

The Laws of Thought and the Power of Thinking

Matthias Haase (Universität Basel)

1. Introduction

Frege taught us to strictly distinguish between the logical and the psychological. This doctrine has deeply influenced the analytic tradition in the philosophy of mind, language and logic. And it was praised, of course, by Wittgenstein, early and late. On closer inspection, however, the way in which Frege frames his anti-psychologism leads to a crack in his system that shows up in a couple of places in his writings. I want to suggest that Wittgenstein's so called 'rule-following considerations' address this difficulty and are intended to show that the 'crack' eventually brings down the whole palace and with it this way of framing the celebrated distinction between the logical and the psychological. Investigating how the rule-following paradox arises within Frege will, I hope, shed some light on the systematic question what might be puzzling about rule-following and which conditions an account would have to meet in order to count as a candidate for a solution.

In "Der Gedanke" Frege writes: "To the grasping of thoughts there must correspond a special mental capacity, the power of thinking".¹ Kant calls the correlative capacity the "power of concepts". In the opening paragraph of his *Logic*, Kant claims that it holds, quite generally, that for each power there are correlative laws governing the exercises of the power.² It would seem to follow that the laws governing the exercises of the power of thinking are what one might call the 'laws of thought'. And one might expect that *these* must be what logic articulates. Frege, however, warns us that the phrase 'laws of thought' is ambiguous. He insists on a strict distinction between the laws of logic and the laws that "govern" how people actually to think. The latter are "psychological laws". The laws of logic belong are "laws of truth": they *don't explain* how we *actually* think; their relation to our activity of thinking is, rather, *normative*. They *prescribe* how we *ought* to think. To suggest otherwise inevitably leads, according to Frege, into the pitfalls of psychologism and the misguided idea that the laws of logic are empirical generalizations over mental processes happening at particular times and places. Now, on the face of it, this way of framing the anti-psychologism seems to have the consequence that the laws that explain our acts of thinking and the laws that figure as a standard for these acts come apart. As I understand it, the rule-

¹ Gottlob Frege, "Der Gedanke: Eine logische Untersuchung", in: *Beiträge zur Philosophie des deutschen Idealismus* 2, 1918-1919, 58-77, here 74.

² Immanuel Kant, *Logik*, §1.

following paradox is about the disastrous consequences of this gap: the normative and the explanatory can *never* come apart if it is to be intelligible how there can be that power to think and judge from which Frege's reflection on the "laws of truth" takes off.

Frege's conception of logic has come under attack in the name of a "formal" account of logic. However, this attack leaves the picture of the relation between the normative and the explanatory intact. This is sometimes hard to see due to the tendency to present Frege as an uncontrolled Platonist who dogmatically postulates queer entities. I will argue that what leads to the gulf between the normative and the explanatory is *not* some peculiar and easily discardable doctrine that declares "thoughts" to be mystic entities that can never *become* efficacious or "active". The issue is more intricate, and it is, as we shall see, not limited to logical laws in the narrow sense, but extends to a general puzzle about the relation between a concept the acts of deploying it judgment.

As I understand it, Wittgenstein's proposed solution of the rule-following paradox is the claim that there is a sense in which the laws of thought that *govern* or *explain* our exercises of the power of thinking *are* the laws that *prescribe* how we *ought* to think. This idea is supposed become available when we realize that the possession of a concepts and the mastery of logic are "abilities" that one acquires by being initiated into a "practice".³ I will not discuss the viability of this solution here. I will focus on the problem. In doing so I will avoid words like "practice", "custom", "institution", "forms of life" and, as much as possible, also "capacity", "power" and "ability". These words are in a sense distracting. They appear to contain the solution, which is, I take it, why at least some of them (or similar words) occur in virtually every treatment on the topic, whether or not it claims to be Wittgensteinian in spirit. Often, however, they remain untheorized – like the words "capacity" and "power" in Frege's text; and it is not always clear how they are supposed to have entered the reflection in the first place.⁴ The whole mystery is to how to hear them in the right way. I will argue that reflection on the role they would have to play in order to provide a candidate for a solution shows that they cannot be taken for granted if one wants to avoid either remaining within the paradox or

³ See *Philosophische Untersuchungen*, §198-202. In the following as *PU*.

⁴ The literature tends to divide two camps when it comes to the role of notions like *ability* and *practice* in the philosophical reflection on concepts. The one group treats these notions as already understood when we get to the theory of judgment and proposes to employ them in a non-circular explanation of judgment and conceptual content. The other group insists that the relevant notions of ability and practice – namely conceptual capacity and linguistic practice – are 'irreducible'. The result is in both cases that the investigation of the notions *ability* and *practice* does not seem to be part of the reflection on conceptual content – in the former case because the task belongs to another part of philosophy or, perhaps, another discipline like psychology or sociology; in the latter case, because questions like 'What is a conceptual capacity?' or 'What is a linguistic practice?' are regarded as *bad* questions. My aim in this paper to suggest that these questions are good questions and that the relevant notions of ability and practice cannot enter the investigation of judgment from the outside. But I will proceed as it were indirectly by focusing on the problem that makes us reach for these notions in the first place.

letting the “laws of truth” collapse into something that one doesn’t call “laws of psychology” anymore, but that would, properly speaking, have to be called “laws of sociology”.

I will begin with a presentation of the ‘crack’ as it shows up in Frege’s writings (§2) and then try to show how it widens to a general puzzle about the relation between a concept and the acts of deploying it in judgment, whether it is the acts of an individual or the acts of a multiplicity of individuals (§§3-4). I will end with remarks on the role of words like “ability” and “practice” (§5).

2. The “laws of truth” and the “laws of taking-to-be-true”

2.1. The normative and the explanatory

Sometimes Frege’s doctrine that thoughts are occupants of a “third realm” different from the realms of the physical and the psychological is presented as a dogmatic and easily discardable reification of “thoughts” as queer entities.⁵ But that is too quick. Frege’s conception of “thoughts” is rooted in his reflection on our activity of judging and the “atemporal” character of ‘is true’. In a nutshell, ‘true’ is said “atemporally”, whereas the phrase ‘taking something to be true’ describes acts and states of individuals in space and time. Something being true is independent of my or your recognizing it as true. That which is true – the “thought” – must thus be independent of what people are actually thinking: the “thought”, as Frege puts it, “needs no owner”.⁶ Only if this is so, Frege insists, is it possible to account for an obvious difference between judgments and “ideas” (“Vorstellungen”). You feeling warm does stand in a relation of contradiction to my feeling cold, even if we sit right next to each other, no more than your liking this vanilla ice cream contradicts my disliking it, even if we try the same cone. By contrast, if we take a position on the questions what temperature the room has and whether this is vanilla ice cream and, our judgments stand in relations of agreement or disagreement. And it seems that they can only do so if we have in view the same thing to be affirmed or denied. We must be able to “share” the *same* “thought”. To begin with, this ‘reification’ of “thoughts” as abstract objects occupying a “third realm” should not seem threatening. After all, these are not objects in *actual reality* (“Wirklichkeit”); but rather ‘objects’ in the sense of ‘whatever can figure as the topic of rational discussion’.

⁵ See, for instance, Michael Dummett, “Frege’s Myth of the Third Realm”, in: *Frege and Other Philosophers*, Oxford: Oxford University Press, 1991, 249-262.. For a critique of this reading and a proper development of the alternative on which I rely in the following see Thomas Ricketts, “Objectivity and Objecthood: Frege’s Metaphysics of Judgment”, in: Haarparanta, L. and Hintikka, J. (eds.), *Frege Synthesized*, Dordrecht: Reidel, 1986, 65-95.

⁶ Frege, “Der Gedanke”, 69.

Just as ‘is true’ is said of ownerless thoughts, so is, of course, ‘follows from’.

Accordingly, the science of logic must study the relations between these ownerless thoughts. Frege’s view of the way in which the laws of logic are related to our acts of thinking flows from this point. On the surface the matter seems straightforward. In the opening paragraph of “Der Gedanke” Frege frames his anti-psychologism in terms of the distinction between two kinds of “laws of thought” that, he says, are easily confused: “laws of taking things to be true” and the “laws of being true”. The former are “psychological laws” that *explain* what people *actually think*: “Whether what you take for true is false or true, your so taking it comes about in accordance with psychological laws.” Logic is concerned not with psychological laws, but rather with the “laws of being true” that pertain to the relations of the objects in the third realm. The way they stand to our acts of thinking is *normative*: “From the laws of being true there follow prescriptions about taking-to-be-true, thinking, judging, inferring”. Frege insists that these two kinds of laws must always be strictly distinguished: “In order to [...] prevent the blurring of the boundary between psychology and logic, I assign to logic the task of discovering the laws of being true, not the laws of taking-to-be-true or of thinking.”⁷

In *Grundgesetze der Arithmetik* Frege points out that the normative character of the relation of a law to the activity of thinking is not specific to logical laws. The point holds for every law that states what is, whether in the realm of physics, chemistry or mathematics: “Any law asserting what is can be conceived as prescribing that one ought to think in conformity with it.” Logic is marked off as being concerned with “the most general laws, which prescribe universally how one ought to think if one is to think at all.”⁸ The mistake of psychologism is to regard the way in which the laws of logic are related to our acts of thinking as akin to the way in which the laws physics are *related to events in the world*:

[...] the expression ‘law of thought’ seduces us into supposing that [the laws of logic] govern thinking in the same way as laws of nature govern events in the external world. In that case they can be nothing but laws of psychology: for thinking is a mental process. And if logic were concerned with these psychological laws it would be part of psychology.⁹

Now, it is clear that if the word ‘governing’ is given the *wrong sense* – namely: “in the same way as laws of nature govern events in the external world” – one ends up in pitfalls of psychologistic logic. The official purpose of Frege’s remark is to resist is the idea that logic could be executed like an empirical science. What he wants to bring out is that it is impossible to *derive* the laws of logic from an empirical investigation of what people actually think. But

⁷ Gottlob Frege, “Der Gedanke”, 58-59.

⁸ Gottlob Frege, *Grundgesetze der Arithmetik, I/II*, Hildesheim: Olms Verlag, 1998, Introduction, XV.

⁹ *Ibid.*, XV.

there is also the opposite danger, which Frege does not mention. If one were to leave the explanation of the activity of thinking *entirely* to the “laws of psychology”, so conceived, then there would be no room for *any sense* in which thinking could be said to be ‘governed’ by the “laws of logic”. After all, it supposed to be illicit to ever blur the line between these two kinds of laws. In consequence, there would be no *explanatory* route whatsoever *from* the “laws of truth” and their “prescriptions” *to* our acts of taking something to be true. However, to hold that there is *no sense* in which the laws of logic are operative or “active” in our thinking and judging would create a problem within Frege’s system. For, we are supposed to arrive through reflection on judging at the very idea of a “third realm” determined by the “laws of truth”. But if such laws could in *no way* be operative or “play a part” in our activity of thinking and judging, that notion could not be *available for reflection*. It follows that if there is to be a science of logic in Frege’s sense, the activity of thinking and judging cannot be left *entirely* to psychology.

It would obviously be a mistake to hold that it is Frege’s official position that the activity of thinking is to be left entirely to psychology. In most places Frege brushes the issue to the side.¹⁰ But his view on matter is already entailed by the fact that the *Begriffsschrift* includes the judgment stroke. And it is clear that psychologism about thinking would simply contradict Frege’s claim that the *fundamental* description of this activity is as an act of “*grasping*” objects belonging to realm of sense. Frege makes this point explicit in a manuscript: the activity of thinking “cannot be completely understood from a purely psychological standpoint”, since with the object of the activity “something comes into view whose nature is no longer mental in the proper sense, namely the thought”. This raises the question how to conceive of this “process”, which now begins to seem quite peculiar – “perhaps the most mysterious of all”, as Frege puts it. It doesn’t seem to fit in any of the three realms, but rather “takes place on the very confines of the mental”. But, once again, the issue is set aside: “we do not need to concern ourselves with it in logic”.¹¹

As it does concerns us here, let’s ask how a logical law or, for that matter, anything belonging to the “third realm” *can* come to play a part in the explanation of our activity of thinking. For it would seem that if the acts of judging and thinking cannot be left entirely to psychology, then it should be possible for something different from “psychological laws” to play a part in the explanation of how they come about. After all, these are, according to Frege,

¹⁰ In the opening paragraph of “Der Gedanke” the question arises: “But may not logical laws also have played a part in this mental process [of judging]?”. The answer is that he doesn’t want to “dispute” this, but that this question doesn’t belong to a logical investigation.

¹¹ See Gottlob Frege, “Logik” (1897), in: *Schriften zur Logik und Sprachphilosophie*, Hamburg: Meiner, 2001, 35-73, here 63-64.

“processes” that located in *reality* (“Wirklichkeit”). By contrast to the occupants of the “third realm”, they belong to a “what happens” and involve individuals situated in space and time. There must thus be *some* “laws” that “govern” them.

Now, importantly it is not that Frege would generally deny an explanatory role to what is in the realm of sense. Once it is established that there are “thoughts” belonging to a realm *different* from the psychological he makes room for the idea of them entering into the explanation of our activity. At the end of “Der Gedanke” he points out that the “third realm” is not cut off from what goes on in actual reality. There is a way for thoughts to, as it were, acquire efficacy in reality, to become “actual” and “active” in a certain sense:

Thoughts are not wholly unactual but their actuality is quite different from the actuality of things. And their action is brought about by a performance of the thinker; without this they would be inactive, at least as far as we can see.¹²

One might put it like this. The idea that “reasons are causes” (in some sense of these words) is *not* alien to Frege. A Fregean “thought” can come to figure as a reason to do something – for instance, affirm another “thought” – and it can become a cause that makes a person do that something. It just that this requires the “performance of the thinker”: the “thought” must be grasped and affirmed by the person if it is to explain the person’s grasping and affirming another “thought”. That is the way in which the occupants of the “third realm” can gain “actuality” and thus play a part in the explanation of our acts of thinking, judging and acting. And it is the only way, “at least as far as we can see”.

2.2. *The laws of logic and the ‘impossibility of our rejecting them’*

It is clear that a lot must already be in place for a subject to be in the position to grasp a “thought”. Any act of grasping involves dividing the thought into its elements and thus introduces connections with other “thoughts”, including other ones grasped and affirmed by the subject.¹³ One can begin to articulate that point by appealing to what is usually called ‘standing beliefs’ and the way they support each other in virtue of the inferential relations between their contents. And if we analyze that system one might get to other more remote and implicit ‘background beliefs’, including beliefs about kinds of inferential relations. But such a procedure must come to an end someplace – in many places, actually. One place is mentioned

¹² Gottlob Frege, “Der Gedanke”, 77.

¹³ See Gottlob Frege, “Über Begriff und Gegenstand”, in: *Vierteljahresschrift für wissenschaftliche Philosophie* 16, 1892, pp. 192-205.

by Frege in answer to the question how the basic laws of logic are related to our recognizing them as true:

[T]he question why and with what right we recognize a logical law as true, logic can only answer by reducing it to another logical law. Where that is not possible, logic can give no answer. Stepping outside logic, one can say: we are forced to make judgments by our own nature and external circumstances, and if we make judgments we cannot reject this law – of identity for example; we must recognize it if we are not to throw our thought into confusion and in the end renounce judgment altogether. I shall neither dispute nor support this view; I shall merely remark that what we have here is not a logical consequence. What is given is not a reason for something's being true but for our taking it to be true. Not only that: this impossibility of our rejecting the law in question hinders us not at all supposing beings who do reject it; where it hinders us is in having doubts whether we or they are right. At least this is true of myself. If other persons presume to acknowledge and doubt a law in the same breath, it seems to me an attempt to jump out of one's own skin against which I can do no more than urgently warn them. Anyone who has once acknowledged a law of truth has by the same token acknowledged a law that prescribes the way in which one ought to judge, no matter where, or when, or by whom the judgment is made.¹⁴

The passage stands in the context of Frege's argument against psychologism.¹⁵ But the way in which Frege proceeds is very revealing for our question about the relation between the normative and the explanatory. The passage begins with Frege specifying what it is to *explain* and *justify*, within logic, why we take a logical law to be true. As long as it is possible to “reduce” the logical law to another logical law one can appeal to an explanatory relation between our standing beliefs that is underwritten by the logical relation between the relevant occupants of the third realm. In this case the explanation will be at the same time a justification. The remainder of the passage proceeds from the acknowledgment that this is not *always* possible. That is to say, in this picture ‘logic cannot take care of itself’, as Wittgenstein would say. On the basic level it cannot explain our thinking in accordance with it; our thinking as we do would have to receive another kind of explanation, if it is to be explained at all. This would involve “stepping outside logic” and appealing to *something else*. Since his interest lies in logic Frege doesn't pursue such an inquiry. But he indicates what kind of thing one would have to invoke in such an explanation: our “being forced to make judgments by our own nature and external circumstances” and the “impossibility of our rejecting the law”.

The significance of this passage is sometimes missed. It has been suggested that the danger of “throwing our thought into confusion” figures as a *justificatory reason* for

¹⁴ Frege, *Grundgesetze der Arithmetik*, XVII.

¹⁵ The argument focuses on a scenario that the psychologist logician would have to regard as possible, namely that there could be creatures, who think according to laws of logic that are fundamentally different from ours. The official aim is show that on closer inspection this scenario dissolves under one's hands: it is not clear what we are supposed to imagine. See James Conant, “The Search for Logical Alien Thought: Descartes, Kant, Frege, and the *Tractatus*”, *Philosophical Topics*, Vol. 20, 1, 1991, 115-180, especially 142-150.

“recognizing the law”.¹⁶ And Frege uses indeed the word “reason” (“Grund”) at this point. But he denies explicitly that there is a “logical” connection that could underwrite a rationalization. Accordingly, the word “reason” (“Grund”) cannot be read as standing for a rational nexus. It signifies another kind of explanation, one that is *not* at the same time a justification. That is to say, there is an “impossibility of our rejecting the law”; we *must* recognize it as true. But the “impossibility” of rejecting it does not derive from the laws of logic; the ‘must’ that *makes* us recognize it is *not logical compulsion*. That we have to recognize the law *without* having a rational reason is, according to Frege, not threatening once we realize that doubt cannot arise. For we have always already judged in this matter: in affirming any thought we have affirmed the laws of logic. This is simply what we do insofar as we are “forced to make judgments by our own nature and external circumstances”. The “impossibility of our rejecting the law” thus also “hinders us in having doubts whether we [...] are right”. Frege inserts “At least this is true of myself” and warns against the “attempt to jump out of one’s own skin” by trying to doubt a law one always already acknowledged in making any judgment.

This should sound familiar the readers of the later Wittgenstein. The insight that it is *impossible* to “dig below bedrock” is not Wittgenstein’s discovery nor is the thought that there *must be* cases in which acting without reason is not acting without right. Frege knew that already. If he had liked slogans, “Reasons must come to an end” could have been one of them. Compare the above passage from the *Grundgesetze* with the famous *PU* §217: “If I have exhausted justifications I have reached bedrock, and my spade is turned. Then I am inclined to say: ‘This is simply what I do.’” On the face of it, it looks like the only thing Wittgenstein came up with in this remark is the vivid wording. If one wanted to learn *that* lesson, it would have been quicker to read Frege’s rather short introduction to the *Grundgesetze der Arithmetik* instead of the quite long and meandering *Philosophische Untersuchungen*.

The problem is that, in itself, the slogan presents no position to rest with. In Frege’s case the appeal to the idea that acting without reason is not always acting without right covers over a gap between the normative and the explanatory that arises out of the strict distinction between the laws that *prescribe* how we *ought* to think and the laws that *explain* how we *actually* think, the heart of his anti-psychologism. If the laws of logic belong to the third realm whose occupants can only become “active” or “actual” via the “performance of a

¹⁶ See Tyler Burge, “Frege on Knowing the Third Realm”, in: *Truth, Thought, Reason. Essays on Frege*, Oxford: Oxford University Press, 2005, 315.

thinker” then it follows that on the most fundamental level there can be no rational connection between “the impossibility of our rejecting the law” and the law being true. In the paragraph preceding the passage I quoted Frege states this explicitly. He says that the sentences ‘It is impossible for people in the year 1893 to recognize an object as different from itself’ and ‘Every object is identical with itself’ express two utterly different laws: “The latter is a law of being true, the former a law of people’s taking-to-be-true. Their content is wholly different, and they are independent of one another; neither can be inferred from the other.” Frege’s target is, of course, again the “psychological logician” and his attempt to conceive of the laws of logic as something that could be derived from an empirical generalization. To emphasize this point Frege inserts a specification of time in the first sentence. But suppose we replace it with one of the following: ‘It is impossible for *human beings* to recognize an object as different from itself’ or ‘It is impossible *for us* to recognize an object as different from itself’. Given Frege’s view these sentence would still express a “psychological law”, utterly different and without any rational connection to the logical law of identity. But if that is so, then it is a mere stroke of luck that we are like that, that there is, for us, that “impossibility”, that “our nature” forces us to judge and prevents us from “throwing our thought into confusion and renouncing judgment altogether”. We are, as it were, blessed creatures, because there is that something – not logic – that forces us to think in conformity with the law of identity.

The same will hold for the “attempt to jump out one’s own skin” by doubting something one takes-to-be-true. For, if there are two ‘laws of identity’ we should expect that there are also two ‘laws of non-contradiction’: the logical law that prescribes how we ought to think and a different and utterly unconnected law that states the impossibility for us to at once affirm and deny the same thing of the same thing.¹⁷ And once again, it will be a mere happy accident that one fits to the other so that the way we are forced to think is in *conformity* with the *logical* law of non-contradiction. But if that is so any drawing of an inference will also be driven by nothing but that non-logical force that prevents us from affirming the opposite of what we just recognized as the conclusion.

In the end Frege has no choice but to leave the activity of thinking and judging entirely to psychology. That leads to the disastrous consequence I mentioned above. If the laws of logic play *no part* in our thinking and judging then the very idea of a “law of truth” is not available for reflection. In consequence, it becomes mysterious how we could ever form the notion that our thinking is subject to any “prescriptions”. There is thus a sense in which

¹⁷ This is precisely what Łukasiewicz argues in his critique of Aristotle. See Jan Łukasiewicz, “On the Principle of Contradiction in Aristotle”, *Review of Metaphysics*, 24, 1971, 485–509. For a discussion see Irad Kimhi, *Thinking and Being: the Two Way Capacity*, (Ms.).

what Frege thought the logician needs to “fear the least” is what he needs to fear the most: “to be reproached with the fact that his statements do not accord with how we think naturally”.¹⁸ The “laws of truth” must be rendered intelligible as something that can *govern* our thinking, if the idea of such laws is to make sense at all. There must be a connection between these laws and how we “think naturally” – a connection, for instance, between the law of identity and what makes it ‘impossible’ *for us* to recognize an object as different from itself. For, if this is not guaranteed the prescriptions that follow from the logical laws are “in danger of becoming empty”, as Wittgenstein might put it.¹⁹ The decisive point is thus not the insight that reasons must come to an end, but rather that they must come to an end in such a way that there is no the gulf between the normative and the explanatory. The task is to render intelligible how this could be.

2.3. Concepts and our understanding them

One might think that the difficulty arises because of Frege’s conception of logic as the most general science, and that the solution is to reject the idea that logic has a subject matter.²⁰ I think that the later Wittgenstein came to see that this threatens to leave the crucial difficulty in place. For an *analogous* problem arises in the case of the question how the meaning of a word or a concept is related to the acts of applying it in an assertion or a judgment. In order to make the issue more tractable, Wittgenstein considers the scenario of a pupil learning to extend a number series according to the principle ($n, n + 2$). This illustration has led to a number of misunderstandings in the literature.²¹ But if we see it in the light of the difficulty that arises

¹⁸ Frege, “Logik”, 65.

¹⁹ See *PU* §107. See also *PU* §81. Positively speaking this is the point Wittgenstein insists on in passages like this: “The laws of logic are indeed the expression of ‘thinking habits’ but also of the habit of *thinking*. That is to say, they can be said to show: how human beings think, and also *what* human beings call ‘thinking’.” Ludwig Wittgenstein, *Bemerkungen über die Grundlagen der Mathematik*, I, §131. In the following cited as *BGM*.

²⁰ That was, of course, Wittgenstein’s strategy in the *Traktatus*. See Cora Diamond, “Throwing Away the Ladder”, *Philosophy*, 62, 1998, 5-27; James Conant, “The Method of the Traktatus”, in: Erich Reck, *From Frege to Wittgenstein: Perspectives on Early Analytical Philosophy*, Oxford: Oxford University Press, 374-462.

²¹ See Ludwig Wittgenstein, *PU*, §§143-155 and §§185-242. It has been suggested, for instance, that the problem is about mathematics or, alternatively, that the problem arises due to unwarranted application of a mathematical model of determinacy to our ordinary concept use. (For the latter see, for instance, Stanley Cavell, “The Argument of the Ordinary: Scenes of Instruction in Wittgenstein and Kripke”, in: *Conditions Handsome and Unhandsome*. Chicago: Chicago University Press, 1990, pp. 64-100.) Focusing on the use of *signs* in the school arithmetical illustration other readers have suggested that the problem is a problem that only arises for the constitution of *linguistic meaning* and not for the constitution of conceptual content. (See, for instance, Colin McGinn, *Wittgenstein on Meaning*, Oxford: Basil Blackwell, 1984, 144 *ff.*) Yet other readers suggest that it is about the epistemological problem how I can come know the meaning what another means with his words. (See Donald Davidson, “The Second Person”, in: *Subjective, Intersubjective, Objective*, Oxford: Oxford University Press, 2001, pp. 107-121, esp. 110-111) I don’t want to deny that all of these questions are addressed in the *Philosophische Untersuchungen*. But I think that they present a distraction when it comes to the task of getting the rule-following *problem* into view. As I see it, the latter problem arises regardless of the preferred view on the relation between conceptual thinking and language and it is independent from any specific position on the difference between our use mathematical functions and the deployment of ordinary concepts.

within Frege, it is, I think, clear that what we are looking at is the most primitive case of the troubling relation between a law that prescribes what we ought to think and the correlated acts of judging. In “grasping” a “thought” the thinker must divide it into its components. The predicative “thought”-component – the predicative “sense” in Fregean terms – will connect the given act of thinking with infinitely many other “thoughts” that can be grasped and taken to be true or false.²² And since we ought to take to be true what is true and false what is false the predicative “thought”-component acts as a kind of “law” that “prescribes” what we ought to think – a “rule”, as Wittgenstein says.

Now, it is important to realize that Wittgenstein does not deny the observation that was Frege’s starting point namely, that ‘is true’ and ‘follows from’ are said “timelessly”. The propositions of mathematics are, as Wittgenstein insists, *not* “empirical propositions” about what people tend to think and do. When we say that applying the operation of addition yields a certain result, “the verb (‘get’, ‘yields’ etc.) is used non-temporally”. (*BGM*, VI, §2) The same holds for the articulation of inferential relations: “When we say: ‘This sentence follows from that one’ here again ‘to follow’ is being used *non-temporally*. (And this shows that the sentence does not express the result of an experiment.)” Because this relation holds “completely independently from any outside happenings” – and *a fortiori* independently of what people are *actually* thinking at any particular place and time – Wittgenstein characterizes it as an “internal relation”.²³ That is the “hardness of the logical ‘must’” that connects a concept – the predicative “sense”-component in Fregean terms – with the infinite series of deployments in judgments that are as they ought to be.

The question that drives the rule-following considerations is how the judging subject is related to the series of “thoughts” to be affirmed and denied. If it is to be ‘impossible’ to at once affirm and deny the same thought and thus possible for one judgment to cancel out the opposed one, the subject’s current judgment must somehow be connected with the judgments she *could* frame by deploying the relevant predicative element. It follows that if an individual

²² There is a further complication since one and the same thought can be divided in different ways. (See Frege, “Über Begriff und Gegenstand”) But we can leave that aside for the present purposes..

²³ See *BGM*, I, §§102-104. The terminology of “internal relations” is, of course, from the *Traktatus*. (See Ludwig Wittgenstein, *Traktatus Logico-Philosophicus*, 4.123.) It therefore peculiar when readers of the later Wittgenstein invoke that notion in order to articulate what they regard as the solution to the rule-following paradox. (See, for instance, Gordon Baker and Peter Hacker, *Skepticism, Rules and Language*, Oxford: Basil Blackwell, 1984, 95-96) I sidestep the discussion of the *Traktatus* here, because I think the difficulty about the relation between concepts and acts of deploying them arises already in Frege’s work. It would be different paper to show why Wittgenstein thought that his early self didn’t do justice to this problem. In a nutshell, one might say that the problem about the way in which “internal relations” are conceived in the *Traktatus* is that it makes it impossible to see how the mastery of a language (or a system of concepts) can be *explanatory* of any true judgment. The impossibility of that idea is addressed in the rule-following considerations culminating in *PU*, §242 that discards Wittgenstein’s early attempt to ensure that “logic can take of itself”.

is to be truly said to deploy a concept in a judgment her act must somehow be firmly placed in such series of potential judgments. Something about her must link her with this space of possibilities, if she is to realize any one of them. One way of putting this point is to say that in judging she must *act on an understanding* of the principle of the series: she must “follow” the “law” or “rule”.

In the literature it is not always easy to keep track of what sense of the words ‘following a rule’ is invoked exactly.²⁴ In order to avoid misunderstandings I will first determine the sense of the phrase ‘following a rule’ that is relevant here by specifying constraints that an account of judging intuitively has to meet (§3). The rule-following paradox will then arise in the succeeding section as a puzzle about how these constraints could all be met (§4).

3. Concepts and acts of judgment

3.1. Generality

As it is traditionally conceived a judgment is an act of thinking *something of something*. Take, for instance, the singular judgments ‘*a* is *F*’ or ‘*a* is *R* to *b*’ as framed by a particular subject. For an act of mind to exhibit the internal structure characteristic of thinking of an object *a* that it is *F* or of two objects that they stand in *R* to each other, the predicative element, ‘*F*__’ or ‘__*R*__’, must be such as to potentially appear in infinitely many other judgments. The

²⁴ Some philosophers simply deny that this phrase characterizes a necessary feature of thinking at all. (See, for instance Davidson, “The Second Person”, 112-113) This dismissal of the topic often goes together with linking the word “rule” right at the outset with Wittgenstein’s talk about “practices” and “customs”. In this way of proceeding the kind of “rule” at issue appears to be social, perhaps something in the order of a ‘convention’. And then it looks like, properly defined, the topic is the task is of specifying how the words of a natural or ‘shared’ language have the meaning its speakers take them to have. In consequence, it seems that the task can be dismissed in one of the following two ways: by denying that thinking is language-dependent or by arguing that the prior notion of language is that of an idiolect rather than a shared language. As a result, the question that seems to have elevated the later Wittgenstein appears to be one that, strictly speaking, doesn’t belong to philosophy, but rather to linguistics. To make it a legitimate object for the philosopher’s worries one first would have to establish the language-dependence of thinking and then the social character of language. I want to avoid these debates here. And I think we can do so by strictly distinguishing the basic sense of “rule” that is required to get the difficulty going from the richer sense this term acquires when Wittgenstein working towards a solution appeals to the idea of practices and social institutions. The *problem* can be presented as puzzle about the *metaphysics of our acts of conceptual thinking*. Accordingly, claims such as that thinking is “language-dependent” or “essentially social” can only enter the dialectic if they contribute to the *solution*. Since I will focus on the difficulty, I won’t say much about language and meaning. When language comes up, then mainly for purposes of elucidation and mostly as a “vehicle of thought” and not as “means communication”. In the specification of an aspect of the generality of concepts the “intersubjective” dimension of thinking will come up. But as far as setting up the *problem* is concerned this involves just the idea of a *possible* second person and the imagined communication might as well operate through *telepathy*.

predicative element is, as Aristotle says, “of such a nature as to be predicated of many subjects”. Kant puts it like this: “[E]very concept must be thought as a representation which is contained in an infinite number of different possible representations (as their common character).”²⁵ The predicative element possesses, as it were, an inward infinity. It is not that other objects must fall under the concept *F*. That might not always be the case. Rather, it must be possible to deploy the predicate in other acts of affirmation or denial: in affirming or denying that *Fb*, *Fc* etc. The concept sets no limit as to how often it can be deployed; it reaches beyond any given act of mind in which it occurs and points to what is in principle an unlimited multiplicity of judgments each of which exemplifying how the concept figures in thinking. As the concept determines the respect in which all the acts are the same it can be said to ‘unite’ this manifold into a series of acts of “going on in the same way”.

As we have seen in the last section, this traditional point can be articulated in the Fregean terminology. But the way I put so far leaves the thinker of the infinitely many potential judgment unspecified. The demand was just that that predicative element must be able to figure other judgments. That doesn’t tell us how the judging subject has to be related to this series of potential judgments if she is to be truly said to deploy the relevant concept in a judgment. Clearly she must be somehow connected with this space of possibilities. When we ascribe the singular judgment that *Fa* to a subject we therefore implicitly refer to what is in principle an infinite series of potential judgments *by that subject* and thus to *something about her* that underwrites our reaching ahead in this way: to her understanding of the predicative element ‘*F*__’ that each of her judgments exemplifies.

Gareth Evans termed *this* the “Generality Constraint”. He expressed it in terms of the “abilities” a judging subject must be credited with: if a subject is to count as framing a judgment involving the concept *F*, her act must be the exercise of a capacity which she can exercise elsewhere, again and again – in deploying *F* in other judgments.²⁶ Since the word ‘ability’ is very loaded I will set it aside and begin with the following formulation: ‘When we ascribe a judgment to a subject we (i) implicitly *point to an infinite set of potential acts by that subject* and (ii) in doing so advert to something which *unites them into series of acts of going on in the same way.*’ It should be clear that this formula is less demanding than what Evans had in mind and, in any case, defines an extremely wide class of relations between a manifold of acts by a subject and whatever can be said to unite them into some sort of

²⁵ Aristotle, *De Interpretatione*, 7; Immanuel Kant, *Kritik der reinen Vernunft*, A25/B40. This is how P.F. Strawson’s formulates it in his *Individuals*: “The idea of a predicate is correlative with that of range of distinguishable individuals of which the predicate can be significantly, though not necessarily truly, affirmed.” (P.F. Strawson, *Individuals*, London: Methuen, 1959, 99.)

²⁶ See Gareth Evans, *The Varieties of Reference*, Oxford: Oxford University Press, 1982, 100-105.

sequence. Let us, then, assemble further constraints that characterize the relevant species of the genus: that peculiar way of implicitly reaching ahead characteristic of the ascription of a judgment. Over and above a broad sense of “generality” there are, I will suggest, five further constraints. Though not uncontroversial, each of them will be familiar from the literature. Together they will determine the sense in which judgments are intuitively acts of “following a rule”.

Wittgenstein’s illustration from school arithmetic is helpful in this connection since it allows us to set aside any special difficulties that arise from the role perception in empirical judgments.²⁷ So let’s return, once again, to the pupil learning to extend a number series according to the formula ‘ $(n, n + 2)$ ’. The student was shown what to do in a few cases and then asked to continue. If he gets the hang of it, we can say that at each step he represents a pair of numbers as satisfying the relational predicate ‘ ζ is next-but-one-after ζ ’. Now, whether his behavior can be seen in this light will not be clear from the first steps he takes. It depends on how he goes on. If it’s haywire after a couple of steps, we will suspect that, all along, he was just randomly writing figures on the paper, or anyway that he hasn’t yet grasped the principle we have in mind. What our initial hesitation brings out is this. When we ascribe the judgment ‘8 is the next-but-one-after 6’ to him, we implicitly reach ahead to his adding, for instance, 2 to 1000, even though he might never actually get that far.

Now, the first thing to notice is that this way of reaching ahead must be distinguished from another sense in which a description of present goings on might be said to point ahead to what is not there yet. If you say of a stone ‘It is rolling down the hill’ or of a person ‘He is walking across the street’, your description implicitly points beyond the present moment and reaches ahead to the next phases of the process you describe. As long as the stone hasn’t arrived at the bottom yet or the person on the other side of the street, there will be, as far you know, infinitely many different trajectories the unfolding movement might take on its way to completion. Consequently, we might say that in describing what is going on here and now you implicitly advert to something – the overarching process or, if you will, the “intention in action” – that determines a range of potential movements that would count as its continuing in the same way. As long as the process is still on the way there are infinitely many possible subordinated phases in view. At the same time, this way of ‘reaching ahead’ clearly has a certain *limit* internal to it. It sets an end by pointing to the *completion* of the process or the

²⁷ The deployment of empirical concepts like *green*, *square* or *fluffy* to objects encountered in experience like *this stone*, *that table* or *this dog* is, of course, more basic than the extending of a number series. But the former cases introduce the vexed question of how the ‘concept’ is joined with the matter ‘given’ through the senses. The arithmetical example enables us to set these complications aside and focus on the joining of the ‘concept’ and with the very act of judgment in which it is in which it ‘figures’.

execution of the intention. As soon as the stone has arrived at bottom of the hill or the person on the other side of the street the item you implicitly advert to has exhausted its power to unite individual operations into a succession of continuing in the same way.²⁸

Similarly, if you say of the pupil as he is writing ‘2, 4, 6, 8...’ that he is doing his homework of writing out the series up to 20 your judgment will implicitly point to the next potential steps of his doing his homework, and it will advert to something that unites all of them: his unfolding action or his “intention in action”. As soon as the action is completed, the intention executed, there is nothing left for it to unite. Still, there is a sense in which your original statement is not done with implicitly ‘reaching ahead’, for it points to *other occasions* in which he might develop another segment of the series. If his present act of mind is to have the determinate content of an intention to develop the series up to any particular point, it must be connected with those acts outside the intended range. Despite your firm knowledge that he will not go on forever and quite reasonably has no intention to do so, your description implicitly reaches ahead to *his* adding 2 to, say, 1000 or any other even number. In taking what he is doing as an act of adding 2 you thus implicitly reach *far beyond* anything he could be said to be in the process of doing or having the intention to do. The item to which you implicitly appeal in doing so must thus be something that sets *no limit*, but unites an *unlimited* series of steps. It doesn’t come to an end in any one of them, but remains unexhausted throughout the whole series. Instead of getting ‘completed’ or ‘executed’ it is something *general* that can get *actualized* or *instantiated*, again and again *ad nauseam*.

This specification of the ‘unlimited’ character of our reaching ahead turns our ‘thin’ formula into a Generality Constraint. Still, it sounds weaker than Evans’ version that includes the word ‘ability’. So let’s take a look what else might be involved.

3.2. Normativity

With our weak Generality Constraint we are still onto a very wide class. One might say that a stone can only be said to be rolling down a hill if it is conceived as the kind of thing that might do that kind of thing. Accordingly, your description of what is happening with *this* one *here and now* will implicitly advert to something that unites all of its potential movements down other hills on other occasions. We can narrow our focus by introducing a further specification of the relation in which our pupil’s acts stand to the relevant “general” item they instantiate. We get the first one when we ask what becomes of the idea that a concept figures

²⁸ On the grammar of this form of reaching ahead see Michael Thompson, “Naive Action Theory”, in: *Life and Action*, Cambridge, Mass.:Harvard University Press, 2008, 85-146.

as some sort of standard when we articulate this aspect in the perspective of *describing* the pupil.

To say that in describing him as adding 2 to 4 we implicitly point ahead to his writing ‘1000, 1002...’ is *not* to say that our description somehow entails some sort of prediction. For, our assumption is not necessarily falsified, if he later goes on to write ‘1000, 1004’. Our peculiar ‘reaching ahead’ leaves space for the possibility that the mistake might be in his act, rather than in our description. The general item that unites the unlimited series of *potential* acts is such that it allows for the pupil’s *actual* doing to remain, in some sense, under the description ‘moving to the next-but-one number’, even though what he actually writes down might be a different one – just as judging that *Fa* might be said to remain under the description ‘listing the objects that fall under *F*’ even though *a* is not *F*, the thought held true consequently false, and the judgment thus not as it ought to be. This non-predictive character of our reaching ahead is the reason why it seems appropriate to say that the concept ζ is *next-but-one-after* ζ figures as some sort of “norm” or “rule” in relation to the acts of deploying it.²⁹ The ‘something general’ we implicitly advert to in describing our pupil as adding acts as some sort of standard or grounds some kind of ‘ought’ or ‘must’ that sorts the steps he actually goes on to take into those that are in the light of it ‘correct’ and those that are ‘incorrect’. The kind of generality we are after is thus one that stands in a *normative relation* to its instances. I will call this the ‘Normativity Constraint’.

3.3. Interpretation

It should be clear that the bare notion of *something general that sorts individual operations into ‘correct’ and ‘incorrect’* is, once again, a very wide category and that not all the species of this genus are relevant here. We can take a further step towards isolating the relevant species if we consider the fact that we reached the idea of the normative by reflecting on what is implicit in our *description* of what the pupil is doing. Contrast this with the following scenario. Say the security rules on a construction site demand that one wears a helmet while laying bricks. You deem that silly and ignore it. I might protest: ‘You ought to!’ Still, I can describe what you are doing without referring to the rule in the light of which I think your

²⁹ See *BGM*, VI, §15: “[W]hy was not this a genuine prediction: ‘If you follow the rule, you will produce this’? Whereas the following is certainly a genuine prediction: ‘If you follow the rule in all conscience, you will...’ The answer is: the first is not a prediction because I might also have said: ‘If you follow the rule, you *must* produce this.’ It is not a prediction if the concept of *following* the rule is so determined, that the result is the criterion for whether the rule was followed.”

manners are offensive. In the case of our critical stance towards the pupil it is impossible to separate the descriptive and the evaluative element in this way. The “rule” in light of which he might be said to make a mistake doesn’t regulate behavior that is intelligible independently of it. In describing him as adding we already refer him to the rule. We might shift to ‘He writes figures on the paper’. But under that description the action cannot be measured by mathematical standards. There is no sense in which one ought to write figures in that order. Writing those figures is only ‘correct’ or ‘incorrect’ *as* the purported development of the number series. Quite in general: stringing together some graphic marks or linking some elements in the head can only be brought under the relevant standard having to do with truth, if it is regarded as the purported representation of something as satisfying a certain predicate or falling under a certain concept. That is to say, our description of the pupil’s act is *logically dependent* on the “rule” in the light of which his behavior is deemed ‘correct’ or ‘incorrect’.³⁰ The standard to which we bring him in the evaluation is somehow internal to our conception of what he is doing: we must appeal to it already in interpreting his behavior as a purported act of adding 2. The relevant relation between the general and the particular must have not only a *normative*, but also an *interpretative* dimension.

3.4. Non-Accidentality

This distinction between two kinds of rules has a further aspect. As far as the security rule is concerned it doesn’t matter whether you wear a helmet *because of this rule*. You might happen to be someone who always wears a helmet, or maybe it struck you that it looks good on the people around and that it would complete your outfit nicely. However it came to be on your head, as long as its there you are safe from reproach.³¹ It is more difficult for the student in class. The math teacher would certainly take back any praise of progress in the learning process, if she formed the suspicion that the pupil just put down those figures that struck him as looking pretty on the paper – perhaps they formed a smiling face. In this case it was a *mere accident* that the numbers he wrote down accorded with the “rule” she had in mind. It wasn’t even that he made a mathematical mistake; as far as class is concerned he was just wasting time. Her original statement ‘He thinks that 8 is next-but-one-after 6’ is falsified: her

³⁰ On the distinction between these two kinds of “rules” see John Rawls “Two Concepts of Rules”, in: *Collected Papers*, Cambridge, Mass.: Harvard University Press, 1999, pp. 20-46. Rawls’ discussion focuses on social rules. But nothing I said so far requires that the rule is *actually* “shared”. For all that is relevant up to this point, the pupil might have figured out adding all by himself.

³¹ It seems that it would require a *further clause* in the formula expressing the rule for it to include the demand that one wears the helmet *because of this rule* – say, as a sign of respect for the company. Controlling whether your manners satisfy *this* rule would obviously be a much more complicated affair – so would describing what it is that you are doing when you act in accordance with it.

implicitly reaching ahead to his writing ‘1000, 1002...’ turned out to be a mere projection. If it is not to be, it must somehow be underwritten by *something about him*.

But what is that ‘something’? It seems it can’t be his intention to execute the series up to a particular point. For, as we have seen, ascribing such an intention to him involves reaching beyond it to steps outside the intended range. For the same reasons, the numbers he wrote down or went through in his head so far can’t figure in the role of that ‘something’. For that will be compatible not only with the rule ‘Add 2’ but also, for instance, with the *bent* rule ‘Add 2 up to 1000, 4 up to 2000, 6 up to 3000 and so on’.³² However many figures he might write down, there will always be an infinite number of patterns with which that sequence accords. If his acts are to be connected in a determinate fashion with any one of them, it is, as Wilfrid Sellars puts it, not enough if they “conform” to the pattern; they must somehow be “governed” by it.³³ There must be a sense in which what pupil is doing is happening *because of the rule*. The general item that unites an unlimited manifold of potential acts can therefore not be something *merely abstract* and “inactive”. If it is to underwrite our reaching ahead such that it is not a mere projection but an implicit aspect of our *truly* describing him as extending the series it must be in some sense *actual in him* such that it can *explain* the steps he takes.

In his original formulation of the Generality Constraint Evans marked this explanatory aspect of the general item we appeal to with the word ‘ability’. Leaving that word aside, the point can be put in our context like this: in taking the act of a subject to exhibit the internal structure characteristic of thinking of *a* that it is *F* we not only take it that she might also affirm that *Fb*, *Fc* etc.; we also hold that there will be “common explanation” for her judging that *Fa* and her judging that *Fb*, *Fc* etc.³⁴ Only if this is so is it not a *mere accident* that her acts exhibit a pattern under which they can be united into a series of going on in the same

³² *PU*, §185.

³³ See Wilfrid Sellars, “Some Reflections on Language Games”, in: *Science, Perception and Reality*, Atascadero, Ca.: Ridgeview Publishing Company, 1963 and 1991, pp. 321-358, esp. 322ff..See also John Haugland’s distinction between “rule-exhibiting” and “rule-governed behaviour” in his “Truth and Rule-Following”, in: *Having Thought*, Cambridge, Mass.: Harvard University Press, 1998, 305-362, esp. 305ff.

³⁴ This formulation departs from Evans, who puts it terms of a “common explanation” not of the subject’s judgments that *Fa* and that *Fb*, but of her *understanding* of ‘*Fa*’ and ‘*Fb*’. (See Evans *The Varieties of Reference*, 101). The simplicity of the example of extending a number series by adding 2, is, I think, designed to allow us to put the complication aside that call for such qualifications. In the case of perceptual judgments it can seem that what explains the judgment, if it is to be knowledge, can only be the *individual thing and its properties*. These, it seems, are the elements in actual reality that can appear in an explanatory role: the thinker judges that the thing is green, ‘because’ the thing is green. It is therefore harder to make out the other explanatory factor at work: the concept that figures as common explanation of a whole range of judgments. In the simple arithmetical case there is only the latter explanatory aspect. In consequence, it becomes clear that what it explains is judgment about which number comes next.

way. This, then, shall be our ‘Non-Accidentality Constraint’: the relation between the general and particular we are after must also have an *explanatory* dimension.

Since the arithmetical illustration allows us to put the complications aside that enter with the role of perception this explanatory role of concepts comes clearly into view. What is explained *primarily* is judgment that is the expression of knowledge – that is, the pupil’s *correct* continuation of the series. If he didn’t get *any* number right, his action would not to be ‘governed’ by the rule. At the same time, the Non-Accidentality Constraint applies even in the case where the pupil makes a mistake. This follows from the Interpretative Constraint: since the acts have to be *interpreted* by reference to the rule in order to be brought under the relevant standard at all, the rule must somehow be operative in *any* act that is *evaluated* in the light of it. A lot must already be place in the subject so that making a mistake can become as a possibility for the subject. And there has be *some* explanatory link between the current act and whatever it is that has to be in place, if this act is to count as a *mistake*, rather than a random writing of figures. So in some *derivative* way the rule must ‘govern’ or ‘explain’ even the acts that deviate from it.

3.5. Transparency

One might think that taken together these three features of the relation in which an individual act can stand to something general – that is, the *normative*, *interpretative* and *explanatory* dimensions of this relation – define the sense in which it might be described as an act of “following a rule”. Accordingly it would hold that a subject *S* is following the rule *R* iff: (i) the sentence expressing *R* contains the description of some logically dependent activities *A*, *B*, *C*; (ii) *R* allows the derivation of ‘ought’ statements that connect *S* with *A*, *B*, *C*; and (iii) *S* is doing *A*, *B*, *C*, *because R*. This formula might define the sense in which mere animals or perhaps even plants could be said to “follow rules”.³⁵ But however useful such a notion of “rule-following” might be for certain purposes, it seems that at least in our case we can’t leave the sense of ‘because’ – that is, the kind of ‘non-accidentality’ – unspecified. Imagine that the pupil’s father discovered that the most efficient way to put his son to sleep is to say out loud: ‘2, 4, 6, 8...’. He might even intend this as a way of teaching adding. Having been submitted to this treatment every night the pattern has gained some sort of actuality in the child: being confronted with numbers the poor fellow now can’t help but recreate the pattern inscribed in his sub-conscious. Given that the father extended the series by adding 2, the concept ξ is next-but-one-after ζ enters into the explanation of the pupil’s present behavior.

³⁵Michael Thompson’s view on the concept of life, for instance, entails that this formula is applicable throughout animate nature. See Thompson, “The Representation of Life”, in: *Life and Action*, 25-82.

But it does so in the wrong way. He might just as well still be drawing a smiling face in numbers.

As a bare minimum, the pupil must take himself to be developing the number series if we are to truly describe him as doing so. And, of course, his ‘I got it!’ might be delusional, and what he is actually doing just a random mess. It is obviously not much progress if he automatically manifests that pattern inscribed in his sub-conscious while accompanying his activity with the thought ‘I’m adding 2’. In order to really count as adding 2, he must somehow *correctly* and *non-accidentally* take himself and thus *know* himself to be manifesting the pattern determined by the concept ζ is next-but-one-after ζ . It is thus not enough that his behavior is just in some way or other explained by the pattern or “rule”. The causal nexus can’t operate behind the subject’s back. It can’t be something hidden from him, something only to be discovered by a psychologist or maybe a neurologist; it must somehow operate through his conception of the ‘because’. The judging subject must be “guided” by his knowledge of the relevant pattern: he must *act on an understanding* of the “rule”.

This further constraint follows from what Michael Dummett calls the “transparency of meaning”.³⁶ Dummett makes the point in the linguistic register by claiming that since asserting is a *self-conscious* activity the speaker must know the sense of the words he uses. Switching back to the register of thinking we can put it this way: since the ‘I think’ must be able to accompany all my judgments I must *know* the content of the elements that figure in my judgment. Given that the Generality Constraint specifies what it is for the act of a subject to exhibit such internal structure, it follows that the three-dimensional relation between the general and the particular it must be *accessible in the first person perspective* of judging. That is to say, in deploying the concept F in a judgment about the object a the subject must implicitly conceive of her activity as something she could do on another occasion as well: for instance, by bringing b under F or by denying that it falls under F . And she must conceive of these acts as having a “common explanation” or springing from a common source – a source, which figures, at the same time, as a standard that sorts them into ‘correct’ and ‘incorrect’. The relation between the general pattern and the acts that exemplify it must somehow be *represented, understood or known* by the acting subject. Shifting to the linguistic register we can express this ‘Transparency Constraint’, as I will call it, like this. In framing the assertion that ‘ a is F ’ the subject’s thinking must implicitly operate, as it were, on two levels at the same time: on the *particular* level on which she predicates ‘ F ’ of ‘ a ’ and on the

³⁶ See Michael Dummett, *The Logical Basis of Metaphysics*, Cambridge, Mass.: Harvard University Press, 1991, 81. See also Sellars’ distinction between “patterned governed” and “rule obeying behaviour” in his “Some Reflections of Language Games”, 324.

general level on which she grasps the “pattern” that she takes her present act to exemplify – the general ‘use’ of the predicate.

3.6. Intersubjectivity

The arithmetical illustration implies, of course, a social situation. And this is how things usually are in ordinary life. But let’s assume for a moment that the pupil somehow figured out it by himself or simply forgot that he was taught. Suppose he is the last man standing. Now that he is alone with the numbers and his writing pad, how does the idea of other subjects enter the perspective of judging? It enters when the ‘normative dimension’ of the described relation between the general and the particular is articulated in the first person perspective. Note that to evaluate a subject’s affirming that *Fa* as ‘incorrect’ is to deploy *F* oneself in a judgment, namely in denying that *Fa*. To conceive of one’s own act of judging that *a* is *F* as the purported instantiation of a general pattern in the light of which it might turn out to be ‘incorrect’ is therefore to entertain the possibility of an opposed judgment. Since I can’t conceive of myself as *at the same time* affirming and denying that *Fa*, conceiving of the possibility of error is conceiving of the possibility of *another subject* deploying the *same* concept in a judgment that is opposed to mine.

Our final specification of the relation between the general and the particular might thus be called the ‘Intersubjectivity Constraint’: an account of judging must be such that it makes space for the possibility of the judgments of two different subjects to stand in a relation of contradiction. This is, of course, just the negative side of the phenomenon that Frege highlights: what I judge can be taken up by another subject; knowledge can be imparted, false opinion can spread. By itself this doesn’t establish, of course, that there can’t be a judging subject unless there are or have been *actual* others. However, it does establish another aspect of the generality of concepts. A concept is not only general in that it sets no limit as to *how often* it can be deployed in a judgment, but also in that it sets no limit as to *by how many subjects* it can be deployed. It thus reaches not only beyond any given act of mind in which it is figured; it also reaches beyond any particular thinker and points to an infinite multiplicity of possible thinkers – all of whose acts it could explain and sort into ‘correct’ and ‘incorrect’. The judging subject has to conceive of her act not only as instantiating a general pattern that is operative *in her*; she must conceive of her act as an instance of something that *can* also be exemplified by the acts of others.

4. A puzzle about concepts

Where the relation between a general pattern or rule *R* and an individual act *A* by a particular subject *S* satisfies these constraints we can say that *S* follows *R* in doing *A* – in the sense of “following a rule” that is intuitively applicable to acts of judgment. None of these constraints is uncontroversial, of course. Each one of them has come under attack in the literature, for different reasons at different times. Here I will consider only one possible ground for rejection: the suspicion that what these constraints, taken together, seem to specify is *unintelligible*. The impression that it is impossible for the relation between the general and the particular to take such a shape is what has come to be called the “rule-following paradox”.

4.1. *The rule-following paradox*

According to the Generality Constraint we don’t describe the pupil as adding 2 to 6 unless we implicitly reach ahead to an infinite series of his potential judgments. And it follows from our Transparency Constraint that *our* description of his act can only reach ahead in this way if his own conception of his act somehow does. That is, it seems our peculiar reaching ahead must be underwritten by *his* doing so in his mind. But what is that supposed to mean? That in understanding the principle of the series his “soul” must implicitly “as it were fly ahead and take all the steps before [he] physically arrived at this or that one”?³⁷ Accordingly, his “soul” would have already done what he could never hope to do with the pen in his hand: extend the series up to infinity. But that seems absurd. The picture of a mind that has always already taken all steps might fit the idea of a divine intellect, but surely not our “discursive intellect”, as Kant calls it: we have to bring objects under concepts.

No doubt, if there is any knowledge to be gained and imparted *something* must, “as it were fly ahead”. And so it seems we have to distinguish between “the rule” that determines the infinite extension of the series and anyone’s activity of judging. With this distinction in place the question arises how our Non-Accidentality Constraint is to be met. How does that general pattern come to be explanatory of my individual acts of mind? As the relevant causal nexus is not supposed to operate ‘behind my back’, it seems that the rule can only acquire its efficacy through my act of grasping it. As Frege would say: what belongs to the realm of sense requires the “performance of a thinker” if it is to have “actuality”. Since that grasping cannot consist in the whole extension of the number series somehow occurring before my inner eye it seems that there also has to be a difference between whatever appears in my mind through such uptake and my projecting the thing so grasped onto the segment at hand. But if my ‘grasping the rule’ is *not yet* my ‘applying it’ what explains, at any point in the series, my

³⁷ See *PU*, § 188.

regarding *this* rather than *that* step as an instance of the pattern? It looks like there needs to be a *further* mental item that relates the rule as it appears in my mind to the present segment of the series.

Wherever a definition is available this train of thought is not very threatening. But a definition will, of course, just introduce further concepts. Given the *general* need for a mediating mental item or “interpretation” of the rule the famous paradox of *PU* §201 arises. Either one holds that if all goes well the interpretation will represent the ‘right’ way to relate the rule to the segment at hand – a way that can also be exhibited in relating it to other segments. Or one supposes that at some point we reach a further mental item that doesn’t mediate, but rather determines what accords with the rule. In the former case one embarks on a regress that makes it impossible to explain how we ever reach judgment. As the interpretation is taken to represent something like a method of projection it gives rise to the very same question as the original representation of the rule: ‘How am I to apply it?’³⁸ In the latter case the regress is stopped by stipulation. Since it is a stipulation one might, just as well, stop it at the first stage and hold that each application of a rule is a decision that determines how it is to be applied. Since, consequently, everything can be made to accord with the rule the distinction between ‘correct’ and ‘incorrect’ loses its sense and the Normativity Constraint can’t be met.

The paradox threatens the intelligibility of the very idea of a finite intellect – in contemporary terms: the very idea of our judgments having determinate content. If we were to accept the paradox it would follow that at least one of our constraints has to be given up.³⁹ It seems fairly clear by now that Wittgenstein does not accept the paradox.⁴⁰ There is less agreement about how he claims to solve and with what success. It has been suggested that it gets solved in the very section in which it is stated. And indeed in the second half of §201 he presents it as a *reductio* of the account of understanding that gives rise to it: “What this shews is that there is a way of grasping a rule which is *not* an *interpretation*, but which is exhibited in what we call ‘following the rule’ and ‘going against it’ in actual cases.” The diagnosis seems

³⁸ This first horn of the dilemma has been articulated in *PU*, §139-141 in relation to mental images. In *PU*, §§198-201 the point is generalized to any kind of mediating mental item.

³⁹ The “skeptical solution” that Saul Kripke famously considers is maybe the most radical in that it rejects the very idea that there are facts of meaning and understanding and tries to retain an independently intelligible notion of the ‘normative’. (See Saul Kripke, *Wittgenstein on Rules and Private Language*, Oxford: Basil Blackwell, 1982.) But each one of our six constraints (including the Generality Constraint itself) has been implicitly or explicitly rejected in some proposed solution to the paradox. In all these cases one has to ask whether the words expressing the constraints that are supposed to be retained can have their intended sense independently of the rejected constraints.

⁴⁰ See Warren Goldfarb, “Kripke on Wittgenstein on Rules”, *The Journal of Philosophy*, LXXXII, 9, 1985, 471-488; John McDowell, “Wittgenstein on Following a Rule”, in: *Mind, Value, and Reality*, Cambridge Mass.: Harvard University Press, 221-262.

to be this. The mistake is to think that the understanding of a concept is the grasping of some general entity that then has to be related in a further step to the case at hand. On the basic level there is no mental act or state of a subject that could be characterized as the *mere* representing of the *unapplied* rule, the *not yet deployed* concept. The *fundamental* way of understanding a concept is the one that is “exhibited” in acts of deploying it: in affirmation and denial, in asking a question, in forming an intention, in giving or executing orders. In other words: the meaning of a word or the content of a concept “lies in its use”.⁴¹ Once we realize this, the regress of interpretations cannot arise.

Put this way the central point doesn’t seem to be an insight that is specific to the later Wittgenstein. That there is no mental act or state of a subject that could be characterized as the mere representing of the content of the unattached predicate is, of course, Frege’s teaching. Words only have meaning in the context of a sentence; predicative “sense” can only appear as the component of a grasped “thought” that was divided into its elements. This “Context Principle”, as it has come to be called, is not Frege’s invention either, of course. Kant uses a version of it to characterize the finite intellect: “the understanding can make no other use of these concepts than that of judging by means of them.”⁴² *That* is the very mark of our finitude: we don’t lay hold of the universals or forms *as such*, but only by judging that objects bear them.

If this piece of traditional philosophical wisdom were all that is required to get done with the alleged puzzle about rule-following one should wonder why a disciple of Frege makes such a big deal of it. However, on the face of it the ‘proper’ picture of the finite intellect appears no less paradoxical than the ‘misguided’ one. According to our constraints, the judging subject must somehow implicitly conceive of her present act as something she could do on another occasion as well – and of all of those acts in which the predicative element might figure as springing from a common source: her understanding of its general

⁴¹ See *PU* §197.

⁴² See Kant, *Kritik der reinen Vernunft*, B 93. See Gottlob Frege, *Grundlagen der Arithmetik*, Stuttgart: Reclam 1987, X. This is not to deny that there are important differences here. But the relevant differences cannot come into view unless one first acknowledges that Wittgenstein’s celebrated slogan “meaning is use” is, *to begin with*, nothing but fancy way of putting something Frege already knew, and the early Wittgenstein never got tired of emphasizing. (See James Conant, “Meaning and Use”, in: *Philosophical Investigations*, 21 (3), 1998, 222-250.) Charles Travis argues that the later Wittgenstein “radicalizes” Frege’s Context Principle in such a way that this amounts to a rejection of the Generality Constraint. (See Charles Travis, “On Constraints of Generality”, *Proceedings of the Aristotelian Society*, vol. 94, 1994, 165-188) According to Travis, where Frege’s Context Principle teaches that words only have meaning in the context of a sentence, Wittgenstein’s radicalization of that principle teaches that sentences only have meaning in the context of a situation. But the Generality Constraint follows from the Context Principle, namely the idea that the basic units of semantic content are structured and that to understand this structure is to conceive of the components as something that could also occur in connection with other elements. To tie an utterance to the concrete situation in which it occurs is to deny the subject an understanding of the situation as structured. If there is to be radicalization of the Context Principle it can’t be a *rejection* of the Generality Constraint, but only a different articulation of the relevant generality.

use. But how is her thinking supposed to implicitly operate on that general level if there is ultimately nothing for her to represent up there? If judging is to be the self-conscious activity of deploying some element that can also occur in infinitely other judgments then the subject must somehow see her present act in that general light. *Something* in her mind must reach beyond her current act. Yet we can't seem to find *anything* that is apt to play this role.

Her understanding of the concept can't be a second order attitude that takes her judgments in which the concept appears as its objects. For it is only in the light of the general pattern that the relevant acts are intelligible as the kind of acts they are. (This was our Interpretation Constraint.) She would thus already have to advert to the alleged 'second order' in conceiving of anything supposedly belonging to the 'first order'. Her understanding of the general pattern must somehow be contained in the act of deploying it. But if so, what exactly is it that she 'implicitly' represents? – Surely neither the whole series of its correct applications, nor the unapplied rule. In the basic case, all a speaker can do in order to explain to someone the general use of a predicate is to offer an example. She might say: "One uses '*F*' like this: one says, for instance, '*a* is *F*'." We have seen that it is not different when we turn to the mental register. As Wittgenstein puts it, it is not that anything else would be available to the speaker if she were to try to explain the use of the predicate to herself.⁴³ On the fundamental level, all I can do to articulate my understanding of a concept to give myself an example of its proper deployment. Questions such as "What does '*F*' mean?" or "What is the content of *F*?" are, in this sense, transparent to the question whether, for instance, *a* is *F*. Of course, the phrase "for instance" contains the whole riddle. For how is the judging subject supposed to conceive of her act *as an example* if she can't represent whatever it is an example of?

There must be a mistake in this way of searching for the subject's 'representation' of the concept. Somehow it must be legitimate to say: "I follow the rule *blindly*", to quote the famous slogan. One way in which one might try to make it 'legitimate' is to reject the Transparency Constraint and deny that the 'because' in 'because of the rule' – that is, the sense of 'Non-Accidentality' – requires a further specification. This amounts to the retreat to some version of the claim that *on the fundamental level of description* the *explanatory* aspect of the nexus that links a person's judgment to *her* other acts in the relevant series of "going on in the same way" is a "causal connection" of a kind that can be found throughout sub-rational nature.⁴⁴ If this move is combined with the insistence that the normative dimension is the

⁴³ See *PU*, §§ 209-210.

⁴⁴ For the present purposes it doesn't matter which vocabulary exactly is chosen for the retreat. All that matters here is the thesis that the *explanatory* aspect of concept possession – that is, the way in which that 'something

‘irreducible’ sense of logical compulsion, then the “prescription” threatens to be turned into a mere projection. That was Frege’s predicament in the case of the fundamental logical laws; it returns in the form of the question how the predicative “thought”-component is related to the series of acts in which a particular subject grasps and either affirms or denies the “thoughts” in which it occurs.

4.2. *The intersubjective version of the puzzle*

Working towards a solution Wittgenstein writes: “It is not possible that there is only one occasion in which only one person followed a rule.”⁴⁵ The remark seems to pertain to the two ways in which I said that a concept that figures in the a judgment framed by a particular subject is *unlimited*: it is unlimited with respect to the *number of other judgments in which this thinker might deployed* it, and it is unlimited with respect to the *number of other subjects who might deploy it*. I have been talking about these two dimensions of the generality of a concept in terms of an infinite series of *potential* other judgments and *potential* other thinkers. Wittgenstein’s remark seems to introduce something in the order of *actuality*. But what are we supposed to think of here? Is suggestion to introduce to a couple of more actual judgments and a couple of more actual thinkers?

It seems fairly obvious that pointing to more acts by the subject won’t help. The appeal to an actual multiplicity of subjects has appeared more promising. In the contemporary discussions the notion of the ‘normative’ often figures in a way that can look very different from how it occurs in Frege. The focus is not on “laws” from which “there follow prescriptions”, but rather on the critical stance we take toward each other in reciprocally interpreting and evaluating our assertions and actions. Donald Davidson, for instance, famously argues that the phenomenon of conceptual content can only be rendered intelligible by appeal to the intersubjective scenario of two people interpreting each other. Similarly, Robert Brandom proposes to shift of the attention away from the relation between a thinker and the “law” or “norm” to the relation of agreement and disagreement between two thinkers ascribing normative statuses to each other. Brandom motivates this approach as follows:

Norms (in the sense of normative statuses) are not objects in causal order. [...] Normative statuses are domesticated by being understood in terms of normative attitudes, which are in the causal order. What is causally efficacious is our practically taking or treating ourselves and each other as having commitments.⁴⁶

about the subject’ provides a “common explanation” of her multiple acts of deploying the predicative element – must be specifiable in terms that are logically independent from the notion of “being logically determined”: whether it is in the language of fundamental physics or biology or in terms of the merely psychological.

⁴⁵ *PU*, §199.

⁴⁶ Robert Brandom, *Making It Explicit*, Cambridge Mass.: Harvard University Press, 1994, 626.

There are obvious and deep differences between the philosophies of Brandom and Davidson. To properly discuss their views would require the introduction of the concepts of language and communication that I set aside here. I just want to point out that despite the differences between them there is a respect in which their thinking about the normative runs parallel, and that what they share in this respect is something that they have in common with Frege. Where Frege claims that for “thoughts” to have “actuality” a “performance of thinker” is required, Davidson holds that it is only through propositional attitudes that “reasons are causes”. What Brandom says is “in the causal order” is the same as what Frege allows into the realm of “actuality”: the “normative attitudes” and not the “norm”, the “takings-to-be-true” and not the “law”. And just like Frege, Davidson and Brandom resist any identification of *what is true* with what the majority of thinkers tend to *take to be true*. By contrast to Frege, both Brandom and Davidson insist that the idea of an objective standard is to be articulated through the perspectival difference between the view of the speaker and the view of the interpreter, between making and ascribing an assertion. Brandom and Davidson differ between each other with respect to the question what vocabulary is to be used in describing our activity of treating each in this way. But as far as our topic is concerned the result seems to be the same in all three cases: the very move that establishes the *objectivity* of the norm turns it into something that can never appear in an explanatory role. The only thing in the order of sense that *can* act as a cause is the subject’s *actual* propositional attitudes. Thus our question arises again: how are we to conceive of that ‘something about the subject’ that figures as a common *explanation* of all her acts of deploying a certain concept.

Due to the shift of focus to the intersubjective perspective of discourse this question appears in another guise. How can two individuals stand in such a relation with each other? When eating vanilla ice-cream gives me pleasure while you find that taste repulsive, our representations just stand beside one another. Now, I might resent you for spoiling my fun by pulling such a disgusted face: mirroring myself in your delight would have increased my enjoyment, which is why I would have approved of your liking it too. But these relations of attunement or disharmony were not the kind of relations of agreement or disagreement we were talking about. If your act is to be something that negates or confirms my act, the same concept must be able to figure in both of our acts. And if we are to enjoy the perspectival difference of a dispute on matters of truth we must know this. For it to be matter of truth between us we must, at a bare minimum, implicitly conceive of the relation in which we stand as something that is *not particular to the two of us*: each of us could stand in this relation with

infinitely many other thinkers. We must implicitly represent the relation that holds between us as pointing in this way beyond the two of us. There must thus be something known to us that can underwrite our reaching ahead and explain the reality of the connection between us. In other words: we must see both of our acts and the relation between them in the general light of the “law” or “rule”. For this reason introducing an actual multiplicity of subject’s is akin to introducing further acts. As Wittgenstein puts it: “It is no use [...] to go back to the concept of agreement, because it is no more certain that one proceeding is in agreement with another, than that it happened in accordance with a rule.” (BGM, VII, § 26)

5. Capacities and practices

In setting up the paradox I have suppressed words like ‘ability’ and ‘practices’ or ‘forms of life’. In consequence, some of the slogans that usually occur in presenting the solution figured as part of the puzzle. By proceeding in this way I hoped to bring out how much of a burden these words must carry. How are they supposed to make all the difference? What role they would have to play follows from the considerations up to this point. The relevant ‘ability’ must be the ‘something about the judging subject’ that underwrites the reaching ahead to an in principle infinite series of her judgments and can figure at the same time as a standard and also as a common explanation for all of them. Her being a participant in a ‘practice’ must be the ‘something about her’ that can figure as the common explanation of the infinitely many possible relations in which she could stand to other thinkers framing judgments that are opposed to or in agreement with hers. And both of these aspects in which a concept is ‘unlimited’ must somehow be known by the judging subject: the relevant abilities and practices must be ‘self-conscious’.

This sounds very much like a program for further philosophical research – not very concrete, but there it is.⁴⁷ In the literature following the later Wittgenstein it is often suggested that such a further investigation of the relevant notions of *ability* and *practice* is not necessary and the demand for it misguided, a symptom of the very “disease” of “wanting to explain” we are supposed to be cured of. On this view, the thrust of Wittgenstein’s rule-following considerations is to bring out that the paradox arises out of a misguided urge for philosophical explanation that we have to overcome. I don’t want to challenge this meta-philosophical view. But I think that the way I set up the puzzle raises doubts about some of the concrete diagnoses that have been proposed.

⁴⁷ For an investigation of the related concepts ‘disposition’ and ‘practice’ in their role moral philosophy see Michael Thompson, “Practical Generality”, in: *Life and Action*, 149-210.

In the context of discussing Saul Kripke's forceful presentation of the material as a "skeptical challenge" it has often been argued that Kripke's way of setting up the paradox rests on the demand for a reductive explanation of meaning.⁴⁸ The proposed diagnosis is that it is only because what can count as a legitimate 'fact' has been artificially restricted to what is describable in merely physical terms that the metaphysics of meaning and understanding can come to seem mysterious. If one takes that this to be the thrust of the rule-following considerations it looks like the lesson to be learned is that meaning and intentionality are 'non-fictional', 'irreducible' and 'intrinsically normative'. Now, as Warren Goldfarb has pointed out, the absence of Frege in this discussion should make one wonder how much this way of framing the issue has to do with what goes on in the *Philosophische Untersuchungen*.⁴⁹ Frege is certainly not driven by *this* kind of reductive project. I have tried to show, to the contrary, that the puzzle arises *within* Frege and thus quite independently of any reductive ambitions. If this is correct, then before we can meaningfully claim that content is 'intrinsically normative', we first need to be able to say what has gone wrong in the Fregean view.

According to another strand in the literature what gets us into trouble is the longing for an 'external grounding' or 'general account' of our ordinary life with meaning. Cora Diamond, for instance, argues that the rule-following paradox arises because we want a *general* answer to the question how an individual judgment can be determinately connected with any one of the infinitely many potential ways to continue in the logical space of all semantic possibilities. But philosophy can give no answer to this question. What we have to accept is that we have always already judged in these matters and that sensible doubt about how to continue can only arise in *concrete situations* that are located *within* our practices. As Goldfarb puts it, the source of the puzzle is the misguided search for a "global grounding of logical compulsion"; the lesson to be learned is, accordingly, that the "distinction between rationality and inclinations [can only be drawn] in particular circumstances".⁵⁰

We have seen that the idea that the chain of questions 'Why?' must come to an end in the acknowledgement that we have always already judged can be found in Frege's philosophy, even though it comes at much a later point in the chain. The crucial difficulty is

⁴⁸ See, for instance, Paul Bohossian, "The Rule-Following Considerations", in: Miller, Alexander and Crispin Wright (eds.), *Rule-Following and Meaning*, Montreal and Kingston, Ithaca: McGill-Queen's University Press, 2002, pp. 141-187; Gary Ebbs, *Rule-following and Realism*, Cambridge, Mass.: Harvard University Press, 1997, 279ff; Barry Stroud, "Mind, Meaning and Practice", in: *Meaning, Understanding, and Practice*, Oxford: Oxford University Press, 2000, pp. 170-192.

⁴⁹ Goldfarb, "Kripke on Wittgenstein on Rules", 475.

⁵⁰ Goldfarb, "Rule-Following Revisited", Ms.. See also Cora Diamond, *Realism and the Realistic Spirit*, Cambridge, Mass.: MIT Press, 1995, 69-70; David Finkelstein, "Wittgenstein on Rules and Platonism", in: Cray, Alice and Rupert Read (eds.), *The New Wittgenstein*, London: Routledge, 2000, pp. 53-73.

to avoid the gulf between the normative and the explanatory that opens up in Frege's philosophy. It is hard to see how insisting on the way in which questions about the rationality of an act are always embedded in particular circumstances can make a difference to this point. For no matter how far down from the philosopher's abstractions we descend into the complex whirl of our ordinary practices the following will hold: if the appeal to the context in which an act is embedded is not to be a rejection of the Generality Constraint, then the particular act cannot be tied to the *concrete* situation in which it occurs; it must reach ahead to an infinite series of acts and situations *of this kind*. There must thus be 'something' that can connect this act with these other acts and situations and figure as a common explanation for all of them. And that 'something' must be known to the subject. Our puzzle was how this could be. If it is to help with our problem the pointing to the wider context of our ordinary practices cannot be a pointing to more stuff available in the particular situation surrounding the act, no matter how rich our description of that stuff might be. If this is right, then we better investigate exactly what it could be that we are pointing to instead.

John McDowell argues that the appeal to "practices" and "customs" doesn't call for *any* further inquiry, since it doesn't carry any explanatory weight: the appeal to these words is nothing but a reminder of "a bit of common sense about following a sign-post [or a rule]", namely that doing so is "not acting on an *interpretation*".⁵¹ On his view, the puzzle about extending a series can arise only if a certain way of thinking is already in place: one must conceive of a person's mind where the understanding of the "principle of the series" is intuitively situated as "populated exclusively with items that, considered in themselves, do not sort things outside the mind, including specifically bits of behaviour, into those that are correct or incorrect in the lights of those items". The first step towards sanity is to see how peculiar it begins to sound when we apply this thesis to "intentionality in general". For that would mean to deny that an "an intention, just as such, is something with which only acting in a specific way would accord" and that "generally, a thought, just as such, is something with which only certain states of affairs would accord". Once we realize how implausible this is as a *general* thesis about intentionality the more restricted claim about the understanding of "the principle of a series" equally looks "quite counter-intuitive." As it thus turns out to rest on a "thesis that we have no reason to accept", the alleged puzzle is "revealed as illusory".⁵²

Once again it seems that Frege is not guilty of the assumption that is said to be the source of the trouble. As I presented it, the assumption that McDowell identifies as an

⁵¹ John McDowell, "Meaning and Intentionality in Wittgenstein's Later Philosophy", in: *Mind, Value and Reality*, 263-278, here 276.

⁵² See McDowell, "Meaning and Intentionality in Wittgenstein's Later Philosophy", 264, 270, 271-272.

externally motivated assumption that gives rise to the apparent puzzle entered the dialectic quite late in the course of trying to answer the reasonable question how to make sense of the Transparency Constraint. The thesis that the relation between the “general pattern” as it is grasped by the subject and her application of it must be mediated by an interpretation was the *consequence* of *one* attempt to specify what the subject ‘grasps’ or ‘represents’ when she understands a concept. The proposal was that she represents the ‘unapplied rule’. This thesis in turn, was motivated by the realization that such understanding surely can’t consist in the subject having implicitly already acknowledged all the true thoughts in which the concept figures. If there is anything to this way of describing the dialectical situation, our puzzle should arise in the material McDowell lays on the table.

The treatment McDowell proposes turns on the analogy between (i) the way an order is related to its executions, an intention to what fulfills it, etc., and (ii) the way in which a concept or the “principle of a series” is related to its applications. However, our specified Generality Constraint suggests that this analogy has to be taken with care. One may ask what distinguishes the judgment of a *rational* animal from the perceptual acts of a *mere* animal. After all, in *some* sense of the words, the perceptual acts of a mere animal can be said to stand in a “relation of accord” to what they “represent”. The traditional answer is that what makes an act the grasping or expressing of specifically *conceptual* content is its *internal structure*. For an act to exhibit this structure is for it to be connected to infinitely many other potential acts that could all be explained by a general pattern that must be *known* to the subject of the act. The relation between this general pattern or “rule” and its “applications” can therefore not be yet another instance of the “relation of accord” characteristic of “intentionality in general”. For it underwrites the specifically *rational* kind of accordance McDowell is talking about. That is to say, our constraints articulate the conditions that have to be met if the acts of a subject are to have the ‘right kind’ of content, namely content that brings them under the peculiar “logical ‘must’” to which the *rational* activities such as framing a judgment, forming an intention, issuing and executing orders are subject.

If the analogy on which McDowell seems to rest his diagnosis were unproblematic we could take the following appearance of the “surface grammar” of our language at face value. The word ‘understanding’ as it occurs in ‘He understands the concept *F*’ seems to function in the same way as when it occurs in ‘He understands the thesis that *a* is *F*’. In the latter case the verb is used to ascribe what on the standard approach is called an ‘attitude’ whose object is a proposition. So if the role is the same we should expect in the former case as well: an attitude and a correlated object. But that assumption was precisely what lead us into trouble: we can’t

seem to find the right ‘object’ of the subject’s ‘implicit’ representation. Our difficulty is thus not that there is some *definite* mental item that seems “normatively inert” if one adopts the wrong perspective. Rather, our puzzle is that on closer inspection it seems unclear *which* ‘item’ we are supposed to be talking about.

Wittgenstein suggests that the riddle about how one can know something without being able to say it (or articulate it in one’s head) can be dissolved when we recognize that the sentence “I know how the word ‘game’ is used” the words ‘know how’ figures in a different role than in the sentence “I know how high the Mont-Blanc is”. When employed in the ascription of concept possession the verbs ‘understanding’ and ‘knowing’ are *not* deployed in order to ascribe an act or state that could be called the adopting or having of a propositional attitude; their role is rather, we are told, akin to the role of ‘being able to’ or ‘having mastered (a technique)’.⁵³

This is, of course, also what Evans suggests in stating *his* Generality Constraint. The understanding of a sentence like ‘*Fa*’ or the grasping of the thought expressed by it results from the joint exercise of “two abilities”: the understanding of the object-term ‘*a*’ and her understanding of the predicate-term ‘*F*’. But then there is that word ‘ability’. Evans renders it as a “single state” that figures as the “common partial explanation” of all the acts of mind in which the relevant element figures.⁵⁴ If it is not to undermine the intended distinction between abilities and their exercises in judgments and beliefs, the word ‘state’ can obviously not be read in the sense in which the knowledge or belief that *Fa* – or, for that matter, that ‘*Fa*’ is true if and only if *Fa* – might be called a ‘mental state’. In a paper that focuses on the linguistic register Evans proposes that the relevant ability is to be conceived as “tacit knowledge” – *not* in the sense of the ‘implicit’ that can become ‘explicit’ upon questioning, but rather in the sense of the “unconscious deployment of information” that is “capable of figuring in an *explanation* of a speaker’s capacity to understand new sentences”. According to the “model” Evans envisions, to ascribe the understanding of a language or the possession of a system of concepts to a subject is to ascribe to her a “set of [interconnected] dispositions”. And he thinks that the correctness of the model is to be decided “by providing a causal, presumably neurophysiologically based explanation of comprehension”.⁵⁵ Now, if one accepts this model and insists at the same time that the norm under which judgments fall cannot be explained in terms of such causal structures, there is, once again, the threat that the

⁵³ See *PU*, §79 and §§146-151.

⁵⁴ See Evans, *The Varieties of Reference*, 101-102.

⁵⁵ See Gareth Evans, “Semantic Theory and Tacit Knowledge”, in: *Collected Papers*, Oxford: Oxford University Press, 2000, pp. 322-342, here 324, 328 and 331.

explanatory nexus that links the multiple acts of the subject to a common source comes apart from what figures as normative standard for these acts.

At the same juncture Michael Dummett rejects any merely “causal” rendering on the ground that it can’t be just any kind of ability: since judging is a self-conscious activity, the correlated capacity must be one that I can only be said to have if I *know that I have it*, and that I can only be said to exercise if I *conceive of my act as an exercise of it*.⁵⁶ But then the vexed question threatens to return. What do I implicitly *represent* in exercising an ability whose actualization involves the subject’s apprehension of the capacity? The solution Dummett envisions is the program to show that the relevant abilities can be “fully manifested” in a “linguistic practice” that is describable without the use of semantic and intentional terms.⁵⁷ The details of the proposal are too complicated to discuss here. I just want to point out that Dummett’s “manifestation condition” that a “full-blooded” theory of meaning has to meet involves the appeal to the view that an ability must be represented in *testable counterfactuals*.⁵⁸ The condition is, of course, meant to apply to the kind of theory that an outside observer of the practice could construct. But we may ask where this view about the semantics of ability-ascriptions leaves us with respect to Dummett’s original claim that possessing a concept is a *special* kind of ability in that the bearer of the ability knows that she has it. The bearer’s knowledge of the ability can surely not consist in grasping the truth of certain counterfactuals. For which counterfactuals would our pupil know in possessing the concept ζ is next-but-one-after ζ ? That he would always write the ‘correct’ number if he were to go at it or that he would write 1002 if he were to get to 1000, and 1004 after that etc.?⁵⁹ The intensional context must change everything. But if so, it seems we made no progress in specifying what it is to possess and exercise the relevant self-conscious capacity

⁵⁶ Michael Dummett, “What Do I Know When I Know a Language?”, in: *The Seas of Language*, Oxford: Oxford University Press, 1993, pp. 94-105.

⁵⁷ Michael Dummett, “Mood, Force and Convention”, “Mood, Force and Convention”, in: *The Seas of Language*, 202-223, here 221.

⁵⁸ Michael Dummett, “What is a Theory of Meaning (II)”, in: *The Seas of Language*, 34-91, 54.

⁵⁹ The counterfactual rendering leads us right back into the paradox: either it won’t be informative about the particular step at hand or there would be an infinite set of counterfactuals representing each step. An analogous difficulty arises for the recent attempt to present the relevant kind of *knowing how* as a species of *knowing that*. Jason Stanley and Timothy Williamson have suggested that ‘I knows how to ϕ ’ can be translated into a statement of the form ‘I know that this is a way for me to ϕ ’. (See Jason Stanley and Timothy Williamson, “Knowing How”, *The Journal of Philosophy*, 98.8, 2001, 411–444) The problem is not whether such knowledge is ‘explicit’ or ‘implicit’; the difficulty is rather that the phrase ‘way to ϕ ’ is ambiguous between the general and the particular that is at the center of our problem. Applied to our scenario the phrase can be read in two ways: either it picks out a particular segment in the number series or it stands for the general way of continuing. In the first case it remains unexplained how representing the particular step can connect the subject to the infinitely many other steps. In the second case the question arises how that general knowledge is to be applied to the segment at hand. This is, of course, no argument against Stanley and Williamson’s analysis; it just shows that if their analysis is correct, it is useless to appeal to the notion of *knowing how* in the context of the rule-following paradox.

Without further elucidation words like ‘ability’ and ‘know how’ seem to get us nowhere we haven’t already been. At this point one might form the suspicion that the source of the puzzle might have something to do with whatever stricture makes it seem that there are only three options available in thinking about the relevant ability: the appeal to testable counterfactuals, the appeal to an underlying ‘categorical basis’ (probably a neurological one), and the appeal to some ‘irreducible’ representational state whose content we can’t seem to capture. I will let this stand as a conjecture. My aim here was only to suggest that we don’t have to imagine a skeptic or, for that matter, misguided urges for reductive explanations and external groundings in order to get puzzled about rule-following. There is a real philosophical problem here – as ‘real’ as these things can get about which Wittgenstein says that they “have the form: ‘I don’t know my way around’” and that they only “disappear” once we have achieved “complete clarity”.⁶⁰

⁶⁰ This paper goes back to material from my doctoral dissertation and inherits many debts especially to my advisors Christoph Menke and Michael Thompson. For very helpful comments and discussion I want thank Matthew Boyle, James Conant, Anton Ford, Wolfram Gobsch, David Hunter, Andrea Kern, Douglas Lavin, Eric Marcus, John McDowell, Richard Moran, Thomas Ricketts and Sebastian Rödl.