Metaphysics Handouts
Revised
10 Nov 2026
07:49

Wittgenstein on following a rule: McDowell

THE ARGUMENT IN OUTLINE. The villain is anti-realism, the thesis that we understand only what we can manifest in behaviour, and since Wright's reading rests on it, it leaves Wittgenstein on the first horn too, against Wright's intentions. No position can be both anti-realist and hospitable to meaning, and reading Wittgenstein as the source of anti-realism is 'a travesty'. In full:

(1) The contractual picture: to learn a word is to be committed to a determinate pattern of use. Everyday objectivity, things being thus and so anyway, needs that pattern to reach the new case independently of our ratification.

(2) Kripke's sceptic: no fact about me constitutes my meaning addition, since whatever is in my mind can be read another way. To concede that is to deny that anyone means anything, and no 'sceptical solution' by way of a community can claw meaning back. The reading also has no room for §201's 'way of grasping a rule which is not an interpretation'.

(3) The paradox is one horn of a dilemma. The other is an interpretation that can't itself be interpreted, the super-rigid machine, the rule as rails. Both rest on the assumption that understanding is always interpretation, and Wittgenstein's aim is to drop it, not to choose a horn. So there are three positions, not two, and Kripke runs two of them together.

(4) Wright keeps the contractual picture and drops only the ratification-independence of the patterns. But approval by everyone is no better a standard than a feeling of satisfaction, as the parody of PI §460 shows. At Wright's basic level there are only vocalisings and feelings of constraint, no norms, so no meaning.

(5) PI §202, obeying a rule is a practice, answers how a grasp of a rule can fail to be an interpretation: a 'blind' reaction is going by a rule when it belongs to a custom. Scylla is understanding as always interpretation, Charybdis a basic level with no norms, and the custom is the middle course, so far 'only a programme'. Blackburn: why not an individual's practice?

(6) Wright's reading hangs on anti-realism, that we understand only what we can manifest in behaviour describable without the language. From such behaviour one can only extrapolate inductively. ⇒ Understanding can't be grasp of patterns, denying ratification-independence obliterates norms.

(7) So meaning requires rejecting anti-realism. We should drop its picture of the behavioural surface, since command of a language gives *direct* contact with a speaker's meaning. That answers Blackburn: we hear a speaker's meaning in his words, a linguistic community is 'a meeting of minds', and so the practice has to be communal.

(8) The considerations attack a picture of truth as pure thought, with meanings taking care of themselves. The middle position keeps patterns that reach new cases independently of our ratification and drops only their 'platonistic autonomy'. ⇒ No position is both anti-realist and hospitable to meaning, and Wittgenstein as the source of anti-realism is 'a travesty'.

THE STAKES. We think of meaning in contractual terms: learning a word commits us to using it in certain determinate ways later. Wright's Wittgenstein says the contract has less in it than we think, since nothing in our understanding fixes in advance what counts as its correct application. Objectivity, in the everyday sense, is things being 'thus and so anyway', and that needs the pattern we grasped to extend to the case in hand independently of the outcome of any investigation. Drop ratification-independent patterns and you drop facts that hold independently of our ratifying them, which McDowell calls 'a kind of idealism'. And when the interlocutor of PI §195 says the future use is somehow present 'in a queer way', Wittgenstein replies that only the phrase 'in a queer way' is wrong. So the target isn't the idea that present understanding carries commitments about the future, but a seductive misconception of it.

KRIPKE'S WITTGENSTEIN. Kripke's sceptic asks what fact about me made it addition that I meant by 'plus'. Nothing answers, since whatever is in my mind I'm free to interpret it differently later. Kripkenstein accepts the paradox and offers a 'sceptical solution': justification conditions in place of truth-conditions, and ascriptions of meaning read as recording acceptance into a linguistic community. If there's no fact that could constitute its being the case that *P*, you can't go on affirming that *P*, and that principle is the engine of the standard arguments against realism, not a realist prejudice. So the concession becomes the denial that anyone means anything, and it's obscure how juggling membership in a community could claw us back. Kripke's reading also has no room for the second paragraph of PI §201 where the reasoning betrays 'a misunderstanding', since we offer one interpretation after another, each contenting us only for a moment. Wittgenstein concludes there that 'there is a way of grasping a rule which is not an interpretation'.

Question 1. Grant the sceptic that no fact constitutes my meaning addition. Is there any sceptical solution that wins meaning back afterwards? Hint: page 228.

THE DILEMMA. The paradox is one horn of a dilemma. Keep the misunderstanding, so that the task is to find a fact constituting my having interpreted what I was taught. Then resisting the paradox drives you to the other horn: an interpretation that can't itself be interpreted, which exactly bridges the gap between instruction and use. That is the 'super-rigid yet (or perhaps we should say "hence") ethereal machine'. With it comes a picture of the rule as rails laid out ahead of us (PI §218), the rule itself settling each step, with training only getting our practice into line. Kripke takes Wittgenstein's attacks on this mythology. But the mythology is wrung from us only because we're caught in the misunderstanding. The attack on it, given that the paradox is intolerable, is an argument against the misunderstanding.

Everything rests on one assumption, that understanding is always interpretation, and Wittgenstein's aim isn't to move us from one horn to the other but to make us drop the assumption and refuse the choice. So there are three positions, not two. Kripke runs two of them together when he equates the paradox of §201 with Wittgenstein's conclusion. That paradox isn't the mild claim that an isolated individual can't make sense of meaning, something a community might repair. It is devastating claim that meaning is an illusion, and the isolated individual belongs to the second horn, not to the way into the abyss.

Question 2. Why must an interpretation that can't be interpreted be 'ethereal'? What would a form of guidance that could never misfire be made of? Hint: 229–230.

Wright's purge. The Wright McDowell critiques is the 1980 Wright of *Wittgenstein on the Foundations of Mathematics*. This is the 'consensualist' view of 'Rule-following and theoretical linguistics' (1989). The consensualist (='Wright'), as McDowell reads him, conceives of the community as subject to no standard beyond its own judgements.

The consensualist keeps the contractual picture and removes one thing, the ratification-independence of the patterns. PI §460 mocks the man who, told 'Bring me a yellow flower', brings the one that gave him a feeling of satisfaction, and McDowell grants that this reads as an attack on idiolectic understanding.

Example 1. Wittgenstein could just as well have written the parody: the man brings the flower that received approval from all the bystanders, and is told that he wasn't asked for the flower everyone else would approve of. A reading on which correctness is communal approval can't say why the second rebuke is as good as the first.

(12-1) This is yellow.

(12-2) This would be called 'yellow' by most speakers of English.

Wittgenstein insists on a radical difference between the significance of (12-1) and that of (12-2), and Wright's reading leaves it a mystery why it should matter so much. The answer that may seem readily available to Wright, that saying (12-2) isn't calling the object yellow, only postpones the serious question: has his picture room for such an act as calling something 'yellow' at all? At the 'the basic level' it contains people vocalising in response to objects, with feelings of constraint, but no shared commitments and nothing with authority over the behaviour. Since meaning is normative, a level with no norms has no meaning in it. Wright hopes to keep a purified normativity in each speaker's openness to communal correction, but once the basic level is in view it claims to be the real truth, and going out of step with one's fellows then looks like what gives us the illusion of being subject to norms, and so of meaning anything.

Question 3. Is correctness relative to a community a thin version of correctness or none at all? Does the best-opinion view of Wright's 'Rule-following and theoretical linguistics', equipped with C-conditions, escape McDowell's critique?

Practice, custom, bedrock. PI §202 opens 'And hence also "obeying a rule" is a practice', and McDowell reads the thesis that obeying a rule is a practice as the answer to the question §201 leaves open, how there can be a way of grasping a rule that isn't an interpretation, with the 'hence also' signalling that inference. The rest of §202, the private language argument in two sentences, comes as a corollary. PI §198 runs the same thought through the sign-post: I was trained to react to it in a particular way, and now I do. The interlocutor objects that this is a causal story, not an account of what going by the sign consists in. The reply is that the training is 'initiation into a custom': without that, the story would indeed be about brute movement and its causes, with nothing to entitle us to speak of following the sign-post.

Example 2. Asked the colour of a flower, I say 'red', and asked whether I might have called the wrong colour red, I say no: the certainty is the rigidity of my measuring-rod (RFM VI-28). How do I know this colour is called 'green'? As I know what 'Help!' means when I'm drowning: that's how I react in this situation (RFM VI-35).

At bedrock, where justifications run out, there seems to be only verbal behaviour and feelings of constraint, no norms, which is Wright's basic level under another name. McDowell's textual basis is PI §289: to use an expression without a justification 'does not mean to use it without right', a sentence whose point is to stop norms leaching out of our picture of bedrock. Following according to the rule is 'fundamental' to our language-game (RFM VI-28): the hard thing isn't to dig down to the ground but to recognise it (RFM VI-31).

Scylla and Charybdis. Scylla is the idea that understanding is always interpretation, which forces the dilemma, and Charybdis is a basic level with no norms, which makes meaning an illusion. The middle course: a performance can be a 'blind' reaction (PI §219) and still be a case of going by a rule, because it belongs to a custom, a practice or an institution.

Until more is said about how communal practice does this, McDowell concedes, this is 'only a programme for a solution'. Blackburn presses the obvious question: if practice does the work, why not the practice of an individual?

Question 4. Can one and the same act be a blind reaction and a case of going by a rule? Give a case outside language where someone has a right without a justification.

Anti-realism, the villain of the piece. Wright's (=consensualist!) reading hinges on a conditional: if possessing a concept were grasping a ratification-independent pattern, there would be no knowing for sure how anyone else understands an expression. Knowledge of idiolectic meaning is an illusion, so by modus tollens possessing a concept isn't grasping a pattern. Wright grounds both premises on anti-realism, the thesis that we understand only what we can distinctively manifest. The conditional rests on its conception of manifesting understanding to others, in behaviour describable without command of the language, such as assenting to a sentence in given circumstances. From such behaviour one can extrapolate only inductively, so understanding can't be grasp of a network of patterns, which is Wright's conclusion. But denying ratification-independence obliterates norms, and openness to correction can't be kept alongside as a second picture, since the norm-free one claims primacy.

So Wright's argument turns on its head: a condition of there being any application for the notion of meaning is that we reject anti-realism. Anti-realism rightly recoils from the idea that another's meaning is a hypothesis about something hidden behind the surface of his behaviour, but it keeps that conception of the surface and tries to find meaning on it. The other way out is to reject the conception: command of the language puts one in direct cognitive contact with a speaker's meaning, which looks like 'X-ray vision' only from inside the rejected picture. This answers Blackburn: shared command of a language lets us know one another's meaning without arriving at it by interpretation, because we hear a speaker's meaning in his words, which is what §201 needed. A linguistic community is bound by 'a meeting of minds', not by matching exteriors. That's why the practice has to be communal and not an individual's. 249 249 253 254

Hence: there can't be a position that is both anti-realist and genuinely hospitable to meaning, and reading Wittgenstein as the source of anti-realism is 'a travesty'. 262

Two realisms. The rule-following considerations attack a picture of genuine truth: a fact is the way things are in themselves, so a true judgement must be an exercise of pure thought, untouched by human nature. The picture lives on the supposition that 'meanings take care of themselves', which lets the judging subject 'shrink to a point of pure thought', whereas a performance is a judgement only by answering to a concept's constraints, which is acting within a communal custom, not being led by an autonomous meaning. 254 255

The picture is a realism in two senses, meaning 'out there' as in the old debate about universals, and an extreme form of the thesis that the facts aren't up to us. But the recoil isn't from truth-conditional semantics, since truth-conditions are given by us in a language we understand, and in '"Diamonds are hard" is true if and only if diamonds are hard' we're just as much involved on the right-hand side as the rule-following considerations say we are, nor from the thesis that the facts aren't up to us, only from its extreme form. So there must be a middle position: understanding is grasping patterns that extend to new cases independently of our ratification, as meaning and objectivity both require, but the constraints lack the 'platonistic autonomy' of the super-rigid machine. Where Wright sees a consensus on calling a new object 'yellow' that is subject to no norms, McDowell says the community 'goes right or wrong' according to whether the object is yellow, and nothing makes its being yellow depend on our ratifying the judgement. 255 255 256 256

Three cases. Let's illustrate McDowell's view by contrasting it with alternatives.

Example 3 (The sum). Asked a sum I've never met, I say that 68 + 57 = 125. Kripke's sceptic: 'What does make 125, rather than 5, the right answer?'

Kripkenstein: nothing. We're entitled to say 125 because we say it unhesitatingly and in step with others, and that's assertibility, not truth.

The consensualist: 125 is right because it's our verdict, and for the community itself there's 'no authority, so no standard to meet'. 226

The Platonist: the addition function fixes 125 for every pair by itself, training only gets our practice into line, and an omniscient God wouldn't need to do arithmetic. 231

Wright 1989: 125 is what best opinion under C-conditions would deliver.

McDowell: 125 is right because 68 + 57 is 125. The pattern I grasped when I learned 'plus' reaches this case ahead of anyone's verdict. If a drug made all of us say 5, we'd all be wrong. But there's no fact behind the arithmetic that fixes the answer with no one doing arithmetic: 'even God can determine something mathematical only by mathematics'. And the sceptic's question, what fact constitutes my having interpreted 'plus' rightly, is the misunderstanding of PI §201, not a question still waiting for a better answer. 222 232 229

The four rivals agree that the question needs an answer in other terms: nothing, agreement, the function, best opinion. McDowell's answer stays inside arithmetic. They call that empty. He calls the demand for a constitutive explanation 'confused'.

Example 4 (The flower). Take PI §460 and the parody again. The man who brings the flower that gave him a feeling of satisfaction is told that isn't what he was asked for. But the man who brings the one all the bystanders approved of is told the same. 234

The consensualist: makes sense of the first rebuke, which is aimed at idiolects, but not of the second. On his view the flower everyone approves of just is the right flower, so he can't say why (12-1) and (12-2) differ so radically. 234

Kripkenstein: wants both rebukes, but having granted that no fact fixes what 'yellow' means he has only our shared inclinations to point to, and McDowell doubts they can carry the second rebuke. 228

McDowell: the second rebuke is as good as the first. The right flower is the yellow one, whoever approves.

Now push it with Wright's sodium street-lighting from Handout 11: a whole village picks flowers under the lamps and everyone approves a white one. Wright 1980: they can't be wrong. Wright 1989: they're wrong because daylight opinion would differ, and that is what makes the flower not yellow. McDowell: they're wrong because the flower isn't yellow. Daylight is how they find out, not what makes it so. So for McDowell the community's agreement isn't the standard even in idealised form. The standard is the flower. 256

Example 5 (Right turn). PI §506: the absent-minded man told 'Right turn!' turns left, clutches his forehead, says 'Oh! right turn', and turns right. Wittgenstein asks what struck him, an interpretation? 239

Scylla: to obey he must have interpreted the order, and any interpretation can be read another way. Regress, then the paradox.

Charybdis: in Blackburn's wooden picture he was trained to react, and a reaction is caused, not faithful to anything. Many wooden persons with propensities to make noises are just more of the same. No norm, so he didn't go wrong. He just moved. 226 253

The consensualist: he went wrong because he went out of step with how the others react to 'Right turn!'.

McDowell: he went wrong because the order said right. He didn't interpret it, and he wasn't merely caused. His corrected turn is a 'blind' reaction that is still going by the rule, because it's a move in a custom he was initiated into. He has a right without a justification (PI §289). Ask him how he knows that 'right' means right, and the answer is RFM VI-35: the way I know what 'Help!' means when I'm drowning. The custom has to be shared, not because the others' approval settles which way is right, but because he can hear what the officer meant in the words, a meeting of minds rather than matching noises.

This is the case that shows the two jobs apart. The community makes his turn a move that can be right or wrong. The order says which.

Question 5. In all three cases McDowell states the standard with the very term in question: '125', 'yellow', 'right'. Take the hardest case, whether '777' occurs somewhere in the decimal expansion of π. Is 'ratification-independent but not autonomous' a different view from the rails, or the rails under another name? Hint: 231–232.