

Following Rules of Nature, not the Pedestrian Muse: Reply to Yamada

Masuhiko Yamada recently has approached the rule-following considerations in what he calls a pedestrian manner. To him, the “...question is: What is left when the fact that *my* arm goes up every five seconds is subtracted from the fact that *I* am following the rule to raise my arm every five seconds?”(Yamada 2010, 284, my emphasis)

This sounds like a really pedestrian question: What constitutes *my* following a pre-established rule? Later we learn that Yamada merely provides necessary condition of doing so. However, in Yamada’s critical discussion of earlier approaches a much grander question looms large: What constitutes a rule, or: What is left when the fact that *one*’s arm goes up every five seconds is subtracted from the fact that *one* is following the rule to raise one’s arm every five seconds? Here the problem is raised how rule-following as a communal practice comes about, how rules come into the world. And this is the question that has puzzled great philosophers from Wittgenstein onwards. My discussion has four parts. (I) I outline what I like in Yamada’s approach. (II) I argue that he does not even answer the pedestrian question. None of his “necessary” conditions is necessary. (III) I show how Yamada’s criticism misses the grand issue because he does insufficiently attend to recent research. He remains captured in an earlier stage of discussion. (IV) I briefly sketch a proposal how to deal with the deeper issue.

(I) There are three things I find convincing in Yamada’s discussion: Firstly, philosophers following Kripke reject the dispositional approach to rule-following (to which Yamada’s belongs) because dispositions are finite. Yamada justly complains: a “...sugar cube has all the infinitely many dispositions needed to unfailingly conform to the rule if you are thrown into a glass of water under normal conditions, dissolve!... there is no difficulty in thinking that a

finite object can have infinitely many dispositions.”(Yamada 2010, 289)¹ Secondly, Yamada points to recursive procedures such as those that are constitutive of the arabic numerals (Yamada 2010, 292). Such procedures contain further resources of dealing with finitude. So Yamada contributes to focusing the grand issue of rule-following on the real problem: What constitutes correctness conditions of following a rule if our dispositions are fallible, i.e. do not perfectly fix them?² Thirdly, Yamada emphasizes that one does not have to explicitly adopt a rule if C, do A! in order to follow it; one does not even have to intend to do so; nor does one have to explicitly believe that condition C is fulfilled.

(II) I outline why I deem Yamada’s approach insufficient to settle the pedestrian question. As recent developments in the philosophy of language show, it is difficult to devise necessary conditions of possessing concepts. For concept possession is a very flexible matter (cf. Williamson 2007, 73-133). Parallel considerations apply to rule-following. These are Yamada’s necessary conditions:

“S observes the rule *if C, do A!* only if:

i) there is a state tokening of σ such that

a) S’s tokening of σ disposes S to do A

b) a) is true because when the dispositions came into place it was the case that if S is in σ , C obtains.

¹ Still one may ask whether we really have suitable dispositions to account for infinitely complex applications.

² However, this has already been done by Schulz (2009, 58). Schulz further argues that Condorcet’s Jury Theorem is suitable to also solve the fallibility issue. Since under fairly common assumptions a community is much more reliable than an individual, in the limiting case of an infinitely large community perfectly reliable, there is no difficulty of dispositions in a community providing an infallible normative standard (Schulz 2009, 59). However, the crucial requirement is that the community be infinitely large lest it be fallible. There is no infinite speech community. So we have no explanation for the existence of actual rules.

and

ii) if S tokens the belief that C, this belief would dispose S to do A.”(Yamada 2010, 300)

None of these conditions fulfils its task.

I start with (ib). There is a dilemma. (ib) is either too strong or too weak. Yamada notes an ambiguity in our everyday notion of rule-following. Sometimes it involves a success component, sometimes not:

“Ordinary language locutions like ‘S follows the rule R’ are ambiguous in an important way. In one sense, following the rule implies that the subject actually conforms to it. Hence the puzzling nature of a statement like this: ‘John follows the rule to brush his teeth at 11 pm. As you can see, he always brushes his teeth at 10.30 pm.’... On the other hand, we can also say ‘John is brushing his teeth at 10.30 pm because daylight savings time has just ended but he has forgotten to set his clock and he follows the rule to brush his teeth at 11 pm.’ ...I will be using the locution ‘S observes the rule R’ for this sense of ‘S follows the rule R’ that does not imply conformity to it. The locution ‘S follows the rule R’ is reserved for the conformity implying sense.”(Yamada 2010, 295)

Observing rules does not require perfect conformity. I completely agree. But now assume that John has newly adopted this rule. Consequently he has brushed his teeth several times at 10.30pm but never at 11pm. He always fails to conform to the rule. Still he may be said to observe it. Yamada accepts the following conditions of observing rules:

“...the agent does not have to be such ...that she unfailling[ly] does A whenever C obtains – errors are possible- but there must be some kind of systematic connection between the obtaining of C and the performance of the output A.”(Yamada 2010, 298)

Consequently, Yamada introduces state σ to abide by the following original requirement:

(Reliability) "...one observes the rule if C, do A! only if there is a state which is such that if one is in that state, C obtains and if one is in that state, that state also disposes one to do A. ... the conditionals are meant to capture a positive correlation."(Yamada 2010, 298)

More precisely:

"The kind of reliability needed here is similar to the one that Sosa (1999) calls safety. A belief is safe only if: in nearby possible worlds in which S believes that p, p is true."(Yamada 2010, 299 ann. 21)

For reasons that do not need to concern us, Yamada then weakens this requirement to (ib). I have the following misgivings. Considering the teeth-brushing example, observing a rule in contrast to following it does not seem to require the kind of systematic correlation envisaged by Yamada. John may always fail to conform to the rule and still observe it. One may reply that at least a systematic correlation obtains between its being 11pm and him brushing his teeth at 10.30 which is explained by the systematic malfunction of his watch. But assume that the watch randomly goes awry. Although John always brushes his teeth when the watch shows "11 pm", effectively he sometimes brushes his teeth at 10.30, sometimes at 10.45, sometimes at 12. Judging from Yamada's presentation of the case, we still might accept that there is a sense in which John is observing the rule. His behaviour can be explained by his observing the rule. Our non-conformity-implying sense of rule-following seems too lenient to require a systematic correlation between some state σ and the obtaining of C. So Yamada's original requirement (reliability) seems too strong.

However, since the requirement of a systematic match with some external condition C has a certain plausibility, I grant (reliability) for the sake of argument. There must be a certain reliability in conforming to the rule. Once we accept it, Yamada's requirement (ib) seems too weak compared to his original requirement, though. It only requires some sort of correlation at some initial time point t . Then we may go completely awry. It is perfectly in tune with (ib) that we totally miss the rule after t . For instance, at time t I adopt the rule to apply "apple" to all and only apples. However, as it happens, at any time after t I call all and only peas apples. Yet once a threshold of reliability is adopted, I cannot count as following the "apple" rule. Yamada may reply that he does not provide sufficient but only necessary conditions of rule-following. However, condition (ib) is motivated only by (reliability): One must conform to a minimum threshold of reliability or safety while observing the rule. Once we drop this requirement, there is no reason to accept condition (ib).

Even if it is accepted that some systematic dispositional correlation must obtain, and that it must obtain only at a suitable time point at which the rule is adopted, Yamada's condition is too strong. Assume that when John adopts the rule of brushing his teeth at 11pm, due to the random malfunction of his watch, there is no systematic correlation between condition C (the time being 11 pm) and any state σ of his. So Yamada's condition is missed. But imagine that as time goes by, John comes to realize that his watch is not reliable and adopts the custom to check it. He begins checking it from time to time and ends replacing it by a more reliable watch. So when John's disposition to brush his teeth at 11pm came into being, there was no systematic correlation between condition C obtaining and his brushing his teeth. Later there was. Yamada may reply that the rule-constitutive disposition σ came into being later. But then he must deny that John observed the rule before he had a reliable disposition. This runs counter to our intuitive verdict as Yamada presents it. Moreover, a gradual process leads from John adopting his rule to him becoming reliable in conforming to it. John's adopting the rule explains his later following it via this process. Surely the most eligible time for the rule-

constitutive state to come into being was the time when he consciously adopted the rule. If Yamada were to suppose otherwise, say, to maintain that state σ only came into being when some threshold of reliability was passed, condition (ib) would appear even less plausible than it appears anyway. How could Yamada on the one hand deny that one observes the rule until one passes some threshold of reliability at some time point and on the other hand drop the reliability requirement for all the time to follow?

Now to requirement ii): If one came to believe that C, one would do A. This requirement is subject to an objection that is reminiscent of finkish dispositions. It is difficult to provide sufficient conditions for belief. But surely one standard case is outright explicit sincere acceptance. However, there can be cases of genuine rule-following in which one would cease to have the disposition to follow the rule when one came to explicitly accept that C. Take Yamada's example of Vaclav who observes the rule "...if there is a government undercover agent nearby, beat the hell out of him!" (Yamada 2010, 301) Assume the *résistance* has trained Vaclav well. He explicitly accepts the above rule. Moreover, he is the most reliable of them all in instinctively spotting the clandestine and sinister air of an undercover agent. Whenever he does so, he immediately beats the agent without bothering about it. So he has reliably beaten many agents over many years. But alas, although Vaclav is impulsive, he is a *penseroso*. Were he to explicitly realize that condition C is fulfilled, were he to think –as he actually never has– “there is a government undercover agent nearby”, he would succumb to a feeling of deep melancholy and contemptuous compassion with a creature so base as to serve the tyrant; briefly, he could not beat him. Still I would say that if anyone does, Vaclav certainly observes the rule of beating government agents. Supposing otherwise would run counter to the spirit of Yamada's approach; one does not have to explicitly endorse a rule to follow it, one does not even have to intend to follow it. Nor does one have to have any particular belief regarding it. So none of Yamada's conditions is necessary.

(III) Now to the grand question: What constitutes a rule? Crispin Wright argues:

“A solution to the problem ... would be to show how suitably circumscribed facts about how one does, will, or would use a particular expression actually constitute the facts how one ought to use it.”(Wright 2002, 110, quoted by Yamada 2010, 288)

Yamada replies:

“The reasoning here is not convincing. We must be careful in distinguishing the rule itself from an agent’s state of following the rule. An agent’s following the rule if C, do A! is constituted by facts about what he does, will or would do, among other things, but the fact that the rule if C, do A! requires doing A when C is decidedly not constituted by what the agent does, will or would do... Missing this is akin to confusing what it takes to believe that there is an odd number of stars with what it takes for it to be true that there is an odd number of stars... There is no mystery here.”(Yamada 2010, 288)

True enough. But Yamada misses the radical point of Wright’s; he misunderstands it as addressing the pedestrian issue.³ Wright does not ask how an agent’s state of following a pre-

³ Yamada’s presentation is also reminiscent of Horwich dismissing the rule-following problem. Horwich accepts Kripke’s result that there is no meaning-constitutive relation R between “dog” and its being true of all and only dogs. But he denies that we need such a relation. If “dog” means dogs, it is true of all and only dogs. This is true “...by definition and explanatorily fundamental” (Horwich 2010, 111). But the problem remains. If it is a constraint on “dog” meaning dogs that it is true of all and only dogs, we need a relation R* between the fallible use of “dog” and its meaning dogs. Kripke’s arguments apply to R* as well.

established rule makes up the rule. He asks how by “if C, do A!” we as a community of speakers come to mean the rule if C, do A!, or how “there is an odd number of stars” comes to mean that and be true iff there is an odd number of stars. Linguistic rules are not facts that are independent of our practice in the way the number of stars is. They are made. Our public words mean what they mean because a community of speakers establishes rules which fix the correctness conditions of our words. On pain of regress, this cannot always be done by resorting to words established before. One can introduce words by other words, for instance “addition” by its meaning “calculating the sum”. But one cannot do so all the way down, or so it seems.⁴ The dispositional account answers that rules making up our words are somehow constituted by our dispositions of using these words in a certain way. And here fallibility comes in again. If we grant that dispositions of use which constitute a rule are fallible, how can a rule emerge by virtue of which some of these dispositions are false? How can dispositions at once be constitutive of and be subject to tracking the rule? This is the problem the Kripkenstein tradition attempts to solve, and Yamada fails to recognize in Wright’s reasoning.

Yamada explicitly does not want to bother about issues of meaning (Yamada 2010, 285). Perhaps we may accept that he restricts himself to the pedestrian question and takes meaning as given. But then he had better not criticize Kripke, Wright, and others; they tackle a different question. Moreover, concerns with meaning cannot be avoided so easily. Yamada concedes that his account is given in overtly intentional terms:

“If our ultimate aim is to give an account of concept-possession in terms of rule-following, one might worry that a non-reductive account leads to a vicious circularity in our theory. But surely, if an attempt to utilize an independently motivated account of rule-following leads to

⁴ To be sure, one may accept rules “all the way down” (Brandom 1992, 44, McDowell 1987).

serious problems in the theory of concept-possession, the right thing to do is to abandon the attempt to understand concept-possession in terms of rule-following.”(Yamada 2010, 309)

Unfortunately, I cannot find Yamada’s account so very well-motivated as to full-heartily sacrifice the obvious connections to questions of concept-possession. After all, concepts are normative. One can misapply them. An approach to rule-following that has nothing to say about this sort of normativity barely deserves our interest. One way for Yamada to evade this reproach is a view as endorsed by Boghossian (1989):

“I used to think it was constitutive of rule-following that it be intentional action, and I used to think that it was implausible that we come to mean things by our mental words because we follow rules in respect of them.”(Boghossian 2005, 194)

Yamada’s taking meaning as given could be understood along the lines of Boghossian taking the meaning of our mental words as given. Intention-based rule-following presupposes mental words which give our intentions content. With mental content at hand, public words can easily be devised. “Dog” means what my mental word DOG means.

However, Boghossian himself has revised his position in his (2005) which Yamada does not attend to:

“If the intention-based model is to capture all instances of rule-following, then it had better also capture my use of the Modus Ponens rule. But ... it is impossible to see how that might be done. If we tried, we would get something like the following:

Intention: If C, do A

Premise:C

Modus Ponens: Whenever both ‘If, C, do A’ and ‘C’ do A

Conclusion: Do A

And now we would be off on a vicious regress. Epistemic rules whose operation cannot be captured by the intention-based model are presupposed by that model itself.”(Boghossian 2005, 195-196)

Here we have the deepest layer of the rule-following considerations. What constitutes the modus ponens (mp) rule? It cannot be our intention. Since Yamada does not take into account Boghossian’s results, he falls back to the state of debate in Boghossian (1989).

(IV) I close with sketching a reaction to Boghossian’s problem: Two questions call for an answer: How do our mental words come to mean what they do? How are we to account for the mp rule? I feel that both questions should be answered along similar lines. There must be some sort of normative stratum which underlies normal rule-following. This stratum is not our work. What constitutes it? If nothing we intentionally do can, few alternatives remain. I can only think of two. The most obvious solution: tokening mental words suitably covaries with the latter’s referents. But as Yamada notes, this alternative requires overdemanding idealizations (Yamada 2010, 291, Boghossian 1989, 539-540). I propose a different solution as indicated by Weatherson in a discussion of vagueness:

“It is hard to identify exactly which features of Ralph’s ‘mental state or neural state or causal relations with his environment’ make it the case that he believes that two plus two equals four, but does not believe that two quus two equals four (I assume Ralph is no philosopher, and so lacks the concept quus). I doubt, for example, that the concept plus has some causal influence over Ralph that the concept quus lacks. But Ralph does have the belief involving plus, and not the belief involving quus. He has this belief not merely in virtue of his mental or neural states, or his causal interactions with his environment, but in virtue of the fact that plus

is a more natural concept than quus, and hence is more eligible to be a constituent of his belief.”(Weatherson 2003, 487)

Alluding to Kripke’s original “quus”-example (Kripke 1982, 22), Weatherson applies to the rule-following problem what Ross Cameron calls the “...Lewis-Sider view that meaning is determined by use plus naturalness.”(Cameron 2010, 283). Reference is fixed by naturally eligible classes of things:

“Only an elite minority are carved at the joints, so that their boundaries are established by objective sameness and difference in nature. Only these elite things and classes are eligible to serve as referents.”(Lewis 1984, 227)

Admittedly this proposal faces many difficulties (cf. Elgin 1995). But these difficulties mainly concern high-level epistemic activity. When we ask basic questions such as how our mental words come to mean something or how mp comes to guide our practice, Lewis’ proposal may fare better.

How does mp come into the picture? Here is Boghossian’s internalist justification of mp:

“Any rules that are written into the possession conditions of a non-defective concept are a fortiori entitling.”(Boghossian 2003, 34)

Since mp is written into the possession conditions of non-defective concepts, it is entitling. Even if Boghossian’s approach to justification is rejected,⁵ there surely is a close connection

⁵ Cf. the criticism in Schechter and Enoch (2006). They propose that we are entitled to any presumption we must make in an endeavour we are rationally required to embark on (2006,

between thought content as determined by naturalness and non-defective concepts. So I propose that mp is basically constituted by relationships among our mental words reflecting relationships of what these words stand for. (If A, B; A; B) in its most basic use reflects nothing but the fact that one may plug in any mental representations A, B such that the premisses are true and achieve a true result B. Further uses of modus ponens, for instance in rules how to act or for mere hypothetical reasoning are derivative.⁶

707). But what is a rational requirement? We may resort to Wright to cash it out: We are entitled to whatever is necessary for our epistemic agency to be successful at all (Wright 2003, 66-67). This sort of entitlement is a weak one, though, because it does nothing to exclude that the presumption is false and that our epistemic endeavour is doomed to fail. For instance, if my mathematical proof cannot be successful unless I make a certain otherwise completely unsubstantiated assumption, am I entitled to this assumption?

⁶ Arguably, rule-following is located in what we, following Sellars, may call the space of reasons. Inferential reasoning is guided by preserving truth and not by causal relationships. Consequently one could doubt that (causal) natural laws provide the right kind of normative constraint on cognitive activity (McDowell 1998, 242). What Lewis' example shows, in contrast, is how natural facts can impose normative correctness conditions on a representational system. Say the mental word DOG corresponding to "dog" cuts nature at the joints iff within a system of other such words, the only natural candidate for it to refer to is the natural kind dog. Now we may apply this word to a cat. Then we have made a mistake because we have violated the normative requirement imposed by nature on the reference of our mental words. In the same vein, when some causal process leads to our failing to conform to mp, we violate the normative requirement imposed on our cognitive activity not by our striving and intending but by the way the world is.

Whatever to make of my proposal, it exemplifies the kind of riddle one should deal with before one outrightly accepts an intentional account of rule-following as Yamada does.

Boghossian, Paul. 1989. The Rule-Following Considerations. *Mind* 98: 507-549.

Boghossian, Paul. 2003. Epistemic Analyticity: A Defense. *Grazer Philosophische Studien* 66: 15-35

Boghossian, Paul. 2005. Rules, Meaning and Intention – Discussion. *Philosophical Studies* 124: 185-197.

Cameron, Ross P. 2010. Vagueness and Naturalness. *Erkenntnis* 72: 281-293.

Elgin, Catherine. 1995. Unnatural Science. *Journal of Philosophy* 92: 289-302.

Horwich, Paul. 2010. *Truth, Meaning, Reality*. Oxford: Oxford University Press.

Kripke, Saul A. 1982. *Wittgenstein on Rules and Private Language: An Elementary Exposition*. Cambridge, Mass: Harvard University Press.

Lewis, David. 1984. Putnam's Paradox. *Australasian Journal of Philosophy* 67: 221-236.

McDowell, John. 1987. In Defense of Modesty. In Michael Dummett, *Contributions to Philosophy*, ed. B. Taylor. Amsterdam: Nijhoff: 59-80.

McDowell, John. 1998. *Mind, Value and Reality*. Cambridge, MA: Harvard University Press.

Schechter, Joshua, Enoch, David. 2006. Meaning and Justification. The Case of Modus Ponens. *Noûs* 40: 687-715.

Schulz, Armin W. 2009. Condorcet and Communitarianism: Boghossian's Fallacious Inference. *Synthese* 166: 55–68.

Weatherson, Brian. 2003. Many Many Problems. *The Philosophical Quarterly* 53: 481-501.

Williamson, Timothy. 2007. *The Philosophy of Philosophy*, Oxford: Blackwell.

Wright, Crispin. 2002. (Anti-)Sceptics Simple and Subtle: G.E. Moore and John McDowell. *Philosophy and Phenomenological Research* 65: 330–348.

Wright, Crispin. 2003. Some Reflections on the Acquisition of Warrant by Inference. In *New Essays on Semantic Externalism and Self-Knowledge*, ed. S. Nuccetelli. Cambridge, Mass.: MIT Press: 57-79.

Yamada, Masuhiro. 2010. Rule Following: A Pedestrian Approach. *Philosophy and Phenomenological Research* 80: 283-311.