

The Picture of a Paradox:  
Rule-Following after Wittgenstein and Beyond

A Thesis Submitted to the College of  
Graduate Studies and Research  
in Partial Fulfillment of the Requirements  
for the Degree of Master of Arts  
in the Department of Philosophy  
University of Saskatchewan  
Saskatoon

By  
Mark Harold Niemi

© Copyright Mark Harold Niemi, June 2008. All rights reserved.

## PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Philosophy  
University of Saskatchewan  
Saskatoon, Saskatchewan S7N 5A5  
Canada

OR

Dean  
College of Graduate Studies and Research  
University of Saskatchewan  
107 Administration Place  
Saskatoon, Saskatchewan S7N 5A2  
Canada

## Abstract

My thesis aims to show that Wittgenstein's view of rule-following involves a misleading picture of the rule. Since he saw the rule as something fundamentally independent of the rule-follower and something with which the rule-follower must comply, he inevitably became entangled in the paradox of compliance: that is, the idea that there must be something other than the rule-follower for rule-following to exist, even though he knew that there was really nothing there to guide the rule-follower or to measure his action. This paradox, dimly expressed within key Wittgensteinian problems relating to how one is able to follow a rule and whether one can follow a rule 'privately', eventually gave rise to the question over the social nature of rules. In that debate, Wittgenstein's commentators vigorously argued, and continue to argue, whether the concept of rule-following presupposes a community of practitioners or not. I argue that this debate itself is misguided, since both sides in this debate take as their starting point a picture of compliance which sees rule and rule-follower as essentially different. In contrast to a 'compliant' picture, I offer a different picture of the rule, which I will call the 'pliant picture of the rule'. I will show that rule and rule-follower are fundamentally the same, and are related to one another, not socially or grammatically, but genealogically. This relationship of identity is in fact exhibited in the relationship between teacher and pupil, when the pupil becomes what his teacher already is through following his teacher. Although compliance can be said to define this relationship initially, it ends with the pupil learning, or becoming, the rule. To conceive of the rule in this way is to avoid the paradox of compliance; that is, it is to go beyond Wittgenstein's picture of a paradox.

## Table of Contents

Permission to Use.....	ii
Abstract.....	iii
Table of Contents.....	iv
Introduction .....	1
Chapter One: ‘Mastery of a Technique’	
1.1 ‘Mastery of a Technique’ and ‘Following the Rule’: A Question of Independence...13	
1.2 The Rule-Follower <i>is</i> the Rule.....	19
1.3 Following a Rule is not Analogous to Obeying an Order.....	26
Chapter Two: The Relation between Teacher and Pupil	
2.1 The Relation between Teacher and Pupil is a Natural Extension of that between Parent and Child .....	33
2.2 Imitation of the Teacher and Correction of the Pupil.....	37
2.3 Agreement between Teacher and Pupil; the Role of the Community.....	40
Chapter Three: Following a Rule ‘Privately’	
3.1 The Difference between Rule-Following with and without Signs .....	53
3.2 The Use and Usefulness of the Sign: PI 258, PI 260, and PI 270 .....	56
3.3 The Arbitrariness and Autonomy of the Sign .....	60
3.4 The Generality of the Sign .....	64

3.5 The Normativity of the Rule and Community Agreement .....	65
3.6 Kripke, Malcolm, and Medina on Normativity and Community Agreement .....	68
3.7 The Misunderstanding of Compliance in Rule-following.....	74
3.8 The Misunderstanding of Internal Relations in Rule-Following.....	75
 Chapter Four: The Picture of a Paradox	
4.1 The Pliant and Compliant Pictures of the Rule .....	82
4.2 The Essence of Rule-Following is Imitation of Action .....	89
4.3 The Essence of the Sign is Metaphor or <i>Identity-in-Difference</i> .....	91
4.4 The Pliant Rule is a Living Genealogical Identity.....	96
 Bibliography.....	 104

## Introduction

In *Remarks on the Foundations of Mathematics* Wittgenstein asks the following question, after noting that a rule allows one to hold by it just as much as it can be said to compel one to act in a certain way:

My question really was: ‘How can one hold by a rule?’ And the picture that might occur to someone here is that of a short bit of handrail, by means of which I am to let myself be guided further than the rail reaches. [But there is nothing there; but there isn’t *nothing* there!] For when I ask ‘How *can* one...’, that means that something here looks paradoxical to me; and so a picture is confusing me. (RFM-V.45)

What Wittgenstein is saying here is that the missing part of handrail in this picture of the rule is not really missing after all; the extension of this short bit of handrail is paradoxically both there and not there. If that missing part of handrail really were missing, one might reasonably ask, *how could that short bit of handrail still function as a rule for us?* In fact, if this is one’s picture of the rule, how could one *not* be puzzled about how one is able to follow a rule? This question itself seems drawn right into such a paradoxical picture of the rule. And the picture, confusing as it is, seems as inescapable as the question it gives rise to. Yet this captivating picture of the rule continued to hold Wittgenstein’s thinking captive, despite his vigilance against being held captive by misleading pictures.

There is, however, more to this picture than meets the eye. Although Wittgenstein could see its paradoxical nature and knew that his question was the result of a confusing picture, he nevertheless could not solve the paradox, because he did not see the *deeper* picture underlying this picture of the rule as a handrail. His answer to his question, therefore, was not to reject this picture outright and to replace it with another, but to work around it by seeing rule-following as something akin to obeying an order. He saw rule-following as the outcome of a *similar* kind of training that both eliminated the possibility of

choosing different courses of action and emphasized the formation of habit, conditioned responses, practice, custom. In his own paradoxical way of stating his position, he finally said: ‘When I follow a rule, I do not choose. I follow the rule *blindly*’ (PI 219). *That* is why the missing part of handrail does not really need to be there: one does not need to see the rule (or hold onto it) in order to choose which way to proceed, because one acts blindly anyway; and because one simply acts the way one has been trained to act, as a matter of course, the missing part is not really missing after all: one’s *training* supplies the missing part of the handrail. Looking at rule-following in this way, Wittgenstein was able to provide a paradoxical answer to the question arising from his paradoxical picture of the rule, that is, ‘How am I able to follow the rule?’—‘I follow the rule *blindly*’. ‘Following the rule *blindly*’ is a paradoxical notion, unlike simply acting blindly—for example, ‘following one’s instincts’—since following is here a *deliberate* action involving something one *chooses* to follow. This is why rule-following is a normative activity, unlike actions done from instinct, since rule-following implies doing something on purpose, correctly or incorrectly. As we shall later see, Wittgenstein and his followers did not and could not abandon this normative sense of ‘following’; they simply switched to a non-normative sense when it suited them. In the case of the rule as handrail, this switch happened when it became obvious that there really was nothing independent of the rule-follower to guide him when following the rule—hence the reliance on custom, inclination, the continuation of past performance, conditioned responses, ‘complete confidence’, or something else amounting to following a rule ‘blindly’. Thus, instead of rejecting this paradoxical picture of the rule and the question that went with it, Wittgenstein simply reaffirmed the picture by giving an equally paradoxical answer to it<sup>1</sup>. His answer was

therefore not a genuine solution to his question (in the sense of ‘dissolving’ it) but the philosophical equivalent of a programmer’s ‘hack’—that is, it was essentially a clever work-around solution to a deeper problem he had yet to grapple with—the problem of *compliance*.

What Wittgenstein had actually stumbled upon with his notion of ‘following the rule *blindly*’ was the paradoxical nature of symbolic representation, that is, the nature of the *sign itself*. Guided by the picture of the rule as handrail, he was merely describing this same picture, although this picture was no longer the picture of the rule at all, but only that of the sign. What he was really describing with this notion is one aspect or moment of what I will call ‘abstraction’. Abstraction is essentially the reduction of a whole to a part through some medium, or its change from independence to dependence—without actually altering what is whole or independent. It is really the nature of metaphor. Such abstraction occurs in perception, for example, when we see or hear something: what is seen or heard has been reduced abstractly *in perception* to the manner in which we perceive it through our eyes and ears. Similarly, abstraction occurs within the institution of slavery, when a free-born and independent human being is reduced abstractly *in society* to an instrument for another and becomes another’s property. In the case of the sign, just as in the case of a tool, a similar abstraction occurs when a mark or sound is reduced abstractly *in language* to something useful, that is, when we do *something else* with it besides merely writing or speaking this sign. ‘Following the rule *blindly*’ was just Wittgenstein’s expression for the mere production of this mark or sound that occurs in such abstraction. He was really just noting that symbolic marks or sounds have no ‘intrinsic meaning’, that is, they have no independent nature *qua sign* determining their production, and so nothing

about this mark or sound can guide me in speaking or writing them. Hence abstraction has a connection to the so-called ‘arbitrariness and autonomy of grammar’. Since abstraction always occurs within some medium of reduction, the medium’s own intrinsic nature is simply ignored yet exploited at the same time—only thereby does it become something else, that is, it now takes on another meaning, another identity—an identity other than what it has by nature.

Abstraction is thus like a second nature, a form of identity imposed on a thing by its usefulness for some end or other. This identity is not ‘intrinsic’ to the thing but ‘extrinsic’; the thing has its ‘abstract’ meaning in virtue of a purpose it now serves for another. A tool is a kind of ‘abstract object’ in this sense, and Aristotle’s famous example of an axe is one such example of abstraction: what an axe is by nature is really the nature of the material out of which it is made—iron or bronze—but its functional use, or the activity of cutting, gives this iron or bronze the identity it properly has as an axe, that is, an axe *is* an *axe* in virtue of the purpose for which this iron or bronze is used. But this functional use does not obliterate the intrinsic identity of the axe—its nature *qua* iron or bronze—but instead exploits this identity, ignoring everything else about the nature of iron or bronze other than what is useful for this end. Iron’s solidity and hardness are useful in an axe, but its colour and taste are not, although it has these also by nature. Abstraction, consequently, ignores the true nature of the axe, the material out of which it is made, except as it relates to the usefulness of this material, and this overlooking of the axe’s true nature is a form of ‘blindness’, signaling the axe’s independence *qua* *axe* from its material substrate. In other words, the axe *qua* *axe* is simply *not* iron or bronze; but *qua* *mere substrate*, it *is* iron or bronze. The axe is thus both iron (or bronze) and not iron (or bronze).

Wittgenstein's answer to his paradoxical picture of the rule, therefore, involved turning a blind eye to the normative nature of the rule while simultaneously acknowledging it, since the picture of the rule as handrail *represents something independent of the rule-follower and his action*. In other words, the handrail in this picture of the rule *just is* the rule's normativity figuratively expressed and then conspicuously overlooked through 'blindness'. The act of following the rule *blindly* is thus a form of 'abstract' action: it is an action simply done, that is, it is something now done for its own sake—an action free of any consequences for the one who acts; in the case of the sign, it is the sign looked at simply as the production of a mark or sound, or, as I will phrase it, it is the *sign qua mere use*. The *sign qua mere use* is thus analogous to the *axe qua material substrate*. In this simplicity of action or mere use resides its *usefulness* under abstraction, just as the slave is made useful to society by acting without regard to his own desires or needs under 'blind obedience' to another. Unless the mere production of a mark or sound is made useful under abstraction, it has no 'meaning', no identity at all, other than as a mark or sound. And because this picture of the rule as handrail is also a picture of compliance, it should be no surprise that compliance itself is a deep-rooted problem for Wittgenstein: 'following the rule *blindly*' is really a public rejection of this handrail without the deeper rejection of compliance. It is the picture of a paradox—the picture of a *rule that is there yet not there*.

Wittgenstein's paradoxical answer may have satisfied some as an answer to the question arising from his paradoxical picture, but the paradox in question goes deeper than what that one question asks. Because his paradoxical picture involves a deeper picture that itself gives rise to the paradox, this deeper picture simply resurfaced in the guise of another central question that Wittgenstein posed—"Is it possible to follow a rule 'privately'?"<sup>2</sup>

Wittgenstein said no, “otherwise thinking one was following a rule would be the same thing as following it.” But this second question and answer gave no hint of a paradox, although it later gave rise to a debate among his commentators that centered on the role of the community in rule-following. This second question now revolved around the picture of the rule as a *measure* and no longer as a *guide*. In fact, it is only because of this later debate that the paradox now had a chance to show itself as the main obstacle to a satisfactory resolution to the question of the role of the community in rule-following, a debate that still continues unabated to this day, over twenty-five years since it resurfaced in full force with the publication of Saul Kripke’s *Wittgenstein: On Rules and Private Language*. One camp in this dispute, the so-called ‘community view’ endorsed by Kripke himself, sees rule-following as essentially requiring a community of rule-followers in order to establish the *normative* nature of rules (the rule as measure), whereas the other camp, the so-called ‘individualist view’, primarily defended by G. P. Baker and P. M. S. Hacker, does not see such a requirement. But both sides *still* agree with what Wittgenstein said in PI 202, namely, that one cannot follow a rule ‘privately’; both sides agree that there must be a *difference* between thinking one is following a rule and following a rule; these two cannot be the same without destroying the concept of normativity that is essential to the concept of rule-following. For them, this also means that the rule-follower must be *different* from the rule, since it is the rule that determines, independently of the rule-follower, whether the rule-follower has followed the rule; it cannot be the rule-follower himself that determines this. But this line of argument essentially presupposes the same deeper picture of the rule that Wittgenstein invoked in RFM-V.45 through the picture of the rule as a guide for the rule-follower; in fact, the rule has also been compared by some to a ‘yardstick’: the

‘yardstick’ is thus, along with the ‘handrail’, a paradigm of the rule<sup>3</sup>. This deeper picture could justifiably be said to be the *compliant picture of the rule*, since it is the picture of something independent of the rule-follower with which the rule-follower must comply or be in accord in order to be said to follow the rule<sup>4</sup>. This picture of compliance, however, leads ultimately to the paradox of compliance: the idea that there must be something other than the rule-follower for there to be rule-following, even though there really is nothing independent of him either to guide him or to measure his action. Both sides in this interminable debate, therefore, merely represent the two sides of the paradox: ‘there is nothing there (there need be *no* community), but there isn’t *nothing* there (there *must* be a community)!’ This deeper paradox thus partly explains why both sides have at least one leg to stand on in this debate and why neither side has won out over the other in any conclusive manner. It never occurred to either side, though, just as it never occurred to Wittgenstein, to question this underlying picture of compliance, since it seems so self-evident that rules and compliance go hand in hand<sup>5</sup>.

My thesis, then, can be succinctly summarized in this way: the compliant picture of the rule sees rule-follower and rule as *different*, such that the rule-follower must act in a way that complies with the rule or is in accord with the rule; it sees this difference as prerequisite for normativity, since normativity is said on this account to presuppose another difference, namely, the difference between what is right and what merely seems right to the rule-follower. This latter difference demands that both rule and rule-follower be different and not the same: “hence it is impossible to follow a rule ‘privately’”. But there is another picture of the rule, however, the *pliant picture of the rule*, that sees rule-follower and rule as the *same*; compliance makes no sense (or no difference) if these two are the same; but

far from abolishing normativity, this identity of the two makes normativity, that is, what is correct and incorrect, part of the rule-follower's *understanding itself*, since what truly makes the rule pliant is the rule-follower's grasp of the *purpose* of the rule. The rule-follower who only understands *how* to follow the rule—that is, 'follow the rule *blindly*'—need not grasp the purpose of the rule: hence the need for the rule to still be different from the rule-follower; one cannot follow a rule 'privately' in *this* sense, if one cannot determine *for oneself* whether one's own action is correct or incorrect; but in the pliant picture of the rule, the rule-follower must *also* grasp the purpose of the rule even more than how he is to follow it; and once he does that, the rule-follower *becomes* the rule: rule-follower and rule are now *one and the same*. My understanding, or grasping of the purpose, of the rule, in other words, *is* the rule in the truest sense. Wittgenstein, I argue, was still working within a compliant picture of the rule—the basis of his paradoxical picture of the rule as a handrail and the presupposition of the claim that one cannot follow a rule 'privately'—just as his commentators still are today. My thesis is an attempt to show a different picture—a pliant picture of the rule.

The implications of these two pictures of the rule are far-reaching and should not be underestimated; indeed, it is my contention that the compliant picture of the rule involves a paradox about rule-following that cannot be resolved in the current debate over the role of the community. For one, both pictures entail different conceptions of independence for the rule-follower. Under the compliant picture, the rule-follower remains dependent on the rule in an important sense, however that rule is understood, so long as rule and rule-follower are said to be different and the rule is conceived here as either a guide or a measure for the rule-follower: it is in fact the *rule* which needs to be truly independent in

this relation and not the rule-follower, since the rule is also supposed to be the *justification* for the rule-follower's action and what he appeals to when he justifies his action (PI 217), and "justification consists in appealing to something independent" (PI 265). The pliant picture of the rule, on the other hand, entails a fundamentally different conception of independence, since if rule and rule-follower are the same, the rule-follower cannot be said to be dependent on the rule at all (except in the initial form of the relation—that between teacher and pupil), and uses the rule as neither a guide nor a measure for his action: it is this identity of the rule-follower with the rule that constitutes the *true* independence of the rule-follower, since he is now no longer dependent on anyone or anything else in order to *be* a rule-follower. The question of independence for the rule-follower, then, is intimately tied to one's particular picture of the rule; and the truth of this independence will reveal to us whether it has any substance at all.

The truth is that the compliant picture of the rule talks about independence—or in Wittgenstein's case, 'mastery of a technique'—in a way that is at odds with true independence and genuine 'mastery'. In this respect, the compliant picture's conception of independence for the rule-follower has similarities to Hegel's discussion of the independent and dependent forms of self-consciousness in the well-known 'master and slave' section of his *Phenomenology of Spirit*. Suffice it to say here that there is also a paradox in Hegel's discussion—one that he explicitly brings out in the reversal of the master's concept of himself—a paradox arising within a relation that is by its nature *unequal* but has as its *raison d'être* the attainment of *equality*, or in Hegel's terminology, 'mutual recognition'. In Hegel's view, this relation is ultimately destined to fail, because the master cannot achieve his intended goal of 'mutual recognition' within it, for the simple reason that the

truth of the master, his certainty of himself as an independent form of self-consciousness, is not another like himself but in fact the slave; his concept of himself in this relation is thus frustrated, because he cannot find himself objectively in this other, and so the master's concept of himself in this relation finds no satisfaction, no corresponding 'object'. And as the master's concept of himself goes, so goes the compliant picture of the rule; for it, too, is destined to fail in the end, and for similar reasons, since the rule of the compliant picture is ultimately an empty rule, a mere sign. The pliant picture of the rule does not produce such paradoxes, because it does not conceive of the relation between rule and rule-follower in the same way: the relation is indeed not equal at first, when the relation is manifested in the asymmetric relation between teacher (the rule) and pupil (the rule-follower); but teacher and pupil are *not* related as master is to slave, since the pupil, though dependent on his teacher, is actually learning to become what his teacher already is, that is, a rule himself. Thus the relation between rule and rule-follower in the pliant picture of the rule is fundamentally 'genealogical' in nature and not merely 'instrumental'—nor even 'social', properly speaking—since the rule-follower, by following the rule for himself, eventually *becomes* the rule, that is, both terms of the relation become the same: this genealogical transformation from difference back into identity reveals the essential nature of rule-following.

## Notes for the Introduction

1. See note 2 of Chapter 1 for more discussion of the paradoxical nature of Wittgenstein's claim that "I follow the rule *blindly*".

2. cf. PI 202. The two most prominent and important questions Wittgenstein asks in his rule-following discussion in *Philosophical Investigations* are "How is one able to follow a rule?" (cf. PI 211, PI 212, PI 217, PI 219) and 'Is it possible to follow a rule *privately*?' (cf. PI 202, PI 258, PI 265, and PI 270). The first question could be said to arise from the picture of the rule as a *guide*, the second from the picture of the rule as a *measure*. Taken together, these peculiar questions are related directly to the deeper paradoxical picture of RFM V.45 and are really just symptoms of the same underlying condition, namely, a distorted picture of rule-following that sees the rule as compliant in nature. The so-called 'paradox of interpretation' discussion—the basis of Kripke's controversial sceptical take on Wittgenstein's view of rule-following—is likewise an outgrowth of the picture of the rule as both a guide and a measure, since new applications of the rule are thought to be like invisible extensions of the rule as guide that now 'hang in the air' as interpretations; and the problem with these interpretations—the idea that *any* extension of the rule can be made out to accord with the rule—only finds its solution in custom or in what *we call* 'following the rule' and 'going against it' in actual cases (PI 199, PI 201). Unfortunately, the invocation of "what *we call* 'following the rule'" as a solution to this problem is as empty and ornamental as our giving definitions merely for the sake of their form (cf. PI 219). The *only genuine* problem of interpretation is encountered within the teacher-pupil relation, when rule and rule-follower are still different from one another as teacher is to pupil; but this problem is easily solved by teaching the pupil how to follow the rule, that is, by having the rule-follower *become* the rule. (The compliant picture of the rule is thus the legacy of the teacher-pupil relation, though this relation is only the initial form that rule-following takes before it matures into fully fledged rule-following.) This view of the problem of interpretation is taken up in Chapter One and Chapter Two.

3. "The rule is the yardstick against which the act is measured." (G.P. Baker and P.M.S. Hacker, *Wittgenstein: Rules, Grammar and Necessity*. Basil Blackwell: Oxford, 1985. p. 151).

4. The word 'compliant' in the expression 'compliant picture of the rule' is used here, not to modify 'picture', but 'rule' in the sense that the picture under investigation is about a rule which is *compliant* in nature, or which involves *compliance*. It is meant to contrast with the word 'pliant' in the expression 'pliant picture of the rule', where the rule is pictured as something that is notably *not* compliant in nature but something now identified with the rule-follower himself—a living being. Later on I will simplify the above expressions by referring to the 'compliant rule' and 'pliant rule': these adjectives are primarily meant to refer to the nature of the two different rules depicted throughout the thesis.

5. Compliance with a rule is not genuinely analogous to compliance with the law or compliance with an order. Judges determine the compliance of citizens with the law and

commanders determine the compliance of subordinates with orders, but no one determines the compliance of rule-followers with rules in a similar way, except when the pupil is learning to follow a rule from his teacher (a special relation in rule-following), or when certain professional societies are established within practices for the purposes of establishing universal standards for its members (such rules are *secondary* rules and are merely derivative in nature); compliance within these societies is *not* the same kind of compliance associated with the rule *per se*. It is a mistake of the community view of rule-following to think of the community as fulfilling such a role.

## Chapter One: 'Mastery of a Technique'

### 1.1 'Mastery of a Technique' and 'Following a Rule': A Question of Independence

Wittgenstein uses the expression 'mastery of a technique' (*eine Technik beherrschen*) several times in the *Philosophical Investigations*. In PI 150 and PI 199, 'mastery of a technique' is connected with the ability to do something, with understanding how to do something: to speak a language, for instance. In PI 692, it is brought in as a criterion of someone meaning such-and-such when he instructed someone else, even though he neither said nor thought explicitly what he meant him to do in every case. These instances of the expression indicate that for Wittgenstein one who has 'mastery of a technique' is simply someone who knows how to speak a language, calculate a math problem, or play a game. In fact, the two key words of the expression, 'mastery' and 'technique', are found in other expressions that mean much the same. For instance, in PI 20 and PI 33 he speaks of 'mastery of this language' and how one 'must already be master of a language in order to understand an ostensive definition'. In PI 31, someone can understand the place of a chessman in a game when shown its shape because he is 'already master of a game'. 'Technique', therefore, is akin in this expression to a language, a game. In PI 125, Wittgenstein says explicitly what he means by 'technique': '...we lay down rules, a technique, for a game...'. Someone who has 'mastery of a technique', then, is someone who knows how to follow the rules of some activity.

The expression 'following a rule', by contrast, reflects—on my account—the fact that one human being can imitate or follow another, and this is what I will claim the expression means or makes reference to in this deeper sense of 'following': rule-following is

essentially the *imitation of action* of one individual by another. At first one individual follows the other, who is the rule for him, and so he himself is in fact learning to become just such a rule, too. By understanding the expression in this way, it is possible to look at rule-following from a different vantage point than the usual one that sees the rule as a yardstick or handrail that measures or guides the rule-follower: the rule-follower *really is* the rule, and the name of ‘rule-follower’ no longer indicates what he does now in relation to another, namely, follow *him*, but rather *how he became who he is*. To fail to see *this* meaning in the expression ‘following a rule’ is in the end to misunderstand in a fundamental way what rule-following is, since this understanding of the expression shows us clearly that the relation between rule and rule-follower is a living relation that only *initially* exists as the difference between one human being and another but *eventually* exists as a genealogical identity, that is, as the identity of the rule-follower with the rule he has learnt to follow—it is no longer a lifeless abstract difference between an independent rule and a rule-follower who must now somehow conform himself to it, using it as either a guide or a measure of his action.

The relation between rule and rule-follower, therefore, is better expressed for us in the living relationship between teacher and pupil, which is the transitional relation between the rule-follower literally speaking and the rule—the transition from rule-follower *to* rule. But attaining so-called ‘mastery of a technique’ for the pupil is not so straightforward as simply requiring that he act as we all do. We are skewed in the wrong direction here by superficially reflecting on our own descriptions of rule-following, the same descriptions that we often use when we are teaching others how to follow a rule. For example, when we talk about the teacher and the pupil, we naturally assume the teacher is instructing the pupil

as we would teach him, although we know the teacher is not infallible. After all, the teacher can be said *by us* to make mistakes, too. But if the teacher is not meant to be like us in any particular example (that is, *we* have described *him* as being in error), then whom do we say the pupil is to follow in such circumstances? His teacher still, of course, since in our description of rule-following there are really only two human beings involved, the teacher and the pupil; and so in this case, if the pupil refuses to accept the instruction or correction of his teacher, we say that he has either failed to understand the rule as the teacher understands it, or failed to understand anything at all, or perhaps he *did* follow the rule after all, despite his teacher's error. In all such cases, we reserve for ourselves, as describers of rule-following, the right to judge the pupil so, and the teacher, too:

independence and mastery are said to belong to the pupil (and teacher) only if *we say* that they do, on the grounds that what we now describe as correct rule-following is what we would instruct anyone else to do, even though our attributions of independence and mastery to the pupil imply that the pupil is no longer dependent on anyone's judgment, even ours, since *we* are not really teaching *those whom we describe* to become anything at all; our descriptions here merely overlap with how *we* would instruct *others* to follow a rule; those who listen to our descriptions of rule-following are now like the pupil being taught how to follow the rule. The 'independence' or 'mastery' attributed to the pupil in our descriptions of rule-following is therefore not genuine independence or mastery but simply conditional forms of independence and mastery—both are still conditioned by *our description* of correct rule-following.

The difficulty, then, in describing the relation between teacher and pupil seems to be in fully recognizing teacher and pupil as *independent like us* and also *equal* in this

independence without subjecting such individuals in our descriptions to the condition that they be the same as us, since the pupil, too, if he were *truly* independent, could now say the same as we do or not, that is, he could say that we are the same as him even if we say we are different, and vice versa; he presumably would *also* have the power of speech, and could give descriptions of his own actions, just as we do. We must therefore bargain away *our own unquestioned* independence and mastery as the price of accepting the independence and mastery of the pupil in following the rule, even in our descriptions of him. We will still describe him as we must do, and say what he does now is correct or not, but his independence and mastery, *if he really has them*, are not something so insubstantial as to turn now merely on the turns of *our descriptions* of him acting correctly or not: the parity found among peers is an acceptance that extends beyond this or that judgment by others; this parity becomes, one might say, an ongoing test of the breaking tolerance for a rule which is no longer compliant in nature but pliant. For the one thing that we have not bargained out of our acceptance of the pupil is the common ground on which we both stand, and this ground is the *purpose* of the rule, that for the sake of which it exists at all: everything else stands by itself or falls on *that* ground. One's ability to grasp *that* is the true test of independence, both for us and for those whom we describe.

It is important, therefore, when describing the interaction between teacher and pupil, to also acknowledge the fundamental difference between ourselves as describers, and hence in some sense as judges of whomever we describe, and our descriptions of the relation between the teacher and his pupil, because we can easily take the pupil or teacher as *our* pupil when we describe him. But the teacher and pupil are really acting in these descriptions only as *proxies* for us: our descriptions of them acting correctly will go only as

far as *we* are prepared to go; and so we say that both are following a rule insofar as they do *what we do*. The danger for us as describers, then, is that through our descriptions of the teacher and the pupil we unwittingly make *ourselves* unimpeachable judges of both and in effect *the rule* against which both are to be measured, and this rather unassuming assumption, as said earlier, only gives us the impression that the rule is really something external and independent of any rule-follower, *both teacher and pupil*. Thus the pupil's independence and mastery, and even the teacher's, are still conditioned on us, on what *we call* 'following the rule'. We fail to understand that, if the teacher and pupil do in fact have 'mastery of a technique' and can also justify their actions in following the rule—and that means they can therefore disagree with our descriptions of them—*only then* do they truly have 'mastery of a technique'. True independence, therefore, cannot consist in someone's *conformity* to what we describe as correct rule-following but only in someone's *equally grasping what another is doing*, and this includes what we as describers do, since in doing so, two human beings have come together in agreement on *equal terms*. Having come to terms with another as an equal or peer is *proof* that independence and mastery have finally been attained. Some nations have gone to war with another expressly to establish such terms of peace as *evidence* that independence has actually been achieved—and most of these even have their own flag to prove it, since a national flag is a symbol of a nation's independence from others.

'Mastery of a technique', however, *seems* to be understood by Wittgenstein as simply the ability to do something correctly. The pupil, for example, "after some efforts on the teacher's part ... continues the series correctly, that is, as we do it. So now we can say he has mastered the system" (PI 145). But the pupil's ability here, as Wittgenstein puts it, is

merely to do it “as we do it”; and his mastery really extends no further than ours, since he has become our ‘pupil’ in virtue of our describing him as we do. The pupil, in other words, is following a rule insofar as he is following us, and whatever independence and mastery he has now is in virtue of his conformity to us and what we do. *We are* the rule for him. But where is the rule for us? And how is *our* independence and mastery connected to this rule? And what do the pupil’s independence and mastery amount to if they are only borrowed from us and are granted to him through us? We must again recognize that the relation between ourselves as describers and the pupil thus described is not a genuine relation between teacher and pupil but an imaginary one, since any pupil thus described will never really become independent of us (since in these described cases he is really just a *reflection* of us—he *is* ‘us’): hence any such ‘mastery of a technique’ attributed to the pupil is not a genuine mastery of anything, insofar as it is just a description of what the pupil is doing when he conforms to what we do.

One might respond by asking *how else are we to describe the teacher and the pupil if not as doing this or that correctly or incorrectly?* But it is not a matter of describing them differently (let alone ‘correctly’) but of acknowledging that these descriptions are masked instructions for rule-following, and so they mislead us in understanding rule-following, unless we are actually *using* these descriptions to teach someone else how to follow a rule that we already understand. If we think of these descriptions, however, as what we would instruct our own pupil to do, that is, as *rules in the form of rule-formulations*, we undermine the notion of independence and mastery that we then attribute to those whom we describe: our descriptions of rule-following will have slipped unawares into *prescriptions* of rule-following, however innocent our descriptions may sound. We then

re-enforce in this way the difference between rule and rule-follower by forcing the rule-follower we describe to ‘agree’ with us as describers, on pain of losing the grammatical title of ‘rule-follower’; consequently, our describing the pupil acting thus-and-so itself inadvertently *preserves the fixed and abstract difference between rule and rule-follower found in our language*, even after the rule-follower is said to have learned the rule from his teacher and no longer needs him; we describers are still the rule and the pupil whom we describe is still the rule-follower—this unspoken abstract relation remains deeply embedded in our thinking about rule-following, and so, too, is the abstract difference that constitutes this relation. In the end, we need a more substantial notion of independence and mastery than offered by Wittgenstein, for without one, certain paradoxes about the rule will continue to dog our discussion of rule-following. We also need, therefore, a more substantial notion of the rule; and in a way I have already given it: the rule is really *a human being qua rule-follower*. The rule-follower *just is* the rule.

## 1.2 The Rule-follower *is* the Rule

It may at first appear objectionable that a mere human being could be the rule, and hence a standard, for someone else, and so transcend every rule-formulation in the sense of being exempt from either being in accord with it or not being in accord with it. This is because we know that human beings are fallible and make mistakes, and so it would be impossible for a *human being* to be exempt in this way; rules hold for everyone at all times, without exception, even the expert, like a law of nature. But this thinking is misguided. Wittgenstein showed in PI 50 that no exceptional quality attaches to the standard object itself, despite its ‘transcending’

nature; it achieves its status because of the role it plays in a given activity<sup>1</sup>. Consider Shakespeare's *Macbeth*, for example: in the fictional world of *Macbeth*, certain things were said and certain things done by various characters. On whose authority do we know this? Shakespeare's, of course, because he wrote the play; Shakespeare is, as it were, an expert on the world of *Macbeth*; he has a privileged status. If a dispute arose among commentators as to what was or was not said by various characters, Shakespeare's play, the standard edition, is the measure by which we judge who is right or wrong. And if Shakespeare were alive, he could also tell us things in his fictional world which are not included in the play. Perhaps he might have written a sequel to the play. Shakespeare himself is entitled to do this: it is the prerogative of authors to be the master of their own worlds of fiction, the undisputed authority, and so the standard. It makes no sense to ask whether Shakespeare is right or wrong in what he says in those works, whether they correspond to some other reality that would vouchsafe their truth. They are 'necessarily true' in the world of *Macbeth*, because Shakespeare said them. And even if we talk of a written text as being the standard, the ground of *its* authority must still be Shakespeare himself, a human being, the man who wrote the text.

In what sense, then, could one say that the rule-follower follows the rule *blindly*?<sup>2</sup> If the rule is really different from the rule-follower, what difference could that rule possibly make if he just follows the rule blindly in the end? What *use* would the rule be for him? What 'essential' *role* would it play in his actions? It is as if the rule itself had somehow been set free of its visible, tangible formulation, only to become something elusive and dark, *yet still there*, like nebulous dark matter. Wittgenstein's expression 'following the rule *blindly*' in PI 219 seems to point out for us the perplexing gap between the rule and its formulation, that is, between the rule-formulation and its extension. But if we are in fact following the rule and not

merely its formulation, then indeed we do seem to be acting ‘blindly’, since only the rule’s formulation is visible and audible to the rule-follower—whatever else is there to follow is neither seen nor heard. Hence there is nothing really *there* to choose from when following the rule. But can one have a rule-formulation without a rule? Wittgenstein comes dangerously close to getting rid of the rule itself by making it *useless* to the rule-follower.

But what if the rule-follower really is the rule after all? The rule-follower suffers no anguish in his decisions, because he has identified himself with his craft and is the drafter of its implementation, the result of which is an expression of his expectations, purpose, and intention. One who has ‘mastery of a technique’ takes orders from no one, obeys no rules in the manner a servile pupil is thought to do. The problem of interpretation cannot even gain a foothold in his case, since there is nothing to interpret; there is only that which bends to *his rule*. It would be better to say that the rule-formulation follows him. Indeed, the pupil does, too. This is because to have mastery of a technique is really to have mastery of one’s environment. For the baker, this environment includes not only the ingredients, but the bakery conditions (e.g., temperature, humidity), equipment, and assistants, if he has any. Over this environment he rules, and this means that he can adjust his conditions to suit himself and to optimize his procedures and results; the rule-follower is precisely *that which is not conditioned in rule-following*. His training is therefore not a passive training or conditioning, since he learns to follow the rule by imitating his teacher, and *imitation is not passive but active*; it is the way the rule-follower learns to grasp the purpose of the rule. The rule-follower, in other words, is the *active principle* in rule-following; and this active principle *just is the rule*.

If we think of the rule-follower as the rule, then, we have a way of looking at the relation between rule and rule-follower that views them as essentially the same. We can see the

expert's action as an expression or *manifestation* of the rule; and if we then think of the novice rule-follower as someone learning to *become* the rule, we can see *his* action as an *interpretation* of the rule. The angle of our perspective here is either from the expert acting (manifesting the rule) or from the non-expert acting (interpreting the rule). For the novice rule-follower, acting on a rule is like giving interpretations of the rule, because we first learn to follow the rule in this way, through imitating the action of the expert and responding to his cues. This in fact is the rationale for training; the pupil is *acting as a mediator of his teacher* through his own imitation of his teacher, since the teacher can be said in an important sense both to initiate the action of the pupil by getting the pupil to imitate him in the first place and to complete it by judging the result to be correct or not: the teacher, as rule for the pupil, at first *guides* the pupil to begin his action and later *measures* the pupil to complete this action; and so the pupil here stands as an *interpreter* of the rule for the teacher until he learns to follow the rule by himself: the pupil's action, in other words, still hangs in the air as an interpretation of the rule until the teacher 'fixes' it for him as either a correct application of the rule or not; and such interpreting of the rule by the pupil is nothing less than his *imitating his teacher*, since the teacher *is* the rule for him. But if the pupil already understands the rule through expertise, he proceeds simply by acting on his expertise and this requires no mediation on his part, that is, he no longer requires the presence of his teacher. His action is now a manifestation of the rule, of his expertise and understanding, that is, of himself, *since he has now become the rule*. And so it is now from the standpoint of the expert that he is then able to justify his action through explanations and rule-formulations, as further evidence of his expertise.

On the community view of rule-following, the perspective of the novice rule-follower, that is to say, the perspective involving interpretations of a rule, is used to argue for the necessity of

other rule-followers who act as the measure for this novice rule-follower; it is an argument for the perspective involving the manifestation of the rule in an action; this community view emphasizes the conclusion Wittgenstein reached in PI 201 in the so-called ‘paradox of interpretation’, namely, that its solution requires that there be *others* to act as a standard by which we can determine whether any given application of a rule is actually in accord with the rule, through what “we call ‘obeying a rule’ and ‘going against it’ in actual cases”<sup>3</sup>. Of course, there is in fact a genuine problem of interpretation when it comes to following a rule. But what the community view fails to see, however, is that one who has attained ‘mastery of a technique’, an expert in his field, has already ‘solved’ the problem of interpretation as it relates to the problem of the novice rule-follower, and his independence is actually founded on the overcoming of this ‘paradox’. As long as one *still* requires others in order to be said to follow a rule, one can never really be an expert or have mastery, but must act as one who follows an order, dependent on another for the determination of correctness: one is continually a pupil, or the slave to another’s rule—you might say that here one never *arrives* at the rule, that is, one never *becomes* the rule. And so the ‘paradox of interpretation’ is really only a problem for the pupil who is learning to follow the rule, a problem that is easily solved by his learning the rule. The conclusion of PI 201, by presupposing a rule to be followed, and hence something that *can* be interpreted, presupposes the teacher or expert rule-follower at the same time, since the teacher or expert really is the ‘rule’ which the pupil then imitates or ‘interprets’: this is hardly a conclusive argument for the necessity of a social community for rule-following to exist; the argument of PI 201 merely reduces to saying that for a rule to be interpreted, it must already actually exist, that is, there must *already be* a rule-follower, someone *other* than the

interpreter—but since there need be *no interpretations* for the teacher, there need be no other rule-followers except *that* rule-follower: rule-follower and rule are now *one and the same*.

If the rule of the expert, that is, the expert qua expert, is the rule itself, then rules *can* act at a distance, as teacher is to pupil. There *can be* an intermediary acting between a rule (the teacher) and its application (the teacher's action), because the rule is really just a human being and this human being's *action* can be imitated by the pupil, and hence *interpreted*. Thus the pupil is this intermediary, since he is the one who acts by imitating his teacher, making his action also the *teacher's* action; we can say this, only because the purpose of the action is still with the teacher and not yet with the pupil.

The rule-formulation or sign, however, confuses our understanding of the relation between the rule and its application, because this sign *can also be interpreted*. But this sign can be interpreted by *any* action, because it can be made to have any use, and *its* purpose will always reside outside itself and in the one who made this sign. The genuine problem of interpretation, however, arises from and is then solved in the relation between teacher and pupil, when the pupil finally grasps the purpose of the rule and unites his action to it; his action is no longer his teacher's, but now his own. The use of a rule-formulation is not a necessary part of this transformation, only an ancillary part. What is often called the 'internal relation' between a rule and its application is just a definition or stipulation otherwise called 'the sign'; but the 'internal relation' between a rule and its application in the pliant picture of the rule is a *genealogical relation*, that is to say, it is the *unity* of the rule-follower's action with its purpose, and this unity *comes to be* through the pupil's imitation of his teacher, that is, through the *external* relation between teacher and pupil that is only now truly called 'compliance with a rule'. This is why I stated earlier that the essence of rule-following is the genealogical

transformation, through the imitation of action, of the difference between rule (teacher qua purpose) and rule-follower (pupil qua action) into the self-identity of the rule-follower (the pupil qua action) with the rule (himself qua purpose). By having the rule-follower become the rule, an internal or self-relation is forged within the rule-follower himself.

Since the rule-follower really is the rule, then, the rule-follower cannot be said to follow the rule *blindly*. The picture drawn here is entirely different from that of Wittgenstein; hence no qualifier is necessary to characterize *how* he follows the rule, blindly or otherwise, for if he follows the rule at all, he already *is* the rule. There is no question or answer concerning *how he is able* to follow it, except by saying perhaps: “he already learnt it.” For Wittgenstein, however, there is still a difference between rule and rule-follower, and so the rule cannot be so easily dispensed with, even if it turns out to be useless for guiding the rule-follower when he follows the rule. *That* is why the expression ‘following the rule *blindly*’ is paradoxical in his case but still necessary, since the expression is also saying that this rule is both there and not there: what is *there* to follow, a rule-formulation or sign, is useless to guide the rule-follower further, and so the rule that is *not there*—the so-called ‘extension’ of the rule qua sign—is what he ends up following—*blindly*. (And what better way is there to follow a rule that is *not there* than to follow it *blindly*!) Despite the fact that this rule he is following is not really there, the rule-follower somehow still follows the rule; the rule-follower in fact now follows the rule because he has been trained in a certain way, that is, he has been *conditioned* to do so. For Wittgenstein, the rule as handrail still works because its missing extension has been reduced to the *conditioning* of the rule-follower; in this conditioning, or ‘training’, he is able to preserve the rule that is different from the rule-follower, even though he now had to think of rule and rule-follower abstractly in order to maintain this difference. This was easy enough—he then

simply *restated* these conditions once again in the form of the sign, although this time the sign was not the standard rule-formulation the rule-follower said he followed, but the sign that fell under *our description* of his training to ‘follow the rule’. The connection here is not just ‘causal’ but also ‘grammatical’. Wittgenstein could not get rid of this abstract rule—the sign itself—and hence the abstract difference between rule and rule-follower, since this difference lay at the heart of his picture of the rule.

### **1.3 Following a Rule is not Analogous to Obeying an Order**

Because Wittgenstein could not get rid of the abstract rule that is followed blindly, his notion of ‘mastery of a technique’ seems in the end to be a limited form of mastery, something closer to what you would expect from a skilled servant who has been dutifully trained: the rule-follower, after all, is said to act blindly (PI 219) and as a matter of course (PI 238)<sup>6</sup>. In fact, his reasons for acting will soon give out, and so he just acts, without reasons (PI 211)<sup>7</sup>, just as he is inclined to say, when all his justifications have been exhausted, “This is simply what I do” (PI 217)<sup>8</sup>. Moreover, his thinking he is following a rule cannot be the same as his following a rule; thus he cannot follow a rule ‘privately’ (PI 202), that is, independently of others<sup>9</sup>. One would think that with such a characterization of the rule-follower, the rule has no useful role to play in rule-following apart from its formulation, but one would be wrong. The rule is still there, but it is there like an order given from a superior: the tendency is to see the rule as though it were analogous to an *order* (PI 206)<sup>10</sup>. The rule imposes itself on the rule-follower just as someone who gives an order imposes the order on its obeyer. The rule, like an order from a superior, requires *compliance* from a rule-follower who is passive by nature.

This last criticism of Wittgenstein's view of rule-following is not based on an isolated passage, but on a pattern of expressions used to characterize rule-following that connect with how he views the nature of the rule itself. At RFM V.32, for example, he says straightforwardly that "the rule conducts you like a gangway with rigid walls....Here is the rule, like an order! And like an order in its effect."<sup>11</sup> He is no doubt expressing a similar view in PI 212, when he says: "When someone whom I am afraid of orders me to continue the series, I act quickly, with perfect certainty, and the lack of reasons does not trouble me." In fact, to speak of following 'blindly' and as a 'matter of course' is closer to the language of obeying an order than that of following a rule, since this language is compatible with acting without purpose. Yet the significant difference between following a rule and obeying an order is that in obeying an order, we need give no thought to the purpose of the order—that belongs to whomever gave the order—our purpose is merely to follow it, that is, to carry out the order. But following a rule should not be characterized in the language of obeying an order, otherwise we merely focus on the mere act and do not see the rule's connection to its purpose. The action in following a rule, like that in obeying an order, becomes an end in itself. In order to establish the meaning of this mere action, then, requires some form of compliance with another. Seeing rule-following as analogous to obeying an order thus re-enforces the picture of compliance.

Following a rule is *not* analogous to obeying an order in this respect: an order is given to someone by someone else; thus the action which is the execution of the order is both initiated and then completed, that is, judged, by someone other than the one carrying out the order, even if the one carrying out the order *knows* what counts as fulfilling the order (the servant must know, after all, what pleases his master). Such an action is closer to that of the pupil who

understands the rule well enough to follow it, but still requires the teacher both to set him upon his appointed task and confirm for him that he has completed this task. Thus the primary objection to any analogy between following a rule and obeying an order is that the difference here is between independence and dependence of action. For the rule-follower, the beginning and end of the action is *with himself*; for the order-obeyer, the beginning and end of the action is *with another*, that is, with the one who gives the order. Characterizing rule-following as analogous to order-obeying is a misleading step towards picturing the rule as an order, and so as something which is independent of the rule-follower and with which he must comply. Indeed, when Wittgenstein makes this analogy explicitly in PI 206, it is an analogy based on our *training* to follow a rule and obey an order; and we fail to see that this training masks the significant difference between the two, because training to follow a rule is a training for independence, whereas training to obey an order is a training for dependence or subordination to another. Training is in fact the only true form of compliance in rule-following.

The objection to this analogy, therefore, is no mere quibble, since the analogy lends support to the idea that rule and rule-follower must be *different* from one another, just as the one giving an order is different from the one who is ordered to obey it. Thus this analogy in large part reinforces the compliant picture of the rule, giving rise to the paradox of compliance. Seeing rule-following as analogous to obeying an order only obscures what is essential in rule-following, namely, the phenomenon of imitation of action. Obeying an order presupposes rule-following skills as does giving orders, since giving orders and obeying them are just rule-following capacities distributed among more than one individual. If there is any instructive analogy here at all, obeying an order is analogous to one stage of following a rule—the pupil's imitation of his teacher. And this analogy can now be given weight through the fact that when

one learns to follow a rule, we do have here a case in which rule-following capacities are being exercised and formed through the interaction of two individuals, one independent (the teacher or parent), the other dependent (the pupil or child). And it is in such a case that we most naturally see ‘orders being given’ and ‘orders being obeyed’. But the asymmetry found in the relationship between teacher and pupil is not fixed or permanent but *transitional for the pupil*, and for this reason, rule-following is prior to order-obeying as whole is to part, since rule-following forms the whole phenomenon of which order-obeying is only a part or stage. In other words, compliance in rule-following is not the *goal* of the pupil’s training to follow a rule, it *is* his training—and this training eventually comes to an end.

#### Notes to Chapter 1:

1. PI 50: “...There is one thing of which one can say neither that it is one metre long, nor that it is not one metre long, and that is the standard metre in Paris.—But this is, of course, not to ascribe any extraordinary property to it, but only to mark its peculiar role in the language-game of measuring with a metre-rule...”

2. cf. PI 219: “‘All the steps are really already taken’ means: I no longer have any choice. The rule, once stamped with a particular meaning, traces the lines along which it is to be followed through the whole of space.—But if something of this sort really were the case, how would it help?

No; my description only made sense if it was to be understood symbolically.—I should have said: *This is how it strikes me.*

When I obey a rule, I do not choose.

I obey the rule *blindly*.”

The last sentence in this section—“I obey the rule *blindly*”—has been interpreted variously by different commentators. Baker and Hacker, for example, understand ‘blindly’ as synonymous with ‘with certitude’ or ‘complete confidence’; Kripke understands it as synonymous with ‘unhesitatingly’. The context here certainly links ‘blindly’ to a lack of choice in how to proceed when following a rule; hence other commentators, like Williams and Medina, see here some form of ‘alternative-blindness’; thus they characterize following a rule as analogous in some sense to ‘blind obedience’, even though Anscombe’s translation of *folgen* as ‘obey’ is misleading at best (‘obey’ is usually the translation of *befolgen*, not *folgen*). But the choice of ‘blindly’ (*blind*) as a word that can also be used to capture all these various senses—confidently, unhesitatingly, unquestioningly, as a matter of course—is symptomatic, I claim, of Wittgenstein’s tendency to see

rule-following as closely analogous to order-obeying. After all, one also obeys an order from a superior unhesitatingly and unquestioningly, that is, with blind obedience; and although there is some overlap between the two, there is a significant difference between obeying an order and following a rule. I only wish to point out here that these same ways of characterizing rule-following are in fact incompatible with an account of rule-following that is distinguished from order-obeying.

It could be argued that it is significant that Wittgenstein chose the word ‘blindly’ instead of others like ‘unhesitatingly’ or ‘unquestioningly’, because ‘blindly’ is a word that connotes a lack of vision, and so it is somewhat paradoxical for him to characterize rule-following, an activity that manifests understanding or knowledge, with the lack of it (‘seeing’ is often a synonym for ‘understanding’ or ‘knowing’; indeed, acting ‘as a matter of course’ (*selbstverständlich*) would seem to underscore such understanding by making what must be done ‘self-evident’). This paradoxical way of putting things in PI 219 is of a piece with his explicitly paradoxical language in RFM-V.46 that compares a rule with a short bit of handrail acting as a guide, beyond which “there is nothing there; but there isn’t nothing there!” Indeed, part of my overall argument is that Wittgenstein never really solved this paradox, although he was aware of its existence—it could also be called the ‘paradox of compliance’—and so the metaphor of acting ‘blindly’ fits his exposition well; this paradox may very well also explain why he chose to see rule-following as closely analogous to order-obeying, since order-obeying also has the aspect of ‘blindness’ to it in the sense of ‘blind obedience’.

Order-obeying does seem to have this one thing going for it, however, in relation to rule-following using signs—in using these signs, one pays no attention to the consequences of using them, that is, their use by the sign-user has no consequences other than their being used in a certain way (a use which also shows at the same time the *autonomy* of the sign). And so, in *this* sense of ‘blind’, we are indeed ‘blind’ to any consequences we may feel coincides with their use, since it is not up to the order-obeyer to determine the purpose of the order—that is determined by the order-giver—he, the order-obeyer, is simply to carry out the order, to follow the order ‘blindly’. For him, the order has *no purpose other than* its mere execution; he expects *no reward other than* a job well done; yet the order-obeyer, like the sign, still is and must be *for another*. This peculiarity of rule-following involving signs is explored more extensively in the discussion of the use and usefulness of the sign in Chapter 3.

3. PI 201: “This was our paradox: no course of action could be determined by a rule, because every course of action can be made out to accord with the rule. The answer was: if everything can be made out to accord with the rule, then it can be made out to conflict with it. And so there would be neither accord nor conflict here.

It can be seen that there is a misunderstanding here from the mere fact that in the course of our argument we give one interpretation after another; as if each one contented us at least for a moment, until we thought of yet another standing behind 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 ‘obeying the rule’ and ‘going against it’ in actual cases.

Hence there is an inclination to say: every action according to the rule is an interpretation. But we ought to restrict the term ‘interpretation’ to the substitution of one expression of the rule for another.”

4. PI 211: “How can he *know* how he is to continue a pattern by himself—whatever instruction you give him?—Well, how do I know?—If that means ‘Have I reasons?’ the answer is: my reasons will soon give out. And then I shall act, without reasons.”

5. In mathematics, as with all rule-governed activities employing signs, the ‘why’ does not dictate the ‘how’ as clearly as it does in baking, for example. A mixing bowl has a ‘why’, that is, a

purpose or function, and a ‘how’, which is the proper way in which it is used that fulfills that function. But there is also *another* ‘how’ for the mixing bowl, and this ‘how’ is the way in which it is manufactured. The ‘how’ in mathematics (e.g. the operation of addition) would be like this latter ‘how’ of the mixing bowl; it is independent from the other ‘how’ that fulfills a practical purpose, and so *this* ‘how’ is also autonomous in an important sense, because there are really two activities, as it were, going on at the same time: the activity that is the use of a sign (including, importantly, its past usage), and the activity which is the use for using such signs. The latter use constitutes the *usefulness* of signs. Thus it may appear as though mathematics needs no ‘why’ in order for it to be mathematics; or put in another way, the normativity in mathematics is not *directly* based on the fact that mathematics “has been found to pay”. But the ‘why’ of mathematics does not determine normativity in mathematics in that way. In the case of the mixing bowl, the material out of which it is made could be said to be its normative basis, since the ideal material will allow it to retain its proper shape so that it can continue to function as it should. Normativity in mathematics likewise retains the proper ‘shape’ of mathematics, such that mathematics can still function in the practical way it does for us, and the activity which retains the proper ‘shape’ of mathematics, its normative activity, is independent from whatever activity that then uses mathematics for its own purposes, which may be many. In short, signs are tools, or rather, the *use or employment* of signs is itself a tool, and unlike a mixing bowl, which has more or less a single purpose and whose shape can only be maintained by a finite number of media, mathematics is not limited to a single purpose, nor is it limited by its medium (its physical medium is ‘pen and paper’, but its functioning media are the ‘uses’ of signs).

This duality of use and usefulness with respect to rules involving signs is why the meaning of rules involving signs *must be overdetermined*. (The physical media of sign-usage, whether ink on paper or vocal sounds, cannot dictate usage or purpose with respect to signs: that is why signs exist at a higher level of abstraction, that is, of arbitrariness and autonomy, than artifacts like mixing bowls; sign-usage is in effect part of the ‘material’ medium of the sign, like the hardness or softness of the material of a mixing bowl made of, respectively, steel or plastic. Think here of the ‘hardness’ of the logical or mathematical ‘must’; think also of Wittgenstein’s talk about the ‘inexorability’ of mathematics.) I discuss this further in Chapter 3 and Chapter 4.

6. PI 238: “The rule can only seem to me to produce all its consequences in advance if I draw them as a *matter of course*. As much as it is a matter of course for me to call this colour ‘blue’. (Criteria for the fact that something is a ‘matter of course’ for me.)”

7. PI 211: “How can he know he is to continue a pattern by himself—whatever instruction I give him?—Well, how do I know?—If that means ‘Have I reasons’ the answer is: my reasons will soon give out. And then I shall act, without reasons.”

8. PI 217: “‘How am I able to obey a rule?’—if this is no a question about causes, then it is about my justification for my following the rule in the way I do.

If I have exhausted the justifications I have reached bedrock, and my spade is turned. Then I am inclined to say: ‘This is simply what I do.’

(Remember that we sometimes demand definitions for the sake not of their content, but of their form. Our requirement is an architectural one; the definition a kind of ornamental coping supporting nothing.)”

9. PI 202: “And hence also ‘obeying a rule’ is a practice. And to *think* one is obeying a rule is not to obey a rule. Hence it is not possible to obey a rule ‘privately’: otherwise thinking one was obeying a rule would be the same thing as obeying it.”

10. PI 206: “Following a rule is analogous to obeying an order. We are trained to do so; we react to an order in a particular way. But what if one person reacts in one way and another in another to the order and the training? Which one is right?”

Suppose you came as an explorer into an unknown country with a language quite strange to you. In what circumstances would you say that the people there gave orders, understood them, obeyed them, rebelled against them, and so on?

The common behaviour of mankind is the system of reference by means of which we interpret an unknown language.”

11. RFM-V.32: “But I mean: the rule conducts you like a gangway with rigid walls. But against this one can surely object that the rule could be interpreted in all sorts of ways.—Here is the rule, like an order! And like an order too in its effect.”

12. Zettel 320: “Why don’t I call cookery rules arbitrary, and why am I tempted to call the rules of grammar arbitrary? Because ‘cookery’ is defined by its end, whereas ‘speaking’ is not. That is why the use of language is in a certain sense autonomous, as cooking and washing are not. You cook badly if you are guided in your cooking by rules other than the right ones; but if you follow other rules than those of chess you are playing another game; and if you follow grammatical rules other than such-and-such ones, that does not mean you say something wrong, no, you are speaking of something else.”

## **Chapter Two: The Relation between Teacher and Pupil**

### **2.1 The Relation between Teacher and Pupil is a Natural Extension of that between Parent and Child**

Some may ask, like Bloor, where experts derive their authority, if not from the consensus of the community?<sup>1</sup> But experts are not experts merely in virtue of being chosen, they are experts in virtue of who they are and what they do in relation to others. They do not in fact derive their status as experts from the community, but derive it analogously, in relation to the non-expert, from the natural relation which exists between parent and offspring: it is an asymmetric relation of dependence. The teacher, who is an expert in a certain relation to another, namely, the pupil, is an extension of this natural relation found throughout the animal world (the parent-offspring relation in animals). As long as the child is dependent on the parent and takes the parent to be its authority and provider, the relation will exist and continue to exist, although it will not exist indefinitely. But an adult is not dependent on the community in the same way; the community is not a parent, and can bestow authority to an individual only in an abstract sense, as is the case where individuals by convention elect one of their own to make unilateral decisions for them all; this conventional kind of authority is arbitrary and does not thereby abolish the authority of those who elect and choose to defer such authority to another of their kind. Authority remains with each of them; and the elected has his abstract authority only on this basis: he cannot strip them of such authority, but they can strip him of his, because they are all peers equally and on a par with the elected, like those who have mastery of a technique. Here the relationship between the elected and the electors is freely chosen and arbitrary, in the sense that anyone of them could now refuse to follow the elected and still have mastery of a technique. In fact, the relationship among peers is reciprocal; techniques can be exchanged or

improved, whereas this is not true of the relation between parent and child: a child will learn a technique in the way the parent practices it (for example, language), even if it is thought to be deficient from another's perspective. For another adult here to correct the parent or to teach otherwise would only *confuse* the child, not necessarily improve his technique. This shows that the parent has a unique status for the child, a status which is not acknowledged among other adults. And eventually, too, the parent will lose this status for the child as the child grows and becomes independent, that is, *another* adult.

The difference between a young child and a mature adult using a language is that if a young child is reprimanded for mistakes, he usually is incapable of disputing with an accomplished language-user whether he is in fact following a rule of language or not; young children are dependent on adults, and so do not ordinarily question their authority. Only with time will they become independent and capable of defending themselves. One of the first ways that they do become independent is by quickly becoming competent language users; they learn language far earlier than they learn other activities, because language itself becomes a necessary tool for learning more sophisticated human activities, like taking up the subjects taught in school, reading and writing, learning skills essential to trades and professions, and so on. Mature adults, however, can also correct one another in matters of language, correcting pronunciation, syntax, the failure to distinguish one word from another; but these corrections are not on the same level as between parent and child. For adults, a standardized language represents an idealized form of the language with which all are thought to comply, but exceptions or variations are permitted in a way that they are not for children learning a language, even if some of the 'mistakes' children make are similar or the same as an adult would make. As adults, we tolerate things other adults do, despite the fact we believe they are wrong or

mistaken. We are justifiably less tolerant with children, because we feel we have the right to teach them as we feel fit, that is, they should do things *as we do them*.

When children learn to speak, for instance, their speaking is largely hit and miss, and it is our responses of understanding and not understanding which is corrective for them in a practical sense, that is, through trial and error they learn how to speak more effectively with us. What shows us that children do not merely learn passively from us by accepting our reactions and corrections as beyond questioning is the fact that they also try to justify their actions, that is, manipulate us, by the same speech we have taught them, since justifying one's action in rule-following is a different skill than the skill one is justifying. Manipulation is clear evidence here of awareness of the power of the sign and the accessibility that the sign provides one to the world at large. The child, as it matures into a competent language-user, no longer feeds off the reactions of the adult in the helpless way in which an infant is suckled by its mother. The reactions of the adult are important and will remain so for the length of the child's education, but there is a weaning process in this relation, because the interaction between parent and child is not completely asymmetric but becomes one which itself primes the child for adult interaction and all the ways in which adults must learn to live with one another, and this invariably involves mutual respect as well as *surreptitious manipulation*: the use of a thing for one's own self-interest separately from the interests of another without this being transparent to the other. By surreptitious manipulation, the child begins to lessen his dependence on the adult, even if he really is still dependent on him and must be; by taking hold of language *for himself*, he no longer is merely given or allowed it at the discretion of another. In a similar way, the master, though dependent materially on the slave and the institution of slavery, can be 'free' of the slave by manipulating him freely and in full transparency, as a tool for his own

purposes. And so, too, for one who has mastery of a technique like language, since the child's ability to give justifications of his action and make similar judgments about another's marks an important turning point in the development of the rule-follower—he can now rationally stand his own ground against the critical judgments of others. This moment therefore marks the end of his training to follow a rule, that is, it is the end of his *compliance with the rule*.

Manipulation of others, like imitation of others, comes naturally to human beings, insofar as neither is learned. Indeed, manipulation may just be a form of imitation, if manipulation is viewed as a child's awareness of the difference between what another says and does, and his exploiting of this difference by imitating any discrepancy between the two for his own advantage. The child would then really be simply imitating the adult, to the extent that he sees the adult's duplicitous purposes in giving justifications and acting. Whether a child manipulates us on this basis, or whether he does so independently of imitation, does not concern us at this point: we simply acknowledge that such manipulation does exist in the child, that this marks his growth towards independence from the adult, and that it is itself not part of the instruction given by the adult, but a natural development within it. Such acknowledgement has implications for our understanding of normativity: normativity is not something that is exclusively found in the relation between parent and child, teacher and pupil—in other words, in a so-called 'social' relation involving compliance—but something that is, more importantly, developing *within* the child or pupil through a genealogical relation to the parent or teacher, as an expression of the child's or pupil's independence; the normativity of the rule *comes to be* through grasping the purpose of the rule, through making an activity *one's own activity, that is, for oneself*. Normativity is thus the expression of self-determination in the pupil or child.

## 2.2 Imitation of the Teacher and Correction of the Pupil

Before children are trained in language-use by correction, they are pre-trained, as it were, by imitating language-users. Imitation, like play, comes naturally to children, and these are the first stages of training in the formal sense. Children can just as easily imitate the correcting behavior of adults, even when they clearly do not know what the correct behavior is. The first sense of a language-game in PI 7, much like a ‘primitive language’, is perhaps like this imitating or playing. Obviously, there is a difference between imitating an activity through imagination and actually doing that activity, and it would be premature to call such behavior ‘rule-following’. But the importance of imitation in rule-following is that it is something we do naturally and is therefore prior to training through correction, indeed to any formal training at all: without it, correction itself would lose its usefulness, since the teacher’s correction of the pupil’s action does not just prompt the student to make another attempt at following the rule, rather the teacher’s example of correcting his own mistakes is itself meant to be imitated by the pupil, so that the pupil learns to correct himself just as the teacher corrected himself. In fact, imitation is so essential to learning to follow a rule, that it actually constitutes his training properly speaking. By imitating his teacher, the pupil is already agreeing with him, that is, complying with his teacher. Imitation is thus the true form of compliance in rule-following; correction is only necessary when imitation at first fails its purpose, and even then, imitation again becomes the only remedy—a second course of action, now involving correction itself.

How much correction is involved in someone learning a language or any other rule-governed activity? Doesn’t imitation, that is, the natural mimicry of the behavior of others, count for much more than such corrections? It must, if someone can learn chess and the rules

of chess from merely watching others play it, as Wittgenstein noted in PI 31. Correction would play no part, unless he were also to observe others being corrected during play; at any rate, his observations and later attempts at imitating the playing of chess can get one so far as to be considered in some sense the master of a game. What underlies and sometimes belies the corrections given by others is their actual behavior, which may or may not be consistent with such corrections, especially when these corrections come in the form of the sign, as judgments of rule-following behaviour. People do, after all, correct others for faults which they themselves routinely commit. Corrections of this sort are like the adage ‘do as I say, not as I do’, and can impede the learning of the pupil, but not necessarily for long. Pupils, like children, can determine for themselves whether teachers are consistent in what they say and do. We say that a pupil or child has matured when he can achieve the stance of a critic himself and judge his teacher. And often pupils and children actually ignore the instructions and corrections of others, if they determine for themselves that there is such a gap between what is said and done; in these cases, what the teacher does will trump what he says. The pupil will simply *act* in the way he *sees* his teacher to act, and this is really just to imitate him. Imitation comes naturally to human beings, and is itself something not taught by another.

What really holds rule-following together is not simply acts of correction by other rule-followers, even teachers, but acts of imitation; and these acts will only bear fruit once the point of these actions has also been grasped, since to grasp the purpose of the rule is itself the purpose of the training, that is, the purpose of the imitation of the teacher’s actions. We overestimate the importance of instruction and correction in learning to follow the rule, because we see such learning as essentially passive and not active; instruction and correction are necessary for the pupil to learn how to follow the rule *blindly*, that is, he needs the proper

*conditioning* brought about by instruction and correction. The pupil must be moulded into a certain shape by the teacher against his own inclinations.

We must be wary, however, of attributing to the teacher in teaching the rule to the pupil a responsibility which does not really belong to him; that is to say, it can be enough that the teacher simply be there for the pupil and do what he normally does in following the rule, since the teacher primarily functions as the rule for the pupil to follow; he does not ‘mould’ the pupil into a rule-follower, although it may look that way in the case of a slow learner who needs constant correction. Jose Medina’s account of Wittgenstein’s view of the matter is unfortunately that of the teacher *drilling* or *inculcating* into the pupil the proper ‘behavioral regularities’ along with a ‘normative attitude’ towards these behaviours (165); and the pupil is merely *submissive* to this drilling; indeed, he has a submissiveness characterized as ‘blind obedience’ to his ‘master’ (163)<sup>2</sup>. In addition, the training of the pupil is meant to produce in him the *illusion* that this training is irrelevant in following the rule (167). But such a characterization of learning is too *forced*: it is as though learning to follow a rule were a form of brainwashing, and the teacher’s responsibility were to *make* the pupil into a blind follower, into a *work of his own hands*, like a living sculpture. On the contrary, the teacher at first needs to stimulate the pupil’s natural interest in learning the rule, so that he *wants to be like his teacher*; and the teacher does this by showing the pupil the purpose or point of the rule, what he can do with it. This is why the pupil needs to be exposed to the purpose of the rule as early as possible, since it is what needs to be grasped most of all and is that for the sake of which the rule even exists; and by grasping this purpose as firmly as possible, the pupil will know that it could have been grasped differently, or perhaps, through some innovation, in a better way—but there is no illusion produced here, no brainwashing. (The pupil’s grasping of the purpose of

the rule *is* his ‘normative attitude’ and not something ‘inculcated’ by his teacher separately from his actions.) The teacher’s instruction and correction of the pupil are important, of course, but these are not the foundation of his teaching; the teacher’s fundamental role is better thought of as *passive* and the pupil’s as *active*, not the reverse, since the pupil is the one who is *becoming* what his teacher already is, and he can only do this if he takes the initiative by imitating another’s actions. The teacher is really just activating by his presence and encouragement the innate capacity of the pupil to follow him, since his pupil is also a *human being like himself*, a living being having its own independence by nature, not a lifeless clump of clay.

We think erroneously that if we only had the correct instruction to give the pupil, he would learn to follow the rule correctly, but it is his natural ability to *follow his teacher* that will determine *that* more than any instruction or correction given: the child can learn language just as well by *listening to others speak*. It looks as though we *had to drill* this instruction into the pupil, because we often feel we have to do that in the case of someone who has little or no natural aptitude for what we are teaching, even though this drilling may often be in vain or counterproductive. One could say this perhaps about teaching and learning: the teacher must not look *down* on his pupil in teaching, rather the pupil must look *up* to his teacher in learning.

### **2.3 Agreement between Teacher and Pupil; Two Different Forms of Community**

‘Rule’ and ‘agreement’ are closely related concepts only with respect to the pupil who is learning the rule. But for the rule-follower, agreement with others no longer really applies, nor does ‘rule-following’ literally speaking. What is important for the rule-follower, though, is

consistency of procedure and consistency of result. These are the tests by which we say that someone has mastered the art of baking, for example. We are concerned that the pupil learn the rule and be in agreement with his teacher, but here learning the rule and eventually mastering it only means that at some point the rule no longer apply to him, that is, the rule of the expert, his teacher. When the pupil himself becomes an expert, he is a practitioner in his own right, and his technique comes with its own rule<sup>3</sup>. This transformation, from pupil to teacher, rule-follower to rule, leaves open the possibility that the new rule-follower may excel even his own teacher, or be innovative in the sense of introducing new rules, or better rules. The point here is that anyone who has mastery of a technique is *eo ipso* a standard by himself, someone capable of teaching others his own technique, his own rules. The rule-follower himself has no superiors, only fellow practitioners, his peers, and with them and in agreement with them, the rules of baking can be standardized publicly. But these standardized rules are really just secondary rules, abstracted from the practices of rule-followers in the form of rule-formulations; they are the common expression of the community of experts in their desire to stabilize their trade or profession and ensure overall consistency of procedure and product. They are like the formal rules of standard English in that no one really observes them strictly in everyday discourse, but they serve instead as an instructional aid in general education and as the formal way in which the experts themselves communicate with one another.

There are, then, two different kinds of rule-following: primary rule-following, namely, that which exists between expert and non-expert (teacher and pupil, parent and child) in a relation of *normativity*, and secondary rule-following, that is, that which exists between experts as peers (adults, teachers) in a relation of *community*. This latter community is what public standards are meant to apply to, and these standards merely stabilize a community of

practitioners and improve its overall practices, since it essentially involves the production of rule-formulations or signs. Nevertheless, these public standards of the community are secondary rules abstracted from its members, the experts (rule-formulations are really just secondary rules or abstract rules). The true source of normativity remains these experts and will always be so, because rule-followers are primarily individual human beings and rule-formulations cannot transcend such individuals and what they do. The community itself is not a rule-follower and cannot be one, no more than humanity is a human being.

The individual and the community do not exist on a par; only individuals exist on a par: the community is made up of such individuals. The community neither acts nor reacts, and consequently cannot be said to make mistakes either. But this is not to define it as infallible, as a standard of normativity, since this concept of community is merely abstract, like the concept of humanity. If the concept of the community is to be meaningful for the concept of rule-following, it must be thought of as the relation of individuals who are already rule-followers themselves; such a community of individuals, if organized formally into professional societies, would function as a buffer against arbitrary, sudden, or dangerous change from within these societies, as long as individuals deferred to such standards or rule-formulations. This community of peers thus stabilizes and solidifies practices which already exist, hinders but does not prevent change, in short, provides an appropriate environment for the attainment of common goals. But the social community's ability to do all this is measured according to the lifespan of individuals and their generations. It is this lifespan of the individual which sets the pace for change allowed for society in general, because we often speak of change in society by reference to generations and the gaps which exist between generations. We freely acknowledge that things such as values, standards of normativity, aesthetic tastes, slang, all

these change, when they do, *between* generations, and so they also must achieve stability *within* each generation: this stability is found *in the form of the individual*, not as part of a whole, but as this whole itself. It is the stability found within generations, that is, within the individual as whole, which plays the important part reflected in the concept of community. Such a conception of community does not think of the community as just a background or context for meaning to be established in our activities, since the individual expert alone determines that; it is something now abstracted from the activities of the individual, a secondary form of rule-following in which compliance makes sense, as long as there are bodies created expressly for the purpose of determining whether members have acted in accordance with the formalized rules or standards of their respective society. Nevertheless, there is no normativity here apart from what the individual actually does; normativity is in the individual—rather, it *is* the individual, whatever *he* does. That the individual can learn to follow rules and in turn teach others, improving his technique through interaction with other experts in his field, is what community makes possible; it is essentially a relation among those who are peers. In the social community, these peers submit themselves to formalized standards for the sake of agreement or conformity, that is, they become a part of something larger, even though these formalized standards mean nothing without their collective submission to another's authority: this is in fact why the slave qua slave is the paradigm social being—he is a whole reduced to a part by becoming an instrument for another through submission.

Community, then, just is these rule-followers in relation to one another as peers, and it is only in a certain kind of relation, namely, the genealogical relation between teacher and pupil, that the kind of normativity sought by the community view of rule-following is established. In the most general sense, we can say that it is one generation which teaches the next, forming the

unbroken links in a chain; but the links are only formed by shaping the individual segments while they are still malleable, according to the shape of the last: once they are cooled and hardened, they become solid and set and can themselves be linked to further segments as their model. But the segments are not fused together nor linked randomly; they are indeed separate and individual, but not separate as isolated pieces, only physically separate as interlinked pieces; in this linked form of independent pieces resides the strength and flexibility of the chain. But the last segment is in a sense the most free of them all, because it is still able to be linked to something yet undetermined, and only *it* is able to determine *that* by *its own* particular shape. For this reason we think of the last link as the most important and decisive in the chain after the first. Of course, an individual piece is conceivable apart from a segment linked to a chain, but the shape of that piece could be almost anything then, and should it have the shape of a linked segment, the shape would nevertheless be no more significant than any *other* shape, because *here* its shape does not have the same role as a linked segment, unless it were somehow the beginning of a *new* chain. The community thus conceived is a *genealogical* community and not a *social* one, and the difference between the two consists in this important distinction: the social community is a community of *parts* in which the community itself functions as the whole through a form of abstraction in which the individual is reduced abstractly in society to a part of society, whereas the genealogical community is a community of self-generating *wholes* who are peers.

Jose Medina (2002) misunderstands the role of the community in rule-following by making the teacher into a *representative* of the community of practitioners, as though the teacher were acting as a mediator of the rule on behalf of a social community in which he is only a part or mere function of this community: the teacher is like the depiction of the rule, one who *speaks*

*for the community*<sup>5</sup>. But the teacher does *not* mediate between the community and the pupil, otherwise we have allowed a gap to exist between the rule and its application through a false intermediary: the *teacher* would in essence be giving an interpretation of the rule that still needs to be ‘fixed’ by the community. Yet if the teacher *is* the rule and nothing comes between him and the pupil, the teacher is the *only* background for the pupil, so to speak, and there is nothing *behind* the teacher, no background behind the background, no ‘community’ acting at a distance. The social community here is a false background, an abstraction. The true background for the pupil is another human being, the teacher or expert in his element. In these cases, when the pupil acts in the presence of the teacher, the *pupil* becomes the mediator of the rule in the sense that the action of the pupil is still done *on behalf of* the teacher, and this action can only be corrected or approved by the teacher, not by the pupil himself. Here it makes sense to call what the pupil does an interpretation of the rule, since the rule is in fact *the teacher* and the rule-follower’s interpretation of the rule is *the pupil’s imitation of his teacher*. Without the teacher present, the pupil would find himself in the paradox of interpretation spoken of in PI 201; and just as there must be an actual rule to interpret and not just interpretations of a ‘rule’, so there must also be a teacher to imitate and not just imitations of a ‘teacher’.

We must take care, therefore, not to misunderstand the role of the word ‘background’ in our account of the teacher’s relation to his pupil. When one speaks loosely of meaning as determined always within some background—for example, a background of practices—the background is meant to fill in the emptiness of the dead sign or action with some content—the *context* supplied by the background. But it is easy to think of the teacher as also belonging to this wider context (for example, the community); perhaps we even need to bring in the

teacher's teacher, too, or something else in which to put the teacher himself in context. We can widen this background as much as we like, however, and attribute all kinds of additional things to this context without really adding any depth to our account, since we are merely connecting one thing to another in our description of this context by simply forming an extended rule-formulation or sign. We must think of the teacher, therefore, not as a part of this background but as this background itself or, better still, as what determines this background, despite the temptation to find the community as this background; after all, we think, isn't the teacher a part of the community?

Unfortunately, thinking of the social community, or 'community agreement', as the 'background' trivializes, indeed dismisses, the one who *actually establishes meaning*, namely, the human being qua expert, the individual, or *the one who acts*. It is always admitted, of course, that the rule-follower must act (e.g., act blindly, without reasons, from inclinations arising from training, etc.), but the rule-follower really becomes absorbed entirely into this background and can safely be ignored other than for his mere utility as an agent of the background, someone who simply acts in a certain way in a certain context because he has been trained by his community to do so; and we think not just of the pupil in this way but of his teacher, too. We have essentially turned the rule-follower into an abstraction, someone who exists only in a context, and that means, he is reduced to the description we give of him. We need not ever appeal to the pupil or the teacher for his reasons or justification or even his expertise, since they have now become irrelevant to our understanding of these actions. The background will tell us—it is the last court of appeal—whatever that background turns out to be in our descriptions. But we forget that it is the rule-follower who first acts in *this* context, against *this* background, in *these* circumstances, just as it is the language-user who first puts

*this* word in *this* context, against *this* background, in *these* circumstances. But the particular background of *this* action, and hence *this* meaning, is then really determined by the rule-follower alone, since the rule-follower, as said earlier, is the one who has mastered his surroundings, and such mastery—this putting things in their proper place and context—is what actually *constitutes* his mastery of a technique. When the rule-follower has gained mastery of *all these circumstances*, and hence mastery of the ‘background’, only then can we say that his actions have a particular meaning and belong to a certain ‘context’. Not only *this* action, then, but *this background, too*, is determined or given meaning by the rule-follower—there only *is* a background to speak of here because there is a rule-follower here<sup>4</sup>. Background and action are thus linked together only *through* the rule-follower as two sides of a coin: hence the ‘background’, ‘context’, ‘circumstances’ of the rule-follower and his actions are not something into which the rule-follower and his actions are placed and that exist as *preconditions* for him acting meaningfully; on the contrary, the rule-follower’s doing this or that is *his* putting a ‘context’, and thus a meaning, around what *he* does. What are called ‘preconditions’ here are simply reformulations of the rule as an extended rule-formulation; such ‘preconditions’ merely point to the *conditioning* of the rule-follower.

My ‘Kantian maneuver’ in rule-following therefore consists in this insight: the rule-follower and his action are not given a meaning by a context, a context is rather given a meaning by the rule-follower and his action, since this context really exists *only* in the form of a *description* of what the rule-follower is doing when he follows the rule, and the rule-follower himself justifies his own action by *giving* such a description, even though he need not give a justification, that is, a description, at all. The ‘context’, the ‘background’, the ‘circumstances’—these are all superfluous, unnecessary in rule-following, like the sign itself,

since they exist *only in the sign and nowhere else*; they are just another manifestation of the rule-follower's expertise—of his *being* the rule himself. Only the context given in a description of an action will pick out the essential from the inessential in rule-following, and so this context merely formulates the rule in an *expanded but identical* form of the rule—and this is in fact the *only* difference between a simple rule-formulation and a context: a context is just a *wider* description of rule-following, not something *deeper*. The community view erroneously thinks that the criteria we as describers use to describe the teacher and the pupil in following the rule constitute the true background of both, but we earlier saw that this unassuming transposition of *our* 'background' onto the teacher and pupil merely expressed *our* expertise or 'mastery of a technique' and thereby failed to acknowledge the independence and mastery, not only of the teacher, but of the pupil, too, even after the pupil was said to have learned the rule from his teacher. The individual, whether teacher or pupil, is looked at as only part of a whole, that is, as part of a wider 'context', something that, by itself, is isolated and private; but if the individual is the only one who can define this 'context', both through wielding signs and performing actions, then *the individual is this whole*. The individual only becomes a part of the whole through *abstraction*, like the slave qua slave, who is merely a part of his master. Consequently, to see both teacher and pupil as part of a larger background or context—'community agreement' in the community view—is to maintain the distance, and hence the difference, between rule and rule-follower, and this is the cardinal mistake of the community view of rule-following.

## Notes for Chapter 2:

1. “Someone is correctly called an authority because everyone else calls them this” (Bloor, 31).

2. “An indispensable condition of initiate learning is that the novice have a *receptive attitude*, that is, that she accept the authority of the teacher *blindly*: ‘Can the pupil contradict and say: ‘How do you know that?’ [...] Well, we do not assume that the pupil can say that’ (RFM VII.26). In fact, we assume that the novice’s attitude will be that of total submission to the authority of the master, of blind obedience to her commands. For, without passive collaboration of the novice, the learning process cannot take place. [...] The role of the novice in initiate learning is to follow the lead of the master” (163).

What Medina fails to appreciate—and Wittgenstein, too—is that the pupil’s following his teacher is not really a passive role but an active one: by conforming his actions to that of another, he makes himself—that is, *becomes*—what the other already is. The relation between teacher and pupil is not like the relation between master and slave, but like that between parent and child—it is a genealogical relation in which the teacher ‘begets’ the pupil, so to speak, since the pupil is the one who *transforms himself* into a teacher by actively conforming to his teacher, that is, by imitating him; and no one teaches him to imitate his teacher—he does this on his own. In fact, the pupil’s inquisitive nature is evidence of his active role in learning to follow the rule; he asks probing questions (“How do you know that?”, “Why do you do that?”), not because he has a perverse attitude, but because he needs most of all to figure out the *purpose* of the rule he is following and how to achieve it.

3. The pupil’s ‘own rules’ are simply the rule-formulations that he uses to explain or justify or describe his own rule, since he is, properly speaking, the rule itself. Thus, the rule-formulation is the self-reflective form of the rule, that is, it is an external expression of the rule-follower’s *understanding* and is not to be conflated with the rule itself.

4. ‘Expertise’ here is not the kind of know-how that is usually the provenance of just the few as opposed to the many, or the specially-talented as opposed to the generally unskilled, but simply a term used to contrast those who are able to do something as opposed to those who have not yet learned to do it: thus the language-user would be considered an ‘expert’ if he has attained enough mastery of a language to be considered competent to speak, for example, an older child. A baker, on the other hand, is still an ‘expert’ in this sense, too, but the skill that is baking is not as versatile or as fundamental as the skill that is speaking, partly because baking has more limited ends than language-use, partly because baking, like most other skills, is learned after we have acquired our language-skills: hence baking is practiced by fewer human beings than language-speaking, since baking is connected to a specific need (e.g. to eat) and one baker can fulfill that need for baker and non-baker alike (e.g. others can eat what the baker produces and therefore they need not learn to bake in order to eat bakery products). The term ‘expert’ is primarily intended, then, as a contrast to ‘non-expert’ in a relation that is paralleled in the parent-child relation and the teacher-pupil relation: it is an asymmetric relation of dependence.

The analogy between language-use and currency-use is fitting, because in both cases, the instruments ‘used’ are symbolic in nature and hence gain ‘value’ or ‘meaning’ by being accepted as instruments in their respective activities: words are used in language, money is

used in commerce. This acceptance on the part of users is their *agreement* in using these instruments in a certain way in order to ‘purchase’ things of tangible and not merely of arbitrary value (e.g. money, which has a certain *arbitrary* value attached to it by convention, can then be used to buy what the baker produces, that is, something that can be eaten and thus has *intrinsic* value for the buyer, just as words can be used to ask him what price are his goods and how many of them he wants). This mutual agreement among language-users and currency-users is necessary in these cases for both systems to work as systems of exchange; consequently, such agreement is necessary only for the purposes of *exchange* and not for the *use of signs*. Thus, in mathematics, if I am to be able to finish your calculations for you, or correct them, we must both agree in our usage of our mathematical symbols; but this does *not* mean that we must *both agree* if I am to finish my own calculations or correct my own errors in some arbitrary mathematical system, as though such a system itself required agreement in order to be said to be rule-governed. That you *could* learn this system, too, does not make agreement any more necessary for using this system, unless, of course, you did in fact learn it and we both interacted using the same system. Agreement, therefore, has nothing to do with the *normativity* of rules, unless agreement also plays a part in the usefulness of these rules, as it does in language-usage, currency-usage, and our mathematics.

5. Medina (p. 163): “The teacher is invested with a special authority by virtue of her competence in the practice at issue. Her mastery of the rule-following practice makes her a representative of the community of rule-followers; and, as such, the teacher has the capacity and authority to bring the behavior of the novice into harmony with the behavior of the rule-following community. The goal of the training is to bring the pupil into the practice, and this is achieved by effecting an ‘agreement in actions’ between the pupil and the teacher (RFM VI.45) and hence, by the same token, between the pupil and the community of practitioners.”

But why should the goal of the training necessitate effecting an agreement between the pupil and the community? Why not just between the pupil and the teacher? Indeed, why should the *goal* of the training be to effect an agreement—or a ‘consensus of action’—of any kind, as Medina later says on page 165? (“However, the creation of these [conditioned behavioral] patterns is only one step toward the ‘consensus of action’ that is the goal of the training.”) Isn’t it enough for the pupil simply to know how to follow the rule, that is, to be able to do something with it on his own? ‘Effecting agreement’ is really just the pupil *imitating* his teacher, and this imitation is just the way in which the pupil becomes like his teacher, that is, becomes the rule himself; such ‘agreement’ is not the *goal* of his training—it *is* his training, since, by imitating his teacher, he already *is* ‘agreeing’ with him. But such agreement with the community only seems to be the goal of his training because we have the teacher acting as the middle term in a relation between the pupil and the community, and so the purpose of the teacher being present, and hence the *raison d’être* of the training, is merely to effect this agreement. It is mistakenly thought that community agreement is a necessary part of the rule’s normative nature, since in the community view, normativity requires such agreement in the form of compliance with the rule. Medina’s ‘contextualized’ view of the rule merely masks this demand for compliance by making community agreement into a decorative background or ‘horizon of understanding’, that is, into something functionally useless in determining such compliance for the rule-follower but somehow still necessary for normativity to make sense; the ‘context’ merely supplants the compliant rule now in disguised form. But then saying that such agreement, while necessary, only ‘constrains but does not determine what

is correct and incorrect' is like saying that the air surrounding a building does not in fact hold up the building as a foundation would; nevertheless, it is necessary to keep the building from falling over—after all, no building can exist in a vacuum! But the *purpose* of the rule by itself establishes the normativity of the rule and is its foundation, while the rule-follower's grasp of this purpose is in fact his 'normative attitude' towards it—and this grasping of the rule's purpose, along with the means to achieve it, is also the *goal of his training*, that is, the goal of his *imitating* his teacher, which ends in his becoming the rule himself. Community agreement determines no part of this, and is thus unnecessary *for determining* what is correct and incorrect in rule-following. Medina has unwittingly undermined the community view by acknowledging that this community cannot determine what is correct or incorrect for the rule-follower; but he nevertheless supports the compliant picture of the rule by affirming at the same time both claims of the paradox of compliance: the community is unnecessary (it does not *determine* normative behaviour), but it *isn't* unnecessary (it *constrains* normative behaviour).

The problem with Medina's understanding of the role of the community—his claim that a consensus of action only 'constrains but does not determine' our normative behaviour—is that he only focuses on the rule as guide but ignores the rule as measure. Yet focusing only on the rule as guide leads directly to the problem of interpretation—that is, how am I to know that *this action* is a continuation of the rule that I learned?—and this problem is supposedly solved in PI 201 by establishing *correct and incorrect* cases of rule-following using "what we call 'following the rule' and 'going against it' from case to case". Here normative criteria are fixed by what *we call* 'correct' and 'incorrect' cases of rule-following; but if these actions-as-criteria are only taken as a *guide* and not as a *measure* of rule-following, then Wittgenstein has not really solved the problem of interpretation, since we *also* learned from our teacher through our training what are called 'correct' and 'incorrect' cases of rule-following. Yet our training by itself cannot overcome the problem of interpretation on Wittgenstein's account; training alone only leads to it, since, even on Medina's own account, community agreement leaves matters 'open-ended'.

Paradoxically, Medina says that "the normative standards for the application of a rule are exhibited in what we (competent rule followers) regard as 'obeying the rule' and 'going against it' in actual cases" (178) and that "a regular way of going on....functions as the standard for judging the correctness of one's actions" (177). But nothing 'functions as the standard for judging the correctness of one's actions' without also *determining for one* what is correct and incorrect, and so what 'we regard' as correct and incorrect cannot merely constrain us—the rule as guide can do *that* by itself—but must also determine our normative behaviour, that is, *judge* whether we have taken proper guidance from the rule. That indeed is what a *standard* does—it is an independent *measure* by which we judge compliance. Widening the 'context' only widens this standard—it does not do away with it. Medina's incoherent view is actually the result of his focusing on the importance of our training and learning, and he draws his conclusion thus: "So, on Wittgenstein's view, the actual practices of ordinary language in which we are trained *constrain but do not determine* our ways of thinking and talking" (187). Here, and elsewhere, he explicitly connects his view of community agreement to our training, that is, to the rule as guide, not as measure, without realizing that training alone, or 'background practices', provides no help in determining the correctness of our actions within the compliant picture of the rule, since the goal of our training, according to Medina's Wittgenstein, is merely to harmonize our action with that of the community's; and so, for the

sake of this compliance, our training can never really come to an end if the community's standards keep changing, and this is just to admit that training alone cannot make the rule-follower truly independent with genuine 'mastery of a technique'; moreover, the rule-follower cannot then judge *his own* compliance with the community, no matter what his training, since, if he could do that, what further need would he have of the community, except as a false crutch? (Just because the rule-follower is able to say "this is what *we* call..." doesn't show his compliance either, since he has merely been taught to say such as a justification of his action, precisely because *there is no independent criterion to which he can directly appeal*—there is no such 'we', no compliant rule, unless this compliant rule is the sign itself, in which case 'compliance' is merely the grammatical projection of a power that the sign does not have; the sign, however, is dispensable in rule-following.) But it would indeed be odd if a teacher could only guide us but never measure us, that is, correct us as we tried to follow him; yet this is exactly what we are told the 'indispensable' role of the community amounts to for us—that it can only guide us. To embrace Medina's weakened version of the community view, by acknowledging the community's uselessness for determining compliance for the rule-follower, while still retaining some use for the community (as a 'constraining' guide, that is), is really to embrace the *paradox of compliance* all the more, which is to say, to drive the incoherence of the compliant view of the rule even deeper into the morass.

## Chapter Three: Following a Rule ‘Privately’

### 3.1 The Difference between Rule-Following with and without Signs

What is interesting about what Wittgenstein says in PI 201 is not only that there is a way of grasping a rule which is *not* an interpretation but that there need not be *any* interpretation in ‘following the rule’, that is, there need be no formal expression of the rule at all, no symbolic representation or sign involved in those rule-governed activities whose rules are not the employment of symbolic representations or signs. Baking, for example, can be learned without learning or using any symbolic representations, whether in natural language or some other system of signs. Since PI 201 showed that any rule-formulation itself needs to be first fixed to some action by an act of definition or stipulation, otherwise the rule-formulation would be ‘dead’, indeed not even a rule-formulation, what this shows in addition is that the use of a rule-formulation is itself really just an intermediary action between an action that stipulates its use and another such action with which it now agrees. But the employment of this intermediary action or sign can be dispensed with, since rule-following only requires that an action agree with another action to maintain its proper ‘shape’, and the rule-formulation here only *depicts* such an action. With language itself, however, the use of the sign just is the activity called ‘language’, and the sign cannot therefore be dispensed with qua mere use—only its self-reflective form can be, that is, the definition we wish to give the sign formally, since this is the explicitly abstract form of the sign with which one reflects; but this shows why language itself is a poor example with which to generalize about rule-following, since we cannot, as it were,

entirely dispense with the sign within language: language, as we said, *just is* the employment of a system of signs.

In the case of baking, we can see this point more clearly. No one would say, for example, that to know how to use the word ‘baking’ implies that one thereby knows how to bake; yet it may be tempting to think, conversely, that to know how to bake implies that one also knows what *we call* ‘baking’, as if an action and the sign that depicts that action somehow amounted to the same thing. The expression “this is what we call ‘baking’” can be used while teaching someone how to bake, but it is not necessary; and if it is used here, one is really learning two different things at the same time, namely, what *is* baking and what *is called* ‘baking’: the one need involve no signs (rule-formulations, interpretations) while the other clearly does.

Obviously, where signs are neither needed nor used to teach a rule-governed activity, the problem of interpretation will not arise as it is expressed in PI 201. It is therefore of fundamental importance to distinguish between rule-following with signs and rule-following without signs. In PI 201 and elsewhere, Wittgenstein does not make this distinction; and so what needs to be first and foremost defined for us in rule-following, according to Wittgenstein, is just the rule-formulation, using an action that ‘*we all call* such-and-such’. But when one simply does what another does, that is, when one imitates another’s action, nothing is ‘defined’ for us by this act, since no sign is involved, but rather an innate, hence unlearned, ability is exercised, namely, our natural capacity to imitate another’s action and grasp its purpose. In the classroom we learn language in part this way: we say what the teacher says, we write what the teacher writes, and *these* are the actions we imitate. If there are any interpretations here, they are really just attempts to imitate such actions.

In the case of rules involving the use of signs, the rule is overdetermined, as Wittgenstein says<sup>2</sup> (RFM VI-16), in the sense that the result is a criterion for following the rule and is not something to be determined by what follows in experience. This is why he says that a mathematical proposition is not to be treated like an empirical proposition, since the result is not something discovered in experience but stipulated through a sign from the beginning. But why must the rule be stipulated from the beginning? Well, perhaps we ought to say that it guarantees the sameness of the result from the sameness of the operation or ‘following the rule’. Baking, for example, gets the same results from the sameness of procedure and ingredients, but this result will be determined from certain facts of nature; hence the rules of baking are to this extent accountable to reality. Mathematics, on the other hand, is fundamentally different, because its rules are not accountable to reality (or derived from reality) in the same way as the rules of baking are. Its results, which are just more signs, need to be stipulated beforehand as what counts as following these rules, as what is the same in following these rules. But how arbitrary are these rules? What is arbitrary and what is not in the sign ‘ $2 + 3 = 5$ ’? Well, that will depend in part, of course, on what we do with this sign, on how we use this sign in our other activities; but these questions just point us back to an activity (not just as stage-setting but as an activity in its own right) that can be done *without these signs*: something quite elementary, no doubt, something like pairing off items, or grouping things into sets of three or four, perhaps by row and column. Could we not say, then, that the rules of mathematics are not arbitrary in just this sense: that we could not do what we actually do with mathematics if its rules had not been grounded in those simpler activities and later defined by them via stipulation? For we could have stipulated any sign to follow ‘ $2 + 3 =$ ’, but whatever was stipulated in that sign had to make contact with a practice *already established*

*without stipulation.* And *that* means that we also learn all the expressions of following the rule in a similar way, that is to say, ‘sameness’, ‘agreement’, ‘rule’, etc., presuppose an activity that can be done without signs at all, linguistic or otherwise.

### **3.2 The Use and Usefulness of the Sign: PI 258, PI 260, and PI 270**

If we can just imitate the mere use of signs, however, would that by itself give the sign a meaning? After all, it seems that imitation alone cannot be enough—to merely write down a sign surely cannot mean anything by itself; there must be something more than just the sign qua mere use. There *is* in fact more. When we follow the rules of baking without signs or rule-formulations, for example, we don’t ‘use’ any rule-formulations at all, but simply act using the equipment of baking. But with language the sign is used in a certain way and this use we identify with the rule itself, and call it ‘the meaning of the sign’. Such identification may also be called the over-determination of the rule, since how we have used the sign has itself become a criterion of using the sign correctly, that is, a criterion of what the sign means. And what this means is that the use of the sign now has a function such that this use has a utility, like the equipment of baking, in doing *something else*. Consequently, when we use a sign, we are also doing something else with it. A sign with a use becomes useful for something other than its mere use as a sign. One might say that the way a sign has been used gives it the shape necessary to do something else with it, just as a mixing bowl must first be manufactured into a certain shape in order to function the way it needs to in the activity called ‘baking’. The sign ‘S’ in PI 258 has a use in the mere sense that when I have a particular sensation, I then write ‘S’ in my diary<sup>1</sup>. But ‘S’ here has no function yet, except perhaps as the name of a sensation,

or the noting of its occurrence; and so it really has no meaning, or rather could mean anything, since there is yet no answer to the question ‘And what of it?’

This contrast between the use and usefulness of the sign may be seen in the orthography of English. The alphabet is a system of signs called ‘letters’ which are supposed to stand for sounds in the language. Consequently, the rules for spelling words in English had the function of allowing readers to pronounce the words they read, and so made recognition of these words easier, that is, recognition of their spoken form. But over time, many words changed in pronunciation, although the written form did not. Thus, ‘gh’ in words like ‘laugh’ or ‘enough’ is pronounced as ‘f’, and in words like ‘ghost’ or ‘ghastly’ is pronounced as ‘g’. What, then, has become of English orthography? Do the letters of English stand for anything anymore or only some do? Have the rules of some merely changed, or are there no rules at all for these? Are we continuing to practice outdated rules, or were these always just pseudo-rules of orthography, unconnected to sounds? Well, there *is* a connection between letters and sounds in English and the written form of many older words have become fossilized, showing how they were pronounced at one time but no longer are today; and by continuing to spell these words in the same way, we are continuing to follow such rules, even though ‘k’ in ‘knight’ is no longer pronounced at all in this language. The rule for spelling ‘k’ in this case is still used but has lost its function as a rule for pronunciation, and so is really phonetically pointless. The spelling for many English words no longer makes sense to us, except in historical context, and so without this function, the rule for pronouncing the letter ‘k’ in ‘knight’ has become defunct in modern English. In short, the rule no longer exists as a rule for anything, but the practice of using it continues, despite its uselessness for pronunciation. Here is where use and meaning can come apart, that is, where practice and function come apart. (Think of the vestigial structures of

body parts in evolutionary terms). It is not just how a sign is used, but also *for what* it is used that determines the meaning of a rule in these cases. And this ‘for what’ or ‘why’ is not of secondary importance but foundational for the establishment of the rule.

The importance of recognizing the use and usefulness of the sign comes out in Wittgenstein’s discussion of so-called ‘private ostensive definition’ in PI 258 and beyond. This discussion is thought to be related to the question raised in PI 202 concerning the possibility of following a rule ‘privately’ and would thus indirectly support the compliant view of the rule by making the rule an independent criterion of correctness for the rule-follower. But the last sentence of PI 258 shows that the case of the sensation diarist discussed in this section is different from the case discussed in PI 202: in PI 258 it makes no sense even to talk of ‘right’; in other words, it is not open for discussion whether I am ‘right’ or ‘wrong’ in using the sign ‘S’. The lack of a criterion of correctness for using this sign only means that we do not yet have the circumstances in which it even makes sense to ask for a criterion of correctness. As PI 260 later informs us, the private diarist is not even making a note of anything, since “a note has a function, and this ‘S’ so far has none.”<sup>2</sup> The interlocutor responds in PI 258 to the lack of a criterion by saying what his criterion is: “whatever is going to seem right to me is right.” And *that* response alone presumably shows why using the word ‘right’ makes no sense. Wittgenstein’s point seems to be that it is not yet appropriate in these circumstances to respond to a lack of a criterion of correctness by looking for one or offering one, but rather to see that it makes no sense to look for one or to offer one. PI 270 will show us that once a use is found for ‘S’, it doesn’t in fact matter if one has recognized the connection correctly or not, that is, the connection purportedly made in PI 258 between writing the sign ‘S’ and my memory of associating this sign with the sensation in question.<sup>3</sup> Consequently, the

attempt in PI 258 to set up a rule or technique for using ‘S’ without the use of ‘S’ itself having a use is “mere show”, just as the lack of an independent criterion of correctness for ‘S’ no longer actually matters.

This case is then obviously different from that in PI 202, since PI 202 shows us more than just the uselessness of giving a criterion of correctness for the use of a rule when the rule itself has no usefulness—it simply affirms that there *must be* such an independent criterion of correctness in rule-following, without which following the rule would collapse into whatever one thinks is following the rule; that is to say, if there is in fact no difference between thinking one is following a rule and following a rule, then one may indeed conclude that there is no criterion for following a rule other than thinking one is doing so, and this would be the same thing as following a rule ‘privately’, since there is now nothing independent of the rule-follower’s thinking by which one could justify one’s action under the rule (cf. PI 265). Consequently, denying the difference between thinking one is following a rule and following a rule appears upsetting to us precisely because it makes the rule in rule-following useless for us in the compliant sense of the rule, and hence destroys the normativity of the rule. But that consequence alone—the undermining of the compliant picture of the rule—need not compel us to affirm this difference, and so seemingly preserve this compliant sense of the rule, since we can still deny this difference matters in the compliant sense of the rule while affirming this difference in *another* sense: thinking one is following a rule indeed still differs from following the rule, but on the straightforward grounds that the latter is an action with consequences for the rule-follower and the former is not, just as the former is really the counterpart of the sign and the latter is not. In other words, PI 258 and PI 270 can be taken together as an argument that *trivializes* the conclusion of PI 202: if saying ‘what seems right to me is right’ means I no

longer have any right to speak of ‘right’ (that is, nothing *follows* from what *seems* right to me) and yet the sign ‘S’ can *still become meaningful* once the sensation called ‘S’ is found to be useful, then this shows that the difference affirmed in PI 202, namely, that between thinking one is following a rule and following a rule, is real but trivial and is not really an argument in support of a compliant picture of the rule. Why? Because it was thought in PI 202 that the rule itself, or a community of rule-followers forming a practice, had to be in some important sense an essential part of my action in following the rule, namely, as something *independent* of the rule-follower by which compliance can be determined, and without this rule or community agreement, my being said to follow the rule would fall apart through incoherence: the sameness of the rule with the rule-follower would undermine the normativity of the rule. But PI 270 then showed us, contrary to Wittgenstein’s intention, that it no more follows that such would happen by following a rule ‘privately’ than that the sign ‘S’ would be meaningless without the independent criterion of correctness sought in PI 258<sup>1</sup>. The usefulness of an action therefore has priority over whatever we say or think about this action, and this makes what we think inconsequential. The only thing undermined by denying the importance of the difference in PI 202 would be the compliant picture of the rule—a picture that is at best paradoxical.

### **3.3 The Arbitrariness and Autonomy of the Sign**

We can perhaps now see why the sign in rule-following is autonomous, just as Wittgenstein insisted it was, since the use and usefulness of the sign can come apart, making use autonomous from its usefulness; but we can also see at the same time why he was misled in rule-following by focusing exclusively on rule-following with signs, since the sign in rule-

following is really dispensable for rule-following and the exclusive analysis of it distorts our overall view of the nature of rule-following, such that we end up seeing it as analogous to obeying an order, where the rule itself, like an order, is something with which the rule-follower complies. Wittgenstein's comparison of grammar with cooking is a good illustration of this point (cf. Zettel 320). Following Wittgenstein, it was thought that rule-following without signs, like baking or cooking, must be fundamentally different from rule-following with signs, like language or games, because the one is defined by its end while the other is not: hence the arbitrariness and autonomy of the rules of language or games. But the fundamental difference spoken of here only relates to the difference between the use and usefulness of the sign; that is, in rule-following with signs, the mere use of the sign itself becomes a tool for action in another activity that is defined by its end; and this is really to say that the mere use of the sign is in fact the only way in which we can make an action a tool by itself, one which, as a tool, *becomes useful for something else*. Hence this action-as-tool has, qua mere action and not qua tool, no definable purpose outside its mere use to give it its 'shape'—maintaining this abstract 'shape', or mere use, would thus be its own purpose in a sense--and this is the reason why the sign is both arbitrary and autonomous *and hence useful*. Wittgenstein no doubt looked at rule-following, that is, rule-following *with signs*, as analogous to obeying an order, because in obeying an order, the action of the order-obeyer is also done without thought to its purpose, since the purpose of the order-obeyer is merely to execute the order and nothing else besides—and he can only do that if he can act independently of whatever consequences for him follow directly from his action; he has simply transferred his action's consequences *for himself* to its consequences *for another*: his action is now a *conditioned form of independence*, like that characterizing the slave in Hegel's master-slave dialectic, since the independence the slave has

attained for himself is nothing other than his ability to act against his own inclinations and self-interests and for those of another. In other words, the natural unity of action and purpose within the order-obeyer has been sundered, just as practice and function came apart in the orthography of certain words in English; but the dismemberment of the obeyer's actions from his own purposes under this unity is not to no purpose, because his actions now still have life by being freely dispersed and grafted onto the purposes of another who unites both in a higher unity, for example, in the institution of slavery. It is in fact the slave who must reply, when asked to justify his action, with "This is simply what I do", since *his* purpose is now simply to act, that is to say, to do an action *for its own sake*, albeit for the sake of another; the action qua action is everything for the slave, just as the master's practical use of the slave qua slave is everything for the master. Thus the action of the order-obeyer is in a sense arbitrary and autonomous, too, provided that he focuses only on the execution of the order and not on the *consequences* of his action under the order. The one who obeys an order is indeed a tool for action for another, namely, for the one who gives the order; yet at the same time, this order-obeyer would not himself be *useful to the order-giver* unless he could function as a tool in *this* sense of being autonomous or independent.

The action-as-tool, or sign-usage, is equally the material out of which a practice is moulded into a certain shape. Just as the nature of iron is not defined by the uses made of it but is rather chosen in manufacturing *for its own particular nature*, the mere use of the sign, too, is not directly defined by the uses made of it; but this mere use of the sign itself won't be *meaningful unless this use is made useful*, since the mere use of a sign has no power to confer meaning; this also means, therefore, that, although the sign's meaning, or identity, is thereby overdetermined through its mere use, such overdetermination is still useless unless the sign

also becomes useful at the same time. After all, we can still spell ‘knight’ with a ‘k’ and simply continue our previous usage to determine this spelling as correct, but this continuation of practice actually makes no real difference for correct spelling, since spelling now has no connection in this case to how the word is pronounced based on its spelling; overdetermination of the sign may itself be arbitrary, but still only in a meaningless way, because whether we actually spell ‘knight’ with a ‘k’ or not no longer really matters—it is now ‘correct’ merely for the sake of form; it is a formal correctness lacking any real content.

We have already seen, for example, that what is overdetermined in a stipulation is the result, but that this result is in fact not a *consequence of an action* but is part of the action itself, that is, it is still an action *with no consequences* for the one who acts. Thus action and result in a stipulation are not united as action is to purpose but are *simply identified through abstraction*, since such a result is neither a purpose nor the fulfillment of a purpose, but is now just part of the sign itself; this abstract action thus *becomes* its own result, its own inconsequential consequence. But action and result can only be united in this way, that is, simply identified in a stipulation through abstraction, within a practice that is already established without such a stipulation, and *that* practice will already have consequences for the rule-follower. Moreover, the sameness of an action that is overdetermined is likewise still dependent on an activity defined by its end and already established without signs, since this abstract sameness in the sign is really just the difference in action that makes no difference, that is, these actions make no difference to the outcome of the action-as-tool; and so the sameness of an action in a rule-governed activity involving signs is determined as the same, not just through the overdetermination of a rule, or through what we call the ‘same’, but through the fact that this stipulated result, whatever it is, actually makes no difference to the outcome of the action-as-

tool, that is to say once again, to its usefulness as an action in another activity: *sameness here is really just the difference that makes no difference*. It is therefore a mistake to see grammar's arbitrariness and autonomy as exclusive of purposes or ends, since sign-usage is really just an action expressing this duality of action—a mere use that is at the same time an action that is actually useful for something other than its mere use<sup>2</sup>.

### 3.4 The Generality of the Sign

The generality of the sign, its 'universal' nature, is directly related to the usefulness of the sign, since the natural unity of an action with its purpose, manifested in nature in so-called 'desire', after being sundered in the case of the sign, is replaced *in the sign* with the arbitrarily determined but essential unity of the use and usefulness of the sign; and this usefulness now becomes the portal through which the action re-establishes its normative connection to some purpose: hence the usefulness of the sign gives the sign its portability throughout a multiplicity of activities defined by different ends. This portability of the sign *just is* the sign's generality or universality; the sign, by becoming useful, thus indirectly comes into contact with purpose, even possible future purposes, as any tool would through its nature qua tool. The universality of the sign, therefore, is not proof of the sign's communal nature but of its instrumental nature; it is the bridge between an arbitrary action (or mere use) and a definite end; and this universality is tied to a community of rule-followers only in the sense that what is portable by nature is also sharable with others, and hence teachable to others. By being made portable in this way, the sign becomes abstract; hence the sign itself is just an *abstract action*.

The portability of any particular sign, however, is merely *externally* connected to any definite end, though the portability of the sign itself is *internally* connected to purpose through its usefulness. The usefulness of the sign, in other words, has the nature of a ‘third thing’ or intermediary; it is the *virtual purpose* of an action that has been elevated to the level of abstraction, since a virtual purpose *is simply the shape of an abstract action or mere use functioning as though it were itself the purpose of that action, even though this action, this virtual purpose, is really just a mere action.* Indeed, a virtual purpose is not really a ‘third thing’, strictly speaking, but simply an action, abstracted from any definite purpose, that can then be used as part of another action, a concrete one with a definite purpose, as though the two actions were really the same action, that is to say, as though the purpose of the abstract action were the purpose of the concrete one. The portability of the sign, based on its usefulness, thus provides the generality that is necessary in order to give justifications of various actions, that is, *to make universal judgments about these actions.* The sign is therefore used as a tool for showing compliance, even though giving a justification of an action is not true compliance; it only seems so because such judgments appear to be independent of the rule-follower himself—but this ‘independence’ of the sign is only a dependent form of independence, like that of the slave.

### **3.5 The Normativity of the Rule and Community Agreement**

The sign or rule-formulation can thus function as a rule only because it has the portability that makes it independent of any activity in which it happens to be found. In the compliant picture of the rule, normativity is established by the rule that is *different* from the rule-follower,

that is, by the rule that is *independent* of the rule follower. But the sign acts as just such a rule, or rather, it *projects itself* as such a rule. This rule is what the rule-follower appeals to when he is asked to justify his action. In the community view of rule-following, this role of the rule is taken over by the community in the form of ‘community agreement’, or a ‘consensus of action’ between rule-followers. This is why it is argued that a community of rule-followers is necessary for rule-following, since a single rule-follower cannot make the distinction for himself between thinking he is following a rule and following a rule; and without this distinction, the normativity of the rule would collapse, leaving *no rule at all*: whatever seemed right to the rule-follower would then be right (PI 258). It is impossible, therefore, to follow a rule ‘privately’ (PI 202). There must still be a rule that is different from, that is, independent of, the rule-follower; compliance demands that it be so.

The advocates of the community view of rule-following nevertheless have two different roles for community agreement in explaining the normativity of the rule. The most demanding is the role just outlined above, where ‘correct’ and ‘incorrect’ are determined by the agreement or disagreement of an action with that of a community of rule-followers. In other words, an action is said to be ‘correct’ on this account if it *agrees* with this consensus, ‘incorrect’ if it does not. This sense of ‘community agreement’ is straightforwardly compliant in nature, for it says that an action is ‘correct’ only if it complies with another action sanctioned as a standard. The community view cannot, however, explain further the normativity of this consensus of action, this action sanctioned as a standard—it simply exists as what we all do. Hence the second role of community agreement in rule-following: it is a *quiet* agreement in action that functions as a *precondition* of anyone following a rule; it is to acknowledge that we all just go on in this same way.

It is necessary in the community view to have these two roles for community agreement, since they both together fulfill the function of the rule set out in the compliant picture of the rule. The rule-follower's dependence on the rule, for example, demands that community agreement make the rule-follower dependent on the community. It does this by looking at the rule-follower as *part of a larger whole*, and thus as an individual essentially isolated and 'private' when looked at by himself; this individual thereby becomes something necessarily 'social', that is, he becomes an *abstraction*, like the slave qua slave. The rule's independence from the rule-follower, on the other hand, makes this same community of rule-followers exempt from such dependence, even though this community itself is made up of these same individuals; it is argued that these collective actions themselves cannot be treated in the same way; they are a *precondition* of normativity, so-called 'background practices'; these actions are what we call 'following the rule' and 'going against it' from case to case (PI 201).

But this 'consensus of action' must now take the form of a sign or rule-formulation, as a *description* of "what we all do when we follow the rule", that is, *community agreement is now also made into an abstraction*. Advocates of the community view are always quick to point out that the agreement they are speaking of is not an agreement in opinions or words but in actions. So be it. It should nevertheless be obvious that if the rule-follower's action is *not* guided or measured by another rule-follower's action, as it actually *is* in the case of teacher and pupil, then it can only be guided or measured by this 'consensus of action', that is, by an *abstract action*. And the only abstract action is an action involving signs—our *saying* what 'following the rule' is. The paradox here is then: this consensus of action is a consensus *only because someone says that it is so*—someone who *already knows* how to follow the rule (cf. PI 224). For the pupil, this someone is his teacher, who *uses this abstract action*—the sign or

rule-formulation—as *an instrument to teach the pupil how to follow the rule using signs*. This ‘consensus of action’, then, is really such a sign, even though we think that it is something *beyond* the teacher himself, something independent of the teacher that the teacher brings the pupil in contact with. But what the teacher is really doing by using these signs—by making references to a ‘context’, a ‘consensus of action’, a ‘rule’—is teaching the pupil at the same time *how to think*, that is, *think abstractly about the rule*, and so *make judgments or give justifications* in rule-following. The abstract difference between rule and rule-follower in the compliant picture of the rule has instead been made here into the abstract difference between community and individual. The amended formulation of this abstract difference is now expressed in the community view by saying that *the rule-follower follows a rule only if he does what we all do in following the rule*.

### **3.6 Kripke, Malcolm, and Medina on Normativity and Community Agreement**

Despite the differences of exposition within the community view of rule-following, the advocates of this view all accept the compliant picture of the rule. And in some form or other, Wittgenstein’s claim in PI 202, that one cannot follow a rule ‘privately’, plays an important part in reaffirming this picture of the rule. In Kripke’s case, for example, PI 201 and PI 202 are taken together as the summation of the so-called ‘private language argument’ (Kripke, 3); for Malcolm, “[a] consideration of fundamental importance, which [Wittgenstein] brings out in PI 202, is the necessary distinction between one’s following a rule, and one’s *thinking* one is following a rule” (Malcolm 2, 28). Similarly, Bloor, after citing PI 202, says: “To believe in a rule is one thing; for there really to be a rule is another. What is it for there to *be* a rule?”

Wittgenstein's answer was clear. The authentic, independent or 'extrinsic', existence of a rule—independent or extrinsic, that is, to any individual consciousness—consists in its being a social institution" (Bloor, 55). And Medina, too, pays tribute to the conclusion of PI 202 by laying out its implications for normativity:

It is important to notice that the claim that "obeying a rule' is a practice" serves to convey two distinct points: first, that to follow a rule is not something we can do *only in thought*, that it must be *public*; and second, that to follow a rule is not something we can do *only in private*, that it must be *social*. The public *and* social character of rule following is substantiated with two different arguments: the Manifestation Argument and the Private Language Argument. These arguments support and elaborate the conclusions of the Paradox of Interpretation, emphasizing the crucial dependence of the normativity of our actions on our actual and shared practices of rule following. (Medina, 178)

For all these advocates of the community view, normativity is necessarily tied to the rule that is different from, and hence independent of, the rule-follower. It is thought that PI 202 makes this point clear by showing the important difference between thinking one is following the rule and following the rule. Like Wittgenstein, they have assumed all along the compliant picture of the rule, only in this case, the rule now takes on the form of community agreement.

Let us then briefly look at the way Kripke brings out this picture of compliance in his account of rule-following. He does this by elaborating the paradox of interpretation as outlined in PI 201 in the form of a sceptical problem, ultimately solving this paradox through community agreement, that is, by having the rule-follower comply with the rule. In his understanding of the paradox of interpretation, there is *no fact* about the rule-follower that establishes whether his present action conforms with his past usage of a sign; hence there is doubt whether he meant anything at all by this sign; meaning cannot gain a foothold here if such a fact cannot be found. What Kripke is really saying, then, is that such a fact would be necessary in order to establish an objective standard with which the rule-follower's action

could be said to comply or not. Searching his own mind for such a ‘fact’, for example, will get the rule-follower nowhere, since “[w]hat can there be in my mind that I make use of when I act in the future? It seems that the entire idea of meaning vanishes into thin air” (Kripke, 22).

Lacking such a standard, that is, such a ‘fact’ about the rule-follower, there can be no normativity, and hence no meaning, for the rule. What is in jeopardy here is not so much the normativity of the rule, however, as the compliant picture of the rule. Kripke’s solution simply finds *another* form of compliance in which the normativity of the rule can be preserved.

A sceptical problem is posed, and a sceptical solution to that problem is given. The solution turns on the idea that each person who claims to be following a rule can be checked by others. Others in the community can check whether the putative rule follower is or is not giving particular responses that they endorse, that agree with their own. The way they check this is, in general, a primitive part of the language game. (Kripke, 101)

Kripke can now connect his solution to the paradox of interpretation with Wittgenstein’s ‘private language argument’ by bringing in a community of rule-followers.

What is really denied [in the private language argument] is what might be called the ‘private model’ of rule following, that the notion of a person following a given rule is to be analyzed simply in terms of facts about the rule follower and the rule follower alone, without reference to his membership in a wider community.

[...]

Our community can assert of any individual that he follows a rule if he passes the tests for rule following applied to any member of the community. (Kripke, 109-110)

Kripke has indeed changed his picture of the rule in his sceptical solution, but only the superficial rendering of this picture. The deeper compliant picture of the rule has remained intact—that form of the picture determining how the community view *really* thinks about the rule. But did these superficial changes to the picture make any real difference to the debate over the social nature of the rule?

Malcolm, for example, thinks that Kripke's analysis of Wittgenstein is wrongheaded, even though he ultimately agrees with Kripke about the social nature of rule-following. In other words, what Kripke got right in his account is his *retention* of the compliant picture of the rule, the view in which the rule is merely substituted by community agreement. With Malcolm, however, we first begin to see the compliant picture of the rule fall apart. In the compliant picture, for instance, the rule is conceived as both a guide and a measure for the rule-follower, since the rule has these dual uses, which go hand in hand with the independent nature of the rule. In Malcolm's account, the rule as guide is attacked, while the rule as measure is vigorously defended. But what is attacked in the notion of the rule as guide is unfortunately only the concept of the rule in which the meaning of the rule *in our own mind* guides us. In this case, being guided by the rule is treated as though it were merely our *thinking* we are following the rule. What Malcolm actually stumbled upon in his criticism of the rule as guide is the *uselessness* of the compliant rule, that is, the uselessness of the sign or rule-formulation to guide us further in following the rule; he shows here a dim awareness that the compliant rule is really not a rule at all, but is somehow connected to our *thinking about the rule*. When he dismisses what we think we are doing in following the rule, he is really just dismissing the role of this sign or rule-formulation in rule-following; he failed to comprehend that the rule *as measure* suffers a similar fate, since it, too, only takes the form of a sign or rule-formulation, albeit as 'community agreement' or a 'consensus of action'.

Malcolm's defense of the rule as measure shows just how powerful the compliant picture of the rule is in the community view of rule-following. He is blind to the fact that the only measure of the rule-follower's action available to the community view—'community agreement'—exists only as the same sign he attacked as useless in guiding the rule-follower in

following the rule, that is, it now exists only as a *description* of what *we all do* in following the rule: this measure is just an abstract action, that is, a mere sign. One can describe what we all do in following the rule as much as one likes—this ‘community agreement’ will still exist only as an extended form of a sign or rule-formulation, expressed by someone who already knows how to follow the rule, thus manifesting *his* expertise. But Malcolm does not see any of this.

When Wittgenstein says that following the rule is a *practice*, I think he means that a person’s actions cannot be in accord with a rule unless they are in conformity with a common way of acting that is displayed in the behaviour of nearly everyone who has had the same training. This means that the concept of following a rule implies the concept of a *community* of rule-followers. When Wittgenstein says, also in PI 202, that one cannot follow a rule ‘privately’, I think he means that the actions of a single individual, whether these actions are private or public, cannot *fix the meaning* of a rule. Wittgenstein relies, in PI 202, on a simple but powerful point, namely that the concept of a *rule* implies a distinction between *following* a rule and *believing* one is following a rule. A person can believe that he is following a rule and be *wrong*.

This distinction could *not* be made if the actions of a single individual could *fix the meaning* of a rule. The meaning of the rule would be exhibited in *whatever* this person did. If it *seemed to him* that he was ‘doing the same thing’ in following the rule as he had done previously, then he *would be* doing the same thing.

We easily see that this is an absurd consequence. This shows that the actions of a single individual, who believes that he is following a rule, must be *measured against* something other than his own actions. (Malcolm 1, 156).

But the only thing this distinction shows is that there is a difference between rule-following *with signs* and rule-following *without signs*, and that the sign is dispensable in rule-following, just as what we think in relation to the rule is dispensable. We can express this point similarly, by saying that thinking about the rule and following the rule are two *different but related* activities—and what, or whether, one thinks about the rule is irrelevant to following the rule. Malcolm only insists on this distinction because *he thinks* that it reaffirms the compliant picture of the rule. He is only half right: it simply reaffirms the *uselessness* of the compliant rule, and with it the uselessness of ‘community agreement’ in

determining the normativity of the rule. Since Malcolm readily concedes that the rule-follower's action is not actually measured against anyone else's action, it can only then be 'measured' against some kind of *common* action—but this 'common action' can only be an *abstract* action, that is, the sign itself. The problem for Malcolm is that he does not really understand what compliance actually is.

Medina's view suffers a similar fate. We have already seen how Medina characterized the pupil in learning to follow the rule: he needs to have an attitude of 'total submission to the authority of the master, of blind obedience to her commands' (Medina 163). Behaving in a regular way is *drilled* into the pupil, "[f]or drilling certainly involves conditioning processes that establish behavioral patterns. However, the creation of these patterns in only one step towards the 'consensus of action' that is the goal of the training" (165). "For Wittgenstein, the process of learning is completed when the learner becomes *blindly* committed to a particular way of acting and talking (cf. RFM VI.7). By the end of the training period, the course of action sanctioned as correct by the teacher is viewed by the learner as the necessary outcome of applying the rule" (166). This way of characterizing the learning of the pupil is in line with the social nature of the rule endorsed by the community view—it is essentially *coercive*. Given the view of the rule-follower as the one who acts now by "instinct", now by "automatic reactions", now by "subjective feelings and tendencies", it is not difficult to see the necessary remedy for these individualistic impulses. In order to become a member of society in good standing, the rule-follower must be reshaped by training into our own image: coercion is simply his rite of initiation. Medina himself calls this process "initiate learning" (163).

Although Medina says that, in his version of the community view, community agreement “only constrains but does not determine” the normative behaviour of the community of rule-followers, his contextualist account of normativity still fully subscribes to the compliant picture of the rule, just as it re-enforces the abstract difference between the community and the individual. It makes little difference whether you now call this rule a ‘context’, or a ‘consensus of action’, or a ‘precondition’ of following the rule. In all these metamorphoses of the compliant rule, this rule still exists only in the form of the sign or rule-formulation, as what we call ‘following the rule’ and ‘going against it’ from case to case. The concept of community agreement itself contains the concept of compliance in the community view, since the agreement discussed here is not the mutual agreement among peers but the asymmetric agreement characteristic of compliance—that is to say, the rule-follower must agree with the community if he is to be a *part* of this community, not the other way around. As we will later see, the compliant rule is not something the rule-follower *actually* complies with anyway, *since the compliant rule is not really a rule at all—it is simply the sign itself.*

### **3.7 The Misunderstanding of Compliance in Rule-following**

The community view is not alone in misunderstanding the nature and role of compliance in rule-following. Baker and Hacker, who champion the individualistic view of rule-following, also misunderstand the concept of compliance with a rule. They make the mistake of thinking that, because the *genesis* of an ability can be separated from the question of whether someone *has* that ability, rule-following does not require more than

one individual for the idea of compliance with a rule. But ‘compliance’ *only makes sense* when there is more than one individual, since compliance appears as the *imitation of another’s action* and is therefore *founded* on the relation between teacher and pupil, if not between parent and child. The expressions ‘following a rule’ and ‘an action which accords with the rule’ would not have the meaning they do save being originally forged *within that relation*. And so the *concept* of ‘following the rule’ already contains that relation in its expressions as does the *concept* of ‘compliance with a rule’; but the relation between teacher and pupil is not the final stage of rule-following; it is really only its initial stage or beginning. Baker and Hacker have therefore also misunderstood the role of community in rule-following, namely, the role of the relation between teacher and pupil in the *expressions* of rule-following. Rule-following is essentially a genealogical phenomenon, founded within a genealogical community; and it is only within *this* community that compliance makes any sense. It is also the only form of compliance that is part of the pliant picture of the rule.

### **3.8 The Misunderstanding of Internal Relations in Rule-Following**

Baker and Hacker say that Wittgenstein’s understanding of the internal relation between a rule and its application is a fundamental insight of his into the nature of rule-following; they go on to say that “[t]he rule and its application are internally related, for we define the concept ‘following this rule’ by reference to *this* result” (RGN, 149). But what Baker and Hacker fail to recognize throughout their discussion of internal relations is that there are really *two different* internal relations at work in the compliant picture of the rule,

since the compliant picture requires the maintenance of the abstract difference between rule and rule-follower. The first internal relation, therefore, concerns the *rule itself*; it is the relation between the rule and its application, or between the rule and acts that are said to be in accord with the rule. We could call this relation ‘the definition of the rule’. The second internal relation concerns the *rule-follower*; it is the relation between the rule-follower’s understanding of the rule and the action which manifests his understanding. We could call this relation ‘the manifestation of the rule’. Why must there be these two different internal relations in the compliant picture? Because compliance with a rule requires that rule and rule-follower be *externally* related and not merely internally related, since the rule must still be an *independent guide and measure* for the rule-follower. What is really happening in the compliant picture is that the rule of the first internal relation, the sign, ends up *externally related* to the action which manifests the rule-follower’s understanding of the rule, that is, the action of the second internal relation: this external relation just is compliance with the rule, since the rule-follower’s action is said to comply with this rule only if it falls under what *we call* ‘following the rule’. This external relation is made possible because of the *difference* between rule-following with signs (the first internal relation concerning criteria of correctness, that is, the definition of the rule) and rule-following without signs (the second internal relation concerning the rule-follower and his actions). The confusion for Baker and Hacker is in not seeing that the ‘acts which accord with the rule’ of the first internal relation are not really acts so much as *descriptions* of these acts, since the description of these acts constitutes the definition of the rule, and a definition, like a description, exists only in the form of a sign or rule-formulation. *That is* why a rule and its application are internally related: one sign is simply defined in terms of

another sign. (Bringing in *ostensive definition* is useless at this point, that is, signs that are defined by actions, since an ostensive definition forges the meaning of the sign and shows the dependence of the sign on an action—the reverse of compliance with a rule. The sign must *already be established or defined* for it to function as a rule in the compliant sense, and this means *the action is now described in signs so as to function as such a rule.*)

Compliance will not work here for Baker and Hacker, however, since not only is the first internal relation dependent on, not independent of, the second internal relation (the first is forged in the action of the second), but the first internal relation is itself just a *different manifestation* of the rule-follower's understanding, that is, of his expertise. This is why, in the pliant picture of the rule, both internal relations are really one and the same, in the sense that rule and rule-follower are one and the same. The sign is really just *another* manifestation of the rule-follower *being* the rule, because it is always the rule-follower alone who states the rule, even though he need not state it at all—his action by itself is sufficient for his following the rule. In other words, the sign is an *external and hence unnecessary* manifestation of the rule-follower's understanding, whereas his action is an *internal and hence necessary* manifestation of this *same* understanding. If this is so, then the sign, whether as a description or definition of the rule, cannot establish the normativity of the rule; and so the description falling under what *we call* 'following the rule' is a *false* measure of the rule-follower's action; it is simply a picture of the rule-follower's action that is drawn by the rule-follower himself. It will be shown later that, because the normativity of the rule is constituted by its meaning or identity, the identity of the sign that is overdetermined by its mere use actually rests upon the deeper identity called 'the purpose of the rule'. It is only this *latter* identity that establishes the normativity of the

rule, for the identity of the sign is itself forged and sustained in this deeper identity. This deeper identity is just the rule-follower himself, a *living genealogical identity*, a human being.

We can now perhaps see why Wittgenstein's followers have overrated the difference between thinking one is following a rule and following a rule when it comes to the normativity of the rule; after all, it is agreed by them that what is right in following the rule cannot simply be what the rule-follower *thinks* is right; the rule must therefore be something *independent* of what the rule-follower thinks he is doing in following the rule, that is, the rule must be independent of the rule-follower himself. But it makes *no difference* what he thinks when he follows the rule, since it is only the connection between the rule-follower's action and his purpose in acting that makes all the difference in rule-following—and *this action is in fact independent of what the rule-follower thinks, but not independent of the rule-follower*. If the rule-follower can learn to follow the rule without the use of signs, then what connection can there possibly be between his thinking that he is following the rule and his following the rule? He cannot even *say* what he is doing, much less what he is *thinking*; he cannot justify his action or appeal to anything outside his doing what he does in following the rule; and even if he were to give an answer when challenged, a proper answer would only be in the form of an action—his actually following the rule. For him to justify his own action by citing a rule-formulation would be as superfluous as appealing to community agreement; it would make *no difference* to his following the rule, since giving justifications is a different activity from the activity being justified. If it does anything, pointing out the difference between thinking one is following the rule and following the rule only makes thinking here trivial, that is, *inconsequential* in rule-

following; and this would be just another way of expressing the view that rule-following using signs is *inessential* for rule-following. In other words, thinking in relation to the rule has *no connection* to the normativity of the rule; thinking is only brought into the discussion of rule-following because of the phenomenon of the sign, as the external form thinking takes concerning the rule. But the sign, as we said, is dispensable in rule-following, just as what we think is. Hence ‘privacy’ is no issue at all in rule-following.

The community and individualist views of rule-following thus amount to the same view in the end. Not only do both subscribe to the compliant picture of the rule, both end up appealing to the sign as this same rule: the community view elevates ‘community agreement’ or ‘consensus of action’ to the status of the compliant rule; the individualist view of Baker and Hacker simply cites the ‘rule’ in the form of a description of what we call ‘following the rule’, as a justification for the way the rule-follower follows the rule. Since this sign is itself an abstraction and simply *the external form abstract thinking takes in relation to the rule*, compliance now becomes the *abstract* compliance of an *abstract individual’s action* with another’s *abstract action*. In the sign alone, therefore, resides the abstract difference between rule and rule-follower, which is at the heart of the compliant picture of the rule. This picture brings out explicitly the essence of the sign: that is, *the sign is that mere use which also becomes what it is not—it is the rule that is there yet not there*. The sign merely expresses this essence to us through the trivial difference it makes between thinking one is following a rule and following a rule, by making itself *different* from following the rule, that is, thinking one is following a rule is *not* following a rule. The compliant rule is thus divided within itself—it says that it is a rule when it is really just a statement of the rule. In truth, there is no such compliant ‘rule’, because there is *no*

*substance* to this picture of the rule; rule and rule-follower are really one and the same; that is to say, the rule is really a *living substance*, a human being. The only *genuine* rule is thus the *pliant rule*.

### Notes for Chapter 3

1. PI 258: “Let us imagine the following case. I want to keep a diary about the recurrence of a certain sensation. To this end I associate it with the sign ‘S’ and write this sign in a calendar for every day on which I have the sensation.—I will remark first of all that a definition of the sign cannot be formulated.—But still I can give myself a kind of ostensive definition.—How? Can I point to the sensation? Not in the ordinary sense. But I speak, or write the sign down, and at the same time I concentrate my attention on the sensation—and so, as it were, point to it inwardly.—But what is this ceremony for? For that is all it seems to be! A definition surely serves to establish the meaning of a sign.—Well, that is done precisely by the concentrating of my attention; for in this way I impress on myself the connexion between the sign and the sensation.—But ‘I impress it on myself’ can only mean: this process brings it about that I remember the connexion *right* in the future. But in the present case I have no criterion of correctness. One would like to say: whatever is going to seem right to me is right. And that only means that here we can’t talk about ‘right’.”

2. PI 260: “‘Well, I *believe* that this is the sensation S again.’—Perhaps you *believe* that you believe it!

Then did the man who made the entry in the calendar make a note of *nothing whatever*?—Don’t consider it a matter of course that a person is making a note of something when he makes a mark—say in a calendar. For a note has a function, and this ‘S’ so far has none.

(One can talk to oneself.—If a person speaks when no one else is present, does that mean he is speaking to himself?)”

3. PI 270: “Let us now imagine a use for the entry of the sign ‘S’ in my diary. I discover that whenever I have a particular sensation a manometer shews that my blood-pressure rises. So I shall be able to say that my blood-pressure is rising without using any apparatus. This is a useful result. And now it seems quite indifferent whether I have recognized the sensation *right* or not. Let us suppose I regularly identify it wrong, it does not matter in the least. And that alone shews that the hypothesis that I make a mistake is mere show. (We as it were turned the knob which looked as if it could be used to turn on some part of the machine; but it was a mere ornament, not connected with the mechanism at all.)

And what is our reason for calling ‘S’ the name of a sensation here? Perhaps the kind of way this sign is employed in this language-game.—And why a ‘particular sensation’, that is, the same one every time? Well, aren’t we supposing that we write ‘S’ every time?”

4. It might be claimed that the manometer, a ‘public object’, functions in PI 270 as the independent criterion of correctness for using the sign ‘S’—thus ruling out a ‘private’ use of

the sign ‘S’—but as will be shown later, the sign ‘S’ only becomes useful, and hence meaningful, once the sensation is found to be useful; and if indeed the *sensation* can be found to be useful, it will still be useful *independently* of the manometer with which it has been correlated, since the readings of the manometer themselves become useful, and likewise meaningful, like the sign ‘S’, only because *the sensation itself is useful*. The manometer perhaps helped in the discovery of the usefulness of the sensation but was not essential for *its* usefulness; on the contrary, both the *manometer and the sign ‘S’* are dispensable in this case, since the one merely *reads* the sensation and the other only *stands for* it.

If it was the manometer that in fact made the sensation ‘public’ in this case, and hence legitimized the sign ‘S’ for our linguistic purposes—wouldn’t that manometer merely confirm for us the ‘privacy’ of the sensation itself? And suppose the manometer hadn’t yet ‘outed’ the sensation in question from its ‘privacy’ by being correlated to it—would that thereby render the sensation useless, and so make the sign ‘S’ meaningless as a result? *How could that follow?* After all, the sign ‘S’ does not stand for the readings of a manometer but for the sensation that is also read by the manometer. The meaning of the sign ‘S’ stands or falls with the sensation, therefore, and not with the readings of the manometer; and this is just to say that the sign ‘S’ and the sensation are internally related just as the readings of the manometer and the sensation are internally related, and the sign ‘S’ is related to the readings of the manometer only *through* the sensation: the internal relations here are not simply grammatical but *physical, too*, since the usefulness of the sensation was *discovered* through its correlation with the manometer, *though it could have just as easily been discovered from the sensation’s own—usefulness*.

5. Kripke, *Wittgenstein: On Rules and Private Language*, p. 110: “Does this mean that Robinson Crusoe, isolated on an island, cannot be said to follow any rules, no matter what he does? I do not see that this follows. What does follow is that *if* we think of Crusoe as following rules, we are taking him into our community and applying our criteria for rule-following to him. The falsity of the private model need not mean that a *physically isolated* individual cannot be said to follow rules; rather that an individual, *considered in isolation* (whether or not he is physically isolated), cannot be said to do so. Remember that Wittgenstein’s theory is one of assertability conditions. Our community can assert of any individual that he follows a rule if he passes the tests for rule following applied to any member of the community.”

## Chapter Four: The Picture of a Paradox

### 4.1 The Compliant and Pliant Pictures of the Rule

The paradox of compliance boils down to this: if you already understand the rule, you cannot be guided by the rule, nor can your action be measured by the rule. Alternatively, it makes no sense to say that one *uses* one's own understanding to guide one's action or to measure one's action. If I say that I use the rule and not my understanding of the rule to do this, then the rule here is something independent of my understanding, and hence independent of the action that manifests my understanding. But then what rule is *that*? The *sign itself*? But if a sign *by itself*, independently of one's understanding of the rule and contrary to its own claim to be a 'rule' for me, can't guide me or measure my action—and it can't, since an action can only conform to another's action, not an abstract action like the sign—then the sign, as a compliant *rule*, is a sham, for it lacks the substantive power of a true compliant 'rule'—its ability to guide and measure—but nevertheless continues to project that power as though it did, substituting itself for the true compliant rule that is the teacher.

A true compliant rule is something by which we are guided and measured *as we act*: hence the pictures of the rule as both a handrail and a yardstick. Measuring itself is a certain kind of activity in which one thing is set against another in order to determine compliance. But when the rule-follower follows the rule, he just acts. There is no measuring of his action, unless someone then asks him for a *justification* of his action. In this case, he is expected to give an answer, that is, he is expected to *say why he is doing*

*what he does*. But the answer he gives is no true measure of his action, nor is it a true guide of his action; and even if this answer were to satisfy the one who asked the question, *his* satisfaction with this answer can be no measure either, since he could have been persuaded by any number of answers. At any rate, we are no more measured or in need of being measured than we are guided or in need of guidance when we are following the rule; the rule-follower, in other words, need not give any justification of his action—he need not say anything in connection to his following the rule. If this is so, then it is also the case that there is no need for ‘compliance with a rule’—not that ‘compliance with a rule’ makes no sense at all, but that it only makes sense in certain circumstances, namely, when someone is learning to follow a rule from someone else. Here the teacher is the rule for the pupil, and both guides and measures his action. The rigidity and constancy of this rule just is the role the expert plays for another in teaching the rule. But for the rule-follower himself, there is no one to guide or measure his actions, and he does not do so for himself, since he need only act on his understanding. The expert himself is a pliant rule; he is a compliant rule only for his pupil. The sign is merely the tool used by the teacher for judging compliance, that is, for teaching the pupil to act as the teacher does.

Without the notion of compliance for the rule-follower there seems to be no room for the connected notions of error and correction in following the rule. But isn’t the rule-follower correcting himself something like the body healing itself? Surely ‘healing’ and ‘health’ are normative notions comparable to ‘correcting’ and ‘following the rule’. We think of the rule as something independent of the rule-follower, and the claim that a human being qua rule-follower could actually be this rule, such that he acts for himself, stands opposed to this view. A rule that corrects itself seems imperfect somehow, something less

than a rule, allowing all kinds of worries to arise about its solidity because of an implied vulnerability to ‘corruption’, ‘error’, ‘illness’. On the contrary, this characterization would just as well show the *robustness* of the rule, its ability to recover from errors in acting, from all the contingencies that surround a functioning, living usage. (After all, the rule *can* sometimes fail). But it is the sign that misleads us in this respect, because we think of *it* as something apart from the rule-follower’s action in following the rule, something that can be used as a guide or measure of this action. The sign’s function here is to justify one’s action, like a written law one appeals to as a defense before a judge, whose externality and independence cast a shadow over everything that we do. But the sign is really a false measure.

Wittgenstein’s picture of the rule is deficient, therefore, because it cannot be used to show rule-following without the idea of compliance. Although he knew about its paradoxical nature, his picture is nevertheless still that of a handrail or yardstick, something independent of the rule-follower which one uses and appeals to as both a guide and a measure for one’s action (something that one also follows *blindly*). The rule-follower’s action is *said* to comply or not to comply with the rule. Another picture, however, paints a different view, namely, that of the rule as something flexible and of a certain shape that bends and twists upon usage but always returns to its original shape. Its strength, not weakness, lies in this flexibility without breaking—thus showing its *robustness* at the same time. Here the rule may be a strong or a weak rule, depending on how much stress it can endure without breaking. This picture of the rule is the *pliant* picture of the rule, as opposed to the *compliant* picture held by Wittgenstein and his followers. Such a picture resembles that of a living organism, whose normal shape we might call its ‘health’. This

organism is subject to all kinds of illness and injury, but nevertheless is able to recover from these by healing itself, by restoring itself to its normal shape or ‘health’. Don’t forget that ‘health’, ‘illness’, ‘healing’ are *normative* concepts, no less than ‘rule’, ‘error’, and ‘correction’. If an organism is able to recover from all these evils while living, and can do so quickly without perishing, we call such strength the *robustness of the organism*. So, too, is it with the pliant rule: it resembles a living organism in virtue of its capacity to restore itself to ‘health’—in fact, it is not just figuratively a living being--it *is* a living being, a *human being*.

It is important to know how each picture of the rule depicts an error. The compliant picture depicts an error in rule-following as an act of non-compliance: to follow the rule incorrectly is to go against the rule or for one’s action not to measure up to the rule. To determine that an error has occurred, then, requires that the rule measure one’s action. The pliant picture, on the other hand, depicts an error in rule-following as a distortion of the rule. An error is just a deviation in the course of one’s activity from the *purpose* of the rule, not an action that fails to measure up against another action. A regular way of acting when following the rule will help show what this distortion looks like, and not every distortion will look the same. The expert rule-follower will know what kinds of errors occur in following the rule, and thus will know what these distortions look like. By grasping the purpose of the rule, the rule-follower has already defined these errors or distortions in a sense, since the rule is now just the rule-follower, and his errors become *his deficiencies in performance*; the rule-follower-cum-rule is thus the *whole phenomenon of rule-following, containing his own errors through his ability to correct himself—that is, to remain whole*.

Why does there seem to be little problem for the rule-follower's action to conform to another's, but there is a problem only when his action is *said to conform to a sign or rule-formulation*? It is because a human being has the innate ability to imitate other human beings and grasp the point of what they are doing. Another human being can be a guide for us and a measure, too, and the problem of interpretation arises only in the case of the pupil imitating his teacher, since any interpretation of a rule is now simply the pupil's imitation of his teacher. In the case of an abstract action or sign, there is no natural, instinctive response in following *that*, except to imitate this same abstract action itself, by repeating the mere use of the sign. All we really need in rule-following is to see someone doing something, and from there we will learn to pick it up. It is hardly surprising that the compliant picture of the rule gives us such difficulty in understanding how one can follow the rule, especially when the sign can be interpreted by various actions. We still mistake the rule for its formulation in signs; we have not yet fully recognized that the ability to give justifications for following the rule, or even just to describe it, and the ability to follow this same rule are different abilities. Whether someone can actually give such a justification does not matter in the end, just as it does not matter what he thinks he is doing when he follows the rule. The same goes for any use of the sign, even under the description of what *we call* 'following the rule'.

Only a human action has the form of a human action, and so to conform one's action to a rule is really just to conform one's action to that of another human being. That is why it makes sense to be guided by a human being and to be measured by a human being. Once again, we are saying that the rule is really just a human being. Hence to conform or to comply with a rule is to form or bend one's action to that of another in order to satisfy him

(conform < L. *conformare*: to make of the same form; comply < It. *complire*: to fulfill, complete). To conform thus is to establish agreement between two beings of the same kind. But to agree is not necessarily to conform or comply, because there is such a thing as mutual agreement, but not mutual compliance. With compliance, one individual is independent of another and this other is dependent on him; as we have already seen, this is essentially the relation between teacher and pupil. Two independent individuals, however, can agree with one another, even though one does not comply with the other; and an independent individual can also be said to agree with himself in this same sense of agreement. For this reason, agreement does not imply compliance or dependence. But it makes little sense to say that one conforms to one's own actions when following the rule, even though one's present action can agree with one's past actions. In the pliant picture of the rule, the only agreement that matters for the rule-follower is the agreement of the rule-follower with himself, since it is a non-compliant form of agreement. This is the form of agreement the teacher has with himself, and the pupil conforms his action to his teacher's, not simply for the sake of agreeing with him in the compliant sense of agreement, but simply in order that he may be able to better grasp the purpose of the rule which then allows him to agree with himself, just as his teacher agrees with himself. Agreement in action is thus the expression of *independence* in the rule-follower, not the expression of *dependence* on others—this agreement with himself says that he now belongs to himself and has become a whole being, that is, he can now act independently of others, since he has grasped the purpose of the rule and made it his purpose; agreement among rule-followers is necessary only as the expression of a common interest or purpose, and even in this case, such agreement cannot be the *goal* of our training to follow the rule—such non-compliant

agreement simply *follows* from our grasping the *same* purpose. At any rate, the independence of the rule-follower is the goal of his training to follow the rule, even when agreement in action implies a common interest or purpose: *true compliance is therefore just our training to become these independent rule-followers*. Grasping the purpose of the rule makes both the rule-follower independent of others and the rule pliant in nature; and only because the rule-follower is made independent in *this* sense is there agreement in action at all, even across individuals.

One section in the *Investigations* that is often used to bolster the idea of compliance is PI 265<sup>1</sup>. In this section, Wittgenstein explores the idea of looking up words in a dictionary which exists only in one's mind and is used to justify the translation of one word by another. This does not make sense, since such a dictionary cannot function as a standard for our justification of a translation, because we need to appeal to something independent and objective, not subjective like this dictionary. In other words, we cannot *test* this dictionary in our memory for *correctness*, we cannot confirm its correctness *independently of our memory*.—But suppose I speak (and hence understand) both languages involved in translating these words from one language into another. “Well, are you saying that in this case you would be using this dictionary in your memory, and so it does make sense after all?” Not at all, because those who are bilingual and translate between languages they already understand do not look up anything in their memory. They simply translate, perhaps trying out different sentences, different words as they go along; but no, they simply translate; they do not look up any words in a dictionary to justify their translation, whether this dictionary is subjective or objective. “And so what is your point? The whole idea of a physical dictionary points to the necessity of having something objective and independent

by which we can justify our translation.” But this is *my* point: we don’t need a dictionary at all if we already understand *both* languages; we need no criterion of correctness independent of us if we can do that: we therefore need no justification either.

#### **4.2 The Essence of Rule-Following is Imitation of Action: The Unity of Action and Purpose**

We naturally think of imitation as a form of representation or depiction, that is, as something reproduced in a medium in some manner, whether linguistic or visual or auditory, in other words, through a sign or picture. But this is not the imitation that occurs when one follows the rule, for nothing here stands in the way of the pupil following his teacher—certainly not a sign—for the pupil’s action is not a depiction of his teacher’s action, but *becomes the same action*, so that the pupil becomes the rule that he is learning to follow. Consequently, there is no need for a medium of depiction here, just as there is no need for signs. And since this is so, the pupil *can and will become* independent of his teacher by becoming a rule-follower himself and a teacher for others, because a depiction is still dependent on what is depicted, but the pupil *depicts nothing by imitating his teacher*; in addition, what he thinks while following the rule no longer matters at all, because all errors, obstacles, and dangers have become part of this rule, too; in fact, they are what give the rule its robustness and show its flexible nature. In short, the rule-follower has all he needs to follow the rule independently of others, and so the rule can only be a pliant rule, since his errors are not errors in thinking but in performance. In fact, his ability to justify his actions or make judgments in rule-following demonstrates his independence of action. To require that the pupil’s action still conform to a sign or rule-formulation, even after he

has learnt to follow the rule, would amount to requiring that his action now conform to a depiction of what his teacher already does, and so make him still dependent on his teacher. This requirement, whether in the form of a ‘context’ that he is placed in, a ‘consensus of action’ that he joins, or a ‘rule’ that he complies with, would also have his actions essentially dependent on the depiction of actions—which is absurd. When an action conforms to another action—for example, when the pupil imitates his teacher—the two actions become the same action, and once the pupil grasps the purpose of this action, both are now the same action *independently of each other*. The compliant rule, which is really just the sign in disguise, is merely this depiction of an action in a foreign medium; and it can only *depict* this action because it is really an abstract action, for an abstract action is simply one that becomes something other than what it is by nature, that is, other than an arbitrary action or mere use. The pupil’s imitation of his teacher’s action, however, does not involve a different medium of action like the sign but the same medium of action, since teacher and pupil are *both human beings*.

By grasping the purpose of the rule that he is learning to follow and unifying his action with this purpose, the pupil not only becomes the rule that he is following, but he also becomes independent of his teacher, that is, he now *belongs to himself*. This is the *true* independence characteristic of individuals within a genealogical community in which the individual is considered whole, a peer among peers. The pupil who is said by us to have learned the rule only because he now does what we all do is the pupil who *still belongs to us*: he is just an abstraction for us, the description of someone who simply acts as we all do in following the rule. But he belongs to us and not to himself only because we depict him through the sign, that is, we still insist that he comply with the rule that is independent of

him, and this compliance is reinforced by our description of him *acting as we all do*. In truth, we still demand that he comply with us, as though we were his teacher. We have not yet comprehended that, because this sign *belongs only to us*, this sign can never be independent of the rule-follower, and so cannot function as the rule that is independent of the rule-follower; in fact, the compliant rule is not a rule at all, but merely the sign itself, a tool for compliance used by the teacher. The sign, as such a tool, constitutes an *insubstantial* form of identity. The only substantial form of identity in rule-following is the pliant rule itself—that is, *the pliant rule is the substantial self-identity of the rule-follower with the rule*.

#### **4.3 The Essence of the Sign is Metaphor or *Identity-in-Difference*: The Unity of Simple Self-Identity and Simple Non-Identity with Itself.**

Metaphor is found in language. One such metaphor that is *in* language says that a certain man is a lion, even though the man is really only a man and not at all a lion. But what does this metaphor *mean*, then, by saying that *the man is a lion*, for a metaphor in language means *something other* than what those words *literally* mean? Has the man leapt audaciously towards the lion, perhaps, sidestepping logic on his way? Has some great logical chasm been traversed that would otherwise be impassible but through metaphor? Not really, since metaphor is not being illogical here, but is simply showing the *depth* of language and not its width or length; metaphor is bringing out a difference *within* language that is constantly being leveled or flattened into *abstract sameness* by the *literal word or sign qua mere use*, that is, by constantly saying that *man is really only man and nothing but man; this is in fact what man is: man-is-man or man-is-himself*. This abstract sameness,

however, is only formal, without real content, and quite meaningless, since it simply presupposes the sameness of the sign qua mere use and then merely stutters this sameness to itself, uncomprehending the deeper sameness in the sign that gives it life, whose meaning has just been stripped down in language to a mere use or bare sign—*the literal word*; this abstract sameness has the form of a ‘this-is-this’ or a ‘this-is-itself’, what could be called *simple self-identity*. But the difference in language brought out by metaphor is a difference *within* language that finds a certain sameness, and hence a certain meaning, by making *no* difference to the meaning of the literal words—by walking on their heads, so to speak, since man is still man, lion is still lion—and *that* difference between man and lion now has more *depth* of meaning than the literal words, precisely because metaphor itself is showing an arbitrariness and autonomy *within* language that is the shadow of the arbitrariness and autonomy of the sign itself; metaphor *in* language is in fact connected to language’s own nature, whose *explicit form of self-reflection* is now found in the more substantial ‘this’-means-this, or the *identity-in-difference* which is *in* language and says that ‘*man*’ means *man*. The *deeper* meaning in all metaphor, then, is that metaphor is really the essence of language and logic, that is, the essence of the sign itself; we can thus call the sign the *original identity-in-difference*; or, to say the same thing but differently, the sign is the *original metaphor*—the metaphor that is not *in* language but *is* language; it is that deeper identity which possesses content, confers meaning, bequeaths knowledge.

Abstract sameness, or simple self-identity, is thus the sign qua mere use that only has the sameness of its form or ‘shape’ by overlooking any substance that goes with it, for its true substance—purpose qua usefulness—is still there and always was, but never simply qua mere use. Thus in the compliant picture of the rule, we see the rule-follower as the one

who simply acts on the rule that is independent of himself and that determines his compliance with it, as though the rule-follower were like this sign qua mere use; consequently, we understand the rule-follower abstractly, too—despite our insistence that we are seeing him concretely in his brute actions, in his observable behaviour—by identifying him only with these simple actions, all the while still seeing the *substance* of his action only in the rule that is independent of him, not yet comprehending that *the rule-follower himself is in fact this substance, that is, the rule itself*, since he now has grasped the purpose of the rule. The compliant rule is merely the insubstantial reflection of this abstract conception of the rule-follower qua mere follower. Indeed, ‘privacy’ in rule-following has been understood as the incoherence of the rule-follower following the rule merely qua follower, since, by abstracting the rule from the rule-follower and his actions and making this rule different from the rule-follower, we find that there cannot now even be sameness in his merely acting the same. Yet this incoherence only appears to be the case, because we have stripped the rule-follower of any substantial existence by placing him in a ‘context’ that defines him; the impossibility of following a rule ‘privately’ is thus the negative consequence of this abstract sameness depicted in the compliant picture of the rule, a picture in which the difference between rule-follower and rule is maintained only by seeing both the rule-follower and rule abstractly from one another: it is the picture of something independent of the rule-follower, an invisible extension of a rail that simply repeats itself in the form of an endlessly repeated self-same sign or rule-formulation—the this-is-this-is-this... (for example, the rule, or sameness, of the arithmetic series 2, 4, 6, 8... is the +2-is-+2-is-+2-is-+2...); we have thus failed to comprehend a *deeper* identity than this abstract identity, this abstract rule. This identity is the *substantial self-identity of the*

*rule-follower with the rule*, that is, a *different* identity from the simple self-identity of the sign qua mere use or the identity-in-difference that is the essence of the sign itself. It is the identity of a living being.

Wittgenstein and his followers, however, had objected to the rule-follower following a rule ‘privately’ essentially because they had thought that simple self-identity itself was being made the identity at the core of the *different applications* of the rule in language or mathematics; and this simple self-identity could never *by itself* establish the meaning of the sign, that is, the sameness of the sign, because simple self-identity, the sign qua mere use, still lacked the substance provided by the rule which had to exist independently of the rule-follower himself and with which his action had to comply. They did not see that once the simple self-identity of the sign qua mere use became identified with rule-follower abstractly conceived, and so collapsed into incoherence, the simple non-identity that said at the same time that the rule-follower is *not* the rule had to be abandoned, too; in fact, Wittgenstein’s argument against following a rule ‘privately’ stretched the abstract difference between rule and rule-follower to the point of incoherence, so that *this abstract difference itself* should have collapsed and *not* the notion of ‘privacy’. If following a rule ‘privately’ is impossible, it is only because both rule and rule-follower are abstractly conceived as different. But Wittgenstein clung to the trivial difference between merely thinking one is following a rule and actually following a rule in order to save the compliant picture of the rule, whose preservation requires the maintenance of this abstract difference. Wittgenstein failed to see that the rule-follower who “just acts, without reasons”, says “This is simply what I do”, and then “follows the rule *blindly*” is as abstract as the simple self-identity of the sign qua mere use. In truth, Wittgenstein’s rule-follower *actually is* this

simple self-identity, since the sign qua mere use *just is the simple abstract action of the rule-follower that is both arbitrary and autonomous*, while the compliant rule is merely the rule which he *thinks* he is following—the phantom rule supposedly independent of this simple self-identity called ‘the rule-follower’ and whose difference from this rule-follower prevents the rule-follower from following a rule ‘privately’: “hence following a rule is a practice” (PI 202), that is, doing what *we call* ‘following the rule’ (PI 201); it is a description of an action that the rule-follower must conform to if he is to follow the rule, an action that still exists only in the form of the sign.

Without the sign in rule-following, then, there would be no abstract sameness or simple self-identity, for language is a form of abstraction and is just the unity of simple self-identity and simple non-identity with itself. In other words, the simple self-identity that says that A is A, and the simple non-identity that says that B is not A, are themselves united in the identity-in-difference that says that B is A. We can state this identity-in-difference explicitly in language by saying that ‘A’ means A. Language itself thus shows once again that it has the nature of metaphor, because this unity of simple self-identity and simple non-identity with itself just means that the sign qua mere use, *simple self-identity*, becomes meaningful only when it *also becomes what it is not*—that is, more than a mere use, *simple non-identity with itself*. We expressed this unity earlier as the unity of the use and usefulness of the sign; it could also be expressed as the unity of the finite and the infinite in the sign, since the sign qua mere use is just a simple abstract action and finite in nature, while what this sign qua mere use is not and what it could be useful for could be *anything* at all—it is simply *not* the sign qua mere use. But the difference here within this

unity is merely an abstract difference that exists *only* in language, since it is really just the difference that constitutes the *meaning* (or identity) of the sign—its *identity-in-difference*.

#### **4.4 The Pliant Rule is a Living Genealogical Identity; Abstraction is the Power of Purposive Negativity**

Whereas the difference between rule and rule-follower manifested in the relation between teacher and pupil is a *living difference* in which what is different *becomes* the same through one becoming what the other already is, the difference that is said to exist between rule and rule-follower in language, and thus in the compliant picture of the rule, is merely the abstract difference made explicit *in language* by the simple non-identity that says that the rule is *not* the rule-follower: this latter difference is the abstract difference that still depends on abstract sameness, whose essential unity with this abstract sameness forms the identity-in-difference of the sign itself. The former difference is the difference between rule and rule-follower manifested through the generation of one rule from another in the relation between teacher and pupil and is thus a *genealogical* difference, one that regenerates genealogical sameness or *substantial self-identity* out of this difference; in other words, the rule-follower, by following the rule for himself and grasping its purpose, *becomes* the rule himself, a rule capable of generating another rule by also being the rule for another. And all this comes down to saying that the compliant rule itself is really just a lifeless abstract identity whose only *defining* distinction is *not* to be the rule-follower and which thus exists *only* in the external form of a rule-formulation or sign, whereas the pliant rule is a *living genealogical identity*—not the sameness of an action lacking its own substance, only finding this substance in another through compliance, but the sameness of

this substance with itself, the rule-follower's identity with the rule he is following, or, in the case of rule-following with signs, the return of the rule-follower from his simple abstraction in the form of an external self-same sign or rule-formulation back into himself *qua purpose*. This is in fact why the sign in rule-following is not an essential part of rule-following, and is hence dispensable for rule-following *per se*, since, qua mere use or as abstract identity, it is simply the difference that makes no difference in the end, or the action that is done for its own sake, albeit for another's purpose; the sign itself really amounts to *nothing* in the end. It is only the substantial self-identity of the rule-follower with the rule, a living genealogical identity, that really matters in rule-following: we call it 'the purpose of the rule', or 'the rule in the fullest sense', or just 'a human being'; this pliant rule is the *higher* rule which says that the Sabbath was made for Man, not Man for the Sabbath, since the Sabbath, like the compliant rule, is a mere observance and shadow of this higher rule; it is a purely formal custom attached to an arbitrary day with no intrinsic significance, except for its mere observer under the Law of Moses. Everything in rule-following, therefore, follows from this purpose; or rather, all roads lead back to it, even in the roundabout that is the sign. And *that* is in fact the *power of purposive negativity*; it is the power which makes something from nothing.

The abstraction in rule-following with signs is thus responsible for the peculiar way in which a rule of mathematics, for example, connects with purpose. Since purpose establishes the normativity of the rule, abstraction must still somehow retain this connection, preserving at the same time the arbitrariness and autonomy of the sign. We have already seen how it can do this through the usefulness of the sign, but we must also take note of how it can do this through the identity of the sign, since the identity of the sign

is just the sign's meaning, that is, that which in fact constitutes the normativity of such a rule. The correctness of a rule of mathematics, for instance, is never said to be based on the *consequences* of using the rule, as if something were true in mathematics only because we have 'found it to pay'. But this is not so far from the truth as one might think, since the stipulation of what is correct in a rule of mathematics takes the form of the abstract sameness of the sign qua mere use, and this abstract sameness itself owes its sameness to the deeper genealogical identity called 'the purpose of the rule'. The correctness of a rule of mathematics, formalized through a stipulation, is thus merely a *formal correctness*, like the formal correctness which says that 'knight' is correctly spelled with a 'k', even though the 'k' here no longer functions as a rule for pronunciation in English and it really makes no difference whether the word is now actually spelled with a 'k' or not. In fact, it was already shown how practice and function came apart in the orthography of English, and this is what has also happened in the case of mathematics: the formal correctness of a rule of mathematics is a correctness divorced from any functional use the production of the sign in mathematics may have had—we simply repeat the production of the sign qua mere use, that is, the simple self-identity of the sign itself: " $2 + 3 = 5$ ". We then ignore the origin of the stipulation that occurred in a practice that already used rules of mathematics without stipulation, as if we were then to say that stipulation alone were enough to establish the correctness of a rule of mathematics and then justified this assertion by pointing out the arbitrariness and autonomy of the sign. But stipulation is not enough. Thus it is with formal correctness in general: there can be no formal correctness without connection to purpose, since the abstract sameness associated with formal correctness already presupposes a prior connection to purpose as its *raison d'être* and later re-establishes this

connection through the usefulness afforded by this abstract sameness; and this must be so *because* of the arbitrariness and autonomy of the sign. And so the purpose of a rule of mathematics is *not part of what is stipulated in a rule of mathematics but what makes such a stipulation even possible*; not only that, but purpose also determines the *character* of the rule itself, what one *does or can do with it*, since the purpose is in an important sense the beginning or origin of the rule—that which *begets* the rule, so to speak. Formal correctness in the use of signs is really the orphan who inherits his mortal character from the earthly purpose that has actually given him life, yet still believes he sprung fully grown and clothed in the form of a deathless being, like Athena from the head of Zeus. But he is an orphan only because his action's natural connection to purpose had been sundered on account of the nature of the sign, leaving him to substitute stipulation for this purpose *as a virtual purpose*, thereby transforming the mere use of the sign into something useful, so that now this formal correctness can be adopted by *another's* purpose and given actual content, that is, a *significant* meaning. Formal correctness, then, as the simple self-identity of a stipulation, is by itself an abstract identity and yet another expression of the compliant rule, since this formalness is also maintained by the equally abstract difference between rule and rule-follower. Purpose alone prevents this formal correctness from completely breaking apart the unity of practice and function in rule-following. What is anathema to linguistic meaning, therefore, is not 'privacy' but purposelessness, since the negativity in the sign which forges symbolic meaning just is a *purposive negativity*, the same negativity in the sign that says that it is what the sign qua mere use is not—that is, it is more than just a 'dead' sign or medium. Purpose thus gives the sign a new kind of life it would otherwise not have—a life essentially moved through this creative negativity.

Show me a rule that is purely formal, and hence truly useless, and I will show you a rule that does not really exist. Such a rule, of course, is the compliant rule, the same rule that *says* that it is a rule but *really is not*, since what it says it is is simply the *grammatical* projection of a power it does not have, that is, the power to guide and measure; its true power, if you will, is the *power of negativity*: it is the rule that is *there but not there*. In fact, *stating* a rule one is following is merely *thinking* one is following a rule. Hence it is not possible to follow a ‘compliant’ rule: otherwise stating a rule one is following would be the same thing as following it.

What language distinguishes through abstraction is not only the difference between rule and rule-follower, however, but also the difference between the rule-follower and his context—the context that here merely substitutes for the rule by conditioning the simple action of the rule-follower, bestowing on his action its proper meaning. But, as we saw earlier, the context of the rule-follower’s action, the ‘circumstances’ or ‘background’ of his performance, is actually dependent on the rule-follower himself and what he does in following the rule; the context is never already there, save the rule-follower already being there. Thus the context, too, is the *simple abstraction of language*, since it exists, like the compliant rule, *only* in the external form of a sign or rule-formulation, as a *description* of what the rule-follower is doing when he “acts according to the rule”; nowhere else but in this sign is there a ‘context’, a ‘background’; the context has *no* independence from the rule-follower it describes, circumscribes, and then finally *prescribes*. What we have really found in the context of the rule-follower’s action is just the rule-follower himself. The ‘context’ is really just another expression of the compliant rule, an abstraction of language that is really just more language—and nothing more. Appealing to this context is like

giving an ostensive definition to oneself: one cannot do *that* unless one *already understands what one is attempting to define*. We only *seem* to get further in our understanding of the normativity of the rule when we look at the context of an action, but we really get no further than what we already understand about the rule, which is to say, we only get out of this context what we have already put into it—we are really just extending the rule-formulation further than it already was. When we then try to ‘decontextualize’ this action, it is not surprising that we think that we have fallen into the incoherence of following a rule ‘privately’; this ‘isolated’ action can no longer have a meaning by itself, no sameness through which we can call it ‘following the rule’, since it now resembles nothing that we know; we have once again removed the substance of the rule from itself, leaving the action lifeless and hollow; but we have actually only made a necessity of something which in truth is unnecessary, that is, we have made an unnecessary distinction that talks about the abstract difference between the rule-follower and his context. Yet no one has ever clearly *said* what a ‘decontextualized’ action would be, how a rule-follower could ever be without a context. But for all that, he still *can’t* be without a context! *That* is what we are indeed inclined to say, since to say otherwise would be to speak nonsense, although we may just as easily be inclined to say that there is no such difference.

This abstract difference between rule and rule-follower, which is the basis of Wittgenstein’s argument against following a rule ‘privately’, has already been drawn into his paradoxical picture of the rule, as *the rule that is both there and not there*. The compliant rule is there *only* in the external form of a rule-formulation or the sign qua mere use, that is, as *simple self-identity*; yet when this sign qua mere use turns out to be really the simple abstract action of the rule-follower himself, the sameness of this sign becomes

identified with the different individual actions of the rule-follower, so that this rule, the sameness of the sign, breaks up into difference and is now forced into becoming what the rule-follower's action is *not*, that is, a form of identity underlying these manifold actions of the rule-follower, extending beyond even the sign that is there. Instead of being forced back into the rule-follower and taking on a substantial form of self-identity with the rule-follower qua purpose, the compliant rule flees into the *simplest and most abstract* form of identity of all, that is, *into nothing at all*—the *missing* part of handrail that is still inseparably connected to the handrail as its *extension*; the compliant rule now takes the form of *simple non-identity with the rule-follower, that is, simple non-identity with itself*. But this compliant rule is really no rule at all, since it lacks any substantial existence of its own; it has no power to guide or to measure the rule-follower—it is not even independent of the rule-follower but is his mere shadow, following him instead of being followed. The only form of identity and difference connected to the compliant rule is thus simple abstraction; and its entire life is spent in this simple abstraction, one form of which is Wittgenstein's picture of a paradox. That elusive and dark rule the rule-follower follows *blindly* does not really exist, because it is not only *not* the rule-follower—it is *not anything at all*. Only through purpose can this 'nothing' then become something else—through the *purposive negativity* that constitutes the sign's identity-in-difference and that is also called 'simple abstraction'. But we *say* that this compliant rule is there and we *think* that it is what we are following when we follow the rule, and it is no wonder: this picture tells us what to say and what to think about the rule. In fact, the sign itself is just such a paradoxical picture, just as the paradox in the picture shares the nature of metaphor, capturing meaning by being both what is there and not there at the same time. And so the

compliant rule is really just metaphor after all—*the picture of a paradox*—and not a living rule like the pliant rule. It was Wittgenstein’s error not to see this compliant rule for what it really is, since he only listened to what the compliant rule told him it was, and it told him it was a rule when it really was not. He should have known better, just as he said he knew that “[a] *picture* held us captive. And we could not get outside it, for it lay in our language and language seemed to repeat it to us inexorably. (PI 115)”

#### Notes for Chapter Four

1. PI 265: “Let us imagine a table (something like a dictionary) that exists only in our imagination. A dictionary can be used to justify the translation of a word X by a word Y. But are we also to call it a justification if such a table is to be looked up only in the imagination? —“Well, yes; then it is a subjective justification.”—But justification consists in appealing to something independent.—“But surely I can appeal from one memory to another. For example, I don’t know if I have remembered the time of departure of a train right and to check it I call to mind how a page of the time-table looked. Isn’t it the same here?”—No; for this process has got to produce a memory which is actually *correct*. If the mental image of the time-table could not itself be tested for correctness, how could it confirm the correctness of the first memory? (As if someone were to buy several copies of the morning paper to assure himself that what it said was true.)

Looking up a table in the imagination is no more looking up a table than the image of the result of an imagined experiment is the result of an experiment.”

## Bibliography

- Baker, G. P., and Hacker, P. M. S. *Scepticism, Rules and Language*. Oxford: Basil Blackwell, 1984.
- Baker, G. P., and Hacker, P. M. S. *Wittgenstein: Rules, Grammar and Necessity* (Vol 2). Oxford: Basil Blackwell, 1985.
- Bloor, David. *Wittgenstein: Rules and Institutions*. New York: Routledge, 2002.
- Hegel, G. W. F. *Phenomenology of Spirit*, translated by A. V. Miller with analysis of the text and foreword by J. N. Findlay. Oxford: Clarendon Press, 1977.
- Kripke, Saul. *Wittgenstein on Rules and Private Language*. Cambridge: Harvard University Press, 1982.
- Malcolm, Norman. *Nothing is Hidden*. Oxford: Basil Blackwell, 1986.
- Medina, Jose. *The Unity of Wittgenstein's Philosophy*. Albany: State University of New York Press, 2002.
- Wittgenstein, Ludwig. *Philosophical Investigations*. translated by G. E. M. Anscombe. Oxford: Basil Blackwell, 1972.
- Wittgenstein, Ludwig. *Remarks on the Foundations of Mathematics*. translated by G. E. M. Anscombe. Oxford: Basil Blackwell, 1978.
- Wittgenstein, Ludwig. *Zettel*. translated by G. E. M. Anscombe. Oxford: Basil Blackwell, 1972.