kind of mental process, and that he takes this picture to contrast with the view of rule-following one gets by thinking in terms of customs and practices of following rules. But it is not congenial to the present interpretation to see Wittgenstein as taking himself to be justified in flat-out asserting that customs and practices are necessary conditions for rule-following. In
fact, I don’t think we need to read this assertion into this passage: in taking issue with the interlocutor, Wittgenstein needn’t be interpreted as embracing the other pole of the dichotomy that the interlocutor articulates—he can be interpreted rather as questioning the aptness of this dichotomy. But I will not try to defend this claim here.

7. Conclusion

Let us recap. We began by noting a striking feature of Wittgenstein’s discussion of rule-following and understanding, one highlighted in recent work by Thomas Nagel: namely, Wittgenstein’s tendency to characterize instances of rule-following in ways that seem, on the face of it, to fit poorly with the idea that following a rule involves acting on one’s understanding of the rule. The idea that rule-following is a sapient activity, an activity that engages the understanding, is surely part of anything that deserves to be labeled our “ordinary” conception of rule-following. If Wittgenstein’s characterizations are genuinely incompatible with this idea, then, there seems no hope of finding him to be faithful in practice to the “quietism” forcefully expressed by many of his programmatic remarks. This situation prompted us to try to get to the roots of our perception of an incompatibility. Where, exactly, is the conflict supposed to lie?

Noting the lack of light shed on this question by treatments of quietism, like McDowell’s, that take their cue from Kripke’s puzzle about the “normative reach” of intentional states, we were led to approach it from a different angle, taking seriously the appearance that the question is one for the philosophy of action. Acting on one’s understanding of a rule is a species of rational action, of acting for a reason. On a plausible view of rational action, this implies that acting on one’s understanding of a rule involves
taking one’s understanding of the rule as a reason, in the normative sense, for what one goes on to do. One implication of this result is that we cannot reduce the fact that a person is following a certain rule to the fact that, in producing the relevant tract of behavior, the person is doing what comes naturally to her. That might anyway seem too attenuated a proposal to take seriously, but as we noted, it is closely related to the dispositionalist view of understanding that Kripke takes as one of his primary targets.

Establishing that the reductive proposal is unacceptable falls short of what we need, however. It is one thing to show that acting on an understanding of a rule cannot consist in its coming naturally to one to produce behavior in accord with that rule. It is another to show that these are mutually exclusive categories, that no case of a person’s acting on her understanding of a rule can be a case of her doing what comes naturally to her. But the forgoing reflections do seem to point in the direction of a basis for this stronger claim. For they can seem to suggest that acting on an understanding of a rule necessarily involves a special kind of mental process, one whose role in producing the action rules out that the agent is just reacting in a way that comes naturally. This process needn’t involve full-dress deliberation on the part of the agent; what seems crucial is rather that the agent’s understanding be understood as guiding her behavior, in the sense of showing her a way to proceed, of instructing or informing her about what she is to do.

But this notion of the special mental process turns out to be impossible to sustain, for picturing the relationship between understanding and behavior as one of informed guidance leads to an insoluble predicament. The fact that you must understand an item in a particular way, and then act on that understanding, in order to be guided by the item appears to generate an endless regress of higher-order understandings, each of which can be acted on only by
acting on the next one up the line. One might hope to prevent the regress by conceiving a state of understanding as an item that ‘logically’ compels understanding of its guidance. But we found that there is nothing intelligible to be made of this proposal.

Nor was there any sensible basis for insisting with Kripke that, despite its incoherence, the ‘guiding’ conception should be regarded as the correct account of what acting on an understanding, were it possible, would have to be. The appropriate conclusion is rather that the ‘guiding’ conception is a mistake, and that the line of thought that led to it has gone astray. In particular, we should reconsider the transition from the considerations about rational action to the idea of the special mental process whose presence is inconsistent with the agent’s doing what comes naturally to her. We saw that Wittgenstein’s appeal to customs and practices of rule-following are plausibly read as a strategy for enabling us to resist this transition.

Obviously, the foregoing does not amount to a proof that the intuition expressed by Nagel about Wittgenstein’s thin characterizations of rule-following is incorrect. In particular, it does not foreclose on the possibility that we may arrive at some other conception of acting on an understanding, one wholly distinct from the ‘guiding’ conception, that both attracts our assent and ratifies the intuition. I do not think we can definitively rule out this possibility. But nor does this seem to me a burden of proof that must be met. That the present dialectic captures what really underlay that intuition in our thinking, and consequently exposes that intuition as mistaken, is supported by two circumstantial but telling pieces of evidence that we have already in effect registered. First, the central elements of the dialectic can plausibly be taken to correspond to important moments in Wittgenstein’s own discussion in the *Investigations*. Second, the conception that emerges first as the culmination and then as the nemesis of the dialectic, the ‘guiding’ conception of acting on an understanding, figures
centrally in the thought of both Kripke and his commentators—albeit with a sufficient lack of clarity that its role has not been well understood.

I believe this the failure of the ‘guiding’ conception, as well the various considerations registered in the course of reaching this conclusion, have important implications for current debates in the philosophy of action and the philosophy of mind. But this matter must be left for another occasion.
REFERENCES


