For *Wittgenstein and Reason,* ed. John Preston; *RATIO* Volume XX no. 4 (December 2007)

Rule-following without Reasons: Wittgenstein's Quietism and the Constitutive Question

CRISPIN WRIGHT

This is a short, and therefore necessarily very incomplete discussion of one of the great questions of modern philosophy. I return to a station at which an interpretative train of thought of mine came to a halt in a paper written almost 20 years ago, about Wittgenstein and Chomsky,[1] hoping to advance a little bit further down the track. The rule-following passages in the *Investigations* and *Remarks on the Foundations of Mathematics* in fact raise a number of distinct (though connected) issues about rules, meaning, objectivity, and reasons, whose conflation is encouraged by the standard caption, "the Rule-following Considerations".[2] So, let me begin by explaining my focus here.

I *The Rule-Following Dilemma*

It is natural to think that in any area of human activity where there is a difference between *correct* and *incorrect* practice, which we achieve is (partly) determined by rules which fix what correct practice consists in, and which in some manner guide our aim. It also seems the merest platitude that wherever there are rules, there have to be *facts* about what their requirements are—and facts, moreover, which we are capable of knowing if the rules are ones whose guidance we are capable of receiving and acting on. Yet, as the rush of Wittgenstein-interpretative literature from the early 1980s onwards amply illustrates, the very idea of *facts about what rules require* seems on examination to raise a clutter of deeply perplexing questions of constitution and epistemic access. It is in the nature of rules, in a wide class of cases, to enjoin determinate mandates, permissions

---

1 Wright [1989]; reprinted in Wright [2001].

2 — the caption was first introduced in my [1980], so I take some responsibility for such conflations as it may have encouraged.

and prohibitions in previously unconsidered types of situation. So much seems to be no more than what is implicit in the idea that rules are things we *follow*: if we follow, then presumably they lead. But *how* do they manage to lead? In order for it to be possible for them to do so, it seems that three interrelated conditions have to be met. First (the *objectivity* condition), they have to issue their requirements independently and in advance of our appreciation of them; otherwise, there is no real leadership. But what kind of fact could it be that, in a context which no one has yet been placed in or considered, such-and-such a response, or course of action, is already what will be required by a particular rule? How in the world can such requirements be constituted? Second (the *relevance* condition), if a rule is to lead us, it has to be *that* rule rather than any other rule whose guidance we are accepting—there have to be facts about the identity of the specific rule we intend to follow. But how can that be? How and when can it have been settled that it is one specific rule in particular which we are following when everything we may so far have said, or explicitly thought, or done would be consistent with its being any of an indefinite number of potentially extensionally divergent rules?[3] And finally (the *epistemological* condition), even if rules are granted the proper independence seemingly demanded by the very idea of leadership, so that the facts about what is in accord with a rule or not really are fixed before any verdict of ours, and even if it is granted that we can once and for all somehow get one such specific, properly independent rule 'in mind', rather than any of an indefinite number of competitors, how can we account for our ability—in very many normal cases effortlessly, even thoughtlessly—to be appropriately sensitive to the specific requirements, case by case, of just that rule? *How* does a rule actually manage to lead us?[4]

These questions can seem both profound and misguided by turns. Certainly, there are tempting deflationary responses. For instance, it may be suggested concerning the objectivity

---

3 This, of course, is the question prioritised by Kripke is his justly celebrated [1981].

4 The overriding concerns about rule-following may thus be presented as an instance of what Christopher Peacocke in his [2000] termed the Integration Challenge: the challenge of "[reconciling] a plausible account of what is involved in the truth of statements of a given kind with a credible account of how we can know those statements, when we do know them."

condition that a rule—or at least, any rule of sufficient generality and definiteness—is nothing if not something that precisely *does* mandate (or allow, or forbid) determinate courses of action in an indefinite range of cases that its practitioners will never have explicitly considered or prepared for. That is just what rules *are*. So there cannot be a legitimate puzzle about *how* a rule does that, or what settles what its requirements are. To ask how it is settled in advance what complies with the rule is like asking how it is settled what shape a particular geometrical figure has. The figure's shape is an *internal* property of it. What settles what shape the figure has is simply its being the figure it is.[5]

Yet the concerns merely reformulate and re-assert themselves. If a (suitably precise and general) rule is—by the very notion of 'rule', as it were—intrinsically such as to carry predeterminate verdicts for an open-ended range of occasions, and if grasping a rule is—by definition—an ability to keep track of those verdicts, step by step, then the prime question merely becomes: what makes it possible for there to *be* such things as rules, so conceived, at all? I can create a geometrical figure by drawing it. But how do I create something which carries pre-determinate instructions for an open range of situations that I do not think about in creating it? What gives it *this* rather than *that* content, when anything I say or do in explaining it will be open to an indefinite variety of conflicting interpretations? How can I make *it*, rather than a competitor, into an object which I intend to follow? And how is its content to be got 'into mind' and so made available to inform the successive responses of those who are to follow it?

Wittgenstein was conspicuously provoked by a certain way of thinking about these issues—perhaps better, a certain way of ignoring them—that he perceived as widespread in ordinary thought about logic and pure mathematics. The tendency is question could fairly be described as that of a kind of cavalier realism. It views logic and mathematics as tracking absolutely hard conceptual structures and interconnections. Discoveries in mathematics are

---

5 John McDowell, for one, makes exactly this response in the context of the corresponding issue concerning intention—see pp. 163-4 of his [1991].

regarded as the unpacking of (in the best case) deep but (always) predeterminate implications of the architecture of our understanding of basic mathematical concepts, as codified in intuitively apprehended axioms. And logical inference, for its part, is seen as the tracing of steps which are, in some sense, —in a favourite target image of his—*already* drawn and which we have no rational option but to acknowledge once presented to us. This way of thinking—it is, of course, Frege's way of thinking—conceives of the requirements of at least logical and mathematical rules as hyper-objective:[6] as somehow constituted quite independently of any propensities for judgement or reaction of ours. So an account seems needed of *how* they are constituted and of how we might reasonably presume ourselves capable of keeping intellectual track of their requirements so conceived.

It was, however—so I propose—a great achievement of Wittgenstein's to grasp the utter generality of this realist tendency and thus to notice that essentially the same way of thinking about the requirements of rules is quietly at work in much ordinary thinking about the mind; specifically, in the seemingly commonsensical yet notoriously troublesome idea that mental states and processes are items of direct acquaintance for their subjects but are strictly inaccessible to others, by whom they are knowable only by (potentially problematical) inference. It is a usually unremarked component in this to find no difficulty with the notion of simple *recognition* of the proper classification of one's own mental states and processes. Privacy is not supposed to be at odds with one's ability to conceptualise and articulate one's mental states for what they are—on the contrary, it is traditionally supposed to go hand in hand with the possibility of a special level of cognitive security in the judgements so articulated. Yet a judgement expressing a putative such recognition, insofar as it can be correct or incorrect, must presumably be a rule-governed response: there has to be a fact about what one *ought* to say of the targeted inner state—a fact about how it *ought* to be categorised, with which one's judgement about it is presumed capable, in the best (normal) case, of correspondence. So again it seems the question has to be

---

6 What Wittgenstein (*Philosophical Investigations* §192) styles "superlative" facts.

faced: what constitutes such a fact?—what can make it the case that, independently of any reaction of mine, the rules of the language in which I give expression to my private mental life *mandate* certain types of description of an episode therein, and *prohibit* others; and what enables me to keep track of such requirements? Or again, if it is made constitutive of rules to carry such requirements and prohibitions: what can make it the case that specific such rules are associated with particular expressions in the language, and how can they be items of awareness for me in such a way that I can recognise what their specific requirements and prohibitions are?

In its most general form, the issue on which I want to focus here can be crystallised around my assent to any particular token statement, expressed in a language I understand, on a particular occasion of use. In order for this assent to be normatively constrained, and hence a candidate to be correct or incorrect, we have to be able to conceive of whatever constitutes its correctness or incorrectness as in some way independent of my disposition to assent. What are the candidates for such a 'requirement-constitutor'? The question confronts us with a broad dilemma. One thought—the *communitarian* response—is that the requirement-constitutor has somehow to be located *within* the propensities for assessment of the case possessed by others in my language community: that for my assent to the sentence in question to be—in the relevant context—in, or out, of line with the requirements imposed by its meaning is, in one way or another, for it to be in or out of line with others' impressions of those requirements. (Of course, this response cannot engage the case of descriptions of one's own mental states, viewed in the Cartesian way.) But the evident awkwardness with this idea is that it seems to reduce the correctness of an assessment to a kind of marching in step, and to exclude all room for the idea of judgements that are true anyway, no matter what we come to say about a question, or whether we even consider it at all. The other—*platonist*—response accordingly demurs: it says that even (hypothetically) shared assessments are constitutively quite independent of the requirements they concern—that even in the far-fetched scenario where a whole speech community assents to a particular utterance, and where everybody is clear-headed, attentive, and generally competent, the communal impression of what ought to be said is one thing and what really ought to be said is

something else: something settled just between the character of the context and prevailing circumstances on the one hand and the meaning (the rules governing the use) of the statement in question on the other—and it should therefore be conceived as a matter on which a consensual verdict, even in the best epistemic circumstances, merely *alights*.

Our dilemma is, then, that prima facie there only seem to be these two options; but that the platonist—constitutional independence—line threatens to raise baffling ontological and epistemological problems; while communitarianism promises to struggle when it comes to recovering basic distinctions on which our ordinary ideas of objectivity, the growth of knowledge and the defeat of superstition seem to depend.

II ***Wittgenstein on the Constitutive Question***

One response to the dilemma is to attempt to find a third way: to work out a conception of rules and rule-governed practices which allows sufficient of a gap between the requirements of a rule and subjects' reactions in any particular case to make sense of the idea of e.g. a whole community's misapplication of a rule they aim to follow, yet which stops short of any spurious, platonised idea of the autonomy of a rule and its requirements. That—specifically a proposal invoking a form of response-dependence—was the direction I took in the paper on Chomsky. But what does Wittgenstein himself think?

Well, it seems clear enough what he regards as the *sort* of considerations that should point us towards the right perspective on the problem. They are the considerations which constituted the last of four themes in the *Investigations* focused on in my [1989], which I there characterised in the following rubric:[7]

> *Language, and all rule-governed institutions, are founded not in our somehow internalising the same strongly autonomous, explanation-transcendent rules, whose*

---

[7] Wright [1989] p. 243-4. I have tinkered with the wording somewhat.

> *requirements we then succeed, more or less, in collectively keeping track of, but in primitive dispositions of agreement in judgement and action.*[8]

One idea rejected here is that the idea of an essential inner process—a cognitive routine—common to all cases of rule-following is mythical. To express the matter dangerously, we need have nothing 'in mind' when we follow rules. The connection between the training in and explanations of rules which we receive and our subsequent practices is no doubt effected in ways which could only be sustained by conscious, thinking, intentional beings; but it is *not* mediated by the internalisation of explanation-transcendent instructions that, in our training, we (something like) guessed at.[9] It is, for epistemological purposes, a *basic* fact about us that ordinary forms of explanation and training do succeed in perpetuating practices of various kinds—that there is a shared uptake, a disposition to concur in novel judgements involving the concepts in question. The mythology of "rules as rails"[10] attempts an explanation of this fact. But the truth is the other way round: it is the basic disposition to agreement which sustains all rules and rule-governed institutions. The requirements which our rules impose upon us would not be violated if there were not this basic agreement; they would not so much as *exist*.

These aspects of Wittgenstein's thought are very familiar from the emphasis placed upon them in the secondary literature and, as the familiar often does, they can seem quite clear. But they are not clear. The difficulty is to stabilise the emphasis on basic propensities of judgement against a drift to a fatal simplification: the idea that the requirements of a rule, in any particular case, are simply *whatever we take them to be*. For if the requirements of the rule are not constituted, as the platonist thinks, independently of our reaction to the case, what can be available to constitute them *but* our reaction? But that idea effectively surrenders the notion of a

---

8 Illustrative passages include *Investigations* §§208-11, §§217-9, and §242; and *Remarks on the Foundations of Mathematics* VI, §39 and §49.

9 *Investigations* §210.

10 *Investigations* §218

requirement altogether. And Wittgenstein in any case explicitly cautions against it as a misreading of his intent.[11] So, what *is* his position?

Again: Wittgenstein says that the requirements of rules exist only within a framework of ongoing institutional activities which depend upon basic human propensities to agree in judgement. But he also reminds himself that such requirements are also, in any particular case, understood as independent of our judgements, supplying standards in terms of which it may be right to regard those judgements, even when agreed, as incorrect. So we have been told what does *not* constitute the requirement of a rule in any particular case: it is *not* constituted by our agreement about the particular case, and it is *not* constituted autonomously, by a rule-as-rail, (our ability to follow which would arguably[12] be epistemologically unaccountable.) But we have not been told what *does* constitute it; all we have been told is that there would simply be no such requirements were it not for the phenomenon of actual, widespread human agreement in judgement. How can he possibly have thought that this was enough?

It is no good searching Wittgenstein's texts for a more concrete positive suggestion about the constitutive question. Indeed his entire later conception of philosophical method seems to be conditioned by a mistrust of such questions. Consensus cannot constitute the requirements of a

---

[11]Thus *Investigations* §241:

> So you are saying that human agreement decides what is true and what is false?" – It is what human beings *say* that is true and false; and they agree in the *language* they use. That is not agreement in opinions but in form of life.

Similarly *Remarks on the Foundations Of Mathematics* VII §40:

> A language game: to bring something *else*; to bring the *same*. Now, we can imagine how it is played.—But how can I explain it to anyone? I can give him this training.—But then how does he know what he is to bring next time as 'the same'—with what justice can I say he has brought the right thing or the wrong?—Of course, I know very well that in certain cases people would turn on me with signs of opposition.
>
> And does this mean e.g. that the definition of "same" would be this: same is what all or most human beings with one voice take for the same?—Of course not.
>
> For of course I don't make use of the agreement of human beings to affirm identity. What criterion do you use, then? None at all.
>
> To use the word without a justification does not mean to use it wrongfully.

[12] Wright [1989] expounds what I take to be Wittgenstein's principal arguments to this effect.

rule because we leave space for—and do, on occasion, actually *make use* of—the notion of a consensus based on ignorance or a mistake. That is a distinction to which our ordinary practices allow content. The thing to guard against is the tendency to erect a mythological picture of the distinction's content, the myth about rule-following as the unaccountable tracking of "superlative"[13] facts. The myth is active in the platonist philosophy of mathematics, and in the Cartesian philosophy of inner experience. So it is important to expose it. But, once exposed, Wittgenstein seems to be saying, it does not need to be *supplanted*:

> Our mistake is to look for an explanation where we ought to look at what happens as a 'proto-phenomenon'. That is, where we ought to have said: this language-game is played.[14]

No further *account* of the distinction—between an agreed move and a correct move—is necessary. Enough has been done when we have pointed out and defused philosophical misunderstandings of our linguistic practices in a way that avoids misdescription of their details. Our discourse of rules and meanings stands on its own feet. Platonism is a misunderstanding of it; but it does not need an alternative, better explanation to shore it up or otherwise account for the various locutions and distinctions which platonism misunderstands.

That, it seems, is his finished view.

III ***Whence the Quietism?***

The question I want to ask is: did Wittgenstein have any sound theoretical basis for this line? He is saying, in effect, that there is no well-conceived issue about the 'constitution' of facts about what rules require, instance by instance, or about what enables us to keep track of such facts. There is no real dilemma between platonist and communitarian views of the matter, and no constructive philosophical work to do by way of attempting to steer between its horns. But what is the ground for this claim—why should we agree with him?

---

13 *Investigations* §192.
14 *Investigations* §654.

The rubric above emphasised the *primitiveness* of our basic dispositions of classification and judgement. By this, I mean something coincident with—as I now propose to understand it—the metaphor of *blindness* that Wittgenstein introduces at *Investigations* §219.[15] Here is the passage in full:

> "All the steps are really already taken" means: I no longer have any choice. The rule, once stamped with a particular meaning, traces the lines along which it is to be followed through the whole of space.
> — But if something of this sort really were the case, how would it help?
> No; my description only made sense if it was to be understood symbolically.—I should have said: *This is how it strikes me*.
> When I obey a rule, I do not choose.
> I obey the rule *blindly*.

What does Wittgenstein mean by saying that we follow rules *blindly*? Clearly, he is thinking of the simplest cases, where nothing takes place which can naturally be regarded as *working out* what a rule requires—cases where one's response seems to be immediate and one can produce no reason for it, no explicit justification. So the cases in point presumably include whatever rules are involved in, for example, the simple classification of colours, or tastes, or Lockean secondary qualities generally, as well as in some of the examples on which Wittgenstein tends to concentrate—judgements about the continuation of certain very simple diagrammatic or arithmetical series.

It would be a mistake, though, to take the point of the metaphor of 'blindness' as concerning the *phenomenology* of such judgements: in effect, as being that in making them (competently), *one is not aware* of any mediating process—of any route to the judgement which one might recover and cite by way of justification for it,—but is just *smitten*, as it were, by the judgement. No doubt that is often so. But it cannot be the whole of the matter. In fact, I do not think it is the point at all. 'Blindness', after all, is a poor metaphor for *immediacy*, and the phenomenology in question is merely that of immediacy. The judgements of a sighted person

---

15 In focusing on this metaphor, I follow a lead of Paul Boghossian (see the references to his work in the Bibliography) who has recently been laying emphasis on it in connection with the epistemology of basic logic, and foreshadowing a connection with the larger discussion of rule-following (see e.g. Boghossian [2003] at p. 237) which is explored in depth in Boghossian [forthcoming]. He may or may not agree with the way I am going to develop it here.

about her local environment will include many immediate ones; they are hardly *blind* on that account. Wittgenstein's point is not (primarily) phenomenological. But then what is it?

Before we take that question head-on, we need to think further about the notion that sets our basic problem: the idea of the *facts* about what a given rule requires, allows or forbids in successive cases. Consider instead the more complex kind of case where one *does* reason one's way to a judgement about the proper application of a rule—for instance the case of Castling in chess. Here we find the following kind of structure of judgement.

> *Rule*: If neither King nor one of its Rooks has moved in the course of the game so far, 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, then one may Castle
>
> *Premise*: In this game neither my King nor this Rook have yet been moved, the squares between them are unoccupied, and ...
>
> *Conclusion*: I may castle now.

Call this the *modus ponens* model of rule-following.[16] The rule is stated in the form of a general conditional. A minor premise states that, in the circumstances in question, the condition articulated in the antecedent of the rule is met. The conclusion derives the mandate, prohibition, or permission concerned. Clearly, the model is of very wide application. In fact, I make so bold as to suggest that it applies, in essentials, in all cases when it is appropriate to think of one's impressions about what is in accordance with a rule as worked out, and when, correlatively, there *are* explicit reasons to be given for those impressions, by citing that working. Notice, however, that it is a feature of the model that one's knowledge of the rule is but *one* ingredient in one's movement to a correct application of the rule. There is a simple holism in operation, broadly akin to that involving belief and desire in the explanation of behaviour. Just as no behaviour, however bizarre, conclusively defeats the ascription to a subject of, say, a particular desire—you can always compensate by making sufficiently radical adjustments in the ascription of beliefs and

---

16 Cf. Wright [1989], p. 256

other desires to her—so no response, however aberrant, in and of itself defeats the claim that a subject correctly understands and intends to follow a particular rule—you can always make compensatory adjustments by ascribing a misapprehension of the *initial conditions* for the application of a rule, as expressed in the minor premise in the modus ponens model. That makes it very easy to see that in cases where the model applies there can be no such thing as (what we may call) *pure* rule-following: that every judgement, or movement, made with the intent of compliance with a rule may go wrong not because the requirements of the rule are mistaken but—quite consistently with correct understanding of the rule—because of misapprehension of relevant features of the circumstances in which the rule is being applied.

Hard on the heels of that thought comes a generalisation: there is no pure rule-following not merely in cases that comfortably fit the modus ponens model but *anywhere*, however simple or basic the rule(s) involved. Even in cases, like the expansion of an arithmetical series, where there might be no *perceptual* input (because one is following the rule 'in one's head'), so no relevant risk of a perceptual mistake, judgements about the correctness, or permissibility of a next step will still depend on *memory*: on not losing track of what one has so far done. To approximate a case of pure rule-following, one would need to consider a rule whose application involves neither perceptual input nor any memory of previous stages, nor even any extended process (of which one might lose track) in executing a single stage—so that each correct application at any stage can be made in a fully informed way without any sensory input or knowledge of anything else one has done or judged. There is no such case. The idea of pure rule-following—rule-following where a correct grasp of the rule is sufficient *tout court* to guarantee correct performance—is chimerical.

So what? Well, a key component of the problem of rule-following outlined in the first section above was the thought that if it is to be appropriate to think of an activity as subject to rule, then there have to be *facts* about what the relevant rules require, or permit, and—if we are to subject our practice to those rules—we have to be in position to know what these facts are. So questions of the constitution and epistemology of such facts seem to be directly raised by the very

idea of following a rule. We want to better understand how facts about the requirements of rules are made and how they are accessible to us. What the essential impurity of rule-following now raises is a question about what exactly *are* the 'facts about the requirements of rules'—what is their canonical form of expression? If there is no pure rule-following, we cannot think of these facts as being the very same as the facts that make *particular* applications of rules, or judgements about what complies with them, correct. For the latter are always contaminated by additional elements—concerning context, or history, or the input *to* which the rule is to be applied.

It may be replied that, at least in cases where the modus ponens model is apt, there are still isolable judgements about what properly belongs to the rule and what belongs to the input provided by a situation in which an application of the rule is at issue. The separation is explicit in the model itself. What properly belongs to the rule corresponds to the conditional major premise in the modus ponens model—which will either be an explicit statement of the rule or a claimed consequence of one—while what corresponds to the situational input will be given by the minor premise. So the issues about constitution and epistemology may therefore intelligibly be focused on the former, or so it would seem. But the evident problem then, of course, is that this way of looking at the matter will not transpose to the basic—'blind'—case. For in basic cases—the very simplest kinds of rule-following, gestured at above—the modus ponens model is inappropriate.

Wright [1989] reached a similar point—

> The ability successfully to follow a rule is thus to be viewed as, at each successive instance, the product of a number of cognitive responses which interact holistically in the production of the proper step. And some of these responses—correctly perceiving the set–up on the chess–board, for instance, or recollecting the expansion of the series to this point—do not strictly pertain to the rule but are possible for subjects who have no inkling of it—

but then observed something further:

> Where R is the rule or set of rules in question, let us call the [other responses]

—the ones that are *not* possible for subjects who have no inkling of the rule —

> *R–informed*. Now, an R–informed response need not be encapsulable in any judgement which the subject can articulate distinct from the output judgement, as it were—the judgement into which his or her R–informed and non-R-informed responses conjointly feed. In that respect, the chess example, in which the R– and non-R-informed components could be respectively explicitly entertained as the major and minor premises for a *modus ponens* step, is untypical. I cannot always have concepts *other* than those whose governing rules I am trying to observe in a particular situation in terms of which I can

> formulate a separate judgement of the input to which these rules are to be applied. So I cannot always extricate and articulate a judgement which, conditionally on such a separate judgement of the input, formulates my impression of the requirements of the rules in a fashion which is neutral with respect to the correctness of my R–uninformed responses to the situation.[17]

So there is not just the *holism* to reckon with, blocking the possibility of pure rule-following. In addition, at least in basic cases, the contribution of grasp of the rule to the responses it informs is *inextricable* from the contribution of one's grasp of the prevailing circumstances. The clean separation effected by the modus ponens model between what belongs to the rule and what belongs to the situation to which it is to be applied is possible only in (relatively complex) cases where the conditions which trigger the application of the rule—those described in the antecedent of the relevant conditional—can be recognised and characterised in innocence of a mastery of the rule. That cannot be the situation in general.

My reaction to this consideration in the [1989] paper was to focus on the—for my purposes then—easier case where one's judgement about the requirements of a rule may be seen as resting upon extricable major and minor premises after the fashion of the modus ponens model.[18] Here, though—where our purpose is to try to get some kind of focus on the impact of

---

17 Wright [1989] pp. 255-6.

18 The thesis I was there aiming to review was that, for the purposes of assessing any potential tension between the 'rule-following considerations' and 'the central project of theoretical linguistics', we should consider the impact of the former upon the status of the judgements—about grammaticality and content—which a systematic syntax and semantics for a natural language will generate concerning each of its strings. And a prime candidate for an encapsulation of that impact was, or so I argued, the thesis that such judgements were *response dependent*, that they failed 'the order of determination test": I wrote

> The test, as so far considered, calls for a class of judgements about which we can raise the question of the relation between best opinion and truth. And the existence of such judgements is just what the inextricability point counsels us not to expect in general. Still, there are extricable cases. The example of castling in chess provided one. And, most significant in the present context, the comprehending response to a novel utterance provides another. Such a response will involve a set of beliefs about the utterance which someone could have who had no understanding of the language in question; but it will also involve a belief about what, modulo the former set, has been said—a paradigm, it would seem, of a rule–informed judgement. Rather than confront the awkwardness presented by inextricability, therefore, let me concentrate for our present purposes on such favourable cases: cases where the acceptability of a rule–informed response can be seen as a matter of the truth of a judgement which the responder may be thought of as making. Our question, then, is: what makes for the truth of such rule–informed judgements?

Wittgenstein's rule-following discussion quite generally—we have no option but to attend instead to the harder range of cases, where the modus ponens model seems to lapse as a framework for the explanation of a rule-governed response, for the want of extricable major and minor premises.

Let us focus on the case of colour. Suppose, undeterred, we stubbornly try to assimilate predications of 'red' to the modus ponens model. The correctness of such a predication is thus to be seen as the progeny of an input condition of a certain character together with a rule associating such inputs precisely with the correctness of the predication:

Rule: If ...x..., it is correct to predicate 'red' of x

Premise: ...x...

Conclusion: It is correct to apply 'red' to x.

To conceive of predications of 'red' as rule-governed in the manner of the model accordingly requires an anterior concept, '...x...', whose satisfaction determines an input as appropriate for the application of the rule. But now it stares us in the face that this concept can hardly be anything other than: *red*! So we get an interesting upshot: the stubborn extension of the modus ponens model to the cases Wittgenstein would seem to have in mind when he speaks of rule-following as 'blind' would demand that we think of linguistic competence in terms, broadly, of the *Augustinian picture of language* with which the *Investigations* begins, and from which it is a journey of recoil. The crucial aspect of the Augustinian picture for our purposes here, of course, is not the confusion of meaning and naming which Wittgenstein himself fastens on in the

---

—and the counselled answer was, roughly: their being made—or coinciding with the judgements that would be made—under conditions of *best* judgement. On this view, the well-formedness/meaning of a compound expression are not self-standing properties of it but are constituted in the very impressions of its well-formedness/meaning which competent speakers form under appropriate conditions. That was the suggested lesson of the rule-following considerations as applied to the Chomskyan/Davidsonian enterprise; and the consequent question was whether such a response-dependent conception of the constitution of the syntactic and semantic properties of whole sentences left room for systematic syntax and semantics as genuinely explanatory empirical theories in the manner their principal architects had conceived.

That, by the way, still seems to me to be a fascinating, under-discussed question. For one recent, helpful discussion of it, see Miller [forthcoming].

immediately succeeding sections of the text, and on which his commentators have largely concentrated. It is the aspect, rather, that is highlighted a little later, at *Investigations* §32:

> ...And now, I think, we can say: Augustine describes the learning of human language as if the child came into a strange country and did not understand the language of the country; that is, as if it already had a language, only not this one. Or again: as if the child could already *think*, only not yet speak. And 'think' would here mean something like 'talk to itself'.

In short, the problem with extending the modus ponens model to cover all rule-following, including that involved in basic cases, is that it calls for a conceptual repertoire *anterior* to an understanding of any particular rule—the conceptual repertoire needed to grasp the input conditions, and the association of them which the rule effects with a certain mandated, prohibited or permissible form of response. From the standpoint of the philosophy of thought and language of the *Investigations*, this is an enormous mistake. With respect to a wide class of concepts, a grasp of them is not anterior to the ability to give them competent linguistic expression but rather *resides in* that very ability. (This need not be a commitment to holding that there is never any sense at all to be made of the idea of thought without language. But it is to repudiate the *general* picture of thought as an activity of the mind which language merely clothes.)

If this is right, then a crucial component of Wittgenstein's thought about rule-following depends upon a perspective which, more than half a century after he put his ideas to paper, seems to have come—one may well feel: regrettably—to appear non-compulsory to many contemporary philosophers of mind: the conception of language not merely as a *medium for the expression* of thought but as—usually, though not exceptionlessly—*enabling* thought: as providing its very raw materials. From this perspective, the modus ponens model *must* lapse for basic cases. Basic cases—where rule-following is 'blind'—are cases where rule-following is *uninformed by anterior reason-giving judgement*—just like the attempts of a blind man to navigate in a strange environment.[19] In such a case one follows a rule 'without reasons' in the precise sense that one's

---

19 Of course, the analogy limps immediately after this point—the movements of the blind man will naturally be hesitant. But in basic rule following "I act quickly, with perfect certainty, and the lack of reasons does not trouble me" (*Investigations* §212.)

judgements about the input condition for correct application of the rule are not informed by the exercise of concepts other than that which the rule concerns—that is, the concept whose expression the rule regulates and grasp of which consists in competence with that very expression. Such a judgement is an ungrounded response in the precise sense that is not to be rationalised by the modus ponens (as I have suggested: the only possible) model—by the picture of rule and input as (potential) components of independent thought. It is still essentially the response of a rational subject, and still to be appraised within the *categories* of rationality—justification and truth. But it is an action for which, it is now tempting to say, the subject has and can have no reason—for the possession of such reasons and their appreciation as such would demand the exercise of an anterior concept, in an independent judgement, of what made the action appropriate.[20]

So here is what seems to be the resulting position. All rule-following involves basic rule-following. And basic—'blind'—rule-following, properly understood, is rule-following without reason—not in the sense of being phenomenologically immediate, or spontaneous in the way in which a good chess player may make a clever move without fully self-consciously rationalising his grounds for it, but in a sense involving the inappropriateness of the modus ponens model. But that model represents the only extant shot—once again, I'm tempted to say, 'the only possible shot'—at the extrication of a class of judgements which would distinctively express the special facts about what rules require that the very idea of normativity seems to call for, and which—if we could somehow extricate them—would provide the necessary focal point for the issues concerning constitution and epistemology at the core of the problematic about rule-following with which we started. So there *has to be something wrong* with that problematic. And what is wrong, one might say, is that in the basic case we do not really *follow*—are not really guided by—anything. The problematic invited us to try to construct an account of what, when we follow a particular rule, constitutes the facts about the direction in which, step by step, it guides us and

how we are able to be responsive to its guidance. But in basic cases the invitation emerges, from the perspective on the matter just adumbrated, as utterly misconceived; for it presupposes a false conception of the sense in which basic rule following is rational. Basic rule following, like all rule following, is rational in the sense that it involves intentionality and a willingness to accept correction in the light of error. But that is not to say that it involves responsiveness to the requirements of the rule, conceived as instructions, as it were, which can feature in thought and rationally inform one's response. The initial problematic—what constitutes the requirements of rules and how are we able to keep track of them—presupposes otherwise.

In summary: To say that in basic cases, we follow rules blindly or without reasons is to say that our moves are uninformed by—are not the rational output of—any appreciation of *facts about what the rules require*. This is, emphatically, not the claim that it is inappropriate ever to describe someone as, say, knowing the rule(s) for the use of 'red', or as knowing what such a rule requires. Rather, it is a caution about how to understand such descriptions—or better: about how *not* to understand them. In any basic case, the lapse of the modus ponens model means that we should not think of knowledge of the requirements of the rule as a state which *rationally underlies* and enables competence, as knowledge of the rule for castling rationally underlies a chess player's successfully restricting the cases where she attempts to castle to situations where it is legal to do so. In basic cases there is no such underlying, rationalising knowledge enabling the competence. *A fortiori* there is no metaphysical issue about the character of the facts it is knowledge of, with platonism and communitarianism presenting the horns of a dilemma. The knowledge *is* the competence. Or so I take Wittgenstein to be saying.

That is why Wittgenstein's own response to his well-argued rejection of platonism is quietist. A non-quietist response would be called for only if platonism had given a bad answer to a good question. Then one would have to try to give a better answer. But the question was bad too. The real error in platonism is not the unsustainability of its sublimated conception of rule-facts, or the vulnerable epistemology that attends the sublimation. Rather the whole conception of rule-following to which it was a response was already an over-rationalisation—an implicit

attempt to impose on rule-following everywhere a rational structure which can only engage the non-basic case.

IV Rational judgements made for no reasons?

There is no doubt that the general tendency of the foregoing discussion goes deeply against the grain. Normal thought envelops even our basic judgements with a *rhetoric* of reasons. I assent to the judgement that something is red. You ask me: what reason do you have to think so? I can perfectly properly answer, "Well, the way it looks, of course". But *how* does the look serve as a reason? In response, one immediately finds oneself thinking in terms of the modus ponens model: "Well, the object has a certain look; it is constitutive of the concept *red* that things that look that way are (defeasibly) appropriately taken to be red. Therefore, the object in question may appropriately be taken to be red." But this is just the model that in basic cases we have discarded.

There are considerations that one may marshal to try to sugar the pill. Suppose I assent to a judgement about something's colour, based purely on its look. In what sense is this assent rational? One can say several things:

That it is an act possible only for a creature that is rational;

That it is an act in the "space of" reasons— open to assessment as correct, responsibly made, and so on;

That it is an act that may in turn contribute to my reasons for (other) acts and judgements. But these considerations are all broadly concerned with the stage-setting and implications of the act: with how I must be regarded if I am credited with that very act, with how it may be appraised, and with what it commits me to. They do nothing, it seems, to restore the idea of the *input-rationality* of basic judgement— to explain how a basic judgement can be made on the basis of reasons which the thinker *has*, how it can be the *product* of a rational response to anterior, reason-giving states.

It is clear where pressure would have to be exerted if the lost ground is to be (even partially) regained. Means would have to be provided to distinguish the conclusion that basic

judgements are judgements made for *no reason that can be captured via the modus ponens model* from the stronger conclusion that they are made for *no reason at all*. I have been suggesting that the modus ponens model supplies the only means whereby facts about the requirements of rules can enter into a subject's reasons for an act, —that it is only in terms of the model that we can make sense of the idea of the rationality of a judgement or action if it is to be conceived as the product of following a rule. If that is right, and if basic judgements are nevertheless input-rational, then the conclusion is clear: we have to understand them as rationalised in a way that takes them *outside* the category of rule following. If basic judgements are judgements made for no reasons involving rational processing of information about the requirements of rules, and if they are nevertheless made for reasons, they are not to be thought of as delivered by the following of rules. (And of course, if they are made for no reasons, then they are not to be thought of as delivered by the following of rules in any case.)

So how might we try to regard them instead? What kind of nature and structure might their reasons have if not that of the modus ponens model? There is one well documented but, as it has proved, vexed proposal. As remarked, it is our normal rhetoric to say that it is looks — more generally, *experiences* — that rationalise our most basic empirical judgements. However, if experiences are to do this, then they have to possess attributes which fit them to do so. What kind of attributes would accomplish that? The modus ponens model will have it that, whatever they are, they will need connection, via a major premise presumed to be already part of the thinker's information, with the appropriateness of the judgement the experience rationalises. If this is to be avoided, there must be no role for such a major premise. Therefore, experiences must have attributes which fit them to rationalise empirical judgements *immediately* — that is, without any kind of rational intermediary processing of thoughts. The idea is then apt to seem compelling that such rationalisation can be accomplished only if we conceive of experience as already *essentially conceptually contentful*: experience has essentially to consist in the reception of appearances that ....., where what fills in the dots is a conceptual content. There has to be such a content because otherwise it is obscure how experience can *rationalise* judgement, rather than merely causally

predispose to it; and the content has to be carried essentially since otherwise the assignment of content to an experience would have to proceed by principles connecting "given", non-contentual characteristics with content — and that would take us straight back to the modus ponens model.

Such is, of course, exactly the conception of experience propounded in John McDowell's *Mind and World*.[21] McDowell's route into it is, familiarly, different: he presents it as a *via media* to avoid the unsatisfactory answers to the question, how can empirical judgements be rational, offered respectively by the (putative) Myth of the (non-conceptual, sensory) Given and Davidsonian coherentism. However, if what I have said is right, McDowell's conception of experience is actually mandatory for any philosopher determined to have it that basic judgements are made for reasons furnished by experience.

It is another question, of course, whether the conception of is stable or satisfactory.[22] Common complaints have concerned its apparent exclusion of the experience of infants and animals and the lack of any foreseeable principled account of which are the contents that experience can carry intrinsically (not presumably, for example, that this object is a geiger-counter...) However, a more urgent question about it now, in the light of the preceding discussion, is *how* experiences come to be fitted out with the conceptual contents which, according to the McDowellian idea, they essentially carry. Labeling the conceptual content of experiences "essential" to them is, for the reasons just gestured at, a forced move. But it is a major concern whether there is any way of making sense of the idea which is not at odds with the broader lessons of Wittgenstein's discussion. For if experience is to be intrinsically such as to rationalise judgement, it must carry the content it does *independently* of judgement. And how is that idea to be sustained except at the cost of crediting the concepts configured in an experience's content with a kind of platonic propensity to self-application?

---

[21] McDowell [1994]

[22] My own previous contributions to the extensive commentary include Wright [1996] and [2002]

The issues waiting in this direction must be material for another discussion.[23]

Arché, University of St. Andrews, and
New York University

cjgw@st-andrews.ac.uk

Bibliography

Paul Boghossian (2000) "Knowledge of Logic" in *New Essays on the A Priori*, Eds. P. Boghossian and C. Peacocke (Oxford: Oxford `University Press) pp. 229-254

(2002) "How are Objective Epistemic Reasons Possible" in *Reason and Nature: Essays in the Theory of Rationality*, eds. J. Bermúdez and A. Millar (Oxford: The Clarendon Press) pp. 15 – 48

(2003)"Blind Reasoning", *Supplement to the Proceedings of the Aristotelian Society* 77, pp. 225-48

(forthcoming) "Blind Rule Following" in *Mind, Meaning and Knowledge: Essays for Crispin Wright Volume I*, ed. Annalisa Coliva (Oxford: Oxford University Press)

Sauk Kripke (1981) Wittgenstein on Rules and Private Language (Oxford: Basil Blackwell)

John McDowell (1991) "Intentionality and Interiority in Wittgenstein" in *Meaning Scepticism,* ed. Klaus Puhl, (Berlin: de Gruyter) pp. 148-69

Alex Miller (forthcoming) Critical Notice of Wright (2001), *International Journal of Philosophical Studies*

Christopher Peacocke, (2000) *Being Known* (Oxford: Oxford University Press.)

Crispin Wright (1980) *Wittgenstein on the Foundations of Mathematics* (London: Duckworth)

(1989) "Wittgenstein's Rule-following Considerations and the Central Project of Theoretical Linguistics", in *Reflections on Chomsky*, ed. A. George, (Oxford and New York: Basil Blackwell), pp.

(1996) "Human Nature?", critical study of John McDowell's *Mind and World*, in the *European Journal of Philosophy* 4, pp. 235-54.

---

[23] Precursors of this material have been presented at various seminars and colloquia including the European Summer School in Analytical Philosophy held at Parma in 2001, the Mind and Language seminar at NYU in spring 2002, and the May meeting of the Scots Philosophical Club and the Reading one-day conference on *Wittgenstein and Reasons* in 2006. Thanks to the discussants on all these occasions. Special thanks to Paul Boghossian, Cyrus Panjvani and John Skorupski for detailed discussion and comments.

(2001) *Rails to Infinity*, (Cambridge, Mass., Harvard.)
(2002) Reprint of "Human Nature" with a new postscript in *Reading McDowell: On Mind and World*, ed. Nicholas Smith (London: Routledge) pp. 140-73.

Wittgenstein's works:
Philosophical Investigations, 3rd edition (Oxford: Basil Blackwell 1967)
Remarks on the Foundations of Mathematics, 3rd edition (Oxford: Basil Blackwell 1981)