Replies Part I: The Rule-Following Considerations and the Normativity of Meaning

Boghossian, Peacocke, Horwich, and Bilgrami

Following a Rule: Two Problems

I first read Wittgenstein's *Remarks on the Foundations of Mathematics* (*RFM*) in the summer of 1965. It took about a week of afternoons in the Fellows' Garden of Trinity College, Cambridge, during a spell of particularly beautiful weather. The effect was to turn my thinking about mathematics upside down. The initial focus of my interest had been on the question—crucial to an understanding of Mathematical Intuitionism, on which I was contemplating a thesis—of how one might coherently conceive of each instance of a general arithmetical proposition, say Goldbach's Conjecture, as determinate in truth-value yet question the validity of the Principle of Bivalence in relation to the general proposition itself. It had never occurred to me to ask what were the presuppositions, or hostages, of the idea that each such instance *is* determinate in truth-value, notwithstanding the human impossibility of an actual verification in all but a paltry initial segment of the countably infinite range of cases involved. But it became clear that the question had, in effect, very vividly occurred to Wittgenstein. I had not at that time looked at the *Philosophical Investigations* at all, and Wittgenstein's editors had (most unhelpfully) omitted from the version of *RFM* then available the sustained discussion of rule-following that, in the third edition, became section VI. But the general tendency of his thinking was nevertheless unmistakable. To regard each instance of Goldbach's Conjecture as determinate in truth-value, irrespective of any ability of ours actually to determine its truth-value—or even, for all but a small initial segment of cases, to come within range of understanding a formulation of the statement concerned—seemed to call for a kind of *objective reach* of meaning that would inevitably relegate our cognitive relationship to it, in understanding, to something not unjustly assimilated to Plato's conception of the strange yet convenient (putative) sensitivity of our minds to the requirements of the Forms. But if, to the contrary, we were to try to regard meaning as wholly grounded in socio-linguistic practices, rather than as somehow driving them from above, how might we make sense of this reach of meaning, to which the

conception of an arbitrary in principle decidable arithmetical statement as already determinate in truth-value was seemingly committed?

The issue is hardly as plain as day, but it struck me then, as now, as raising some of the most profound questions to which analytical philosophers are challenged to respond, fully deserving of the preoccupation—indeed, obsession—that it became for many philosophers in the last three decades of the twentieth century. It is rare, perhaps unheard of, for an important philosophical concern to spring, originally minted, from the mind of a single thinker, but I share Saul Kripke's opinion that in his musings on the issues in this vicinity, Wittgenstein in effect invented a new philosophical problem, or cluster of problems, for which—like all the deepest philosophical problems—it is extremely difficult both to capture exactly the right formulation and to fix on a settled response.

In my own early work, including *Wittgenstein on the Foundations of Mathematics*, the problem shaped itself as concerning the idea of what I termed *investigation-independence*: the idea that the meaning of a statement can, in collaboration with the facts, take care of its truth-value in a way that is constitutively independent of any propensity of judgment of ours. If one combines that, at first sight platitudinous and anodyne thought with the second apparent truism that we give our statements meaning not in the sense of annexing them to Platonic contents that already sustain investigation-independence but by constituting their meanings in our communicative practices, then we immediately confront the difficulty, prominent also in Kripke's discussion,[1] of explaining how it is that anything we do, or can do, amounts to the constitution of unique meanings, autonomously so conceived. This concern is certainly an important element in Wittgenstein's own discussion, but it should be noted that it is first and foremost a concern about linguistic content—though it spreads to psychological content and the relation between, for instance, the formation of an intention and what is subsequently adjudged to fulfil it. It is happily captioned as a problem about *rule-following* only insofar as the contents of rules, too, like all contents, come within its shadow—or because one is thinking of the meaning of an expression as akin to a rule for its correct use.

There is, however, a second aspect of the 'rule-following considerations' that only later, and gradually, began to distinguish itself for me. And this does indeed have to do with rule-*following* per se. The two aspects are contrasted in my paper on Chomsky[2] by the questions:

How can I tell which rule I (used to) follow? (Or: how is the rule that I used to follow to be identified?)

versus:

How can I tell what the rule I grasp requires of me *here*?

---

[1] Kripke 1982.   [2] Wright 1989, p. 243.

In that paper I suggested that Wittgenstein's thought about a regress of interpretations (*Philosophical Investigations* §§198–201) belongs with a difficulty attending the second question, rather than—as on Kripke's account—with the first. This difficulty, put in the most general terms is the following: if the transition from a rule 'in mind'—let it be, for the sake of argument, a fully determinate, if you like Platonic rule—to its (correct) application in a given particular context is a *mediated* transition: one bridged by supplementary cognitive activity, then that activity better not itself involve further rule-following, on pain of making all rule-following involve further rule-following, and hence regressive and impossible.

The distinctness of this issue is somewhat obscured in Wittgenstein's own exposition in those particular paragraphs, where the mediating activity is represented as one of *interpretation*, so that it is the question of the determinacy of the content of the rule that still seems to be in play. Nevertheless the second aspect is different; and crucially so. It will be my main focus in the following remarks.

Throughout the mid-to-late 1980s Kripke's book generated an avalanche of commentary, mostly expressly concerned only with a character called 'KW', and mostly ignoring most of the other commentary. Everyone seemed to want to define a personal escape route from the 'Skeptical Paradox'. After my paper on Chomsky, I felt I'd had my fill of the debate and took a sabbatical from the topic for a double septennium. It was a series of conversations and exchanges with Paul Boghossian[3] on Lewis Carroll and the epistemology of basic inference that stimulated me to think afresh about the issues and I returned to them in a paper I presented at the NYU Mind and Language seminar in 2002.[4] The principal new development in my thinking in that paper was a much clearer focus on the second question. The last part of the paper canvassed what is there termed the *Modus Ponens* model of rule-following, according to which a rule is conceived as a general conditional prescribing a mandate—that is, a requirement, permission or prohibition—for an action A:

If P, then !A

on the condition P, and the following of the rule as essentially an *inferential* transition from a state of cognition of the satisfaction, in the particular circumstances, of its antecedent P to a detached recognition of (or action upon) the mandate concerned. Evidently, the operation of such a rule is going to require an anterior capacity of judgment whether or not P. But if such is the basic form that anything properly regarded as rule-following has to take—the suggestion of the paper was that it is hard to see any alternative—and if we propose to think of all competent linguistic practice as consisting in rule-following, the question looms how to conceive of the rules

---

[3] See e.g. Boghossian 2000, 2001. Relevant papers of mine are Wright 2001, and Wright 2004.
[4] "What is Wittgenstein's Point in the Rule-following Discussion: Five Themes", referred to by both Boghossian and Paul Horwich in their contributions to this volume. A descendant of this paper was subsequently published as Wright 2007.

governing the most basic and primitive expressions of the language, or indeed any expressions an accurate statement of whose satisfaction-conditions would have to take a homophonic form. Suppose, for example we conceive of our respective uses of 'pain' as governed by, if by any rules, then, inter alia, by the rule:

If in pain, then assent to "I am in pain".

Then the cost of thinking of competent operation with such expressions as a form of rule-following is going to be that grasp of the concepts they express must be viewed as cognitively *anterior* to mastery of the relevant expression—exactly the conception of linguistic competence, resting on a prior capacity for fully articulated thought, which is expressed in the passage from St Augustine with which the *Philosophical Investigations* begins and from which, as I put it, that text is a "journey of recoil".

The conclusion is a dilemma: either linguistic competence is not, *au fond*, a matter of rule-following, or some quite other model of rule-following is demanded. Christopher Peacocke's and Paul Horwich's chapters in this volume may be seen as attempting to outline (very different) responses that embrace the second horn. I'll come to them shortly. But the problem is not confined to understanding the sense, if any, in which basic linguistic competence can be rule-informed. Indeed, as Paul Boghossian's essay forcefully brings out, it seems to embrace not merely basic linguistic competence but rational thought itself.[5]

The Regresses

To bring the issues into focus, let us step back and try to say at an intuitive level what it is to follow any rule. Suppose you are playing chess and wondering whether now is a good time to castle. Is it permissible to do so? The rule for castling says, roughly:

If neither the king nor one of its rooks has so far moved, and if the squares between them are unoccupied, and if neither the king nor any of those squares is in check to an opposing piece, and if the king and the rook in question are on the same rank,[6] then one may castle, that is, move the king two spaces towards the rook and . . .

Your *following* this rule, presumably, will consist in your recognizing what it has to say about your present situation in the game and then acting in the light of that information. So, for example, you reflect:

(Minor premise)   In this game so far, neither my king nor queen's rook have yet been moved, the squares between them are unoccupied, . . . etc.

And you draw the conclusion:

I may castle now.

---

[5] See also Boghossian 2008.
[6] Exercise for reader: do you see why it is necessary to include this condition?

Several points are salient. First, as in the Modus Ponens model, the rule is given as a general conditional content, whose antecedent portrays the kind of circumstances (the *triggering* circumstances) in which the permission articulated in its consequent is activated. Second, the ability to *follow* the rule involves knowing what this conditional content is, and attempting to uphold it—to avoid making moves that "break" it, as we are wont to say—in the course of the game. Of course you may succeed in this attempt without in any way overtly distinguishing yourself from someone who is ignorant of the rule for castling but can otherwise play chess, namely by never availing yourself of the permission that the rule grants. So—a familiar but vital point—we may not equate following a rule with acting in the manner that it requires: conforming to a rule is one thing, following it something different.[7] Third—or so it is natural to think, though this will prove to be a crucial and controversial point—following the rule essentially involves (very elementary, perhaps unreflective) *reasoning*. The rule supplies the major premise for what is, in effect, a step of modus ponens, the minor premise being supplied by your observation of the board configuration and memory of the game so far. Fourth, the output of this minor feat of information-processing is a judgment—"I may castle now"—on which you may or may not act. That you are following the castling rule explains your making this judgment by giving your reasons for it; and if you act on the judgment, that too will be rationally explained by your following the rule in question. Finally, the rule is normative in the sense that both the output judgment, and any action upon it, are assessable as *correct* or not in the light of the rule: you may be mistaken in thinking that the trigger conditions are met, or you may correctly judge that you may now castle but bodge the actual move (maybe you inadvertently move the king only one space), but in either case, the propriety of describing your move as incorrect depends on bringing it within the purview of the rule—that is, on viewing it as an attempt to follow the rule.

Here are the key points again. Rules are general conditionals, connecting a trigger condition with a mandate. Following a rule requires somehow *accepting* its content and subjecting one's practice in relevant respects to its requirements. What one then does is assessable as *correct* or not according as it squares with the output of the rule for the circumstances in which it is done. This output is a matter of entailment, and following the rule accordingly involves *somehow* tracking this entailment—so presumably inference. Rule-following is rational activity: that one is following a particular rule contributes to the *reasons* one has for doing what one does, and in part rationally *explains* one's performance.

I said "so presumably inference". But that is going to be a key issue. Provided we stall on any firm commitment to the role of inference, this model—the italics in the previous paragraph lend themselves to the acronym, "the ACRE model"—is effectively that proposed by Boghossian in his present chapter. And it surely hits off one very

---

[7] Note that I do not say, "something more". You can be following a rule yet fail to conform to it, as a result of a mistake, inattention, and so on.

intuitive notion of rule-following. But how extensively can it be coherently applied? All believing, we tend to think, is assessable as rational or not. More, it's intuitive[8] to suppose that all believing is potentially rational or not at the point of ingress, so to speak: that a belief's being rationally managed is one question and its being rationally entered into is another, and that each of these dimensions of constraint engages any particular belief of any particular thinker. What Boghossian calls the "Rulish" picture of belief essentially has it that both dimensions of rationality are matters of successful rule-following. Entry-rationality, in particular, will consist in the following of correct epistemic rules in the formation of the belief. Can this idea be sustained under the aegis of the ACRE model of what rule-following consists in?

Well, we have noted that there will be a bind on any account whereby the transition from any rule to its (correct) application in a given particular context is a *mediated* transition: one bridged by supplementary cognitive activity involving further rule-following. Wittgenstein's "regress of interpretations" is one illustration of this pattern and now we have set ourselves up for another, as Boghossian notes.[9] For the ACRE model, harmless and truistic though it seems at first blush, *does* seem to require that rule-following essentially involves inference. How are we to understand the model except like this: that a rule connects with the practices it rationalizes and explains by providing a general instruction—hence a premise—for conclusions about the mandates it issues in particular sets of circumstances, so that following it involves drawing and recognizing such conclusions. But drawing conclusions is inference, and inference, surely, *is itself a kind of rule-following*—specifically, the following of rules of inference! In following the rule for castling at a particular situation in the game of chess, I in effect reason from the rule as premise, together with a description in relevant respects of the circumstances of the game, to a conclusion about whether I may castle at that particular juncture. But in drawing this conclusion, I also follow the rule of modus ponens—and according to the ACRE model interpreted as incorporating inference, this secondary act of rule-following will itself involve reasoning from a set of premises, one of them a general statement of the rule of modus ponens, to a conclusion—presumably, that the drawing of the original conclusion is admissible—distinct from the original conclusion about the admissibility of castling. And now this inference in turn, once assimilated to an act of rule-following and subjected to the ACRE model so understood, calls for a further inference—still with a statement of the rule for modus ponens as its major premise, but with a differing minor premise and a further novel conclusion, viz. that the drawing of the conclusion of the second inference is admissible. And so on. The regress is vicious, since the inferences concerned all involve the drawing of different conclusions. In brief: if, as may seem wholly natural, even unavoidable, it is built into the ACRE model that following a rule always involves inference, and if inference is always a form of following a rule, then it follows that any act of rule-following requires another distinct from it, and not relevantly easier than it. Moreover every act of

---

[8]   Though of course denied by some distinguished philosophers, including Quine and Davidson.
[9]   This volume, pp. 40–1; see also Boghossian 2008, p. 492 and following.

rule-following in the series thus generated is distinct from all those that precede it in the series. That makes following a rule an example of a "supertask". We cannot perform such tasks.

Call that the *Inference* problem. It's bad, but it is not the full extent of the woes. As we in effect already noted in connection with the Modus Ponens model and St Augustine, there is also a *Minor Premise problem*. The ACRE model, inferentially interpreted, involves that an epistemic rule connects with a judgment it rationalizes and explains via a minor premise, that its trigger condition obtains. However the model won't *rationalize* the output belief unless this premise too is accepted in whatever sense the rule itself is accepted. If my judgment that it is legitimate to castle now is to be rationalized and explained by my following the rule for castling in terms of the ACRE model, then I need to be credited with a judgment about the trigger conditions—the board configuration and the history of the game—as well as an acceptance of the rule. But any old judgment? The Rulish picture has it that all rational belief is belief formed by following basic epistemic rules. Yet however epistemically virtuous the rules concerned, the quality of the beliefs that they deliver must obviously be hostage to the quality of one's judgments of the truth of the minor premises to which they are applied. It is only if the minor premises are themselves rationally accepted that the application of the rules will guarantee the rationality of the beliefs thereby formed. But if acceptance of the minor premise is to be rational, and if all rational judgment is conceived as entered into in the mode depicted by the ACRE model, specialized to epistemic norms, then another regress is launched: to arrive at any rational judgment, I must first make another, independent rational judgment.

So this nexus of problems is the second aspect of the 'rule-following considerations' that I referred to—a second fundamental form of paradox, additional to and independent of the problem of scepticism concerning the determinacy of rules emphasized in Kripke's exegesis and my own first book. The core problem is that three very natural assumptions about what rule-following consists in, and about inference, viz.

> that the essence of rule-following is captured by the ACRE model;
> that the ACRE model requires that the application of rules everywhere involves inference; and
> that inference is itself a form of rule-following;

are jointly inconsistent with the very possibility of finite beings' following a rule—*any* rule. In consequence, if rational belief-formation is a matter of the following of (correct, or rationally acceptable) epistemic rules, then rational belief-formation is also impossible. And if basic linguistic competence is a matter of the following of rules that constitute the proper use of the language concerned, then there is no such thing as basic linguistic competence.

A less hysterical conclusion is this: if each of those three theses is faithful to *our actual concept* of rule-following, then it will indeed be impossible to follow a rule, of any kind, so long as doing so is thus conceived. The challenge, then, would be to explain what to

give up, how to refashion our concept in such a way as to conserve at least something of the central role we ascribe to it in linguistic competence, intentional action, rational belief-formation and management, and the subjection of multifarious of our practices to standards of correctness.

At the conclusion of his present chapter, Boghossian tentatively suggests that, natural though it undoubtedly is, the idea to expunge may be that applying a rule has to be a matter of inference from the rule.[10] I agree that we must leave no stone unturned in the presence of paradox, but that suggestion seems to me too difficult. How can someone be properly said to be *following* a general rule through a series of cases if she is not in effect instantiating its content to them in what amount to successive inferential steps? Boghossian himself was surely right to stress that a rule-follower must in some sense *accept* any rule that she is following, and surely in following it she must somehow *process its instructions* in such a way as to bring it to bear on new specific cases? Is there any intelligible model of such processing that does not involve inference? Of course, how exactly we should understand the notion of inference is itself a vexed issue. But the point is not whether we call this "processing of instructions" inference or not, but that, however described, it cannot be itself any form of rule-following or we will in effect reinstate the first of the two regresses. Nor can it in general involve processing a judgment that the trigger condition for the rule is satisfied or—at least on the Rulish picture of rational belief—we will reinstate the second of the two regresses.

A more plausible direction is that the mistake—or the thing we need to revise—is the supposition that inference is everywhere and essentially a kind of rule-following. In that case, we can allow the inferential interpretation of the ACRE model and avoid the first regress. But still the point will remain that in order to lead to rational belief, any inference—whether it involves rule-following or not—will still require rationally accepted premises. So more would still need to be said to staunch the second regress. And if inference is not a matter of rule-following, then presumably the Rulish picture of rational belief will be wrong for some kinds of inferentially formed belief. So in what sense will such beliefs be rational? And in any case, if inference is not (always) a case of rule-following, what is it when it is not?

Neither Paul Horwich nor Christopher Peacocke directly addresses this cluster of problems in their chapters, but the views they respectively defend represent nonetheless two of the most salient directions of possible reaction. Horwich would be content, I surmise, with the ACRE model as an account of what he terms *explicit* rule-following, but he contrasts that with the wide range of cases—par excellence, rational belief-formation and language mastery—where normative constraints seem to be in play and we are attracted to the idea of a kind of *implicit* rule-following. And of this idea he

---

[10] That rule-following might, in the most basic case, consist in actions for which reason is somehow directly provided by a state of acceptance whose content is that of a (general) rule, without intermediary inferential processing—that such a connection may be a primitive ingredient in rational action—was apparently mooted by Tyler Burge in conversation and is the hesitant parting shot in Boghossian 2008.

moves very quickly to a deflated account, whereby the ACRE model's core idea of *acceptance* of the rule—of the performances in question as sourced in an information state that generates mandates over performance—is simply dropped in favor of a conception of performance as governed by ideal laws only in the manner in which patterns in nature are governed by covering laws in natural science. Stones do not carry the information conveyed by a statement of the laws of gravitation to which their motion conforms. If acceptance is essential to rule-following, then Horwich offers regularity, not rule-following. The idea that the performances concerned are nevertheless subject to normative constraint is grounded, in Horwich's view, not in the existence of mandates issuing from internalized rules but in our customary practice of holding them open to appraisal, qualification, withdrawal, and apology. This critical aspect of our practice is likewise subject no doubt to meaning-constituting covering laws; but is not informed by the content of those laws in the mode of explicit rule-following.

Horwich's primary concern in his present contribution is with meaning, rather than rational belief. But one can foresee the outline of the corresponding view of rational belief. If entry-rational beliefs are properly required to be beliefs sustained by the processing of reasons for them, under the aegis of correct epistemic norms whose contents are represented in the mind, then the view will be that we should give up on the notion and accept that most of our beliefs are, simply not entry-rational. Their formation may still present an example of implicit rule-following in Horwich's deflated sense: the patterns of their formation may be subject to natural law, and their acceptance and rejection held subject to criticism and correction. But the rationality of a belief will not in general equate with its being *formed* in the light of rational constraints.

On such a view, the ground is cut from under the regressive paradoxes. I shall later touch on the matter of whether it involves surrender of anything that we might understandably wish to keep.

Tacit Knowledge

Peacocke's proposal is quite another thing. While his discussion is primarily directed at the other half, so to speak, of the rule-following problem: that connected with meaning- and content-scepticism, and at the development of a framework in which investigation-independence can be saved without any epistemically problematic Platonization of rules, meaning, and concepts, the apparatus he outlines is nevertheless a prototype of a second salient form of response to the collapse of the generalized ACRE model. This form of response is to propose that the generalized model fails not by dint of misconstruing entry-rational belief as an effect of our processing the requirements of norms of rational belief-formation, somehow internalized, but in a misunderstanding—encapsulated in the exposition of the model in personal-level terms—of the nature of the information-states involved and the thinker's relation to them. For Peacocke, the crucial step to a coherent view of the matter is their reconstrual as states of *tacit knowledge*.

One crux here lies in the interpretation of 'tacit knowledge'. In an early exchange with myself about rule-following and semantic theory,[11] Gareth Evans too took recourse to the idea of competent speakers' states of tacit knowledge of the axioms of a semantic theory for their language. But for Evans, talk of such states stood to be interpreted dispositionally, specifically in terms of dispositions to understand whole sentences in certain ways. A speaker's tacit knowledge, for example, that "London" refers to London was identified with a disposition to construe the truth-conditions of sentences containing "London" in referential position in certain ways. Semantic theory, on Evans' view in that paper, was to be construed as a purely descriptive enterprise, specifically as the project of giving systematic, compositional syntax-based descriptions of speakers' pattern of interpretation of (novel) sentences in their language. This conception was precisely intended to contrast with the project of systematic formulation of a body of information whose realization in the cognitive architecture of a speaker would somehow explain his mastery of the language concerned—the conception of semantic theory to which, in my half of the exchange, I had worried was destined for a losing contest with Wittgenstein's discussions of rules.

It is otherwise with Peacocke. On his account—moving now to rational judgment rather than language mastery—the invocation of tacit knowledge is precisely aimed at explanation where Evans was content with description.[12] The invocation of tacit knowledge is to assist with the explanation both of what entry-rationality consists in and of how thinkers are enabled to arrive at entry-rational judgments.

According to Peacocke, concepts are identified by their "fundamental reference rules" and grasp of a concept consists in tacit knowledge of such a rule for it. For example, the fundamental reference rule for the observational concept *round* is, he offers, that something falls under it just if it has the same shape as things are represented as having when they are experienced as round.[13] In moving to an entry-rational non-inferential

---

[11] Evans 1981.

[12] I am simplifying some of the subtlety in Evans' discussion. Evans was clear that we should understand the disposition associated with a semantic axiom as a *single* state lying at the causal source of its manifestations in a speaker's linguistic practice; and clear too that in order to legitimate thinking of the relevant dispositions that way, it was crucial that there should be additional evidence, going beyond the patterns in their construal of whole sentences, for speakers' possession of such dispositional states. He anticipated that such additional evidence would be forthcoming from patterns of acquisition (on learning) and decay (say, in brain-damaged subjects) of whole-sentence competences. The important point is, however, that even with dispositions robustly so conceived, the theory of meaning, as conceived in his paper, is still a descriptive project, with no answer essayed to the question, "How are speakers able to parse novel utterances?" It is a further step—indeed, a jump—to the thought that the states concerned should be conceived specifically as *information-bearing* states, rather than so-far unspecified states with the appropriate causal powers. (Some of the issues here connect with the contemporary debate concerning knowledge-how and information that has been stimulated by Stanley and Williamson 2001.)

[13] There are, naturally, major methodological questions about how such rules are to be identified. Peacocke offers several examples: the fundamental reference-rule for the perceptual demonstrative *that F*, for instance, is alleged to be that where "the apparently perceived object is given in way W in perception", a thinker's tokening of it will refer to the unique object perceived in way W by the thinker at the relevant time. These proposals seem, by comparison with the familiar kind of homophone risk-free formulations of 1970s semantic theory, to involve substantial reflective analysis, executed in the armchair, and one might

judgment—say, that [that object is round]—a thinker *draws on* his tacit knowledge of the fundamental reference rules for the concepts involved, possibly in combination with further information, thereby arriving at an appreciation that a certain state or states—perhaps an experience of the object as round—gives reasons for thinking that the relevant judgment is true.[14] This 'drawing' and 'arriving at an appreciation' is psychologically real, but it is not, Peacocke is quite explicit, a matter of conscious inference or thought. "There is a subpersonal transition from the content of the tacit knowledge to the appreciation that a certain kind of perceptual experience makes a perceptual application of the concept *round* rational in the given circumstances."[15]

In sum: the judgment that P is entry-rational when it is made on the basis of a state S that is recognized to provide reason for P by a suitable derivation from the fundamental reference rules for the concepts in P, together perhaps with other information. The relevant S may consist in other beliefs, and the judgment that P may then be inferential.[16] But it may also consist in an episode of sense experience, or of apparent memory, or perhaps a 'finding obvious'.

One immediate question is how this kind of model is supposed help with issues of content-scepticism and investigation-independence. Obviously it cannot do so unless the place occupied by a fundamental reference rule in the kind of information processing that Peacocke envisages is filled by something whose content is determined in ways that are proof against the familiar sceptical arguments. Yet about this crucial issue Peacocke's discussion impresses as more than a little casual. He is right, of course, that the content of a subpersonal state is not something that we need to constitute by linguistic practice, nor fix by our thoughts. But the question still arises, what factors *do* contribute to the fixation of the content of such a state, and how determinate a content are they empowered to fix? Peacocke raises the issue only in the closing paragraphs of his essay, and briskly canvasses the idea that an account based on the causal explanatory role of such states has "reasonable prospects". One might well be pessimistic about that. A cardinal part of the causal explanatory role of such states, if Peacocke is right, is in explaining the formation of content-bearing attitudes and the utterance of content-expressing sentences. It is quite unclear how any account of the causal explanatory roles of the relevant tacit knowledge states could succeed in fixing determinate contents for them if this aspect of their roles was to be left out of account. But, if it is to be taken into account, how is any path back from these explananda to the content of the underlying

---

wonder if that methodology can be reconciled with their role as the putative contents of *tacitly* known states. Such concerns are taken up in Peacocke 2008 but I cannot pursue the issues here.

[14] "The content of the tacitly known reference rule itself also contributes to an explanation, non-causal, of why the experience of an object as round is a good reason for the judgment that the perceptually given object is round, when the thinker is entitled to take experience at face value." This volume, pp. 55–6.

[15] This volume, p. 56.

[16] It won't be, though, if first-order beliefs can rationalize their second-order counterparts—a kind of model of intentional psychological self-knowledge that Peacocke himself has long advocated.

tacit knowledge states supposed to work unless *fully determinate* conceptual contents are already ascribed to the former?

I won't pursue this matter further here, however. The issue I wish to consider in more detail is how Peacocke's model might help with the regresses. Obviously enough the Minor Premise problem is addressed immediately: the model is explicit that the triggers for rational movement to new beliefs need not themselves be beliefs, nor any states for which the question of antecedent reason arises in the way required to launch the Minor Premise problem. But the situation with the Inference problem is less clear. While the ACRE model represents the cognitive processes involved in rule-following in personal terms—the states of mind and movements of thought it posits are those of a fully intentional subject—there is no immediate prospect of assistance with the Inference problem merely in the suggestion that we should recast the model in terms of subpersonal processes and states. As Boghossian observes, if subpersonal rule-following involves subpersonal inference—as on Peacocke's account it seems it does—and subpersonal inference is itself a form of sub-personal rule-following, then the threat of the regress remains. Subpersons—or anyway, finite populations of subpersons—are no better at supertasks than persons are. The Inference problem will not be helped by retaining the *structure* of the processes of rule-following that the ACRE model postulates and merely insisting that the model must sometimes be realized subpersonally rather then personally. If going subpersonal can help, it will be because it somehow allows us to refashion that structure.

So, does it? The question turns on the extent to which the subpersonal operations involved in the acquisition of rational belief on the kind of account proposed by Peacocke are themselves to qualify as rational, and what it can legitimately mean so to describe them. Let's prescind for a moment from the fundamental reference rules central to Peacocke's account, and consider the matter in the most abstract terms. Let S_1 be a state of *registration* (subpersonal acceptance, or programming) of the rule:

If P, then !A

where A is a type of action and '!A' expresses some form of mandate for a performance of A. Let S_2 be, correspondingly, a state of registration of an instance of the trigger circumstance, P.

Suppose Agent starts in these two subpersonal states (and is in no other state which is material to the rationality of his doing A) and then does do A. Does he act rationally? In one—deflated—sense, yes, of course: he does something for which he has reason, something that is rationally sanctioned by his subpersonal states of information. But that is not the sense of 'rationally' that interests Peacocke. Peacocke requires that in order to qualify as rational, specifically: as the rational formation of a belief, an action must be *sourced* in the information states that provide its mandate, even if the sourcing is to operate entirely at the subpersonal level. The question is whether this requirement can be fleshed out without generating an analogue of the Inference problem.

Let's go carefully. For Agent's token of A to be so sourced, it needs to be an action *on* an information state S_3 of his cognitive system which registers that there is the mandate:

!A. (No doubt that notion—what constitutes action *on* a mandating information state—may prove troublesome under philosophical pressure, but I am going to take it for granted here.) The question is: how did Agent *get into* S_3? Well, if he did so rationally, it must have been that he in some sense 'transitioned' from S_1 and S_2 to S_3. So let $A^\star$ be his token of that transitional act type, and ask: was $A^\star$ performed rationally? Again, in the deflated sense, yes. The movement from information states S_1 and S_2 to S_3 is licensed by the entailment of the content of S_3, viz. [!A], by the contents of S_1 and S_2, viz. [If P, then !A] and [P]. But again, the requirement is not merely that the transition, $A^\star$, be one which is licensed by an entailment, and is in that sense of a rational type in the context, but that Agent's *token* of it be rational: that it have been performed *because of* that entailment. And that in turn requires that there have been an information state S_4 of Agent's cognitive system which registers that there is a mandate—!$A^\star$—for the transition $A^\star$, and that the token transition have been an action on *that* information state. Suppose so. Then the question is: how did Agent *get into* S_4? Well, if he did so rationally, it must have been that he in some sense transitioned from ulterior states whose contents entail the content of S_4. Let $A^{\star\star}$ be Agent's token of that transitional act type and ask: was $A^{\star\star}$ performed rationally? . . .

So there is still a regress. Going subpersonal doesn't help, at least on the assumptions made. Those assumptions were: (i) that any token action of the system, including in particular any particular transition from one set of information states to another, is rationally performed only if mandated by a suitably specific information state; (ii) that the achievement of such a specific information state will in general be possible only by a transition from states encoding relevant (general) information to the specific information state in question. Under these assumptions, one token rational transition demands the prior performance of another . . .

It does not seem to me likely that (ii) is negotiable. If that is right, then going subpersonal can assist with the Inference problem only if we relax the requirement that a movement of the system counts as rational only if it is made on a suitably mandating specific information state. We have to allow that some of the informational transitions in a rational system are merely *licensed*, in the sense above—that they merely *conform to* rational pattern—rather than being *mandated*, in sense of: made because the system has entered into an information state that encodes the relevant license, on which it then acts. In a nutshell, a rational system must, at some level exhibit mere regularities of rationality—brute rationality—rather than any subpersonal analogue of personal rule-following.

This brings us to the brink of what should be a much fuller discussion, which I cannot undertake here, but let me pursue the matter just a little. It seems to me that the immediate effect of the foregoing reflections is to blow away the motivation for the thoroughgoing rationalist position that Peacocke attempts to take. We have to acknowledge that psychological representations, whether personal or subpersonal, of principles which we would like to regard as rules of rational belief-formation—rules of logical inference, and the kind of ampliative rules of transition which philosophers have been tempted to postulate for perceptual justification, induction, and perhaps

self-knowledge—cannot coherently be regarded as always playing an indispensable part in the cognitive economy of a rational thinker who moves to rational beliefs in the ways they prescribe. Peacocke's model of rational belief-formation involves play with principles additional to these, viz. his fundamental reference rules for the concepts involved in a judgment which a rational thinker allegedly processes in arriving at an 'appreciation', as he puts it, that a certain state constitutes sufficient reason for acceptance of the judgment. This appreciation should be thought of as, in our terms, an information state mandating acceptance of the judgment. But since we know that some of the rational transitions involved in the accomplishment of such a state must be *brutely rational*, without the source of mandates achieved by processing via representations of the relevant rules, why believe, or require, that it is any different with fundamental reference rules? Why should they feature in the cognitive architecture of a rational thinker, as items of his tacit knowledge, rather than merely encode patterns of belief-formation which, no doubt as a result of training, he is disposed to display? Why does there have to be any cognitive state mandating the making of the judgment at all?

The doubt tends in the direction of the sort of view I take to be Horwich's, where the conception of a rational thinker as one whose judgments are characteristically reached by processing of information states encoding the requirements of rationality is supplanted by that of a thinker who is merely characteristically disposed to judge what it is rational for him to judge, for who-knows-what cognitive psychological or any other cause. And clearly, purely as far as the thinker's performance is concerned, there is not going to be anything to choose between a Peacockian form of explanation of his judgments, and another—*austere dispositional*—model which will continue to find work for the idea of registration of trigger conditions, but, rather than view the consequent transition to belief as driven by an apparatus of cognitive psychological states encoding the content of logical, epistemic, and fundamental reference rules, views it simply as the exercise of a bare disposition to conformity with the rules concerned. Once again: dispositions in general do not call for explanation by means of states of a system conceived as encoding the content of rules prescribing the kind of performance that the disposition in question is a disposition to execute. Why should it be any different with the causation of what we regard as rational belief-formations?

What Peacocke's programme appears to need is a well-motivated distinction in the respective roles to be played by epistemic norms and fundamental reference rules. It has to be acknowledged, on pain of the regress, that in order to contribute to the rational formation of beliefs, transitions in the cognitive processing of a rational thinker do not need to be mandated by specific information states sourced in tacit knowledge of appropriate epistemic norms. I take that to be proved by the reflections above. But it can still consistently be insisted that specific judgments made at the end of the line, as it were, do require a mandating information state, and that it must be achieved by suitable processing of fundamental reference rules if it is to subserve genuinely rational belief. But *how* might this distinction be motivated?

There is, indeed, a further concern: on Peacocke's view, the rationality of a thinker's judgments now becomes a question of the structure of the *actual scientific explanation* of his making them, no matter how impeccably rational they impress, in ordinary intellectual commerce, as being. Once we recognize that we are stuck with the idea that certain kinds of informationally unconstrained tendencies are an essential part of rational judgment, it seems to be an open scientific question how far that phenomenon—absence of informational mandate—extends. The austere dispositional view says: why not all the way? To insist that that answer be resisted, is to endorse a conception of rational judgment as not merely an ability but an ability that needs to be grounded in a certain way in order to count—as if adroit shot selection at tennis had to be grounded in a certain kind of cognitive processing before it counted as a genuine skill. This, it may be contended, is to confuse the (possible) explanation with the datum to be explained. Peacocke's conception leaves our rationality—contrast: our behaving *as if* rational—hostage to empirical scientific (cognitive psychological) fortune.

But there is also, it needs to be said, a very significant cost to the drift towards austerity. Indeed, I think the dialectical situation here brings out real tensions in our basic thinking about human rationality. In other spheres of good performance—for example, acting well ethically—the contrast between doing the right thing and doing it for the right reasons is paramount. There is all the moral difference in the world between action for the right reasons and action of the same type, in the same context, on quite other or no good reasons. It is an analogue in the epistemic sphere—in the ethics of belief—of this idealized conception of fully responsible moral agency that is the casualty of the regressive argument. The austere dispositional view embraces the limitation. But what is lost thereby is akin, in the doxastic sphere, to loss of the kind of value associated with courage in contrast to recklessness. That is a major loss, and it would be a prize worth having if there were some principled way to limit the damage.

The Normativity of Meaning

The normativity of meaning is a lynchpin of Kripke's Wittgenstein's 'Skeptical Argument'.[17] It is, for example, because meaning is supposedly normative, and because that is taken to involve that understanding involves sensitivity to norms, that a dispositional account of meaning—the natural avenue to explore after the apparent failure of antecedent use and mental facts to determine meaning—is allegedly foreclosed.[18]

---

[17] I shall use the American spelling, and the corresponding 'Skeptical Solution', whenever referring to the dialectic in Kripke's exegesis.

[18] It is true that the Kripkean argument also makes play with the finitude of our dispositions, in contrast with the presumed infinitary determinacy of meaning, manifest in, for example, the determinacy in truth-value of the infinite range of quantifier free arithmetical equalities. But this is actually just the same point again. For meanings to be determinate in infinitely many cases is for them to contribute towards the determination of the truth-values of infinitely many sentences, irrespective of our dispositions of assessment of those sentences. The truth-values so determined will then supply a standard of correctness for those dispositions.

Dispositions can at best determine what we are inclined to say; meanings, by contrast, determine what we *ought* to say.

The claim has been widely accepted since Kripke's discussion. But it is not incontrovertible. Certainly, the concept of linguistic meaning is inextricably linked with that of *correctness* in linguistic practice. It is only utterances that have meaning for which there is any such thing as correct or incorrect use; and whether an utterance is correct or not depends on what it means and how matters in the world stand in relevant respects. But this is not enough to ensure that meaning is normative in more than a deflated sense. A conception of normativity serving merely to respect these correctness platitudes needn't amount to anything that would pre-empt dispositional or other forms of naturalistic construal. Indeed, the Skeptical Solution itself, whereby the letter of the platitudes is (as intended, anyway) saved, illustrates the point. The idea of the normativity of meaning presupposed in the Skeptical Argument is a full-blown conception, according to which grasping a meaning is a state of information involving internalization of a *standard* of correctness, a principle that can be used in conjunction with an appreciation of the relevant facts to *determine* whether or not (there is reason to regard) a token utterance is true. So regarded, meaning stands to correct linguistic performance precisely in the role of a rule to a practice it governs. The moral of the preceding section of these remarks, however, was that it is incoherent to regard the operation of standards of correctness as everywhere involving their internalization in determinate information states of practitioners. That conclusion was forced for the case of rational belief-formation. But in parallel, there has to be space for a conception of what it is to understand an expression which centralizes the ability to make correct use of it without viewing it as informed by an internalized rule determining what is and isn't correct use.[19] The normativity of meaning, on the full-blown interpretation, is no platitude. It is contestable, and is indeed contested in his present chapter and elsewhere by Paul Horwich.

Meaning-Intentions as Degenerate

Let me, though, before turning to Horwich's proposals, respond to the very different challenge to the normativity of meaning developed by Akeel Bilgrami. In commentary on Kripke's book written shortly after its appearance,[20] I proposed that the letter of the Skeptical Argument might be addressed simply by invoking our intuitively immediate

---

[19] It is a corollary of this that Kripke's Skeptical Argument, if successful, undercuts, in a way, its own objection to dispositional views. It is the full-blown interpretation of the normativity of meaning that is beyond dispositional construal. Once that interpretation is abandoned, there is space for broadly dispositional accounts that place normativity not at the source but in the character of the practices—retraction, criticism, endorsement, and so on—that the meaning-constituting dispositions are dispositions to engage in. If that is correct, then the distinctively non-factualist tenor of the Skeptical Solution is unmotivated. Although he finds no use for the notion of a disposition in his proposals, I take Horwich's proposal to make essentially this move.

[20] Wright 1984.

self-knowledge, in the ordinary case, of our own intentions—and hence that the real issue raised, when one is challenged to adduce facts in virtue of which one means (or meant) one thing by an expression, rather than another, is to understand better the nature of immediate (or recollected) intentional self-knowledge and how it is feasible. Bilgrami accepts this broad thought, identifying or assimilating what it is to mean an expression in a particular way with a kind of intention—what Boghossian in his contribution to this volume calls the Intention View[21]—at least for the purposes of his present essay. Normativity comes into question, in Bilgrami's view, not because the assimilation of meaning to an intention, or any other intentional state, is in question, but because the relevant intentions, the ones that look best suited to determine meaning, are, Bilgrami contends, a "degenerate" case of intention, and in effect *non-normative*.

A natural candidate for a meaning-constituting intention for, say, a predicate would be the intention to respect a certain satisfaction-condition in one's use of it—for instance, to apply "square", or to acknowledge it as correctly applied, just to square things. That is an intention one could fail to live up to. But as a proposal for what it is to mean *square* by "square", there is, as Bilgrami observes, a difficulty with it.[22] We do not want to make it a condition of understanding "square" that a speaker intend to be *honest*; or—if this is something he could anyway strictly intend—to make no mistakes. Even a persistent failure to apply, or to acknowledge as correct applications of "square" to square things need not enjoin that a speaker has failed to use "square" with the meaning: *square*. The explanations of the aberrant uses may lie elsewhere. If meaning so-and-so is an intention-like state, it is, at any rate, not the intention to use the expression correctly, or to comply with certain standards, or uphold certain regularities in one's use of it but rather, as an approximation, to accept that one's uses are properly assessable by those standards, or open to assessment as correct or not according to whether they maintain those patterns. Meaning *square* by "square" is thus better assimilated to the intention to allow that one's uses of "square" are *answerable* to the

---

[21] Boghossian argues that the ultimate objection to any Intention View is that it must succumb to the Inference problem. This needs a more extended discussion, but let me observe that it cannot in general be correct to think of action upon intention as involving any kind of inference. Whatever the conclusion of such an inference, one would still have to act *on* that; there simply has to be some notion of inferentially unmediated action on an intentional state if ordinary psychological explanations are ever to serve to rationalize an agent's behavior. And once that is granted, why cannot ordinary intentions be such directly action-connected states?

Boghossian's point, however, is not so easily finessed. For in the case of meaning-constituting intentions, the contents concerned will presumably be *conditional and general*: the intentions in question will be intentions to abide by a rule, or maintain a certain pattern in one's linguistic practice, and it is unclear what sort of model might be given of action upon these intentions without involving processing of triggers and simple inference.

I am inclined to stand by the proposal that the problem raised by Kripke's skeptic must, insofar as linguistic practice is rational, intentional activity, ultimately concern our knowledge of the intentional states that underlie linguistic performance, including meaning if that is one such state. But I agree with Boghossian that simply falling back on a comparison between meanings and, presumably, general intentions holds no evident promise of help with the regresses.

[22] The relevant intention is, of course, to be construed as conditional: roughly, that if the question arises, then to apply, or acknowledge the application . . . etc. One would need to be very busy otherwise.

relevant satisfaction-condition than to have them conform to it. It is a *standard-determining* intention, not an intention to comply with standards.

   This point seems very well taken. But now Bilgrami suggests that the adjustment is crucial in the following respect. Its effect is to take us from intentions which an agent may, notwithstanding effort to the contrary, fail to comply with to intentions which there is no such thing as *genuinely having yet failing to comply with*. And surely in that case, he argues, we are no longer dealing with anything properly regarded as normative. Anything that constitutes a norm has to be something that one can fail to live up to.

   It might be wondered whether an intention that is degenerate in this sense—an intention the mere having of which on a particular occasion ensures that one lives up to it on that occasion[23]—could be any real intention at all; if not, then the real tendency of the point, if sustained, is that the original assimilation of meaning and intention is basically misconceived. But Bilgrami goes in the other direction. Meaning, in his view, *is* an intentional state; but it is not a normative state.[24]

   The suggestion that meaning-intentions are peculiarly non-normative is intriguing and worth further scrutiny. Bilgrami's argument for it recapitulates core themes from his *Belief and Meaning*.[25] How might I intend my use of an expression to be answerable to a certain satisfaction-condition and yet it fail to be so answerable? More generally, how might I intend a certain performance to be answerable to certain standards of evaluation and yet fail to make it so? Suppose I set myself to cook a chicken chasseur successfully by Cordon Bleu standards (minimal passing level). I might produce so bizarre and inappropriate a performance that it falls outside the range of evaluability by the standards in question (I deep-fry an old inner tube in castor oil, say). Or I may not know what Cordon Bleu standards involve, so that I am in no position to set them for myself. Bilgrami envisages two corresponding ways in which my linguistic performance might fail to be answerable to a satisfaction-condition that I intend it to answer to. I might misspeak myself, producing a malapropism, or nonsense word, and so produce a performance that is not evaluable by the satisfaction-condition I intended for my uses of an expression I thereby fail to use. Or—the case on which he expends most attention—I may not

---

[23] The reference to the occasion is essential, of course. There is no difficulty, in the cases concerned, with the idea of a change of mind.

[24] The reader might wonder whether, if Bilgrami is right, there are not resources here to deflect those of Kripke's arguments against dispositional construals of meaning which presuppose its normativity. For if meaning is a non-normative mode of intending, there is no force in the charge that the much-vaunted normativity of meaning goes missing on any dispositional construal. It is missing anyway.

   But there is some doubt about the dialectical relevance of this thought. If someone is content to accept some form of Intention View of meaning, he is presumably not going to be drawn to a dispositional account in any case.

[25] Bilgrami 1992.

know, for semantic externalist reasons, what that satisfaction-condition actually is, so that it is not by that standard that I intend my uses of the expression to be evaluable as correct or not. Nevertheless, according to the usual externalist contention, it will be in fact to the externally determined standard that my uses will be accountable.[26]

Bilgrami's response to the second case is to repudiate the externalism that supposedly opens the gap. The burden of the greater part of the argument of his essay is that any form of semantic externalism strong enough to render one's own meanings opaque to one, so that the intention to use an expression with a particular satisfaction-condition may be out of accord with the actual satisfaction-condition that, in one's mouth, it has—any such semantic externalism will be impotent to draw the proper and necessary distinction between irrational belief on the one hand and beliefs that are wayward through ignorance or error about a posteriori matters on the other. In his view, only a conception of meaning that, in the fashion of Frege's notion of sense, renders it transparent to the thinker can properly subserve that distinction.

To engage this claim in any depth would take me far beyond the pardonable scope of these comments. Let me merely anticipate the externalist reply that the distinction that Bilgrami rightly demands be safeguarded marches in step with the distinction between glitches in one's belief system that are apprehensible a priori and those whose disclosure requires a posteriori investigation. If irrationality is essentially an a priori detectable vice, then when content is fixed non-transparently to a thinker, the presence of certain contradictory beliefs in his system need reflect no irrationality. But no valid transition is apparent from that concession to the conclusion that none of the vices of irrationality, including the presence of contradictions, will any longer be a priori detectable in a system of belief involving externally determined contents. So unless we are dealing with a form of externalism whose effect is to undercut the a priori altogether, the expectable upshot will be at most to redraw the boundaries of irrational belief management, rather than to obliterate them.

---

[26] For what it is worth, I have some difficulty seeing that either type of case really does frustrate the kind of intention that is relevant. If I suffer a slip of the tongue, or other misspeaking, I fail to use words that I intended to use. That isn't the same thing as failing to implement the standing *conditional* intention to have my use of those words answer to certain satisfaction-conditions. To frustrate that conditional intention, it has to be that I do utter the intended words but that they somehow fail to be answerable to the relevant condition. As for the externalist case, the question is: what is the satisfaction-condition that I intend my uses to answer to? If it is an externally—socially or causally—determined standard, then presumably my intention succeeds. (Ignorance of a standard need be no barrier to an intention to be held answerable to it.) If it is an internally determined, 'narrow' or idiosyncratic condition, why should others' holding my use accountable to the external standard count as frustrating that intention? The impact of externalism in this case seems to be, not to provide a way that the meaning-determining intention can be frustrated after all, but rather to bar its qualifying as meaning-determining.

Let it be, though, that Bilgrami is right in his principal contention: the thought that to intend that one's uses of "square" be answerable to the satisfaction-condition: x is square, is not normative since not an intention that one can fail to live up to. I think we should pause before drawing the conclusion that this undercuts the normativity of meaning, even when meaning-fixing intentions are construed as standard-setting intentions. There is an underlying issue here about the sense, if any, in which intention, even regular non-degenerate intention, is properly described as normative in any case. That one goes into a context with a certain intention does not per se set a standard of *correctness* for one's behavior in that context. A failure to act on the intention may reflect a change of mind, or distraction, or some misapprehension about what it takes to implement the intention in the circumstances, or some puzzling form of akrasia. But there is no clear sense in which, in any of these cases, one acts incorrectly. If assimilating meaning to intention is, inter alia, meant to provide for the normativity of meaning in a sense relating to the setting of standards for correct linguistic practice, it is the *contents* of the relevant intentions, rather than merely their intentionality, so to speak, that is going to do the work. What is required to provide for correctness and incorrectness is not—or not immediately—that the intentions be ones one can have yet fail to live up to, but rather that the standards they are intentions to answer to be ones that one can fail to live up to. The intentions concerned can be degenerate in Bilgrami's sense and yet deliver all the normativity that we could want. That will be the case if, even if the relevant intentions cannot themselves fail to be lived up to, the standards which they are intentions to answer to can.

There is more to say. I am sure I am not the only philosopher to have been worried by the looseness of the notion of 'the normativity of meaning' as it has featured in these debates over the last thirty years. When some parameter is described as normative over a certain practice, the generic idea is that it contributes towards determining some notion of good practice; but 'good' may be filled out in ways that involve any or all of correctness, value, or support by reasons. Thus each of rules (correctness), morality (value), and intentional states like belief and desire (reasons) are said to be normative. This is a landscape littered with tissues of confusion. The mis-assimilation, noted by Bilgrami, of meaning *square* by "square" to the intention—crudely—to apply the word just to square things betrays a conflation of the first two notions (the setting of a standard is confused with the desirability, presumably derived in this instance from the value of truth, of incurring positive evaluations by that standard). It is a further question, to be broached shortly, whether meaning—that is, the state of understanding an expression in a particular way—should be considered as contributory to a subject's reasons for her linguistic actions. Belief is normative in that way; but it has nothing especially to do with correctness, or value.

Meanings as Fixed by Natural Law

On Horwich's view, the Skeptical Argument goes wrong in the assumption that because meaning is a normative concept, meanings—if they exist—have to be available

to speakers as guides and justifiers of their uses of language. For Horwich, as noted, the meaning-constituting facts reside in the ideal laws governing speakers' uses of expressions. These are natural laws, broadly comparable in status to those of any natural science. It is true, and important, that in linguistic practice we in many respects act as if we were following known rules constitutive of correct 'play': we criticize and correct ourselves and others, we qualify and reformulate what we say, we stick to it or retract it. But this is, I take Horwich to say, a basic propensity. These corrective aspects of our practice are "fundamental to our language-game"[27] and while, with all other aspects, they are to be explained in terms of relevant meaning-constituting fundamental regularities, the template of this explanation is not that of intentional rule-following; and the details of what we find acceptable practice are not in general grounded in psychological states personally or subpersonally encoding the relevant laws.

Horwich's views are complex and are developed more extensively elsewhere.[28] He recognizes, of course, that his stance involves revision of many of our preconceptions about meaning and would doubtless respond that one effect of the Kripkean paradox is to teach us that such preconceptions *have* to be given up. The relevant questions are therefore whether the broad approach can be fit for purpose—that is, whether it restores a notion of determinacy of meaning that is proof against further sceptical challenge—and whether the preconceptions whose sacrifice it involves are, more than preconceptions, essential ingredients in the notion. I will briefly canvass one doubt on each score. They share a common root.

The common root is the point that language is intentional, rational activity. The sayings and writings that constitute public linguistic practice are *actions*, up for rational explanation by the citation of intentional states of the language user. That is enough to call into question whether we should expect any simple regularities in speakers' public linguistic practices of the kind Horwich's account rests upon. Holistic patterns of explanation will operate: what speakers are prepared to say will depend not just on meaning but variable collateral beliefs and desires, and their meanings will consequently surface in their linguistic behavior in the essentially variable and indirect way that, for example, their beliefs and desires do. The answer to the question, how will someone use "square" who means by it: *square*—like the answer to the question: how will he behave if he thinks the ice is thin—can only be: *ça dépend* (on what he thinks about all kinds of things, what he wants, and how he understands other expressions).[29]

---

[27] Compare Wittgenstein *RFM* (VI, 28).

[28] See especially Horwich 2005. The normativity meaning is focused on in Ch. 5.

[29] The reader should not get the impression that Horwich himself passes this holism by. See for instance Horwich 2005, p. 41. But he believes its impact is qualified if we focus on activity at the level of thought, and in a *language of thought*. At this level, considerations like what a speaker wants to communicate and how he wants to go about it will not impinge, and the patterns assumed by his use of expressions will consequently be uncomplicated by such factors. But that does not affect the basic point. For even someone's patterns of (involuntary) acceptance of an 'internal sentence' will still be complicated by lots of other factors, including background information, weighting of evidence, risk aversion, and collateral mistakes, which may vary even though his understanding of the sentence does not.

Now obviously, if there are no expectable straightforward regularities in the use of an expression which competent speakers—or thinkers—generally can be expected to exhibit in common and which might be encapsulated in some foreseeable kind of ideal governing law,[30] then it is forlorn to hope that determinacy of meaning might be grounded in the specifics of such underlying governing laws. Absent any reason to think that the use of an expression allows of *any* purely natural scientific systematic description, the only model we have of a systematic description of linguistic competence is that of the familiar kind of theory that assigns *semantic* significance to words and modes of sentence composition from the start. So Horwich, the concern is, doesn't do enough to support the basic assumption of his view that there actually *are* any ideal laws governing our use of words, fit to constitute determinate meanings.[31]

The idea that a thinker's understanding of an expression combines with other aspects of her intentional psychology to determine her *reasons* for employing it one way or another is deeply engrained in our intuitive thought about these matters. It belongs with it that rational linguistic practice is informed by an *awareness* of meaning. Horwich's putative ideal laws of use, by contrast, are nothing of which we are aware. He anticipates this point, and finesses it by the counter that, in his view, we will be self-aware of meanings only where their application can be assimilated to cases of explicit rule-following in his sense.[32] It is a datum, on the other hand, that where we are at best merely *implicitly* following rules, we are not aware of what the rules are. But this view of the matter may seem better adapted to a speaker's relation to, for example, the grammar of her language than to her understanding of individual expressions. The grammar is identified theoretically on the basis of competent speakers' individual impressions of grammaticality and ungrammaticality. A thinker's grammar will indeed shape the detail of her linguistic behavior, notwithstanding the fact that what it shapes is intentional action, subject to the holistic variations just noted. But—the crucial point—the patterns that the grammar provides a theoretical systematization of are invariant under those variations. What sentence I am prepared to produce, or assent to, in given circumstances may vary with my beliefs and desires; but the patterns consistent with grammaticality in my language will not so vary but will, ideally, be manifest, whatever sentence is concerned. Grammar does not interact holistically with aspects of my personal psychology in shaping my linguistic behavior in the way they interact with each other. By contrast, my understanding of "square" does, as normally conceived, so interact. It is not given to me as a hypothesis best explaining my impressions of acceptable and unacceptable uses of "square". Rather, as least as ordinarily conceived, I *choose* my words, and do so in the light of knowledge about what I am trying

---

[30] Horwich does not say how he understands the notion of a *governing* law, but we can take it that such a law should at least subsume a range of regularities in the phenomena it 'governs'.

[31] Horwich's grounds for this assumption are of course developed in other writings; see especially 2005, Chapter 2.

[32] See §8 of his essay, at p. 83 of this volume.

to achieve, beliefs about how things stand circumstantially in relevant respects, and knowledge as to what my chosen words will say.

In sum. So long as we stand by the idea that the way a speaker understands an expression contributes to the explanation of her production and reception of speech acts as a reason-giving factor, alongside belief, desire, and intention, there is little cause to expect the kind of regularities of linguistic behavior that could underwrite Horwich's ideal governing laws, or support their candidacy to constitute determinate meanings. Conversely, the putative meanings constituted by such laws cannot be items for ordinary psychological self-knowledge, any more than grammar can, and it is consequently deeply revisionary of our ordinary ideas about rational linguistic agency to propose that it is such laws that constitute the determinacy of our meanings.

I'll return to this latter objection after further review of some of the issues concerning self-knowledge of meaning in Part II of these Replies.

Bibliography

Bilgrami, A. 1992 *Belief and Meaning*, Oxford, Blackwell.

Boghossian, P. A. 2000 "Knowledge of Logic", in P. A. Boghossian and C. Peacocke (eds.) *New Essays on the A Priori*, Oxford, Clarendon Press, pp. 229–54.

Boghossian, P. A. 2001 "How are Objective Epistemic Reasons Possible?", *Philosophical Studies* 106/1–2, pp. 1–40.

Boghossian, P. A. 2008 "Epistemic Rules", *Journal of Philosophy* 105/9, pp. 472–500.

Evans, G. 1981 "Semantic Thery and Tacit Knowledge", in S. Holtzman and C. Leich (eds.). *Wittgenstein: To Follow A Rule*, London, Routledge & Kegan Paul, pp. 118–37.

Horwich, P. 2005 *Reflections on Meaning*, Oxford, Oxford University Press.

Kripke, S. 1982 *Wittgenstein on Rules and Private Language*, Cambridge, Mass., Harvard University Press.

Peacocke, C. 2008 *Truly Understood*, Oxford, Oxford University Press.

Stanley, J. and Williamson, T. 2001 "Knowing How", *Journal of Philosophy* 98/8, pp. 411–44.

Wittgenstein, L. 1956 *Remarks on the Foundations of Mathematics*, 3rd edn (Oxford: Basil Blackwell 1981).

Wright, C. 1980 *Wittgenstein on the Foundations of Mathematics*, Cambridge, Mass., Harvard University Press.

Wright, C. 1984 "Kripke's Account of the Argument Against Private Language", *Journal of Philosophy*, 81, pp. 759–78.

Wright, C. 1989 "Wittgenstein's Rule-Following Considerations and the Central Problem of Theoretical Linguistics", in A. George (ed.) *Reflections on Chomsky*, Oxford, Basil Blackwell, pp. 233–64.

Wright, C. 2001 "On Basic Logical Knowledge", *Philosophical Studies*, 106, pp. 41–85, reprinted in J. Bermudez and A. Millar (eds.) *Reason and Nature*, Oxford, Clarendon Press, pp. 49–84.

Wright, C. 2004 "Intuition, Entitlement and the Epistemology of Logical Laws", *Dialectica* 58, pp. 155–75.

Wright, C. 2007 "Rule-Following Without Reasons: Wittgenstein's Quietism and the Constitutive Question", *Ratio* 20/4, pp. 481–502.