DID WITTGENSTEIN FOLLOW THE RULES? (OR WAS HE GUIDED BY THEM?)

Abstract

In *The Blue Book* Wittgenstein recognized two possible accounts of what it means to follow a rule: (1) acting in accordance with it; (2) the rule’s playing a part in my activity of following it. Wittgenstein initially rejected (1) because it seems to provide only a causal account of rule-learning and rule-following. He opted for (2), which posed a problem because he had given up his earlier belief in the direct accessibility of rules in experience. They could now be accessed only via their external codifications. Hence Wittgenstein’s problem was not an epistemological problem of how to follow a rule, but the question as to how a rule as an external object can guide my actions. This is essentially the same as to how the blueprint of a machine determines its movements. Wittgenstein never solved his problem, and reverted back to the account (1). This led him to assign entire language games a conceptual priority over their rules. This priority assumption rests on non-trivial assumptions and involves Wittgenstein in serious problems.

One of the most intriguing aspects of Wittgenstein’s philosophy is his discussion of rules and of their role in guiding our actions. The best known series of his remarks on this topic is the so-called rule-following discussion in the *Philosophical Investigations* I, secs. 143-242. Related remarks are found elsewhere, for instance in the *Remarks on the Foundations of Mathematics* Part I, especially secs. 1-5. (Also ibid. Part VII, secs. 26-27, 66-67.) Wittgenstein’s ideas of rules have been almost universally misunderstood. Philosophers refer routinely to his “rule-following discussion”, but his very phrase epitomizes a misunderstanding. In reality Wittgenstein’s discussion is not about how we consciously follow a rule. His problem is not whether I can know what rule I am following, whether I can know whether I am following a rule correctly, or any similar skeptical or other epistemological problem. One way of seeing this is to note that following a rule typically is for Wittgenstein a primary language-game. And as I will show in another paper (Hintikka, forthcoming(a)), according to Wittgenstein epistemic notions such as knowledge, doubt and certainty cannot be used in primary language-
games. And if so, it is according to Wittgenstein a mistake to raise any epistemological questions concerning rule following.

Instead of trying to diagnose sundry misinterpretations of Wittgenstein’s remarks on rules, it is more constructive to outline the correct way of reading him. If Wittgenstein’s problem was not epistemological, what was it?

One major difficulty in interpreting Wittgenstein usually is that he seldom spells out what the problems are that he is trying to solve. (See here Hintikka 1998.) In the case of rule-following we are fortunate in that Wittgenstein himself formulates the problem he is facing in so many words. In The Blue Book pp.12-17 he asks: What does it mean to follow a rule? How can we teach someone to follow a rule? Wittgenstein envisages two possible answers. First, following a rule might mean simply acting in accordance with it. Then all there is to the teaching of a rule is the process of training the learner to act in a certain way. In other words, the acquisition of a rule must be thought of as a mere drill. For instance, in the case of language learning such a drill causes the learner to associate a word on paper with a certain sound and with a certain meaning. In general, the learner is somehow brought to associate an expression of the rule with a certain action.

In The Blue Book Wittgenstein rejects this conception of rule following. He thinks that it only amounts to a causal story of learning rules which does not involve any real teaching in the sense of conveying an insight to the learner. Instead, Wittgenstein says that in genuine rule-following the rule must play an actual role in the process of following it. But how is this requirement to be understood?
What he says is that in genuine rule-following a rule

is itself involved in the processes of understanding, obeying, etc;

“involved”, however, meaning that the expression of this rule forms part
of these processes.  (Op.cit., pp. 12-13)

Wittgenstein further explains his view by saying that on it

[t]he rule is involved in the understanding, obeying, etc.

in the sense that

the symbol of the rule forms a part of the calculation.

But what does this amount to? Wittgenstein gives a number of examples. One of them is following an algebraically expressed rule of calculating in arithmetic. In such a case we can think that what happens is that the calculator literally replaces algebraic symbols by numerals and then carries out the intended operations. Here the symbolic expression of the rule literally forms a part of the calculation.

Another example is especially interesting. It is outlined by Wittgenstein in The Blue Book p. 14 and in Philosophical Investigations I, secs. 50-53. This example concerns the operation of color-words. Wittgenstein envisages a language community whose members use color words by means of a color chart. In order to identify the color of a certain object a member of this community compares it with different samples of the chart. After having found the one that closest matches the color, he or she reads the name associated with that sample on the chart. That name can then lead to further action, for
instance to the obeying of a verbal order in which the color word in question figures importantly.

We can try to bring this scenario closer to real life by assuming that the role of a physical color chart can be taken over by remembered color images. The purpose of Wittgenstein’s thought-experiment here is to illustrate how the color chart can play a role in the concrete use of color vocabulary. The chart is involved in the kind of use Wittgenstein envisages. It is literally part (an instrument) of the color language.

Needless to say, this is not how Wittgenstein thought that our actual color concepts operate. The color-chart people were for Wittgenstein a thought-experiment calculated to throw our use of the words into a sharper perspective. His own views will be discussed below.

Thus according to the view espoused by Wittgenstein in *The Blue Book* the rule must somehow be present to me whenever I follow it, because it must actually guide what I am doing, just as much as a color chart must be present when a member of Wittgenstein’s imaginary tribe uses color words.

In the sense illustrated by such examples the rule must literally guide my actions when I am following it. But now Wittgenstein faces a new problem. How can the mere symbolic expression of a rule guide my thinking and my actions? There would not be any such problem if the rule were a phenomenological entity present in my experience. But how can an inanimate physical object, for instance a symbolic formula on paper, guide my action and my thinking? As Wittgenstein puts it, “a rule, so far as it interests us, does not act at a distance.” (*The Blue Book*, p. 14.) Hence Wittgenstein’s problem
about rules is not how I follow a rule but how the rule guides me when I follow it. The problem is not about rule following but about being guided by a rule. In other words, what his problem is, is not whether I know the rule that I am following or whether that rule is somehow underdetermined by evidence. Wittgenstein is not a “rule skeptic.” He even admits situations in which an explicit symbolic expression of a rule is given to me. Wittgenstein’s robust respect for the language games we actually play should be enough to drive home this point. Surely in the language-games we actually play when carrying out numerical or symbolic computations the symbolic expression of a rule is typically a rock-bottom conclusive answer to the question as to which rule I am to follow. Wittgenstein’s question is therefore not whether the symbolic expression specifies the right rule, but how it determines my actions when I follow it. We might call this problem the guidance problem instead of the rule-following problem.

Wittgenstein’s own main examples should be enough to settle this point. What does he use as his paradigmatic example of rule-following? Reading (in the sense of reading aloud). (See e.g. *Philosophical Investigations* I, secs. 156-171; *The Brown Book* pp. 119-125) Now what is the rule that guides me in reading aloud? Clearly the alphabet. Now when I recite a poem, should I as a philosophical follower of Wittgenstein entertain even as a thought-experiment skeptical doubts whether I know my alphabet? To a philosopher who thinks that I should, one can only respond: read what Wittgenstein says. What Wittgenstein is not doing is to throw skeptical aspersions on my knowledge of the alphabet, which is proverbially thought of as the most trivial part of our conceptual repertoire, not even as a thought-experiment, but to ask: How does that knowledge guide me in my reading? Obviously I am not conscious of the alphabet when I read. Am I
subconsciously reciting the alphabet when I am reading? If not, how is my being able to utter the right sounds based on my knowledge of the ABC?

In order to understand why the guidance problem was so serious for Wittgenstein, a number of observations are in order. First, why should Wittgenstein be thinking in terms of being guided by a rule as an external object in the physical world, for instance by the symbolic formulation of a rule vs. paper? Such questions can only be answered by reference to the historical and genetic context of his thinking. The massive change that had taken place in Wittgenstein’s thinking in October 1929 is switch from the primacy of phenomenological languages to the primacy of everyday languages. (See here Hintikka, forthcoming (b).) After this change, what our words of our language primarily refer to according to him are denizens of the everyday world of external objects not given to me directly in my immediate consciousness. Hence in talking about rules we presumably have to refer primarily to their codifications in external reality. But then they are not immediately present to my consciousness and the question as to how they guide my actions becomes relevant.

This does not mean that Wittgenstein had become a behaviorist who does not acknowledge the reality of the internal world. He acknowledges readily that rule-following can take place in thought. What he is maintaining is that in principle the rule-following in external reality is the primary case conceptually. In other words, what is crucial is merely that we must always be able to think of rule-following as happening in external reality. Mental rule-following then becomes an extension of the primary external use, like doing arithmetic in one’s head. As Wittgenstein writes in The Blue Book p. 13:
As we are not interested in where the processes thinking, calculating, take place, we can for our purpose imagine the calculation being done entirely on paper. We are not concerned with the difference: internal external.

In other words,

we could perfectly well, for out purposes, replace every process of imagining by looking at an object or by painting, drawing or modeling: and every process of speaking to oneself by speaking aloud or by writing.

(*The Blue Book*, p. 4.)

Wittgenstein returns to the same problem area in the *Philosophical Investigations* I, secs. 172-178 where he discusses the experience of being guided. The upshot of the discussion is that even if we sometimes have such experiences they are not what it means to be guided by a rule. An experience does “not contain the essence of guiding” (sec. 178). What follows from this and what prompts the guidance problem is thus that we do no any longer according to Wittgenstein have direct access to rules in our thinking. The remarkable thing is that in his earlier thought Wittgenstein had in fact believed in the possibility of such immediate acquaintance with rules. In MS 116 he confesses that

[Earlier I thought at one time that grammatical rules are an explanation of what I experience on one occasion when I use the word. They are as it were consequences or expressions of the properties which I momentarily experience when I understand the word.]
Did Wittgenstein have in mind the impossibility of such direct acquaintance with rules when he formulated his problems in *The Blue Book*? Or was he led to a realization of this impossibility by reflecting on the question he posed in *The Blue Book*, pp. 12-17? There are some indications that the latter may have been the case. This is suggested by a passage similar to our quoted one but at the same time interestingly different. This passage occurs in MS 213 (“The Big Typescript”, written probably in 1933), p. 160. There the view he later in MS 116 says that he had rejected is presented as a prima facie tempting but fallacious consequence of his own “exposition of grammatical rules”. Hence Wittgenstein is obviously thinking of views he had held shortly before putting together MS 213. Now Wittgenstein had given such an exposition shortly before the composition of MS 213 precisely in *The Blue Book* passage (pp. 12-17) that was discussed above.

In looking for further confirmation of this insight concerning the nature of Wittgenstein’s problem about rules, one is confronted by a veritable embarrassment of riches. For instance, no unprejudiced reader of the so-called rule-following discussion in the *Philosophical Investigations* I, secs. 143-242, can fail to be struck by Wittgenstein’s harping on the question of whether and if so how rules “guide” us in our activities, especially in comparison with the few passages where epistemological concepts make their appearance. Indeed, the most conspicuous examples where Wittgenstein’s evokes the notion of knowledge in the “rule following” discussion concern situations where a rule-follower says, “Now I know how to go on.” (Sec e.g. secs. 151-155, 179-181.) It is not hard to see why he felt called upon to discuss such examples. This is not because he wants to raise the skeptical question whether the speaker really knows the rule, but
because they seem to show that one can have direct access in one’s thought to the rule. As was pointed out earlier, such an access was the very thing Wittgenstein now found impossible, and his remarks on “Now I know” were calculated to maintain this impossibility, not to argue for skepticism.

Perhaps even more striking is the discussion that has been published in the volume Remarks on the Foundations of Mathematics, Part I. This discussion was not intended by Wittgenstein to concern only mathematics. It was a part of an early version of Philosophical Investigations, indeed the part that corresponds to the “rule-following discussion” of the published versions. In that discussion, one finds plenty of remarks on how mathematical rules guide our action but little or no concern with our mathematical knowledge. Wittgenstein might have thought that mathematical truths are not absolute, but his discussion of mathematical rules is not calculated to make or to buttress that claim.

One particular but important case of rules concerns the rules of logic. In this case, Wittgenstein takes his problem to concern the nature of logical necessity or, as he once expresses it, “the hardness of the logical must”. Not unexpectedly, he is deeply puzzled by, and in some sense skeptical of, the alleged necessity of logical rules. At one point he says that he can understand what geometrical necessity amounts to, but that he cannot understand what logical necessity is.

In a superficial sense, Wittgenstein is here a prey of the same mistake as a lot of other philosophers. Contrary to widespread views, what are known as the rules (sic) of inference in twentieth-century logic are not rules of inference in the sense that they would
tell what a reasoner should do or what human reasoners actually do. They are neither prescriptive nor descriptive, but permissive. They tell us what inferences one may draw without committing a fallacy. Even though Wittgenstein’s elusive idea of logical necessity is thus a chimera, his construing it in the way he did nevertheless illustrates his concern with the problem as to how rules determine a rule-follower’s behavior.

The sense in which valid logical inferences are necessary is model-theoretical, not phenomenological. What is necessary is for a conclusion to be true — that is, to be realized in reality — if its premises are. But from this it does not follow that whoever contemplates the premises must also think of the conclusion. This “does not follow” nevertheless does not include thinkers like Aristotle who thought that to think of something is to realize it formally in one’s mind. For other thinkers a thought was a picture realized in the mind. In both cases, what necessarily accompanies such a form or picture must also be present in one’s mind. Accordingly I strongly suspect that Wittgenstein’s misapprehension was inspired by his picture theory of language and thinking.

We are here dealing with a problem whose significance extends way beyond Wittgenstein’s philosophy. He was for instance perfectly right about most of the rules of logic that one finds in textbooks and philosophers have in mind when they think about logic. Not only do they not express any necessity of thinking of the conclusion if one thinks of the premises. They do not show in any reasonable sense why they express necessities even in a model-theoretical sense. This poses a very real problem to analysts of logic. Contemporary philosophers would be well advised to take this problem of Wittgenstein’s much more seriously than most of them seem to have done.
There are exceptions, however. In the case of logical rules — at least in the case of logical rules of first-order logic. We can in fact find in contemporary logical theory an example of what could be an answer to Wittgenstein’s guidance problem in a special but important case. Suppose that the rules of first-order logic are formulated as tableau rules in the sense of Evert Beth (1955) or in an equivalent tree-building form. Then an attempt to show that a first-order statement G is a logical consequent of F amounts to an attempt to envisage a scenario in which F is true but G false. If this attempt inevitably leads to a contradiction, it can be concluded that G indeed follows from F. What is crucial here, the necessity of this consequence relation is made obvious directly by the formal proof.

The example of logical rules suggests that there might after all be an epistemic element in Wittgenstein’s problem about rules. But it is not a question of whether we know the rule we are following, but whether we are aware of it. In the case of the tableau rules, it is redundant to ask how they guide an attempted counterexample description. That they do so, and how they do so, is obvious as soon as we realize what we are doing. Hence Wittgenstein’s problem about rules is not the problem of knowing how to follow a rule or knowing which rule we are following. It is being aware of what we are doing when we follow a rule.

Wittgenstein’s guidance problem is in fact a special case of his wider and perhaps deeper problem of the relation of formal rules to their concrete meaning. Wittgenstein had on one hand a keen sense of purely formal language-games, such as we have in logic and in mathematics. (Cf. here Hintikka 2004.) On the other hand, he was attuned to the role the notions so formalized in the applications of logic and mathematics. At one point he envisages an application gone wrong.
Cut down all these trees! — But don’t you understand what “all” means? (He had left one standing.) *(Remarks on the Foundations of Mathematics I, sec. 10.)*

What Wittgenstein did not understand is how the formal activities are related to their applications. As we might put it, he thought that he did not really understand what he was doing when he was following a rule like the rules of logic and mathematics or rules of language, for that matter.

One corollary to my interpretation of Wittgenstein’s problem about rules is that the problem has nothing to do with intentionality. Wittgenstein could have raised the same problem about machines instead of humans – and, in fact, did so. How is a machine guided by its blueprint? This question is parallel with the question: How does a rule guide my actions when I follow it? Accordingly, Wittgenstein does not ask, as a later thinker might do, “Does a calculating machine think?”, but “Does a calculating machine really *calculate*?” (op. cit. V, sec. 2). Wittgenstein discusses how the motions of a machine are “guided” by its blueprint in *Philosophical Investigations* I, secs. 193-194 among other places. He asserts the parity of the machine’s “rule-following” with a human’s in his inimitable way in *Remarks on the Foundations of Mathematics* IV, sec. 20:

If calculating looks to us like the action of a machine, it is the *human being* doing the calculation that is the machine.

Wittgenstein’s use of a machine analogy is by no means accidental. Even much earlier he already had made occasional use of machine analogies to obviate epistemological
questions related to the rule following problems. In two sets of notes on logic dictated in October 1913 and April 1914 Wittgenstein bluntly stated that logical rules in the right logical notation are being followed ‘automatically’. (I am grateful to Dr. DePellegrin for calling my attention to these passages.) They are not followed by a metaphysical self or by the self of empirical psychology. However, while he later returned to the machine metaphor in various formulations and examined many of them in the mid-thirties, obviously in the hope that the problem of being guided by a rule might be easier to solve in the case of machines than in the case of humans, Wittgenstein was not able to proceed beyond his previously shaped conception of rule. He eventually abandoned the machine metaphor and, with it, his searches for an account of being guided by a rule. (The role of machine analogies in Wittgenstein is treated in depth by Enzo De Pellegrin in his forthcoming monograph on this subject. See also De Pellegrin 2000.) Moreover, Wittgenstein resigned to providing merely an account of how one is trained to act in accordance with a rule and, thereby, acknowledged the conceptual primacy of entire language-games over their rules.

The tableau method is not the only instance where logical theory has produced a solution to Wittgenstein’s problem at least in special cases. Applied to computers, his question amounts to asking: how does the software of a computer determine its actual operations? But this is no longer a philosophical question for us. It is not a genuine intellectual puzzle for us. Or, in so far as it is, it is answered by pointing out that the circuitry of a computer constitutes in the simplest case a model of Boolean algebra. Wittgenstein’s problem is no longer alive. We do not ask: Do calculating machines
really calculate? The development of computer technology has pre-empted Wittgenstein’s question.

It is important to realize that this is not a mere historical point. The analogy between switching circuits and Boolean algebras constitutes a genuine explanation of how computer compute, an explanation that ought to have satisfied Wittgenstein.

From the vantage point of this interpretation of Wittgenstein’s problem about rules we can understand the development of his philosophy, not only the development of his ideas about rules but the changes in his overall view of how language works. Wittgenstein tried very hard to solve the guidance problem. It would take us too far afield to attempt an exhaustive chronicle of his struggles. A few characteristic features of his thoughts on rules must nevertheless be mentioned in any case.

For one thing, Wittgenstein did not find his way to the examples of self-explanatory rule-following described above. This was of course virtually impossible for historical reasons. The tableau method and the tree method were introduced four years after Wittgenstein’s death. And even though the structure of switching circuits is the same as the Tractarian theory of truth functions, Wittgenstein does not seem to have paid much attention to the theory of switching circuits.

In general, Wittgenstein did not master the question of how the blueprint of a machine determines its motions. He seems to have been particularly impressed by puzzle examples where a plausible-looking machine description turns out to be impossible to realize in actual hardware. In our day and age, he perhaps might have emphasized the difficulty of program verification.
For all his heroic efforts, Wittgenstein’s struggles with the guidance problem ended in a defeat. He initially tried to find an account of how a rule guides my actions, but ended up saying that such an account is impossible. All that we can say of rule-following is that it is acting in conformity with the rule. This means that he in effect returned to the view he rejected in the *The Blue Book*, viz. the view that the teaching of a rule is a drill except that he apparently no longer thinks of this process in merely causal terms. What Wittgenstein says is the following, speaking in the first sentence in the voice of an imaginary interlocutor:

But that [training] is only to give a causal connexion; to tell how it has come about that we now go by the sign-post, not what this going-by-the- sign-post consists in. On the contrary; I have further indicated that a person goes by the sign-post only in so far as there is a regular use of sign-posts, a custom. (*Philosophical Investigations* I, sec. 198)

Sign-posts here play the role of rules.

Wittgenstein’s falling back on the account he rejected in *The Blue Book* does not mean that he is defining “rule” by reference to regularity. (See *Philosophical Investigations* I, sec. 208.) For to understand the notion means according to Wittgenstein to understand those practices that constitute rule-following. This is an instance of the general view of his that I have expressed by saying that for him language-games are conceptually primary with respect to their rules. Also, a learning process becomes part of the language in the same sense as the external manifestation of a rule according to his
earlier view. When he subsequently asks whether I can know the rule I am following, he is not inquiring into my ability to grasp an entity called the rule. He is highlighting his newly acquired view that in an important sense in my actions and my experiences there is no rule consciously guiding them.

On a semantical level, this development meant a switch of priorities. Wittgenstein first evoked the idea of language-game to highlight the nature of language as a rule-governed process. Rules were the basic semantical reality. “You can’t get behind the rules, because there isn’t any behind” (Philosophical Grammar p. 244). However, after Wittgenstein’s change of priorities language-games become the ground floor of semantics. One is not taught a language-game by teaching its rules. One is trained (“drilled”) to play the language-game. In principle one can understand the rules of a language-game only by being trained to master the entire game. It is in this sense that we according to Wittgenstein follow a rule “blindly”.

By the same token language does not any longer operate according to Wittgenstein by means of criteria. Wittgenstein continues to use the word occasionally in his late philosophy, but it has to be taken in a loose sense. The only criterion for the use of a word is strictly speaking the entire language-game which is its logical home. Wittgenstein’s later views have sometimes been characterized as a “logic of criteria”. This is diametrically opposed to the truth, as far as Wittgenstein’s views after 1937 or so are concerned. For instance, when Wittgenstein introduces the idea of “family resemblance”, he is not offering it as a deep new insight into the way our language works, but merely as a pedagogical warning against the view that language operates by means of criteria.
This transition from the primacy of rules to the primacy of language-games is the most important development in Wittgenstein’s views in the middle to late thirties. I have discussed it earlier, but I have not emphasized the role of the rule problem as its most important driving force. It is too large a topic to be analyzed and documented here. Suffice it here to examine some of its aspects.

Wittgenstein’s comments on linguistic rules serve again to illustrate his ideas. For after having given up his attempts to explain the phenomenon of being guided by rules Wittgenstein sees language as operating in the last analysis by means of drilled (or innate) reactions to what our words stand for. He distinguishes for instance between two uses of the word “rabbit”. (See *Philosophical Investigations* II, p. 197.) In one of them, I am asked “What animal is that?” and I answer, with all due deliberation, “A rabbit.”. In the other, an animal jumps out from the bushes and I cry, “A rabbit!”’. It is the second use that is thought of by Wittgenstein as being fundamental. It is the same kind of use that is found in primary language-games with internal sensations. Such a use is a part of a language-game that relies primarily on reactions to sensations.

For instance, Wittgenstein says in so many words that color words operate in the same way as words for internal sensations (*Remarks on Color* III, secs. 71, 72). As the so-called “private language argument” (*Philosophical Investigations*, secs. 243 - 304) shows, the semantics of such words relies on our spontaneous reactions to them. Likewise, color words receive their meaning from the fact the language user is trained to react to colors in a certain way, in this instance by means of a verbal response. Thus Wittgenstein’s contrast between the color chart example and the idea of drilled responses
to colors illustrates the development of Wittgenstein’s thinking that I have been discussing here.

The color language example opens at the same time a critical perspective on Wittgenstein’s later semantics. One of them concerns the relation of Wittgenstein’s semantical models to the actual facts of the psychology of language and language learning, in this case facts of the neuropsychology of color language. Is our actual understanding and use of color words more like consulting remembered color samples or responding in a certain way to color stimuli? I believe that in this way we can in fact evaluate Wittgenstein’s ideas. For instance, can he distinguish color blindness from color agnosia? In color agnosia a patient has a normal color vision, but cannot use color concepts, name colors or point to a particular color when asked to do so. In brief, a color agnostic behaves like a member of a Wittgensteinian color chart using tribe who has lost his chart. Can this phenomenon be reconciled with passages like Remarks on Color, sec. 112? Wittgenstein is there talking about the difference between a normal and a color-blind person:

> The one can learn a language-game that the other cannot. And indeed this must be what constitutes color-blindness of all kinds.

The physiology and psychology of color perception is a complicated matter. It is nevertheless obvious that the Wittgensteinean trained response idea is too simple to do
justice to our actual use of color terms. The telling question is whether it is possible to lose color concepts even though one’s color sensation remains normal. There is definite evidence to suggest that this is possible. Yet on Wittgenstein’s view it cannot be possible, for a subject presumably can be trained to respond differently to different sensations. This conceptual analysis of color-blindness simply is not adequate.

There are other problems with Wittgenstein’s view of the primacy of language-games over their rules. This primacy of language-games over their rules was also instrumental in leading Wittgenstein to his view that epistemic concepts cannot be part of these primary language-games that form the interface of language and reality. Rule-following constitutes a good example. In order to doubt meaningfully whether I am acting in accordance with a rule, I must be able to ascertain in some cases that I am following it correctly. But if acting in accordance with a rule means simply that I have been trained to do so, it does not make any sense to ask whether I have really been trained to follow it.

More generally speaking, if rule following is a primary language game, then such notions as doubt, certainty, and even mistake do not naturally apply in it. There is in fact something conceptually awkward about saying that a trained dog makes a mistake in doing its tricks. It either does it or does not do it. But if so, Wittgenstein’s ideas run into trouble in this direction, too. The most interesting examples of rule following involve a complex series of operations. And then it makes sense not only to speak of mistakes but to locate a mistake within that complex. In a calculation according to an algebraic formula one can make a mistake in this or that step and in carrying out this or that operation. A member of the color chart tribe may make a mistake in comparing the color
of an object with different samples or in reading the name of the color represented by a sample. Such instances of rule-following are not “blind” in the sense of being merely trained responses to a situation.

There admittedly are some rules that apparently are applied blindly in Wittgenstein’s sense, for instance the multiplication table. But even there it makes a difference whether I know what I am doing. Suppose I suddenly have a momentary mental lapse and wonder what $7 \times 6$ is. But I need not panic. I can say to myself: $5 \times 6$ is surely 30, and $2 \times 6$ is 12. Hence $7 \times 6 = 30 + 12 = 42$. What is relevant here is of course not how I recover a momentarily lost bit of the multiplication table, but the fact that I could not even try to do so if I were not aware what I am doing. Here we have an example of the kind of awareness that might be called informed rule-following. Wittgenstein obviously thinks it is not characteristic of rock-bottom rule-following. This seems to me mistaken. Such an informed awareness of what I am doing normally remains tacit, but it can none the less be present. For one thing, informed rule-following plays a crucial role in the teaching of mathematics, contrary to what Wittgenstein suggests. An illuminating example of the psychological difference between following rules blindly and following them with an awareness of what one is doing, is for instance found in Kibel (1992). We are here approaching the problem whether a purely mechanized proof should be considered a real proof. (See here e.g. MacKenzie 2001.)

What those observations show is among other things that Wittgenstein’s arithmetical examples are not very representative illustrations of his ideas. It seems to me that he was to some extent aware of this difficulty. As was mentioned earlier, Wittgenstein’s discussion of rule following published in the volume Remarks on the
Foundations of Mathematics was originally a counterpart of the rule following discussion of the published version of Philosophical Investigations. Why did he thus replace it by a different discussion? The answer is that the Remarks discussion was largely in terms of mathematical examples. It looks as if Wittgenstein became aware of the awkwardness of such examples and replaced many of them by nonmathematical ones.

Wittgenstein’s switch from learning rules and criteria to being trained to play a language game in accordance with them leads him to a new problem. How can one teach a language learner to connect a certain expression with a certain kind of object or event? There does not seem to be any problem here as long as the object or event in question can be witnessed both by the teacher and the learner. But how can such teaching take place if the object in question is not accessible to the teacher? A case in point might be the teaching of our vocabulary for internal sensations to a child. The question as to how such teaching can succeed is for all practical (and theoretical) purposes the problem Wittgenstein addresses in his so-called private language argument. The reason Wittgenstein indulges in this famous argument is thus the fact that our discourse about internal sensations seems to constitute a counter-example to his view of the primacy of language games over their rules. Like so many arguments of Wittgenstein’s, the “private language argument” therefore does not initially amount to a new insight or a separate line of thought, but is part and parcel of Wittgenstein’s argumentation for his own central views. It is an attempted refutation of a prima facie counter-example. This explains why the “private language argument” enters the center stage of Wittgenstein’s philosophy relatively late. It does not occur in the earliest versions of the Philosophical Investigations.
Does this refutation succeed? It is significant that Wittgenstein formulates it in terms of language learning. What Wittgenstein says is the following:

How do words *refer* to sensations? — There doesn’t seem to be any problem here; Don’t we talk about sensations every day, and give them names? But how is the connexion between the name and the thing named set up? — of the word “pain” for example. Here is one possibility; words are connected with the primitive, the natural, expressions of the sensation and used in their place. A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain behavior. (*Philsoophical Investigations* I, sec. 244.)

This diagnosis of how our vocabulary of internal sensations may seem plausible. However, it reveals an important presupposition of Wittgenstein’s account. He believes in the reality of internal sensations. As he once puts it,

There are internal and external concepts, internal and external ways of looking at people. There are even internal and external facts — as there are e.g. physical and mathematical facts. But they do not stand next to each other like different kinds of plants. For what I said sounds as if one had said: All these different facts occur in nature. And what’s wrong about
that? … The internal is not only connected with the external by experience, but also logically. (MS 173, pp. 72-73)

But in order for the replacement of spontaneous pain behavior by verbal expressions of pain to yield a way of referring to the real internal happening, it must be assumed that a similar kind of pain behavior goes together with a similar kind of pain experience. Wittgenstein’s account of the meaning of our discourse about pains is an account of a referential semantics of this discourse only if there in fact is a correlation between different pains and their expressions. Such a correlation is a presupposition of our vocabulary and syntax of pains. And the same holds for all our discourse about internal sensations.

Wittgenstein’s line of thought thus leads him to recognize important factual presuppositions of the semantics of our language. Such factual presuppositions are an important theme of Wittgenstein’s philosophy in the *Philosophical Investigations* and elsewhere in his late writings, for instance in *On Certainty*. On the one hand, he emphasizes that for our interpersonal use of language it is only the external manifestations (e.g. the pain behavior) that matter. This is the point of his discussion of the sensation S in *Philosophical Investigations* I, secs. 258-261, 270. But such a smooth intercourse serves a representational function only on the assumption of behavioral uniformity. Wittgenstein acknowledges this conundrum in a striking statement which according to his own principles should not even be expressible in language:
The assumption would thus be possible — though unverifiable —
that one section of mankind had one sensation of red and another

I see reasons to think that Wittgenstein’s idea about the conceptual primacy of
language games over their rules is not the right diagnosis of the problem situation.
However, this does not reduce the importance of his discussion about rules or that of the
changes in his views for appreciating the central ideas of his philosophy.
Wittgenstein’s works are referred to in the usual way.


Hintikka, Jaakko, forthcoming (a) “Wittgenstein on Knowledge and Skepticism.”.

Hintikka, Jaakko forthcoming (b) “The Crash of the Philosophy of the Tractatus”.


Professor Jaakko Hintikka

Boston University
Department of Philosophy
745 Commonwealth Avenue
Boston, MA 02215 USA
hintikka@bu.edu