

## Rule-Following Skepticism, Properly So Called

In Conant and Kern (eds.), *Varieties of Skepticism*, DeGruyter, 2014.

Jason Bridges  
University of Chicago

Saul Kripke famously argued that there are no facts about what people mean by the words and sentences they utter. His ‘skeptical paradox’, as he called it, has generated an enormous secondary literature. Indeed, it’s not an exaggeration to say that it gave birth to a new subfield of the philosophy of mind and language. Despite the degree of attention, however, I believe that the real source and character of Kripke’s skeptical doubts have never come into clear view. Previous commentary on Kripke has failed to grasp that the doubts fundamentally concern the possibility of a person’s *following* a rule. That might seem an absurd charge, given that everyone knows that Kripke conceives the skeptical paradox as a development of Wittgenstein’s remarks on rule following. But the point of the charge is that previous interpretations of the paradox go astray because they miss that its central target is the idea that person’s performance might be *based upon* a rule—that a rule might be *her reason* for proceeding as she does. The skeptical paradox belongs to the philosophy of rational explanation, of explanations that account for what people do or think by citing their reasons for doing or thinking so.

That this feature of the paradox has been overlooked is partly Kripke’s fault. His exposition of the skeptical argument intertwines two strands without adequately distinguishing them. The secondary literature has picked up on one of these strands. But the material constituting this strand cannot make sense of crucial steps in the discussion; thus commentators are led to find disappointingly large gaps in the skeptical line of thought. After an analysis of this exegetical state of affairs in the first few sections of this paper, I shall focus on developing the other strand.

The exercise will have a number of positive results: it will yield a more satisfying understanding of Kripke's text, it will uncover a skeptical argument of interest and depth, and it will allow for a richer engagement between the skeptical argument and Wittgenstein's own discussions of rule following.

## 1. The prescriptive and evaluative interpretations of Kripke's normativity thesis

Kripke's skeptical argument, on any reading, hinges on the idea that the relationship between the meaning of an expression and the use of that expression is "normative". The standard interpretations of Kripke's talk of the normativity of meaning take their primary cue from the following passage:

Suppose I do mean addition by '+'. What is the relation of this supposition to the question of how I will respond to the problem '68+57'? The dispositionalist gives a *descriptive* account of this relation: if '+' meant addition, then I will answer '125'. But this is not the proper account of the relation, which is *normative*, not descriptive. The point is *not* that, if I meant addition by '+', I *will* answer '125', but that, if I intend to accord with my past meaning of '+', I *should* answer '125'. Computational error, finiteness of my capacity, and other disturbing factors may lead me not to be disposed to respond as I *should*, but if so, I have not acted in accordance with my intentions (Kripke, 1982, p. 37).<sup>1</sup>

Some commentators interpret this passage as proposing the following thesis: what a word means has implications for how a speaker of the language ought to use the word. Allan Gibbard, for example, writes:

The claim I draw from the passage is this: If I mean something by a term, this has implications for what I should do when I use the term, or answer questions couched using the term. The 'should' here is normative: What I mean determines what I *ought* to do. The crux of the slogan that meaning is normative, then, might be another slogan: that *means* implies *ought*. To use...Kripke's example, from statements saying what I mean by the plus sign and other arithmetical terms and constructions, it will follow that I *ought* to answer "125" when asked, "What's

---

<sup>1</sup> Henceforth, this text will be cited by page number only.

68+57?” Not that I necessarily *will* answer “125”, but that is what I ought to answer” (Gibbard, 1994, p. 100).<sup>2</sup>

On Gibbard’s reading, then, Kripke’s claim is that the meanings of my words determine how I ought to use them. Call the view that the meaning of a word determines how one ought to use it *the prescriptive view of meaning*, and call the claim that this view captures what Kripke intends in speaking of the normativity of meaning the *prescriptive interpretation*. Gibbard applauds the prescriptive view of meaning, and suggests it opens the attractive possibility that a word’s having the meaning that it does will prove reducible to facts about how the words ought to be used. Other commentators, such as Paul Horwich (1998, Ch. 8), are less sympathetic, but agree that Kripke must have something along these lines in mind.

But there seems to me very good reason to reject Gibbard’s reading of the passage. *Pace* Gibbard, Kripke does not say that, given what I mean or have meant by “+”, I ought to answer “125”. What he says, rather, is that if I *intend* to accord with what I have meant by “+”, I ought to answer “125”. This distinction is not pedantic: it marks the difference between a hypothetical and categorical imperative, between an “ought” claim whose holding of a person depends on her possession of certain aims, intentions or desires, and an “ought” claim whose holding of a person is not so dependent. Gibbard ascribes to Kripke the principle, “What I mean determines what I *ought* to do.” If Kripke had said or implied that facts about meaning give rise to categorical imperatives, the ascription would be apt. Given that Kripke just asserts a hypothetical imperative, however, the principle fits the text only if understood as shorthand for the following: “Facts about what I mean determine what I ought to do contingent on my possession of particular aims, desires or intentions.” This is hardly a noteworthy thesis, for facts about *anything* can determine what I ought to do contingent on my possession of particular aims, desires or intentions. If I

---

<sup>2</sup> Gibbard actually uses the example of answering the question “What’s 5 + 2?” For ease of exposition I have substituted Kripke’s example.

want clearer vision, I ought to wear corrective lenses. Obviously the possibility of this kind of remark should not motivate a thesis of the normativity of optics.<sup>3</sup>

If the prescriptive view of meaning is glossed in terms of hypothetical imperatives, it is vacuous; and if it is glossed in terms of categorical imperatives, the prescriptive interpretation has no textual support in the passage under consideration. A charitable reading of Kripke should find the point of his distinction between the normative and descriptive to lie elsewhere.

Such a reading suggests itself when we note that “ought” and “should” are not the only terms in the passage that might reasonably be regarded as normative; “accord” is another. Consider again Kripke’s hypothetical imperative: “if I intend to accord with what I have meant by ‘+’, I should answer ‘125’.” Compare the following remark: “If you intend to accord with the rules of the Charleston, at this juncture you should step before you kick.” This remark obviously doesn’t presuppose that facts about the rules of the Charleston have implications for what you ought to do now irrespective of your particular aims and intentions. What it does seem to imply, however, is that in light of its rules, there is such a thing as a *correct* way (and such a thing as an *incorrect* way) to dance the Charleston. It seems open to read Kripke’s hypothetical imperative simply as a way of making the analogous point about meaning. The availability of that particular imperative shows that in light of the meaning of a linguistic expression, there are right and wrong ways to use that expression. The meaning of an expression, like the rules of a dance, provides a standard, or norm, against which performance is determined as correct or incorrect. In this sense, the relationship between meaning and use is normative.

---

<sup>3</sup> Oddly, many commentators miss the hypothetical character of Kripke’s imperatives. The most striking version of this oversight is in Boghossian (2003). Boghossian’s quotation of the passage (Boghossian, 2003, p. 31) stops before the final mention of intentions, and, more remarkably, omits the phrase, “if I intend to accord with my past meaning of ‘+’”, with no ellipsis to indicate the omission. This omission sets the stage for Boghossian to write later in the paper, “Kripke says: ‘If I mean addition by ‘+’, then it doesn’t follow that I will say that ‘68 + 57 = 125’, but only that I ought to say that it does’” (p. 39), and to proceed to criticize the second part of this claim on the ground that whether one ought to give the correct answer will be contingent on one’s intentions.

Call the claim that the meaning of an expression determines uses of that expression as correct or incorrect *the evaluative view of meaning*, and call the claim that this view captures what Kripke means in speaking of the normativity of meaning *the evaluative interpretation*. The evaluative interpretation has been endorsed by a few commentators, including Paul Boghossian (1989, but see also 2003), John McDowell (1998) and Hannah Ginsborg (forthcoming). In addition to better fitting the text, the evaluative interpretation has a further advantage: unlike the categorical version of the prescriptive view, the evaluative view of meaning seems plausible on its face. At the very least, it can be said to have the support of commonsense, in the sense that the great majority of language speakers would immediately assent to it. Nearly everyone takes it as obvious that words can be used correctly and incorrectly, and that the meaning of a word determines which is which.<sup>4</sup>

Some philosophers appear to believe that an item can sort behavior into correct or incorrect only if it has categorical implications for what one ought to do.<sup>5</sup> This would collapse the distinction I have drawn between the evaluative and prescriptive views. But the belief is simply wrong. I might take an hour right now and invent a new dance. If I do so, I will have brought

---

<sup>4</sup> Some lexicographers and social commentators appear to set themselves against commonsense here, for they appear to regard the evaluative view of meaning as an expression of elitism. We have long accepted that there is no one “correct way” to dress or otherwise comport oneself; why should we persist in insisting, in the face of the great diversity of dialects, that there is a “correct way” to speak English—say, the way people educated at high-end colleges do? (For citations of this line of reasoning, see Wallace (2001).) But as an objection to the evaluative view, this confuses two dimensions along which linguistic performances might be evaluated for correctness: whether a word is used with its proper meaning, and whether a word is used correctly in light of its meaning—that is, in light of whatever meaning it has on the speaker’s lips, be it proper or improper. The meaning it has on your lips may be special to the particular dialect you speak; it may even be special to your idiolect. But if we are genuinely talking about meaning, then it will determine correct and incorrect ways for you to use that word. Two children may decide to communicate in a secret code where, say, “rain” means snow and “snow” means rain. If one of them on some occasion forgets that that is what they are doing and uses “rain” when intending to speak of rain, then insofar as it is still true to say that in this context the word “rain” on her lips means snow, it is also true to say that on this occasion she is using the word incorrectly in light of that meaning. Now, perhaps one might want to claim that the children’s code assigns the ‘incorrect’ or ‘improper’ meaning to the word “rain”. But that is a different matter. And it is only the latter sort of claim that, at least in some contexts, may merit the charge of elitism.

<sup>5</sup> This view appears to be common ground between Gibbard and Boghossian in their debate in Gibbard (2003) and Boghossian (2003). However, I have difficulty finding a consistent interpretation of Boghossian’s essay.

into being a new dimension along which human actions can be judged as correct or incorrect: namely, as performances of that dance. I will have, in that perfectly legitimate sense, introduced a new norm for judging potential human performances correct or incorrect. But the mere fact that the rules of my dance are norms in this sense obviously does not give them any categorical power, not even a *prima facie* power, to dictate what anyone ought to do. We should count ourselves lucky that there is no implication from an item's having the power to sort our performances into correct and incorrect to its having the power to determine, categorically, what we ought to do. As this example shows, the former power can be had very cheaply.

The evaluative interpretation yields an appealingly straightforward explanation of Kripke's dissatisfaction with reductive accounts of meaning. On this interpretation, the gist of Kripke's objection to dispositionalism, for example, would be this: that while one's behaving in a certain way can be inconsistent with the hypothesis that one possesses a certain disposition, there is no sense in saying one's behavior fails to accord with that disposition. Dispositions do not sort behavior into correct and incorrect. But meanings do. Thus one's meaning something by an expression cannot be identified simply with one's possessing a particular disposition in the use of that expression. Analogous objections can easily be constructed for the other reductive accounts Kripke considers. Whatever the ultimate merits of these objections, they certainly seem to have some *prima facie* force.

## **2. The evaluative interpretation and anti-reductionism**

I do not wish to deny that the evaluative view of meaning is at work in Kripke's discussion of his skeptical paradox. On the contrary, I think it clear that he objects to dispositionalism and the other reductive accounts of meaning at least partly on the ground that they cannot

accommodate the evaluative view. But if we suppose, à la the evaluative interpretation, that the evaluative view exhausts the import of Kripke's appeals to the "normative" relationship between meaning and use, we will be unable to make sense of his dissatisfaction with *anti-reductionism* about meaning. After disposing of the various reductive proposals he considers, Kripke concludes the skeptical argument with a brief consideration of the possibility that states of meaning things by one's words are "*sui generis*" and "primitive", and hence that semantic facts cannot be reduced to facts specifiable in non-semantic terms. Kripke's response is simply to complain that the proposal "leaves the nature of this postulated primitive state...completely mysterious" (p. 51). Why does Kripke think this?

There are two possibilities. The first is that Kripke assumes some version of semantic naturalism—the view that all facts about semantic content must be reducible to facts expressible in non-intentional, non-semantic terms.<sup>6</sup> This view sets a clear research agenda for the philosophy of mind: to explain possession of semantic content and correlative states and properties in non-intentional, non-semantic terms. And it licenses suspicion about the existence of a putative kind of content-involving state if, in light of repeated failed efforts to construct a satisfactory reductive account of what it would be to occupy such a state, the prospects for such an account come to seem dim. If Kripke's discussion proceeded from an assumption of semantic naturalism, then it would be quite clear why, after a wholly unsuccessful series of attempts to produce a reductive account of states of meaning things by words, he should be impatient with the rearguard insistence that we do not after all need such an account to vindicate our belief in the existence of such states.

But the claim that semantic naturalism is a presupposition of Kripke's skeptical argument is extremely dubious. For one thing, Kripke never explicitly registers a commitment to semantic

---

<sup>6</sup> For an example of this interpretation, see Horwich (1984, p. 164).

naturalism in his presentation of the skeptical paradox. Indeed, he gives no hint at all that such an assumption plays any role in motivating the skeptical conclusion. Moreover, one might well wonder where the assumption came from. Kripke himself is no friend of semantic naturalism or the constructive philosophical projects it licenses.<sup>7</sup> And surely he is not such a poor interpreter as to suppose that Wittgenstein is committed to semantic naturalism either. But if neither Kripke nor Wittgenstein is a semantic naturalist, how did it come about that Kripkgenstein is such a knee-jerk one?<sup>8</sup>

We are left with the second possibility. This is that Kripke takes himself to have brought to our attention, in the course of his exposition of the skeptical argument, certain seemingly problematic or paradoxical features of the envisioned states of meaning, features so seemingly problematic, in fact, that we cannot justifiably countenance the existence of anything possessing those features until we have produced a satisfactory philosophical account of what it would be for something to possess them. If Kripke were to believe he had shown this, then it would be quite understandable why he should feel justified in dismissing the claim that states of meaning are *sui generis* and hence not to be subject to analysis or explanation. When a compelling doubt has been raised about the very intelligibility of a supposed property or state of affairs, it is no defense to style it as *sui generis*. If we simply cannot make sense of the imagined state of affairs or property, the insistence on irreducibility is a smoke screen.

---

<sup>7</sup> That Kripke (1980) served as a foundational text in the development of various ‘naturalistic’ accounts of reference and content seems to me to show that one of the central lessons of that book was not absorbed.

<sup>8</sup> An analogous point applies to Robert Brandom’s claim that the skeptical argument presupposes what we might call normative naturalism, according to which all normative facts must be reducible to facts specifiable in non-normative terms. In Brandom’s words, “Kripke’s skeptic implicitly restricts the vocabulary used to specify [past] use to nonnormative vocabulary. He thus assumes that normative specifications of proprieties of concept use are in principle intelligible only if they can be reduced without remainder to specifications of nonnormative properties of concept use” (Brandom, 2001, p. 605). But the ascription of normative naturalism to Kripke’s skeptic is just as unsatisfying, and for just the same reasons, as the ascription of semantic naturalism. (People who like bad puns—not me, certainly—will be tempted to call this version of Kripke’s skeptical persona Brankenstein. For his argument, if perhaps not quite monstrous, is nonetheless a pitiable creature.)

What is the problematic feature of states of meaning that Kripke thinks he has brought to our attention? Evidently it has something to do with normativity. For the claim that meaning is normative is the only positive point about meaning that Kripke makes.<sup>9</sup> Similarly, that an account of states of meaning must provide for the normativity of meaning is the only substantive condition that Kripke explicitly imposes on solutions to his skeptical paradox.<sup>10</sup> And so if Kripke believes he has revealed something paradoxical in the very idea of a state of meaning, he must take it to be an implication of the idea that states of meaning are normative.

But if there is nothing more behind Kripke's talk of the normativity of meaning than the thesis I labeled the evaluative view, the obvious question is why Kripke should think he has identified an especially problematic feature of meaning. Is the power to sort performances into correct and incorrect really so paradoxical or mysterious? What's the mystery? What's the paradox? Kripke does not say. It is easy to sympathize with Crispin Wright's initial reaction to Kripke's treatment of the anti-reductionist proposal. Wright (1984, pp. 775-777) objects that it is our ordinary concepts of contentful mental states that ascribe normativity and correlative properties to them, and that we need to be given some substantive reason, which Kripke does not provide, for supposing that our ordinary concepts are not in order as they stand. Philosophers who share Wright's sense of the flimsiness of Kripke's case against anti-reductionism include Boghossian (1989) and Donald Davidson (1992).

It would be disappointing if Kripke's celebrated skeptical argument turned out to rest on an unmotivated suspicion of the very idea of a state that sorts behavior into correct and incorrect.

But I do not think it does. As we will shortly see, the real culprit here is the going interpretation

---

<sup>9</sup> It is true that Kripke objects to specific accounts of meaning by appealing to further purported facts about meaning, in particular, that the meaning of an expression determines its correct use in an infinite number of possible circumstances, and that people are disposed to use their expressions incorrectly some of the time. But Kripke takes these objections to be applications of the basic normativity objection. See pp. 24, 37, 52fn.

<sup>10</sup> p. 11. Note that here Kripke states the normativity condition in terms of talk of justification; this is a crucial point to which I will shortly return.

of the claim that meaning is normative. A central element of Kripke's treatment of normativity, and the real basis for his skepticism, is quite different from anything we have yet considered.

### 3. The evaluative interpretation and the justification requirement

Early on in his discussion Kripke announces an essential condition on a satisfactory answer to his skeptical paradox:

An answer to the sceptic must satisfy two conditions. First, it must give an account of what fact it is (about my mental state) that constitutes my meaning plus, not quus. But further, there is a condition that any putative candidate for such a fact must satisfy. It must, in some sense, show how I am justified in giving the answer '125' to '68+57' (p. 11).

Here Kripke says that a satisfactory response to the skeptic must show how it can be the case that a particular use of a linguistic expression (in this case, a response to a question involving that expression) is *justified*. This requirement, which I will call the *justification requirement*, Kripke presents as a condition that must be satisfied by any putative "account of what fact it is (about my mental state) that constitutes my meaning plus". Given that he states no other requirement on an "account of what fact it is (about my mental state) that constitutes my meaning plus", we can infer that the justificatory role of meaning is the primary topic of the skeptical argument. Elsewhere he indicates that it is this justificatory role that he has in mind when he speaks of meaning as normative. For example, in the paragraph immediately following the passage about the "normative" relationship between use and meaning that was our focus in section 1, Kripke summarizes the thrust of that passage thusly: "Precisely the fact that our answer to the question of which function I meant is *justificatory* of my present response is ignored in the dispositional account and leads to all its difficulties" (p. 37, his emphasis).

According to the evaluative interpretation, Kripke demands of an account of meaning that it provide for the capacity of an expression's meaning to determine uses of that expression as correct or incorrect. Adherents to that interpretation presumably suppose that the justification requirement is just another way of formulating the same demand. And indeed, that "125" is the correct answer to the question "What is  $68+57$ ?" can, in certain contexts, be all that one means in saying that this answer is justified. In general, if we are engaging in a kind of activity that essentially involves certain standards of performance—doing addition, dancing the Charleston, using a language—then we might assert that a given response or performance is "justified" as a way of expressing our judgment that the response or performance accords with the relevant standards.

However, a closer look at Kripke's text reveals that the evaluative interpretation cannot in fact account for the particular significance of the skeptic's demands for "justification". Consider the following passages:

Ordinarily, I suppose that, in computing ' $68+57$ ' as I do, I do not simply make an unjustified leap in the dark. I follow directions I previously gave myself that uniquely determine that in this new instance I should say '125' (p. 10).

The sceptic argues that when I answered '125' to the problem ' $68+57$ ', my answer was an unjustified leap in the dark (p. 15).

How can I justify my present application of such a rule, when a sceptic could easily interpret it so as to yield any of an indefinite number of other results? It seems that my application of it is an unjustified stab in the dark. I apply the rule blindly (p. 17).

Here the issue raised by the skeptic is said to concern whether, when an ostensibly mathematically competent person computes what is ostensibly an addition problem, her answer is an "unjustified leap/stab in the dark". Now, suppose her answer was correct, and so justified in the sense of talk of what is "justified" mentioned above. Does it follow that her answer was not a stab in the dark? Certainly not: one can take a stab in the dark and happen to hit on the right

answer. It seems obvious, however, that in speaking of an “unjustified stab in the dark”, Kripke does not mean to point to a contrast between two kinds of stabs in the dark: those that are justified (in virtue of being correct) and those that are not. Rather, Kripke’s characterizing the answer as “unjustified” is meant to go hand in hand with his characterizing it as a “stab in the dark”: what makes the answer unjustified just *is* its being a stab (or leap) in the dark. And so it would not cease to be unjustified, in the relevant sense, if it happened to be correct.

It follows that the kind of justification provided for by the evaluative interpretation—according to which the meaning of ‘+’ determines answers as “justified” simply in virtue of determining them as correct or incorrect—is not the sort of justification with which the skeptic, as portrayed in these passages, is concerned. Thus the evaluative interpretation is incorrect. Its interpretation of Kripke’s talk of normativity cannot after all make sense of the requirement intended by Kripke when he says that answers to the skeptical argument must identify a fact that will “show how I am justified”.

#### **4. Rule following, justification and rational explanation**

So what does Kripke mean when he requires that a candidate for the fact that constitutes his meaning plus by “+” be such as to “show that he is justified in answering ‘125’ to ‘68+57’”? We can see what he has in mind by registering some simple points about the idea of rule following.

Before doing so, I should note that from this point forward, I shall be couching my discussion of the skeptical argument largely in terms of rules as opposed to meanings. Kripke himself moves freely between talk of meanings and talk of rules; he evidently regards the use of language as a rule-following activity, with the relevant rules constituted by the meanings of the linguistic expressions. It is because I doubt that Kripke is right about this, and because I

nonetheless see Kripke's skeptic as setting forth an important and interesting puzzle about rule following as such (regardless of whether the use of meaningful expressions counts as a kind of rule following), that I will key my subsequent exposition to rules rather than meanings.<sup>11</sup> I shall say a bit more about this at the end of this section.

Start with the undeniable point that following a rule is not just a matter of behaving in accord with the rule (cf. Wittgenstein, 1960, p. 13). 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 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 calculating the Fibonacci sequence—that by itself would not imply that I was following that rule. I may be doing something wholly unrelated to the rule—perhaps I'm recording business expenses, or just listing my favorite integers—and happen to notice the correspondence.

What more is required if my performance is to count as a case of my following the rule for calculating the Fibonacci sequence? At least the beginning of an answer seems obvious: I am following the rule only if the rule is *my reason* for doing what I do. This is to imply, of course, that there is an explanatory connection between the rule and what I do. But not just any explanatory connection: what is at stake when we speak of a person's reason for doing something is not merely a reason in the thin sense in which anything that explains something counts as a reason for it. If last night's heavy wind explains why the tree in my yard fell over, then the wind was the

---

<sup>11</sup> Proponents of the evaluative interpretation don't run into this difficulty. Meanings and rules are both undeniably normative in the sense identified by the evaluative view: they both sort performances into correct and incorrect. So there is no particular need to distinguish between them for purposes of engaging the text when it is understood in terms of the evaluative interpretation. What is questionable—and matters, on the interpretation of the skeptical argument I will defend—is whether a user of language follows meanings, in the way a rule-follower follows a rule.

reason the tree fell over. But it was not the *tree's* reason for falling over—the tree had no reasons. By contrast, if I am following the rule for the Fibonacci sequence, then the rule is *my* reason for proceeding as I do—to put it another way, it is the reason *for* or *upon* which I proceed. Rule following is an exercise of the human capacity of *rationality*, of the capacity to recognize and act (or form attitudes) for reasons.

Let us call an explanation of an action or attitude that cites a person's reasons for that attitude or action a *rational explanation*. One familiar and plausible thought about the content of rational explanations is this: when we say that a person S's reason for A-ing was B, we imply that S took B to be a reason for her to A, and that this fact about S explains her A-ing.<sup>12</sup> This analysis uses the word “reason” in what philosophers of action call its “normative” sense: a reason in the normative sense (or ‘normative reason’, for short) is a consideration that justifies, or at least counts toward justifying, that which it is a reason for.<sup>13</sup> Here nothing especially freighted or demanding is meant by talk of justification; the same point might be made by saying that a normative reason is a consideration that supports, recommends, or counts in favor of S's A-ing.<sup>14</sup> Putting these points together, if I say that Rachel's reason for moving to Toronto was the fact that her mother lives there, then I am saying that Rachel moved to Toronto because she took something to count in favor of her doing so, namely, that her mother lives there. More generally, the view at which we have arrived is that rational explanations explain actions and attitudes in terms of the subject's conception of what supports or counts in favor of them.

---

<sup>12</sup> For elaborations of this thought, see Grice (2005, pp. 40ff), Darwall (1983, pp. 32ff), and Stroud (1979), among many others.

<sup>13</sup> B may count toward justifying S's A-ing without flat-out justifying it if, for example, there are stronger reasons against S's A-ing. I will ignore this irrelevant complication in what follows.

<sup>14</sup> Contra Millar (2004, Ch. 2), who argues for restricting talk of “justifying” actions only to special cases. In other cases, he prefers to say, not that a reason for an action “justifies” that action, but that it “recommends” it (p. 59), or that it “represent[s] the action as being favoured in some way” (p. 58). Certainly no one would deny that there are distinctions to be drawn among kinds of reasons for actions. But I would argue that nothing Millar says on this score adequately motivates his fastidiousness about the use of “justification”.

Anyone familiar with the voluminous philosophical literature on rational explanation will know that there are a great many controversies about the view of rational explanation just briefly outlined. Does the idea that acting for a reason entails perceiving a normative reason for what one does unrealistically cast people as hyper-vigilant seekers after the good? What role, if any, does this view leave for desires in the rational explanation of actions? Exactly how does a person's taking B to be a reason for A-ing explain her A-ing: do we have here a case of a belief's causing behavior, or is the explanation rather best understood as non-causal, perhaps even as non-psychological? Is the involvement of normativity in a rational explanation limited to the beliefs about normative reasons such an explanation ascribes to the subject? Or is there a stronger sense in which rational explanations are normative, suggested perhaps by the Davidsonian thought that one can find such an explanation intelligible only to the extent that one shares, or at least finds reasonable, the beliefs about normative reasons ascribed by the explanation to the subject? When, as we say, S's reason for A-ing is B, is her reason best understood as the fact or proposition that B, as her belief that B, or as her belief that the fact that B counts in favor of her A-ing? And so on.

Fortunately, we don't need get into any of these issues here. The skeptical argument, as we will see, can be understood as a very general challenge to this conception of rational explanation, at least as that conception manifests in characterizations of people as following rule. Understanding this challenge will not require taking a stand on any of these matters.

Let's return to the justification requirement: Kripke's demand that our "account of what fact it is (about my mental state) that constitutes my meaning plus...show how I am justified in giving the answer '125' to '68+57'". As we've noted, one might hear this merely as requiring that our account show how Kripke is *correct* in giving the answer '125' to '68+57'. But we've seen that the reading does not square with Kripke's repeated indications that the issue raised by the skeptic is

whether Kripke's answer was an "unjustified stab in the dark". Showing that Kripke's answer was correct is one thing; showing that it was not a stab in the dark is another.

What would it take to show that Kripke's answer was not a stab in the dark? To say that his answer was a stab in the dark is to say that it was a guess. And to say that it was a guess is to say that Kripke had no reason for giving that answer as opposed to another. This suggests that the justification requirement is best construed, in application to this particular case, as the demand that our account reveal Kripke's reason for giving the answer '125'. According to the views of rational explanation and rule following just put on the table, giving Kripke's reason for his answer will involve portraying his answer as explained by his taking the rule for addition to justify that answer. The justification requirement, as I think we ought to construe it, is the demand that our account of what constitutes Kripke's meaning plus by "+" entitle ourselves to this portrayal of Kripke's performance, in which his answer is explained by the justification he takes the rule to provide for it.<sup>15</sup> What will give the requirement bite is a skeptical argument purporting to show that, despite initial appearances, we cannot after all understand how his answer might be explained in this way.

Support for this interpretation of the requirement does not lie merely in the quotations about stabs and leaps in the dark. Once one knows what to look for, it becomes obvious that throughout Kripke's initial "intuitive" presentation of the skeptical paradox, and at several other places in the book (notably, as we will later see, his discussion of anti-reductionism), he is articulating a puzzle targeted to the thought that in giving his answer to the "+" problem, a would-be adder is following a rule, and in so doing, engaging in a rational performance, a

---

<sup>15</sup> In suggesting that Kripke's demand to show the rule-follower's justification for proceeding as she does amounts to a demand to explain her performance in terms of the justification she takes herself to have for it, I am trading on a familiar and (in this context) harmless ambiguity in ordinary use of the word "justification", akin to that between "reason" in the sense of a normative reason and "reason" in the sense of a person's reason for doing something. (Compare: "She told me her justification for leaving, but I told her she had no justification.")

performance based upon a reason. Consider, for example, this passage from the second half of the book:

Almost all of us hesitatingly produce the answer ‘125’ when asked for the sum of 68 and 57, without any thought to the theoretical possibility that a quus-like rule might have been appropriate! And we do so without justification. Of course, if asked why we said ‘125’, most of us will say that we added 8 and 7 to get 15, that we put down 5 and carried 1 and so on. But then, what will we say if asked why we ‘carried’ as we do? Might our past intention not have been that ‘carry’ meant *quarry*; where to ‘quarry’ is ...? The entire point of the sceptical argument is that ultimately we reach a level where we act without any reason in terms of which we can justify our action. We act unhesitatingly but *blindly* (p. 87, ellipsis in original).

The question with which the skeptical argument is concerned is “why we said ‘125’”. This is an explanatory question; it asks why we performed a particular action. The import of the skeptical argument, as here represented, is that what we are inclined to say in answer to this question does not have the substance we credit it. We think we can explain why we answered ‘125’ by portraying ourselves as following a rule, involving adding and carrying and so forth. This explanation is supposed to show our reason for giving that answer—to show that we gave that answer because we took the stated rule to justify our doing so. But as Kripke sees it, the explanation we give in terms of the rule, despite initial appearances, does not succeed. For it prompts a further question about our reasons for proceeding as we do, which can then be pressed in turn, *ad infinitum*. The upshot of this regress is that we cannot after all explain our answer in terms of our reasons for it. The only option is to conclude that there was no reason for which we acted as we did. “The entire point” of the skeptical argument is that “ultimately we reach a level where we act without any reason in terms of which we can justify our action.” We do not act for a reason; we act “blindly”.

The next question for us, obviously, is why exactly a regress is supposed to arise here. But before turning to it, I want to briefly connect some of the dots of the discussion to this point.

First, there is the matter of normativity. The passages I've been discussing, which together link normativity to the justification requirement and the justification requirement to the rational character of rule following, motivate a third interpretation of Kripke's talk of normativity: to say that the relationship between a rule and its application is normative is to say that when a person follows a rule, the rule provides her reason for proceeding as she does. Despite some infelicitous historical connotations of the word "rationalism", I will call the view that the rule is the rule-follower's reason for proceeding as she does the *rationalist view of rules*, and the claim that Kripke's talk of normativity is intended to express this view the *rationalist interpretation* of such talk.<sup>16</sup> The thrust of this section is that the rationalist interpretation is the right one, at least with respect to one significant strand in the exposition of the skeptical argument. And what makes it sensible to associate the rationalist view with talk of normativity is the internal connection, discussed above, between the idea of the reason for which a person acts and the idea of a normative reason.

Second, there is the matter of the irrealist character of the skeptic's conclusion. On the present interpretation, the immediate conclusion of the skeptical argument is that we cannot make sense of the idea that we follow rules. To put it another way, there is no such thing as rule following. What the skeptic actually concludes, however, is that there is no such thing as rules. It might seem that there is an illicit inferential leap here: that it is possible for there to be rules despite the fact that the idea that we follow them turns out to be untenable. Perhaps being capable of being followed is not an essential property of rules. I will not argue on the skeptic's behalf against this possibility, although pressing this objection to the skeptical argument seems to

---

<sup>16</sup> One might deny that a rule can be a person's reason for doing or thinking something, on the ground that only proposition-like items can be a person's reasons and that a rule is not proposition-like. (For example, one might think that a rule is best modeled not on an assertion, but on a command: "Step, then kick," "Add two to the previous number.") Given this claim, the rationalist view of rules would need to be reformulated thusly: when a person follows a rule, her reason, at each step, for doing what she does is that doing so at that step accords with the rule. It would make no difference for my purposes if the view was reformulated thusly, but I will stick with the simpler formulation in the text. (Thanks to Jon Ellis and Hannah Ginsborg for raising this issue.)

me likely to be a singularly unrewarding exercise. If one wants to insist on the possibility of unfollowable rules, we may on the skeptic's behalf revise his conclusion to dismiss only rule following and not rules as such. That result is devastating enough.

Finally, there is the matter of meaning. Throughout the book, Kripke focuses on the example of the '+' problem. In discussing the problem, he shifts seamlessly between talk of rules and talk of meanings. Now, it is certainly an ordinary thought, a thought that any skeptical engagement with our ordinary understanding of our practices is entitled to view as a component of that understanding, that a person who computes an addition problem is following a rule. It is also an ordinary thought, in this sense, that it is in virtue of the meaning of '+' that the rule for addition is the relevant rule for answering the question, '68+57?'—in particular, it is in virtue of the fact, as Kripke would put it, that '+' means plus. But it is quite another matter to hold that the meaning of '+' is *itself* a rule, a rule that we follow in responding to problems posed with that expression. A fortiori, it is quite another matter to hold that the meanings of linguistic expressions in general are sensibly construed as rules that we follow in speaking a language. It seems to me that if one wants to claim that our ordinary understanding of linguistic practices is committed to this thesis, then one must do some work to show that that is so. An adequate assessment of this claim would require engagement with an array of difficult matters, and I won't attempt it here.<sup>17</sup> Suffice to say that, although we have seen the rationalist view of rules to be

---

<sup>17</sup> The basic question is whether it is right to think that the meanings of linguistic expressions serve as reasons for which we use those expressions as we do. Certainly, the use of language is a rational activity: we generally have reasons for saying the things we do. But it is consistent with this point to suppose that for a fluent speaker of the language, the only descriptions under which her utterances (or responses to utterances) will be intentional (at least in typical cases), and hence the only descriptions under which she will have reasons for those utterances (or responses), are descriptions that cite the utterances' content. On this picture, the question what Kripke's reason was for answering '125' in response to '68+57?' when interpreted strictly, may not have an answer. The answerable question in this vicinity may rather be what Kripke's reason was for answering 125 when asked what 68+57 is. It would certainly be natural to answer this question by citing a rule as Kripke's reason, but the rule will not be a rule for the use of a linguistic expression.

plausible, we cannot show the analogous *rationalist view of meaning* to be similarly plausible simply by assimilating, without explanation or defense, meanings to rules.

But my interest is in the skeptic's argument against the possibility of rule following, and the interest of that argument is not diminished even if we reject the assimilation of meanings to rules. If, for example, the skeptic can show that Kripke is not following a rule when he gives his answer to the "+" problem, that would be a remarkable and disturbing result, whatever the precise implications of this result for meaning. Since I want to stay focused on this issue, I will for purposes of argument grant Kripke his view that the use of meaningful expressions is a species of rule following.

## **5. The skeptical regress and the situatedness of explanation**

What, exactly, is the nature of the regress that the skeptic takes to undermine our rational explanation of Kripke's response to the "+" problem? The line of thought articulated in the passage quoted might be construed as follows. We offer an explanation of Kripke's response by citing a rule as his reason for that response. But this explanation only prompts further questions about Kripke's reasons for giving that response. Answers to those questions are met with further such questions. Eventually, we "reach a level" where Kripke's reasons give out, where there's nothing further to be said about his reasons. This result gives the lie to our original explanation. It shows that Kripke did not after all have a reason for his response; rather he acted "blindly".

If the skeptical argument is to be worth our time, it must contain premises and inferences not made explicit by this characterization. For to think that the argument, thus characterized, is valid would be to make a fairly uninteresting mistake. Granted, in the face of a claim about a person's reasons for doing or thinking something, it's always possible to formulate a further

question about the person's reasons. And granted, if we keep pressing such questions, sooner or later we'll get to one we can't answer, and about which we may perhaps be inclined to say that it has no answer. But of course it is true of *explanation in general* that one can always respond to an explanation by raising questions about considerations appealed to or presupposed or implied by that explanation. And of course it's true that if we keep pressing such questions, we'll eventually reach a question we can't answer, and indeed such that we don't have a clear idea of what an answer might be. What we need to ask in any such case is this: is the further question well taken? That is, does the adequacy of the original explanation—its capacity to explain that which it is intended to explain—depend upon our being able to provide, or even on there being in principle, an answer to this question? The mere fact that one can *formulate* the question does not suffice to show that it is well taken in this sense. What the possibility of formulation shows at best is only, as is anyway obvious, that every explanation takes something for granted, that it assumes something that it does not itself explain. That a three-year-old keeps asking "Why?" no matter what his parent says does not demonstrate that the explanations proffered to him stand in need of additional content. Rather, it demonstrates that the child does not yet have the knowledge or sensibility to appreciate these explanations, situated as they inevitably are against an assumed background of what Wittgenstein calls "agreement in judgments" (1973, §242).<sup>18</sup>

Consider again the scenario discussed in the previous section. I list 1, 1, 2, 3, 5, 8, 13, etc. on a piece of paper. Suppose you ask a third party why I'm doing that, and are told that that I'm following the rule for the Fibonacci sequence. On the rationalist view of rules, this

---

<sup>18</sup> Of course, there is such a thing as a putative explanation taking too much for granted, and so failing to adequately explain. Unfortunately there is no algorithm for determining whether an explanation does say "enough". Explanations aim to yield understanding, or better, to provide people with the opportunity to understand (an opportunity they may fail to take). Whether a given explanation counts as doing its part to achieve this aim is not something we can determine absent an engagement with the discursive context in which it is, or might be, offered. For a classic discussion of the role of context in determining "explanatory relevance", see Van Fraassen, 1980, Ch. 5. (Note that all subsequent references to Wittgenstein, 1973, will be by section number only.)

characterization of my behavior offers a rational explanation of it: it says that at each step, my reason for listing the number I do is the rule for the Fibonacci sequence. This answer may prompt you to ask further questions, for example, why I'm following the rule for the Fibonacci sequence to begin with. (This question may take a particular shape for you depending on your view of relevant features of the situation. "Why is he following that rule when I explicitly asked him to calculate the Perrin sequence?" "Why is he sitting there calculating sequences while his wife angrily packs her bags in the next room?") I've noted that the mere possibility of pressing such questions does not show that the original explanation is inadequate pending answers to these questions. Perhaps the original explanation *is* inadequate; the point is that we need a better reason for thinking so than the fact that rational explanation, like all other kinds of explanation, has to start somewhere.

But the questions I just imagined you asking seek explanations that are 'upstream', as we might put it, of the original explanation of my activity; such explanations would provide insight into the larger interests, considerations or motivations that prompt my rule-following behavior. More pertinent in the present context is the possibility of seeking explanations that are 'downstream' of the original explanation. The rule for producing the Fibonacci sequence may be quasi-formally characterized thusly: to list the  $n^{\text{th}}$  number in the sequence, list 1 if  $n=1$  or  $n=2$ , and otherwise list the number that is the sum of the  $(n-1)^{\text{th}}$  and  $(n-2)^{\text{th}}$  numbers. Obviously, applying this rule for a given  $n>2$  will require as input the previous two numbers on the list. And so if it is said that I am following the rule for the Fibonacci sequence in listing 4181 at the twentieth step, a fuller statement of the reason for which I am thereby said to list 4181 is that it is the rule coupled with the fact that (we may assume) the previous two numbers on the list were 1597 and 2594. Now, it's possible to imagine someone asking why I took the rule, coupled with that input, to yield that result. Presumably, my reason for taking this to be so was that the sum of

1597 and 2584 is 4181. But what was my reason for believing that the sum of 1597 and 2584 is 4181? (We could imagine a novice in arithmetic wanting an answer to this question.) Probably there is something to be said here concerning my application of the rule for addition. Sooner or later, however, we will reach an arithmetical belief on my part such that even I would be hard-pressed to reconstruct some further reason for which I hold it—beyond that (as I might say) I remember that it's true. In this sense, my reasons will give out. But so far, this shows nothing at all. Everything turns on whether the further questions one might think to ask here—e.g., “What is your reason for taking it that your remembering this putative fact is a reason to believe it to obtain?” —is well taken. If we see the skeptical argument as appealing only to the abstract possibility of responding to rational explanations with further downstream questions about reasons, then we must conclude that it is entirely devoid of force.

## 6. The guidance conception of understanding

We need, then, to find an interpretation of the skeptical argument that has it resting on more than this. I think the materials for such an interpretation are present in the text. But the key premise is not made as clear as one might like, either in its motivation or in its role in the skeptical argument.

What is this premise? Kripke repeatedly associates the justification requirement with a further idea. Indeed, he seems almost to run them together. Consider the following passage from the discussion of dispositionalism:

So it does seem that a dispositional account misconceives the sceptic's problem—to find a past fact that *justifies* my present response. As a candidate for a 'fact' that determines what I mean, it fails to satisfy the basic condition on such a candidate stressed above on p. 11, that it should *tell* me what I ought to do in each instance” (p. 24).

The skeptic challenges us “to find a past fact that *justifies* my present response.” As we have seen, what he challenges is the legitimacy of our picture of Kripke as following a rule, and hence of the rule as providing Kripke’s reason for his response. Here that challenge is transformed, without comment, into the condition that we find a fact that will “*tell* me what I ought to do in each instance.” Language to similar effect pervades Kripke’s exposition. Here are a few examples:

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 (p. 17).

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 (p. 11).

The answer to the sceptic’s problem, ‘What tells me how I am to apply a given rule in a new case?’ must come from something outside any images or ‘qualitative’ mental states...[Even] if there were a special experience of ‘meaning’ addition by ‘plus’, analogous to a headache, it would not have the properties that a state of meaning addition by ‘plus’ ought to have—it would not tell me what to do in new cases (p. 43).

Sometimes when I have contemplated the [skeptical paradox], I have had something of an eerie feeling. 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...But when I concentrate on what is now in my mind, what instructions can be found there? How can I be said to be acting on the basis of these instructions when I act in the future?...To say that there is a general rule in my mind that tells me how to add in the future is only to throw the problem back on to other rules that also seem to be given only in terms of finitely many cases. What can there be in my mind that I make use of when I act in the future? It seems the entire idea of meaning vanishes into thin air (pp. 21-22).

If I am to follow a rule, something must “instruct” me how to go on. It must “guide” me. It must provide “directions”. It must “tell me how I am to apply [the] rule in new cases”. And what is the item that is to guide or instruct me? It is “something in my mind”. Elsewhere, Kripke characterizes it as “my mental state” (p. 11). In particular, it is his “state” of “meaning addition by ‘+’” (p. 51). In a similar vein, it is his “grasp” of the “rule” (p. 7), or his “understanding” (p. 48). Given that the skeptical argument applies to rule following in general,

and given that I do not wish to get sidetracked by the question of whether linguistic meanings are sensibly thought of as rules that we follow, I will tend to favor talk of understanding and grasping rules rather than of states of meaning. I will say, then, that what is supposed to guide a rule-follower, for Kripke, is the mental state that constitutes her understanding or grasp of the rule. For Kripke, one species of this state is the state of meaning something by a linguistic expression.

The common kernel expressed by these various locutions is the idea that, when a person follows a rule (at least as we “normally” think of this activity—i.e., in advance of acquaintance with the skeptical argument), there is something in her mind that she consults, something that she refers to, at each step. This item communicates information to her about how to proceed, and she so proceeds. Imagine a person navigating a hedge maze. In her pocket is a guide—perhaps a map, perhaps a set of instructions. At each intersection she takes the guide out of her pocket, consults it, and proceeds on the basis of the information thus gleaned. On the vision of rule following suggested by Kripke’s formulations, rule following is just like this, with the difference that the rule-follower’s guide is in her mind rather than her pocket. It is in fact the mental item that constitutes her grasp, her understanding of the rule.

I will call this idea *the guidance conception of understanding*, though this label comes with two caveats. First, the word “conception” can have connotations of determinacy and clarity that are out of place here. What I am calling the guidance conception of understanding is not a well thought-out philosophical theory of anything; it is rather a picture our grip upon which is largely intuitive and analogical. Indeed, as we shall see, the main point of the skeptical argument is that we will run into insoluble difficulties when we try to think through the conception more carefully. The second caveat is that we sometimes speak of one item’s “guiding” another and mean only that the first item influences the second, with no implication that this influence was mediated by any uptake or comprehension on the part of the item being guided. Thus we speak, for example,

of guiding a ship into harbor. But the relation between subject and understanding posited in the guidance 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 the subject acquires information about how to proceed, information upon which she then acts. What is at stake is not the guidance of brute causal influence, but what we might call *informed* guidance. Kripke's talk of instruction, direction, showing and telling all avoid this ambiguity. But I prefer talk of guidance for its greater generality, and so will stick with that term.

Why should we be attracted to the guidance conception, as Kripke implies that we are? Kripke seems to suggest that the phenomenology of rule following favors the conception: even after being convinced by the skeptical argument, he can't help but "feel confident" that there's something in his mind instructing him how to go on. Now undoubtedly, there are occasions of rule following when one will think of oneself as guided or directed by something in one's mind. For example, if you are trying to follow a rule whose application is complex and which you have not fully got the hang of, you may bring to mind instructions, images, and so on that were used as aids in your training. You may then find it natural to say these images or instructions guide or direct your application, or that they tell you how to proceed. But in other cases, the conception is less obviously apt. For example, if you are following a rule that is second nature to you, or one whose application is exceedingly straightforward (imagine a rule for producing a simple numerical sequence), will you have any sense at all that there is something in your mind that actively instructs or directs you in how to go on? At the very least, the phenomenology is much less clear-cut in such cases.

Is there a motivation for the guidance conception other than the unconvincing appeal to phenomenology? Kripke does not tell us. But I think it's possible to get at least some sense of why the conception has such allure for Kripke by thinking through a natural reaction to remarks

by Wittgenstein in the part of the *Philosophical Investigations* that prompt Kripke's own skeptical reflections.

Wittgenstein is prone to a certain kind of characterization of what goes on when a person follows a rule. He speaks, for example, of a rule follower as doing what comes naturally (§185), as just reacting as he was trained (§198), and as applying the rule as a matter of course (§238). The thrust of these characterizations, and others like them, is that in typical cases, a rule follower is merely responding to the circumstances, without reflection or deliberation, in whatever way he has been trained or habituated. Now, I think it's easy to get the sense that Wittgenstein's characterizations of an ostensible rule-follower are in fact incompatible with viewing him as genuinely following a rule. What the characterizations might seem to conflict with in particular is the very thing I have been harping on here—the rational character of rule following, the fact that a rule-follower proceeds as she does because she takes the rule to provide a reason for doing so. It might seem that there is a failure of fit between the idea that a person is acting in light of her conception of what the rule gives her reason to do and the idea that, in proceeding as she does, she is just doing what comes naturally to her in light of her training. Wittgenstein himself is happy to encapsulate his vision of what goes on in rule following with the remark that we “apply the rule *blindly*” (§219). But if our application of the rule is truly “blind”, one might want to say, then we are not really following the rule after all. For rule following is not blind; it is informed by a sense of what, at each step, the rule gives us reason to do.

This worry about Wittgenstein's characterizations of putative instances of rule following is nicely articulated by Thomas Nagel (1996). Nagel is troubled by these characterizations, which he calls Wittgenstein's “facial descriptions of our practices” (1996, p. 52). To Nagel, the facial 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)” (1996, p. 48). So viewed the practices “lose their meaning”; they appear as mere “impotent rituals” (1996, pp. 51, 53). And 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 what the relevant rules dictate.

Suppose we share Nagel’s intuitive sense of the incompatibility of Wittgenstein’s “facial descriptions” of behavior with conceiving that behavior as rule-following behavior. Can we fill out this intuition further? What, exactly, is the nature of the incompatibility? A natural first thought here is that acting for a reason is, or involves, a mental process, the presence of which gives the lie to the facial descriptions. But this thought doesn’t take us very far: the bare notion of a mental process is too indefinite to do any real work here.<sup>19</sup> We need to specify a particular kind of mental process, one whose incompatibility with the descriptions at issue is evident. Nor can we do with the most obvious candidate, and cast the process in question as one of thinking, in the sense of active deliberation or reflection. Granted, a person who proceeds on the basis of careful deliberation is not aptly described as, say, applying the rule as a matter of course or as just reacting as she was trained. The problem is that many cases of rule following (including the kinds of cases upon which Wittgenstein tends to focus) obviously do not involve a process of thinking in the relevant sense. It would seriously over-intellectualize matters to insist that, in following a simple and familiar rule, a competent adult must engage in anything akin to deliberation or reflection.<sup>20</sup>

---

<sup>19</sup> The later Wittgenstein is often read as denying that understanding, and like phenomena, are mental states or processes. But it seriously misreads his intentions to interpret him as ever flatly giving out such a denial. It’s 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 he is offering this advice and why.

<sup>20</sup> Compare Wittgenstein (1978, §VII-60): “One follows the rule mechanically. Hence one compares it with a mechanism. ‘Mechanical’—that means: without thinking. But entirely without thinking? Without reflecting.”

But even in simple cases, it might appear that we can still conceive the behavior in question as *guided* by the rule-follower's understanding of the rule. Being guided, being instructed or shown or told how to proceed, needn't involve any deliberation or reflection on one's part: if someone or something tells me, "Write down '5' now," I can grasp that instruction, and follow suit, without any intervening deliberation at all. And the proposed mental process meets our other desideratum as well: viewing a person as being guided in her performance comports very poorly with viewing that person as just doing what comes naturally, or just reacting as she was trained. Indeed, to do what comes naturally is precisely not to rely on any directions, instructions or 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.

The appeal of the guidance conception for Kripke, I propose, is a function of its apparent capacity to underwrite the intuition articulated by Nagel. It seems to give flesh to the thought that following a rule, as a species of rational activity, must involve a mental process that is not in play when a creature just reacts to the circumstances in whatever way comes naturally to it.

There is room for further examination of why this thought can so easily get a grip on the philosophical imagination. But before we return to this matter in section 9, I want to consider how incorporating the guidance conception as premise of the skeptical argument enables us to make satisfying sense of the skeptical charge of regress and the rejection of anti-reductionism.

## **7. The skeptical regress, properly construed**

The problem for the guidance conception of understanding flows from an exceedingly simple point: if an item is to successfully tell you something, or direct you or instruct you in some particular way, then you must understand what it tells you, or what it directs or instructs you to

do. Recall the map or set of directions in the pocket of the person navigating the hedge maze: that map or set of directions obviously cannot show the person the way through the maze if she does not understand it as showing her anything. And of course the point is perfectly general. To conceive an item as showing or telling a person something, or as instructing or directing her in some way, is to conceive that item as communicating something—which is to say, some content—to her. And successful communication with a person requires two contributions, one from the communicator, and one from the communicatee. The contribution from the former is to express or otherwise make available the relevant content. And the contribution from the latter is to take that content in—to grasp, comprehend, understand it.<sup>21</sup>

But if the point is perfectly general, then it applies in particular to items in the mind that purport to guide or direct one. And so if Kripke's grasp of the rule for calculating “+” problems consists in the presence of, as he puts it, “something in my mind...that *instructs* me what I ought to do in all future cases”, then we must posit a further state of understanding on the part of Kripke: namely, his understanding of the instruction provided by the item in the mind that constitutes his grasp of the “+” rule. The question we now face is this: how are we to construe the role of this second-order state of understanding in helping to explain Kripke's subsequent performances? More specifically, if we are to conceive of the first-order understanding of the rule on the model of an item that instructs the rule-follower in how to proceed, are we to conceive the second-order understanding in the same way?

Neither answer is satisfactory. If the answer is no, then we are left with, as Kripke and Wittgenstein would put it, a “level” at which his performance is “blind”, and this conflicts with the original motivation for the guidance conception. Let's switch for a moment to the example of my following the rule for producing the Fibonacci sequence. The aim of the guidance

---

<sup>21</sup> Obviously the notion of communication I'm here appealing to is not the notion treated in information theory à la Shannon and Dretske.

conception is to ward off a description of a rule-follower like me as applying a rule “blindly”. It seeks to do so by ruling out the possibility that in writing each number that I do, I am simply doing what comes naturally to me, as a result of my training, in those circumstances—such circumstances including, in this case, the numbers I’ve previously written down. So we posit something in my mind that instructs me what number to write down at each step. If the presence of this item is to be of any use to me, I must understand the instruction it provides. But if my understanding of the instruction does not itself consist of, or involve, an item in my mind that guides me in applying the instruction, then we have not after all ruled out the possibility that I am merely reacting as I was trained to do in the circumstances. All we’ve done is introduce a further circumstance, namely, the presence of the item in mind that constitutes my understanding of the rule. Nothing we’ve said so far blocks the following thought: that in writing each number that I do, I’m simply doing what comes naturally to me given the combined presence of those numbers on the page and that item in my mind. So the original motivation for the guidance conception is left unsatisfied.

On the other hand, if the answer is yes, then we are embarked on a regress. Suppose my understanding of the instruction provided by the item that constitutes my understanding of the Fibonacci rule is itself constituted by an item that instructs me how to proceed. If it is to instruct me, I must understand the instruction it provides. We will need then to posit a third-order state of understanding on my part: an understanding of the instruction provided by the mental item that constitutes my understanding of the instruction provided by the mental item that constitutes my understanding of the Fibonacci rule. If we allow this regress to begin, then, as Kripke’s skeptic sees, it cannot be stopped. At each stage of the regress, we come to an item that is supposed to instruct me how to apply the preceding item, but is unable to do so unless a further

item has instructed me how to apply it. The result is an endless chain of ‘dead’ items, each awaiting a breath of ‘life’ from the next item down the line, and so none receiving any.<sup>22</sup>

Arguments of this general thrust, more or less compacted in presentation, occur in several places in *Philosophical Investigations*. To take an occurrence that is less frequently discussed than others, consider the following:

How is he to know what colour he is to pick out when he hears “red”?—Quite simple: he is to take the colour whose image occurs to him when he hears the word.—But how is he to know which colour it is ‘whose image occurs to him’? Is a further criterion needed for that? (§239)

Think of a game in which one of the rules is to identify a sample of red from among an array of color samples when “red” is uttered. Wittgenstein’s interlocutor suggests that the participant’s grasp of this rule might be partly constituted by an association between that word and an image of a color in his mind. It would be a mistake to take the point of Wittgenstein’s response to be that such an association could not possibly be of use in following the rule. Certainly such an insistence would be hard to swallow. Who is Wittgenstein to say what might or might not be of use to a person in following a rule? “Red” is perhaps not a good example, as most of us apply that word as a matter of course. But imagine a native English speaker who’s previously acquired little of its extensive vocabulary for speaking of shades (I am such a speaker) and is now being trained, perhaps preparatory to working as a salesperson in a clothing store, in the use of words like “vermilion”, “fuchsia” and “cerise”. Here it’s easy to imagine that his application of those words might be mediated, at least initially, by his calling to mind appropriate images—perhaps memories of samples he’s been shown.

There is no reason to take Wittgenstein to deny this possibility. What he points out is simply that a mental image cannot show the person that he is to pick out a certain color unless he understands it as showing him that. If, then, the interlocutor’s motivation for positing the image

---

<sup>22</sup> For the life/death imagery, see Wittgenstein (1960, p. 4).

was a worry, however inchoate, to the effect that a person *must* have something in his mind that shows him how to apply the rule if he is to count as genuinely following the rule, we are faced with a dilemma. For what shows him how to apply the image that is to show him how to apply the rule? If we posit a further “criterion”, regress threatens. If, on the other hand, we deny that anything shows him how to apply the image, the question presses whether we have arrived at a picture of what goes on when a person applies a rule that quells the interlocutor’s original worry or merely defers it.

Kripke’s own account of the skeptical regress also depends upon the guidance conception, as we can see if we attempt first to understand what he says in motivating the regress without appealing to the conception. Consider Kripke’s treatment of the regress in the initial “intuitive” exposition of the skeptical paradox. The crucial step occurs after Kripke has noted that his past answers to computational problems expressed using “+” are consistent with “+” standing for a function other than addition, such as ‘quaddition’, in which eventuality “68+57” would denote 5. (This possibility is consistent with Kripke’s past answers because, by supposition, Kripke has never previously been confronted the problem “68+57”.) Kripke anticipates that the reader will protest that this point on its own does not undermine Kripke’s claim that he meant *plus* by “+” in the past. For what fixed Kripke’s past meaning is surely not just his finite set of answers to problems couched with “+” or “plus”, but the general rule he employed for solving such problems. Kripke’s imagined interlocutor suggests that the rule can be explained as follows: “Suppose we wish to add  $x$  and  $y$ . Take a huge bunch of marbles. First count out  $x$  marbles in one heap. Then count out  $y$  marbles in another. Put the two heaps together and count out the number of marbles in the union thus formed” (p. 15).

Kripke is unimpressed with this protest. His response begins this way:

Despite the initial plausibility of this objection, the sceptic's response is all too obvious. True if 'count', as I used the word in the past, referred to the act of counting...then 'plus' must have stood for addition. But I applied 'count', like 'plus', to only finitely many cases. Thus the sceptic can question my present interpretation of my past usage of 'count' as he did with 'plus'. In particular, he can claim that by 'count' I formerly meant *quount*... (p. 16).

For all I've said so far, this response is a non-sequitur. The response challenges an assumption about Kripke's past use of the word "count". But Kripke's interlocutor, as I just presented him, said nothing about Kripke's past use of that word. His proposal was limited to Kripke's past use of "+" (and "plus"): he suggested that Kripke had previously understood a certain rule to govern the use of that expression. It is true that in stating this rule, the interlocutor himself used the word "count". But the meanings of the interlocutor's expressions are not now in question. Indeed, it is an explicit ground rule of the skeptical argument that we—those formulating and discussing the paradox—can take for granted that we know what our own words mean: "the sceptic, provisionally, is not questioning my *present* use of the word 'plus'...Not only does he agree with me on this, he conducts the entire debate with me in my language as I *presently* use it" (p. 12).<sup>23</sup>

What's going on here? The answer is that Kripke understands the interlocutor not only as suggesting a rule that Kripke followed in solving "+" problems, but as offering a particular conception of what his grasp of the rule consisted in. Kripke states the general idea this way: "I learned—and internalized instructions for—a *rule* which determines how addition is to be continued" (p. 15). The hyphenated bit is essential: the point of the interlocutor's suggestion, as Kripke interprets it, is not to restate the basic thought that Kripke is following a rule in responding to the "+" problem, but to identify the item in Kripke's mind that instructs him how to apply the rule. And so when the interlocutor proposes a statement of the rule Kripke

---

<sup>23</sup> The point of couching the skeptical puzzle in this way, as a challenge about past use, is to avoid worries about "whether the discussion is taking place 'both inside and outside language'" (p. 12).

previously followed, he is understood as suggesting that Kripke “explicitly gave [himself]” this “set of directions”—that they are “engraved on [his] mind as on a slate” (p. 15). With the interlocutor’s proposal thus understood, Kripke’s rejoinder is well-taken. If the set of directions Kripke previously gave himself is to be of any use in helping him to following the rule, he must understand it. But in what does his understanding of those directions consist? If it does not consist in any mental item that shows him how to understand, how to apply, the directions, then the interlocutor’s proposal does not in the end insulate Kripke from the charge that he follows the rule “blindly”. But if we hold that his understanding of the directions does consist in such an item, all the previous difficulties recur.

## **8. Anti-reductionism and logical compulsion**

We’re at last in a position to return to the matter we left in section 2: Kripke’s puzzlingly brusque rejection of the view that mental states and occurrences are irreducible. The rejection becomes explicable when we register two points. First, the treatment of anti-reductionism, like the initial exposition of the skeptical paradox, centers not around the evaluative view of meaning, as commentators have assumed, but around the rationalist view of meaning—the view, recall, that meanings, à la rules, serve as reasons for our using meaningful expressions as we do. Second, and relatedly, Kripke’s understanding of rule following is controlled by his subscription to the guidance conception. Kripke recognizes, as we’ve seen, that this conception of rule following yields a regress. But he remains committed to the conditional premise that if there were such a thing as rule following, it would have to conform to the model of the guidance conception. The upshot is that Kripke can make nothing of the appeal to the irreducibility of the mental but an attempt to protect the possibility of rule following, conceived in terms of the guidance

conception, by insisting that any worries about it are improper because states of understanding and meaning are *sui generis* and hence not to be subject to examination. It is no wonder that Kripke finds this maneuver unconvincing.

But there is in fact a more specific point to be made about Kripke's response to anti-reductionism: in an abbreviated fashion, it recapitulates an important element of Wittgenstein's treatment of rule following. That element is Wittgenstein's criticism of the idea of "logical compulsion".

When we are in the grip of the guidance conception, it can be tempting to think that mental items are special, in precisely the way that is needed to block the regress. As we noted, there are two contributions to a scene of guidance or instruction: one from the instructor and one from the instructed. The contribution of the former is to provide or make available the instructing content; the contribution of the latter is to understand it. But perhaps the perception that this second contribution is always required stems from modeling the guidance provided by one's understanding too closely on the guidance provided by material things in the outside world. The mental sphere is unique in a great many ways, after all. Suppose in particular that we can make out the following thought: that the mental item you consult when you follow a rule differs from maps, written directions, stop signs and other physical objects in that its power to communicate information to you is not dependent on your hitting upon the right understanding of what it expresses. We might put the thought this way: the mental item provides for its own understanding. The proper understanding of the direction it provides is built into the item itself. Hence this understanding is not something *you* must contribute to the transaction: it is part of the package the item presents you.

Here, it might seem, we have the key for preventing the regress. The regress is set in motion by the fact that the mental item constituting your understanding of a rule, if it is to guide you,

must itself be understood. The question arises what constitutes your understanding of that mental item; the structure then replicates ad infinitum. But if the understanding of what the mental item tells you is somehow built into the item itself, so that the sheer presence of the item in the mind guarantees proper understanding, then the question of what constitutes your understanding of that item has already been answered: it consists in the presence of the original item. And so, it might seem, the regress cannot get started.

Such an item would be one that it is logically impossible for you to understand in any but one way. It is in this context that 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 different order. In the case of the picture that forces its application upon you, the question of how you are disposed to understand the picture does not so much as arise, for the proper understanding of the picture is, as it were, part of the picture itself. We can no more isolate something about which the question of understanding might arise than we can isolate, say, a triangle from its angles.<sup>24</sup>

It's hard to fill out this idea more fully, because however tempting it may be—and for Wittgenstein, I believe, it belongs among a set of ideas about the nature of mental processes that

---

<sup>24</sup> It's important not to run together this idea with the view that contentful mental states have their contents essentially or intrinsically. On that view, insofar as a belief that  $68+57=125$  can be said to be an "item in the mind" of the person who has it, it is an item identified and individuated by its being a belief with that very content. But the idea we are now considering is that of a mental item whose conveying or communicating to you a specific content is an essential property of it. To the extent that we construe mental states as intrinsically contentful entities—and so as not aptly conceived on analogy to signs, images or linguistic expressions—we will be disinclined to construe our relationship to them in terms of communication or guidance to begin with. (This last point warrants elaboration, which I cannot here provide.)

we can find well nigh irresistible—it falls apart under scrutiny. There is simply no making sense of the idea of an item, mental or otherwise, 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. An item cannot tell a person something unless she understands it to tell her that, and there is no getting around the fact that her understanding it this way is something she must bring to the table, not something that the item can itself provide for. No matter what we place before a person, no matter how rich or extensive the materials we display, there can be no guarantee that the person will take those materials to say or otherwise convey what we intend, or indeed to convey anything 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 as something that we onlookers exhibit to her; it is something she has in her mind. But nothing can circumvent the dependence of guidance upon a person’s uptake, and so 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 of, say, a stop sign may be arbitrary, the ways in which we understand 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 use a bit of green paper as a representation 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 the ‘psychological compulsion’ is here (if it pleases us to put it this way); it cannot vindicate the confused idea of ‘logical compulsion’.

Let’s return to Kripke. Since Kripke takes the guidance conception for granted, he assumes that the anti-reductionist does too. But the anti-reductionist also believes that the skeptical argument can be defused simply by casting our grasp of a rule as a *sui generis* mental state. At the core of the skeptical argument is, of course, the regress. From Kripke’s perspective, then, the anti-reductionist is committed to all of the following propositions: that a rule-follower’s grasp of a rule consists in an item in his mind that tells him how to proceed, that we needn’t worry that this picture generates a regress, and that the reason we needn’t worry is that the mental item in question has a *sui generis* character. As we have just seen, Wittgenstein discusses a view that precisely encompasses these three propositions, and that is the view that a rule-follower’s grasp of a rule consists in an item in his mind that guides his performances by ‘logically compelling’ him to understand it in a particular way. Given all this, and given that Kripke understands his work as an elaboration of Wittgenstein’s rule-following remarks, it would be unsurprising if Kripke should take the anti-reductionist proposal to be tantamount to the idea that states of meaning and understanding logically compel our understanding of them.

And indeed he does. He presses his charge of mystery-mongering against the anti-reductionist thusly: “Can we conceive of a finite state which *could* not be interpreted in a quus-like way? How could that be? The proposal I am now considering brushes such questions under the rug, since the nature of the supposed ‘state’ is left mysterious” (pp. 52-53). To speak of a state

that could not but be interpreted in one way is to speak of an item that logically compels a particular understanding of it. Kripke is right to doubt the possibility of such an item. As soon as we think of our relationship to an item as one of our interpreting or understanding it—as we do when we think of the item as directing, guiding, or instructing us—we cannot avoid the fact that the way in which we understand that item is a contingent fact about us.

To the extent that commentators pick up on Kripke’s protest about the idea of a state that could only be interpreted in one way, they fail to grasp its significance. Boghossian’s reaction to the protest is a good example. He writes, “[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” (1989, p. 542). Boghossian here equates the idea of a contentful state with the idea of a “state” that “contains information about the correct applicability of the sign.” Perhaps for Boghossian talk of “containing information” *is* just a fanciful way of talking about possession of content, in which case his gloss on Kripke’s objection overlooks the crucial structuring role played by the guidance conception. On the other hand, if we take Boghossian’s talk of “containing information” literally, then the question we should ask is: to whom, or what, is the information “contained” in the “state” to be conveyed? If it is to be conveyed to the rule-follower, in order to guide her application of the rule, then she must grasp that information. To suppose that Kripke has no “independent reason” for finding this proposal problematic is to entirely miss the strand of the skeptical argument that has been the focus of this paper.

## 9. The appeal to practice

The skeptical argument occupies, of course, only half of Kripke's book. The second half is given over to exposition of a "skeptical solution" to the "paradox"—i.e., an account of the significance of our ordinary talk about meaning that is consistent with the conclusion of the skeptical argument, and hence does not presuppose any facts about what we mean by our words. The key ideas seem to be, first, that a satisfactory understanding of the role of semantic talk in our lives can be provided by describing the conditions under which members of our linguistic community are licensed by that community to utter such sentences as, e.g., "Kripke means plus by 'plus'" and "Kripke's answer to '68 plus 57' was incorrect", and second, that we can adequately describe these conditions without supposing that there is any fact of the matter about what Kripke, or anyone else, means by the expressions they use.<sup>25</sup>

In its reliance on the communal setting of our utterances, Kripke takes his skeptical solution to be in the spirit of Wittgenstein's own frequent appeals to customs and practices. But as many commentators have pointed out, it's quite a stretch to read Wittgenstein as endorsing a skeptical solution, in Kripke's sense, to a puzzle about meaning or rule following. It's true that in the famous passage in §201, Wittgenstein speaks of a "paradox" about rule following, and that he takes one upshot of reflection on this paradox to be that "'obeying a rule' is a practice" (§202). But Wittgenstein's attitude to the putative paradox is manifestly not that the perception of a paradox is correct and that we must learn to live with the consequences, but rather that the perception is predicated on a "misunderstanding" (§201). Wittgenstein evidently believes that thinking of rule following in terms of customs and practices can help us to dissolve the perception of paradox, not to make peace with it.

---

<sup>25</sup> Both of these ideas seem to me quite wrong-headed, but that is not my concern here.

With the argument for the skeptical conclusion interpreted as I have here, how can the appeal to customs and practices work to counter it? The argument turns on the soundness of the intuition that there is an incompatibility between conceiving a person as just reacting to the circumstances in the way she was trained and conceiving her as following a rule, with its attendant implication that the rule is her reason for going on as she does. It is this intuition that prompts the guidance conception. And it is precisely this intuition, I believe, that Wittgenstein, in reminding us of the possibility of conceiving our rule-following behavior in terms of customs and practices, means to undercut.

Wittgenstein's basic point 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 rational 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, hence of rational activity.

For Wittgenstein, the relevant "complicated surroundings" encompass a great deal. 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 characterize this background of human activity in terms of talk of practices, customs, institutions and uses. There is no indication that he invests this talk with any special or technical meaning. Indeed, it would be counterproductive for him to do so, for he wants it to be an uncontroversial claim that each of us participates in numerous customs, practices and institutions. As ordinarily used, these terms speak of common activities—of things people typically or regularly do. And they often suggest an element of conventionality: walking as such is not naturally described as a custom, but walking on the right side of a sidewalk or hallway is (at least in the United States). The conventional element of customs and practices is in turn often characterizable in terms of rules.

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.

“Then can whatever I do be brought into accord with the rule?”—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).<sup>26</sup>

The first paragraph raises the regress problem for the guidance 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

---

<sup>26</sup> Following the practice of McDowell (1998), I have put quotes around the first sentence of the final paragraph to indicate the interlocutor’s voice. My treatment of this passage is strongly influenced by McDowell’s on a more substantive level as well.

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 guidance conception is to underwrite the thought that what is going on here is precisely a case of *going by* a sign—which is to say, a case of *following* a rule—rather than a brute, non-rational 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 fail on this score. But Wittgenstein takes this response to miss his point. 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 refraining from 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 following a rule, and so of her acting for a reason. In citing our training, Wittgenstein seeks to remind us that we have been initiated into such practices, practices of following rules.

As this example illustrates, practices and customs are especially useful for Wittgenstein’s purposes 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 are features of practices that everyone takes for granted. But taken at face value, 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 when the sign says “walk” and waiting when it says “don’t walk” does not require any deliberation,

reflection or guidance; one just reacts in the pertinent situation as one was trained. But those who participate in this practice are following a particular rule; it just *is* a practice of following rules. We might put Wittgenstein's suggested therapy this way: if you find yourself doubting that any given tract of behavior can satisfy both of the descriptions, "following a rule" and "doing what comes naturally", then remind yourself that the behavior may well satisfy a third description, namely that by which we characterize some specific rule-following practice.

I suggested that the characteristic of rule following that makes these descriptions seem incompatible is that it is a species of rational activity, in which the rule-follower's performance is explained, at each step, by the justification she takes the rule to provide. In §198 and similar passages, Wittgenstein points us toward the recognition that this perception of incompatibility is not justified by our ordinary thought and discourse, by the ways in which think and talk about relevant phenomena in everyday life. For on the one hand, we do ordinarily conceive rule following, even of simple rules, as a form of rational activity. That is to say, on appropriate occasions we cite traffic rules, arithmetical rules, and so on, as people's reasons for doing what they do. At the same time, we are perfectly aware that competent participants in familiar rule-following practices are often just doing what comes naturally to them in the circumstances. Indeed, we generally take the effortless, natural, non-deliberative character of a rule-follower's responses as reflective of her *mastery* of the rules. (It is precisely this effortless, non-deliberative quality that we aim to inculcate in those we train in the relevant rule-following practices.) And so the idea that rule following involves a special mental process—a mental process which, if not strictly speaking deliberation or reflection, is nonetheless like deliberation or reflection in virtue of rendering inappropriate Wittgenstein's 'facial descriptions' of the activity—certainly does not come from observation of a divide between the behavior we ordinarily classify as done for a

reason and the behavior we ordinarily classify as a natural reaction to the situation (or as a product of habituation, as done as a matter of course, etc.). There is no such divide.

How effective is the appeal to practice, so construed, as a response to the skeptical line of thought? It seems to me that the point Wittgenstein is making is a good one. The idea that rule following must involve a special mental process, one whose presence removes rule-following behavior from the realm of natural reactions, may be an idea we find attractive during philosophical reflection. It is perhaps in the context of such reflection that one is most likely to “feel confident”, with Kripke, that there must be something in our mind that that instructs us how to proceed when we follow a rule. But this idea finds no support from observation of the cases in which, in ordinary life, we think of people as participating in given rule-following practices. In particular, that a given person is evidently just doing what comes naturally to her in the circumstances is no bar to our ordinarily thinking of her as, say, obeying a walk sign, producing a simple arithmetical sequence, and so on.

At the same time, it seems to me that the appeal, if it is to be effective, must go hand-in-hand with a diagnosis of why the idea of the special mental process should be so tempting in the context of philosophical reflection. If the intuition articulated by Nagel is in tension with the judgments and reactions we have when, as in everyday life, we see a person’s actions against the background of the “whole hurly-burly of human actions”, then why are philosophers as insightful as Kripke and Nagel so gripped by that intuition? Presumably there is a fairly deep-seated mindset at work here. Until we know what that mindset is well enough to take the measure of it, no reminder about ordinary practices and customs is likely to block its influence.

This is a topic on which Wittgenstein could say much more than he does. He does make some diagnostic moves in *Philosophical Investigations* and elsewhere. For example, there is the discussion in the *Blue Book* of the philosopher’s “preoccupation with the method of science”,

understood as “the method of reducing the explanation of natural phenomena to the smallest possible number of primitive natural laws” (1960, p. 18). This preoccupation has “made the philosopher dismiss as irrelevant the concrete cases, which alone could have helped him to understand the usage of the general term” (1960, pp. 19-20). We might, in the present case, see this preoccupation as leading the philosopher to posit an underlying mental mechanism that unifies and explains the varied rule-following practices that we observe and participate in every day. Whereas for Wittgenstein it is our familiarity with the “concrete cases”—going by a sign, calculating, performing a dance, and so on—that constitutes our understanding of what it is to follow a rule.

But a diagnosis at this high level of generality is, I think, not fully satisfying. For it seems to me a natural suspicion that it is specifically the rational character of rule following—its location in the “space of reasons” as Sellars and McDowell say—that gives the philosopher trouble. It is plausible to suspect, in other words, that the intuition is at bottom an instance of the familiar difficulty of reconciling two conceptions of ourselves: as rational beings, and as natural beings. Wittgenstein’s hypothesis of a favoritism for the general over the concrete does not really make anything of this.

It would be good to follow out the suspicion just mentioned, but I will not do so here. This paper is already long enough.

## **10. Conclusion**

Let’s recap. Kripke’s talk of the normative relationship between meaning and use may be partially explicable as an expression of the evaluative view of meaning—the view that uses of linguistic expressions can either accord or fail to accord with their meanings. But that

interpretation of his normativity talk cannot explain his attitude toward anti-reductionism.

Moreover, Kripke ties the normativity talk to the justification requirement he imposes upon accounts of meanings, and that requirement, given that it is articulated as the demand to show that Kripke's responses to "+" problems are not mere stabs in the dark, cannot be understood in terms of the evaluative view. It flows rather from three premises. The first is the rationalist view of rules: to characterize a person as following a rule is to offer a rational explanation of her performance, with the rule serving as her reason for proceeding as she does. The second is the idea that rational explanations of what people do explain their performances in terms of their own conception of what justifies their proceeding as they do. The third is the view that use of language involves following rules laid down by the meanings of the expressions.

Now, Kripke assumes that viewing rule following as a species of rational activity (i.e., as performances susceptible to rational explanation) will require taking on board the guidance conception, according to which a rule-follower's grasp of a rule involves a mental item that guides her applications of the rule. This assumption can be understood as arising from our intuitive sense of the incompatibility of conceiving a rule-follower as acting on the basis of the justification she takes the rule to provide and conceiving her, à la Wittgenstein, as just doing what comes naturally to her in the circumstances given her training and background. But the guidance conception generates a dilemma. If we do not reapply the guidance conception at the level of the person's understanding of the guidance provided by the mental item constituting her understanding of the rule, then the conception turns out just to temporarily defer a description of the putative rule-follower as doing what comes naturally to her in the circumstances. Hence it does not after all provide an account of rule following that validates the aforementioned intuition of incompatibility. If on the other hand we do reapply the conception at this level, the dilemma simply reappears at the next level up. The pull of the conception can be so strong that we may

be led to a vision of occupants of the mental realm as regress-proof, in virtue of their ‘logically compelling’ our understanding of the guidance they provide. But this vision proves unintelligible. The skeptic concludes that rule following, and the application of linguistic meanings in particular, is impossible.

Is the argument for this conclusion sound? Suppose we accept, as I’ve suggested we should, the premise that rule following is a form of rational activity, coupled with the view of rational explanation that casts such explanations as appealing to the subject’s own conception of what justifies her performances. Suppose we accept, as I’ve been doing for purposes of exposition, that Kripke is right to assimilate the use of meaningful expressions to the following of rules. And suppose we accept, as I’ve argued we must, that the guidance conception generates the dilemma just summarized, and that the idea of ‘logical compulsion’, being a confusion, cannot help resolve the dilemma.

One issue then remains outstanding. Is the intuition that I have portrayed as the motivation for the guidance conception sound? Are, as Nagel worries, Wittgenstein’s “facial descriptions” of putative rule-following behavior incompatible with conceiving the person as genuinely following a rule? Wittgenstein’s appeal to the customs and practices that surround everyday episodes of rule following, and inform our ordinary understanding of those episodes, aims to challenge this intuition. That appeal would benefit from supplementation with a diagnosis of why the intuition, in the context of philosophical reflection, can seem so compelling. Wittgenstein offers something along these lines, but there is room for a good deal more.

### References

- Boghossian, Paul. (1989), “The Rule-Following Considerations”, *Mind*, 98: 507-549.
- Boghossian, Paul. (2003), “The Normativity of Content”, *Philosophical Issues*, 13: 31-45.

- Brandom, Robert. (2001), "Modality, Normativity, and Intentionality", *Philosophy and Phenomenological Research*, 63: 587-609.
- Darwall, Stephen. (1983), *Impartial Reason*. Ithaca: Cornell University Press.
- Davidson, Donald. (1992), "The Second Person", *Midwest Studies in Philosophy*, 17: 255-267.
- Gibbard, Allan. (1994), "Meaning and Normativity," *Philosophical Issues*, 5: 95-115.
- Gibbard, Allan. (2003), "Thoughts and Norms", *Philosophical Issues*, 13: 83-98.
- Ginsborg, Hannah (forthcoming), "Primitive Normativity and Skepticism about Rules", forthcoming in *The Journal of Philosophy*.
- Grice, Paul. (2005), *Aspects of Reason*. Oxford: Oxford University Press.
- Horwich, Paul. (1984), "Critical Notice on *Wittgenstein on Rules and Private Language*", *Philosophy of Science*, 51: 163-71.
- Horwich, Paul. (1998), *Meaning*. Oxford: Clarendon Press.
- Kripke, Saul. (1980), *Naming and Necessity*. Cambridge: Harvard University Press.
- Kripke, Saul. (1982), *Wittgenstein on Rules and Private Language*. Cambridge: Harvard University Press.
- McDowell, John. (1998), "Meaning and Intentionality in Wittgenstein's Later Philosophy", in his *Mind, Value and Reality*. Cambridge: Harvard University Press.
- Millar, Alan. (2004), *Understanding People*. Oxford: Oxford University Press.
- Nagel, Thomas. (1996), *The Last Word*. Oxford: Oxford University Press.
- Stroud, Barry. (1979), "Inference, Belief and Understanding", *Mind* 88: 179-196.
- Van Fraassen, Bas. (1980), *The Scientific Image*. Oxford: Oxford University Press.
- Wallace, David Foster. (2001). "Tense Present: Democracy, English and the Wars over Usage", *Harper's Magazine*: 39-58.
- Wittgenstein, Ludwig. (1960), *The Blue and Brown Books*. New York: Harper and Row.
- Wittgenstein, Ludwig. (1967). *Zettel*. Berkeley: University of California Press.
- Wittgenstein, Ludwig. (1978), *Remarks on the Foundations of Mathematics*, revised edition. Cambridge: MIT Press.

Wittgenstein, Ludwig. (1973), *Philosophical Investigations* (third edition). Prentice Hall.

Wright, Crispin. (1984), “Kripke’s Account of the Argument Against Private Language”, *Journal of Philosophy*, 81: 759-78.