JoLMA

The Journal for the Philosophy
of Language, Mind and the Arts

Vol. 5 – Special issue   October 2024

**Edizioni
Ca'Foscari**

JoLMA
The Journal for the Philosophy of Language, Mind and the Arts

Editor-in-Chief
Luigi Perissinotto

JoLMA
The Journal for the Philosophy of Language, Mind and the Arts

Six-monthly Journal

Table of Contents

Perspectives on Wittgenstein's *Philosophical Investigations*

edited by Luigi Perissinotto and Elena Valeri

The *Philosophical Investigations* and Its Seventieth Anniversary

Luigi Perissinotto
Università Ca' Foscari Venezia, Italy

Elena Valeri
Università Ca' Foscari Venezia, Italy

Ludwig Wittgenstein's *Philosophical Investigations* (*Philosophische Untersuchungen* in German) were published by Basil Blackwell on 1 May 1953, just over two years after their author's death in Cambridge on 29 April 1951. Ferruccio Rossi-Landi, who was in Oxford at the time, has left us a vivid account of the climate of excitement with which, at least among philosophers and aspiring philosophers at Oxford and Cambridge, the publication of the *Philosophical Investigations* was greeted when it appeared in a bilingual edition edited by G. E. M. Anscombe and Rush Rhees, with the original German text and the English translation facing it, thanks to G. E. M. Anscombe, one of Wittgenstein's favourite pupils and, together with Rhees himself and G. H. von Wright, one of the three literary executors appointed by Wittgenstein before his death. Thus, Rossi-Landi recalls: "On 1 May 1953, along with many hundreds of other people in Oxford, I awoke with a particular feeling of anticipation, ate my breakfast in a hurry and ran to Basil Blackwell's in Broad Street to be there when it opened. As had been announced, the first copies of Ludwig Wittgenstein's *Philosophische Untersuchungen* lay shining in the windows. They were bound in dark blue cloth, as befitted the austerity of their contents; but, almost as if to encourage our hopes, they were wrapped in a pale green dust jacket bearing only the title of the facing English translation, *Philosophical Investigations*" (Rossi-Landi 2002, 185).

The history of the interpretations of the *Philosophical Investigations* began on that same day, as Rossi-Landi again reminds us. In fact, on the afternoon of that same 1 May, Anscombe gave a packed lecture in which, in addition to talking about his tradition and the various errors it contained, she also pointed out that one of the merits of the of the *Philosophical Investigations* was "the extremely punctilious, individualising, never generalisable nature of every single observation of the book". "Even now", writes Rossi-Landi, "I can almost hear her tone of voice as she said emphatically, 'what Wittgenstein says in one point should never be connected with what he says in another point', or words to that effect" (Rossi-Land 2002, 186; first published in Italian in Rossi-Landi 1968).

From that 1 May 1953 and over the following decades, the fame and influence of the *Philosophical Investigations* and of their author, hitherto known only as the author of the *Tractatus Logico-Philosophicus* (Wittgenstein 1922) and for the lectures, courses and seminars held at Cambridge from 1930 to 1949, gradually grew to the point that, at the end of the last century, a survey of North American philosophers declared them, perhaps not without some exaggeration, to be the most important philosophical text of the twentieth century (see D. Lackey 1999, 331-2).

This is not the place to reconstruct the history of the *Philosophical Investigations* and their interpretations. Here we can limit ourselves to recalling, first of all, the two questions that have always accompanied the *Philosophical Investigations* and on which an impressive number of essays and books have been written. The first concerns the relationship between the *Philosophical Investigations* and the *Tractatus Logico-Philosophicus*, an issue that Wittgenstein himself brought to the fore when he wrote in the *Preface* that at a certain point he had come to the conclusion that it would be better to publish 'his "old ideas" (the *Tractatus Logico-Philosophicus*) together with the new ones (the *Philosophical Investigations*), believing that "the latter could be seen in the right light only by contrast with and against the background of my older way of thinking". And this was all the more true because, ever since he had returned to philosophy sixteen years earlier (in 1929), he could not fail to recognise that he had made "grave mistakes" (PI, 4) in his first book. For a long time, these remarks of Wittgenstein's were read in a one-sided way, mainly by insisting that he had spoken of "grave mistakes" and "contrast". Later on, things changed not only because the path that led Wittgenstein from the *Tractatus Logico-Philosophicus* to the *Philosophical Investigations* was better and more thoroughly known, but also because several interpreters (especially the so-called "neo-Wittgensteinians'), perhaps just as one-sidedly, emphasised what seemed to them to be strong elements of continuity between the first and second book (see Diamond 1991 and Crary, Rupert 2000).

The second question also has its origin in the *Preface*, particularly where Wittgenstein observes that his book "is really just an album", collecting "a number of sketches of landscapes which were made in the course of [...] long and meandering journeys" through "a wide field of thought" (PI, 2-4). Now, on the one hand, Wittgenstein seems to suggest that this album character depends on the "very nature" of his investigation; on the other, he seems to acknowledge that it is also a consequence of his inability to write a book in which thoughts pass "from one subject to another in a natural, smooth sequence" (PI, 2). In this case, too, the scholars are divided between those who think that Wittgenstein was unable to write the book (in the traditional sense of the term) that he hoped to write (see Hilmy 1989) and those – beginning with Anscombe, as we have seen – who believe that the form of the *Philosophical Investigations* corresponds to Wittgenstein's way of practising philosophy (see, p. e., Pichler 2004). This question, moreover, is closely linked to another and even more crucial one, namely whether the *Philosophical Investigations* should be read in the light of the indications given by their author in the many observations that he devotes to philosophy and its aims and methods (see, in particular PI, §§89-133), or whether it would be better, as many analytic philosophers believe, to set aside these indications and the anti-theoretical and anti-systematic attitude that they express, and to start looking in the *Philosophical Investigations* for theses to discuss, theories (or sketches of theories) to verify and arguments (or sketches of arguments) to evaluate.

As can easily be seen, these two issues are part of (and embedded in) the complex history of the interpretation of the *Philosophical Investigations*, which has been marked by many phases and turning points that deserve careful and close investigation. For example, after an initial phase dominated by interpretations, mainly by Wittgenstein's students, in which the *Philosophical Investigations* was read as the source and inspiration of the so-called "ordinary language philosophy" and as the clearest example of an anti-metaphysical and therapeutic conception of philosophy, there was a long period (roughly coinciding with the last three decades of the last century) in which many Wittgenstein scholars (and others) engaged with the interpretation that Saul Kripke had given to the *Philosophical Investigations* in his 1982 book (Kripke 1982), and in particular of the sections on rule-following and the so-called "private language argument", to the point where it sometimes seemed that Kripke's interpretation was more important than Wittgenstein's text itself. It should be noted that this Kripkean season, which had at its centre a Wittgenstein engaged in posing problems, producing arguments and seeking solutions, marked the closest proximity between the author of the *Philosophical Investigations* and analytic philosophy.

At the end of the last century, a new and different phase (we are not saying "better" or "worse" here) of Wittgensteinian criticism opened up, in which the stage was largely occupied by Cora Diamond, James Conant and, following in their footsteps, by an increasingly numerous (and fierce) group of neo-Wittgensteinians. This phase also saw the rediscovery of interpretations that had been on the fringes of the history of the *Philosophical Investigations*, notably Stanley Cavell's reading (1979), or the new readings of the late G. P. Baker (2006). The most striking novelty of the neo-Wittgensteinians is that they have decidedly opted for a "therapeutic" and "anti-metaphysical" Wittgenstein, without any distinction between *Tracatus Logico-Philosophicus* and *Philosophical Investigations*, albeit in a different sense from that which was present in the readings of the 1950s and 1960s. As Alice Crary points out from the very first page of her *Introduction* to *The New Wittgenstein*, all neo-Wittgensteinians agree that "Wittgenstein's primary aim in philosophy is [...] a *therapeutic* one" (Crary/Rupert 2000, 1).

Obviously, the above is a very partial sketch or outline of the history of the *Philosophical Investigations*. It would be easy to point to many readings and approaches that do not fit in with those just mentioned, and to recall the many debates that have plumbed this or that aspect, this or that passage. To take just one example, consider how much space has been devoted in the literature to the meaning of the reference to Augustine with which §1 opens, or to the many controversial readings that have been made of that "language [...] meant to serve for communication between a builder A and an assistant B" (PI, §2) that Wittgenstein introduces in §2, together with what he will continue to call "the method of §2". But, if we want to broaden our view, we could also recall how the *Philosophical Investigations* have been read and used outside the Wittgensteinian or analytic environment, by thinkers such as Jean-François Lyotard (1979) or Jürgen Habermas (1985). In short, the seventy years of the *Philosophical Investigations* have been rich, intense and complex and Wittgenstein's book has often proved capable, as he hoped, of stimulating many to think for themselves (see PI, 4).

The purpose of this special issue of JoLMA is not to make an (impossible) evaluation of seventy years of philosophical engagement with the *Philosophical Investigations*. What we have set out to do is to give space to a number of scholars, from different backgrounds and with different perspectives , who, over the decades, have addressed the philosophy and philosophical method of the *Philosophical Investigations* in different ways. We have tried, as far as possible, to privilege the diversity of voices, not favouring any particular line of interpretation and not worrying about being faithful to unlikely "orthodoxies". The result, we believe, is an issue that can make a significant contribution to a better understanding of the *Philosophical*

Investigations, but also to the state of contemporary philosophising, or at least of that philosophising that finds in Wittgenstein an important point of reference. Even with regard to the topics to be dealt with, the authors were given complete freedom to choose the subject that best corresponded to their research interests and that they felt could shed some light on the *Philosophical Investigations* and, in some cases, on its influence and presence in later philosophy. The result is an issue that is not merely a container for different essays, but has, as we hoped, has a character that we would like to call "polyphonic".

During the long gestation of this issue, Marjorie Perloff passed away. We would like to recall here the kindness with which she agreed to contribute to this issue, and the beautiful and intense email exchanges we had about Wittgenstein and the many projects she still had in mind. This issue is dedicated to her.

References

Baker, G.P. (2006). *Wittgenstein's Method: Neglected Aspects*. Oxford: Wiley-Blackwell.

Cavell, S. (1979). *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy*. Oxford: Oxford University Press.

Crary, A.; Read, R. (eds) (2000). *The New Wittgenstein*. London; New York: Routledge.

Diamond, C. (1991). *The Realistic Spirit: Wittgenstein, Philosophy and the Mind*. Cambridge MA: The MIT Press.

Habermas, J. (1985). *Zur Logik der Sozialwissenschaften*. Frankfurt a.M.: Suhrkamp.

Hilmy, S. (1989). *The Later Wittgenstein: The Emergence of a New Philosophical Method*. Oxford: Blackwell.

Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.

Lackey, D. (1999). "What are the Modern Classics? The Baruch Poll of Great Philosophy in the Twentieth Century". *Philosophical Forum*, 4, 329-66.

Lyotard, J.-F. (1979). *La condition postmoderne: Rapport sur le savoir*. Paris: Les Editions de Minuit.

Pichler, A. (2002). *Wittgensteins Philosophische Untersuchungen: Vom Buch zum Album*. Amsterdam: Rodopi.

Rossi-Landi F. (1968). *Il linguaggio come lavoro e come mercato*. Milano: Bompiani, 77-127.

Rossi-Landi, F. (2002). "Towards a Marxian Use of Wittgenstein". G. Kitching, G.; Pleasant, N. (eds), *Marx and Wittgenstein: Knowledge, Morality and Politics*. London; New York: Routledge, 185-212.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. London: Kegan Paul, Trench, Trubner & Co.

Wittgenstein, L. (2009a). *Philosophical Investigations*. Revised 4th edition by P.M.S. Hacker and J. Schulte; translated by G.E.M. Anscombe, P.M.S. Hacker and J. Schulte. Chichester: Wiley-Blackwell. First edition: *Philosophical Investigations*. Translated by G.E.M. Anscombe. Oxford: Basil Blackwell 1953. [PI]

Home Language and Philosophers' Language

Lars Hertzberg
Åbo Akademi University, Finland

Abstract   This paper investigates whether Wittgenstein can be considered an 'ordinary language philosopher'. A central role in his thinking is that of what may be called our 'home language' – the language we bring along in coming to do philosophy. The intelligibility of philosophers' language depends on its relation to the home language. This is the central point of *Philosophical Investigations* § 116. Traditional philosophical 'uses' of a word like 'knowledge' have a problematic relation to our customary uses of the word. In consequence, traditional philosophers have sometimes lost the grip on how such words are actually used in human interaction.

Keywords   Ordinary Language. Wittgenstein. Language Use. Philosophers' Language. Knowledge.

Summary   1 Introduction. – 2 Philosophers' Use. – 3 Bringing Words Back. – 4 Conclusion.

1 Introduction

Ludwig Wittgenstein's later work is often held to be the main origin of the philosophical movement known as 'ordinary language philosophy'. I here wish to explore in what sense he can be called an ordinary language philosopher.[1]

When Wittgenstein uses expressions like 'ordinary language', 'ordinary sense', 'ordinary ways of speaking' (*gewöhnliche…*), in the *Philosophical Investigations* (PI),[2] he seems to have different contrasts in mind. A few times, the contrast is simply between, on the one hand, an imaginary form of speech thought up by Wittgenstein for the occasion, and on the other hand customary ways of speaking: thus, in PI, § 19 it is between the one-word commands used in the imaginary builders' game and the customary way of formulating commands; in PI, § 60 it is between someone who refers to an object by listing its parts (he asks for a broom and the stick fitted into it) rather than, as we normally do, to the composite object (the broomstick); in PI, § 243 it is between someone who gives voice to his feelings and moods in a language only he can understand, and the customary way of talking about feelings and moods in a shared language.

At other times, the contrast has to do with the notion of a philosopher stipulating a form of speaking which is, in some sense, assumed to be more adequate than the customary ones: thus, in PI, § 39 the suggestion is that a *name* ought really only to refer to something *simple* (not composite), and thus what we customarily call names are not really names in the strict sense of the word. In PI, §§ 81 and 98 Wittgenstein speaks about the idea that our customary ways of speaking ought to be replaced by a 'perfect' language, that is, presumably, a language in which the logical relations between propositions are supposedly mirrored in their physical form. In PI, § 402 he speaks about the notion that our customary ways of speaking fail to describe things 'as they really are'; thus presupposing the idea that the way we refer to things may or may not correspond to the way reality is constituted. Here, the contrast is between customary ways of speaking and ways of speaking that are, in some sense or another, thought to be philosophically superior.

When the philosophical value of concentrating on ordinary language is debated, the issue is often regarded as a matter of choosing

---

[1] Among fairly recent discussions of ordinary language and philosophy I should like to mention Hanfling 2000, Levi 2000, Baz 2012 and parts of Cockburn 2022. These works are helpful elucidations of the field, and I find myself largely in agreement with the thoughts expressed in them, though I also have some points of disagreement. I find Levi's work particularly incisive on the issue of ordinary language. For reviews of Baz's book, see Levi 2014, Hertzberg 2016.

[2] All references to the *Philosophical Investigations* are to Part I.

between different objects of study. Critics of ordinary language philosophy often allege that a concern with customary forms of speech is intellectually uninteresting or culturally conservative. There is no reason, it is argued, to investigate any but the most sophisticated forms of language currently in use. Thus, it will be thought that it is more fruitful to focus on uses of language accompanying, say, the latest advances in natural science than to concentrate on the conventional talk of ordinary citizens. At other times, again, the issue may be thought of in terms of the goals of philosophical activity. While ordinary language philosophers 'simply attend to the use of words', it is more important to get clear about the reality those words refer to: not just to ask, "How do we use the word 'real'?" but "What *is* the nature of reality?", not "How do we use the word 'know'?" but "What *is* it to know things?", etc. What is to be sought for is the essence of reality, knowledge, the self, the proposition, and so forth. Or then again, the goal of the activity may be thought to be to replace what are seen as the shifting, ambiguous and vague forms of everyday speech – for the purpose of philosophical inquiry if not in everyday life – with a logically exact language in which each well-formed sentence has determinate sense. Bertrand Russell, in a well-known critique, wrote:

> I [...] am persuaded that common speech is full of vagueness and inaccuracy [...]. Everybody agrees that physics and chemistry and medicine each require a language which is not that of everyday life. I fail to see why philosophy, alone, should be forbidden to make a similar approach towards precision and accuracy. (1959, 178)

Contrary to this, J.L. Austin – who is regarded as another originator of ordinary language philosophy besides Wittgenstein – saw a particular value in the study of customary forms of expression:

> [O]ur common stock of words embodies all the distinctions men have found worth drawing, and the connexions they have found worth marking, in the lifetimes of many generations: these surely are likely to be more numerous, more sound, since they have stood up to the long test of the survival of the fittest, and more subtle, at least in all ordinary and reasonably practical matters, than any that you or I are likely to think up in our armchairs of an afternoon [...]. (1970, 182)

Things may not be as straightforward as Austin makes them out to be here. Human life-forms are subject to constant change, and we can hardly think of the evolution of language as linear progress towards ever more useful vocabularies. The situations in which our common words are used may vary greatly over time and context: a

distinction that has stood the test of time in one context may sit awkwardly in another.

I shall get back to the idea of looking for essences further on. When it comes to the idea of linguistic reforms, Wittgenstein does not reject it outright. In PI, § 132 he writes:

> We want to establish an order in our knowledge of the use of language: an order for a particular purpose; one out of many possible orders; not *the* order. For this purpose we shall again and again *emphasize* distinctions which our ordinary forms of language easily make us overlook. This may make it appear as if we saw it as our task to reform language.

> Such a reform for particular practical purposes, an improvement in our terminology designed to prevent misunderstandings in practice, may well be possible. But these are not the cases we are dealing with. The confusions which occupy us arise when language is, as it were, idling, not when it is doing work.

What Wittgenstein is questioning is the idea of a wholesale reconstruction of our language. In PI, § 98 he writes:

> On the one hand, it is clear that every sentence in our language 'is in order as it is'. That is to say, we are not *striving after* an ideal, as if our ordinary vague sentences had not yet got a quite unexceptionable sense, and a perfect language still had to be constructed by us. – On the other hand, it seems clear that where there is sense there must be perfect order. – So there must be perfect order even in the vaguest sentence.

The idea that a language might be in need of wholesale improvement is problematic. Our sentences normally function in the contexts in which they are uttered. However, a limited reform for practical purposes might very well be called for in a given case: thus, it might be found that some part of the vocabulary employed in the context of a specific activity such as astronomy or car repairs is confusing in some respects, and that it needs to be replaced by one that is more transparent. However, the capacity for undertaking such a task is primarily to be found with those involved in the activity – it does not off-hand appear to be an occupation for which philosophers are particularly well suited (conceivably, philosophers might be in a position to contribute, say, when it comes to legal terminology, or to the vocabulary of the human sciences).

The abjuring of customary forms of expression should not be accepted without detailed examination of any alleged problems. Besides, unless we have a clear understanding of the very forms of

expression we use in introducing an allegedly superior language form, our deficient understanding will simply be transplanted into the new language.[3] Hence attention to our customary forms of expression will be necessary in any case.

Let me formulate a truism: philosophical discussion is carried out by means of words, so: in a language. Then where does that language come from? The language in which we come to do philosophy is the language we inhabit, the language we have learnt to speak and understand in living a life with other people. This 'home language' will of course comprise more than the everyday language we all share: it may include the language of specific areas of concern, such as religion, politics, the law, or science. The language we start out with will no doubt be modified in the process of doing philosophy: new forms of expression may gradually become commonplace, professional terms may be introduced, but this too will take place with the home language as a starting point. We have no choice where to start.

In knowing her language the philosopher knows herself. She has grown into her language. Some words may be unfamiliar and she will try to master their use, but language as such is not a skill she is trying to master. Of course, you may raise the question whether she is actually using this or that word in the customary way. You may tell her that in using the word the way she does she is liable to be misunderstood. Suppose she says "I was really annoyed by that waiter", and you ask "Don't you simply mean you were irritated?". Her response may be, "No, I really was annoyed", or "Oh yes, I actually meant to say I was irritated" or perhaps "I never thought about the distinction between annoyance and irritation". In the last case you may try to explain the difference to her, maybe by giving examples of how the two words are used or making clear to her how the words differ in the way they would sit in the context. In accepting our instruction she finds her way back to the language she means to be speaking. (On the other hand, she might insist that the attempt to distinguish the two is pointless and that she is not planning to heed the difference between the words – in which case you may simply shrug your shoulders and wish her good luck.)

---

3 Austin, too, recognised the potential need for linguistic reform. He writes: "[O]rdinary language is *not* the last word: in principle it can everywhere be supplemented and improved upon and superseded. Only remember, it *is* the *first* word" (1970, 183).

2 Philosophers' Use

The language we bring to philosophy, our home language, is not brought along as an object of study (at least not in the ordinary sense), nor as a model for the correct use of language. Philosophy is not of course, say, the study of English. (It would be tempting to call our language a philosopher's tool, as Austin does, although this metaphor is misleading since there is no separate material on which this tool is to be applied.) In a sense, I should like to argue, philosophy has no 'object', and I believe this was Wittgenstein's position. The language we speak becomes a point of focus when, in reflecting, some forms of expression appear to give rise to intractable problems, as when it seems that the word 'I' has no meaning or that any claim to know something is always erroneous. As soon as we agree on how those expressions are used, the problems vanish. (*What* expressions will give rise to problems varies with the language in question.) This point is being made in one of the remarks that are the most frequently quoted in discussing Wittgenstein and ordinary language, PI, § 116:[4]

> When philosophers use a word – 'knowledge', 'being', 'object', 'I', 'proposition', 'sentence',[5] 'name' – and try to grasp the *essence* of the thing, one must always ask oneself: is the word ever actually used in this way in the language in which it is at home? –
>
> What *we* do is to bring words back from their metaphysical to their everyday use.[6]

I would like to make three points about this remark. The first concerns Wittgenstein's use of the word 'philosophers'. What he is referring to here are obviously traditional philosophers – the kinds of philosopher who are the target of his criticism throughout the *Philosophical Investigations*. He evidently excludes himself from this group.

---

[4]   For a meticulous discussion of Wittgenstein's use of the word 'metaphysical' in this remark, see Baker 2006.

[5]   The Hacker-Schulte version (PI) of the translation of the *Philosophical Investigations* gives two words, 'proposition' and 'sentence' where the German has only one: '*Satz*'. This is a response to the vexing translation problem arising from the fact that the German uses only one word for both.

[6]   In German: "Wenn die Philosophen ein Wort gebrauchen – »Wissen«, »Sein«, »Gegenstand«, »Ich«, »Satz«, »Name« – und das *Wesen* des Dings zu erfassen trachten, muß man sich immer fragen: Wird denn dieses Wort in der Sprache, in der es seine Heimat hat, je tatsächlich so gebraucht? – /*Wir* führen die Wörter von ihrer metaphysischen, wieder auf ihre alltägliche Verwendung zurück".

The second and third points require more discussion. The second point concerns the word 'use'. The third concerns the form of the first paragraph of the remark. I shall address each of these points in turn.

Wittgenstein apparently juxtaposes two kinds of use of the words he lists, the (traditional) philosophers' use and the use made of them in the language in which they are at home. However, it is not clear exactly what it means to speak about 'the way philosophers use a word' or *how* we are supposed to compare these.

The notion of a philosophical use also occurs in the last paragraph of PI, § 38:

> Naming seems to be a *strange* connection of a word with an object. – And such a strange connection really obtains, particularly when a philosopher tries to fathom *the* relation between name and what is named by staring at an object in front of him and repeating a name or even the word 'this' innumerable times. For philosophical problems arise when language *goes on holiday*. And *then* we may indeed imagine naming to be some remarkable mental act, as it were the baptism of an object. And we can also say the word 'this' *to* the object, as it were *address* the object as 'this' – *a strange use of this word, which perhaps occurs only when philosophizing.*[7]

Different types of use are also contrasted in PI, § 117:

> I am told: "You understand this expression, don't you? Well then – I'm using it with the meaning you're familiar with". As if the meaning were an aura the word brings along with it and retains in every kind of use.
>
> If, for example, someone says that the sentence "This is here" (saying which he points to an object in front of him) makes sense to him, then he should ask himself in what special circumstances this sentence is actually used. There it does make sense.[8]

On the face of it Wittgenstein's formulation, in contrasting actual uses with various other uses, is contradictory. How are we to understand this? The philosopher imagined in Wittgenstein's remark seems to assume that the sentence uttered is meaningful because it consists of familiar words and its syntax is familiar. However, I would suggest that it is the other way round: unless the speaker's utterance makes sense to us, we have no way of telling how the individual words are to be taken.

---

[7] Last italics mine.

[8] Also PI, § 412 and OC, § 10.

Suppose someone, say, in the course of trying to install a stereo system, points to a wire and says, "This is here"; his assistant might ask, "I'm sorry, which do you mean?" or "Where did you say it is?". However, when a philosopher simply ponders the sentence "This is here?", there does not seem to be any room for his interlocutor to ask which object he means or which place he is referring to. There is no distinction between understanding what the speaker is saying and not understanding it. In that way, the case differs from an 'actual use' of the words.

How then are we to understand the distinction between 'philosophical' and 'actual' uses of words? It is not as if the philosophical and the customary use could be compared like two nomenclatures: the aim of the traditional philosophers is not just to propose a different range of application for our words as if it were a question, say, of different ways of carving up the colour spectrum or different ways of classifying birds. Rather, I would suggest, we are up against *different senses of the word* 'use', though this distinction is not explicitly marked by Wittgenstein; apparently he did not worry about the unclarity.[9]

It is hard to find words by which to mark this distinction, since the word 'use' seems to cover a variety of aspects of linguistic expression. I would like to suggest that what the phrase 'actual use' seems to hint at might be called instances of 'making use of a word' or of 'putting a word to use' – as opposed to a word appearing or occurring in a sentence.

Consider, for instance, how we may make use of the word 'know' and its cognates. I may use the word in an attributive sense, as in "He knows who stole the money", say, as a preface to saying (in one type of case) "so you may ask him" or (in a different type of case) "so you don't need to tell him".[10] Or I may use it to claim knowledge: "*I know* who stole the money – trust me!", or as a declaration: "I know who stole the money: it was…", or as an admission, "Yes, I've known it all along but I didn't want to say", etc.

Now consider, on the other hand, the following passages from philosophical texts, chosen more or less at random, and yet, I hope, possible to recognise as representative of the sorts of thing philosophers are liable to say in discussing issues of knowledge:

Whatever the process and the means may be by which knowledge reaches its objects, there is one that reaches them directly and forms the ultimate material of all thought, viz. intuition. (Kant 1966, 21)

---

9 The same is true of the German. Wittgenstein uses two words here: '*Gebrauch*' and '*Verwendung*', but he seems to employ them interchangeably.

10 Hanfling has a useful discussion of situations in which we attribute knowledge to someone (2000, 94-110).

Our knowledge springs from two fundamental sources of our soul; the first receives representations [...], the second is the power of knowing an object by these representations. (Kant 1966, 44)

We must recognize that when we know something we either do, or by reflecting can, know that our condition is one of knowing that thing, while when we believe something, we either do or can know that our condition is one of believing and not of knowing: so that we cannot mistake belief for knowledge or vice versa. (Prichard 1967, 63)

Since one condition of knowledge is truth, it follows that no belief constitutes knowledge unless it is true. Thus, if our justification fails to guarantee the truth of what we believe, then it may leave us with a false belief. In that case, we lack knowledge. So justification sufficient to ensure us knowledge must guarantee the truth of what we believe. (Lehrer 1974, 79)

When there is some chance that a man is in error, that his belief is incorrect, then there is some uncertainty, however slight, and he does not know for certain that what he claims is true. (Lehrer 1974, 239)

When you *know* that something is so, the thing is absolutely *clear* to you. Thus, no further experience *could possibly clarify* the matter as far as you are concerned. Nothing that could turn up could make it even the least bit clearer to you that the thing is so. (Unger 1975, 141)

I would suggest that there is a clear contrast between the appearance the words 'knowledge' and 'knowing' make in these passages, and the instances given above of making use of the words as in attributing, claiming or admitting knowledge. The distinction should be kept distinct from the classical one between the mention and the use of a word: one typical sign of a word being mentioned is its being put in quotation marks. (I am bypassing for now the problems attaching to the use-mention distinction.) It is true that philosophers in discussing knowledge will occasionally mention rather than use the words 'know' and 'knowledge'; however, in the instances quoted here they are not just mentioning the words – they are purporting to talk *about knowledge*. They are advancing 'philosophical theses', which according to Wittgenstein cannot be done (PI, § 128). Possibly, in some cases when the word 'know' is employed rather than simply mentioned in a philosophical text, the writer's purpose may nevertheless be to say something general about the ways we may make use of the word 'know'. Traditionally, however, philosophers have been

taken to wish to say something about *knowledge as such* rather than talk about uses of the word 'knowledge'. This seems obvious, say, in the quotation from Kant: he is making a remark about the 'faculty of knowledge' and how it functions. In other cases, again, it may not be clear whether a remark is intended as an assertion about uses of the word 'knowledge' or about knowledge as such.

A conception of what it means to write philosophically about knowledge is expressed in the following passage from Keith Lehrer's book *Knowledge*:

> A theory of knowledge need not be a theory about the meaning of epistemic words any more than it need be a theory about how people come to know what they do. Instead, it may be one explaining what conditions must be satisfied and how they may be satisfied in order for a person to know something. When we specify those conditions and explain how they are satisfied, then we shall have a theory of knowledge. An analogy should be helpful at this point. Suppose a man says that there are only two kinds of theories about physical mass. Either a theory of matter is a theory about the meaning of 'mass' and semantically related physical terms, or it is a theory about how something comes to have mass. This dichotomy would be rejected on the grounds that it leaves out the critical question of what mass is, or to put it another way, it leaves out the question of what conditions must be satisfied for something to have a given mass.

> A theoretician in physics might be concerned with precisely the question of what conditions are necessary and sufficient for an object to have mass, or more precisely, to have a mass of n, where 'n' is a variable that would be replaced by a number. Similarly, a philosopher might be concerned with precisely the question of what conditions are necessary and sufficient for a man to have knowledge, or, more precisely, to know that p or that S is true, where 'p' is a variable that would be replaced by a declarative sentence and 'S' by the name of a sentence. (Lehrer 1974, 5ff.)

Lehrer is arguing that there is such a thing as identifying conditions for knowing which are independent of the question how the word 'know' is used. He appears to assume that we may measure our customary ways of making use of the word 'know' against the nature of knowledge as such.[11] Thus, we might imagine cases in which we

---

11 Austin, surprisingly, hints at such a view when he writes: "[W]ords are not [...] facts or things: we need therefore to prise them off the world, to hold them apart from and against it, so that we can realize their inadequacies and arbitrariness, and can relook

would not make use of the word 'know', but where the conditions for attributing knowledge to someone would nevertheless be at hand, and thus an assertion of the form "N.N. knows that p" would be true – would be part of a complete description of how things are in the world. (G.E. Moore holding up a hand and saying "I know this is a hand" might be a case in point.) And on the other hand it might turn out that some of the cases in which a person would customarily be said to know something or other do not in fact fulfil the conditions for being called knowledge, and hence the customary use would be seen to be erroneous. A radical sceptic (such as Unger) would claim that this is true of all attributions of knowledge and knowledge claims, while those who adhere to more limited forms of scepticism would argue that it is true only where the knowledge in question concerns future events, the past, or other people's thoughts and feelings (conceivably the sceptic may add that it is acceptable for practical purposes to attribute knowledge to someone in such circumstances, although the attribution would not be strictly correct).

The comparison of theories of knowledge with theories about mass is not illuminating. Conceivably Lehrer is regarding mass here under the model of a substance like water. There is that which we commonly take to be water, but there is also a chemical formula specifying what water is. Given that, there is a possibility that something taken to be water under the normal criteria is actually some other chemical compound. Applying that to the case of knowledge, there is the possibility that some instances of what to all intents and purposes appears to be a case of a person knowing something are in fact something else. Perhaps, in accordance with Hilary Putman's twin earth thought experiment, there might be twin earth 'knowledge' which is not knowledge at all, though it coincides with what we call knowledge in its manifestations. However, putting it this way, Lehrer's suggestion sounds like a weird fantasy.

Anja Weiberg (forthcoming) has drawn attention to Wittgenstein's use of the word 'subliming' in criticising philosophers' tendency to use models which give a distorted picture of the actual use of words. She quotes his *Remarks on the Philosophy of Psychology*, vol. 2:

> Some will say that my talk about the concept of knowledge is irrelevant, since the concept as understood by philosophers, while indeed it does not agree with the concept as it is used in everyday speech, still is an important and interesting one, created by a sublimation

---

at the world without blinkers" (1970, 182). The picture drawn here is quite problematic. Austin, I am tempted to say, is at his strongest when he practices his skill of taking note of specific verbal nuances and distinctions, not when he is giving an account of the kind of activity philosophy is or might be.

from the ordinary, rather uninteresting one. But the philosophical concept was derived from the ordinary one through all sorts of misunderstandings, and it strengthens these misunderstandings. It is in no way interesting, except as a warning. (RPP, II, § 289)

We may think that the philosopher's 'subliming' of the use of the word 'knowledge' somehow has the power of deepening our understanding of what knowledge is. But there is no basis for such a belief. (One might ponder why philosophers should consider their own – actually made up – conceptions about our words more 'interesting' than the customary ones.)

3 Bringing Words Back

Presumably, it is this form of reasoning that Wittgenstein is rejecting in PI, § 116. This brings us to the third point I wish to make about this remark. I think readers of Wittgenstein have frequently overlooked the fact that the first sentence of PI, § 116 has the form of a question, not a prescription. Wittgenstein means to remind us of something, not to prohibit certain forms of expression. To say that words like those he mentions must never be employed (must never appear) in any way that deviates from that in which they are used "in the language in which [they are] at home" (PI, § 116) (even apart from the point made about the word 'use' above) would be pointless. It would of course be futile to try to prohibit people from deciding to use words in any way they like (though they may have to explain their use if they wish to make themselves understood).

Now, as I was arguing above, what we are to compare here are not really different ways of making use of the same words, but the use we make of certain words on the one hand, and philosophers' alleged claims about the things talked about on the other hand. The philosopher's point is dependent on recognising that in presenting her conception of knowledge, say, she means to be talking about 'the same thing' that is involved in our making use of the word 'knowledge'. The tension arises because the philosophers' claims are supposed to have consequences for the use we make of those words.

Quite often, the conclusion the philosopher ends up with will have the form of the assertion that customary uses of the word in question are illegitimate. Thus, we are told that we cannot claim to have knowledge of some fact unless the possibility of us being mistaken is excluded. Normally, I may say "I know where the car is", without allowing for the possibility, say, that I may misremember where I parked it, or that it may have been towed. So the philosopher's claim would be that I do not really 'know' where the car is (he may concede, however, that the way I was using the word 'know' here is all right for everyday purposes).

'Ordinary language philosophers' are often accused of wishing to make our customary ways of speaking normative for philosophical language – but as it turns out, it is the philosophers who are trying to make their ideas about our words normative for customary speech.

How do philosophers arrive at the claims they make, say, about our inability to know things? Such claims are normally arrived at by a series of steps which starts with making observations about some customary occurrences of the word in question, and then gradually reaches a point where some quite extraordinary assertions are made about what the word means or about the conditions for using it.

In PI, § 116 Wittgenstein seems to be urging us to look back and take note of the long distance we have travelled from the ways the word enters into our customary conversations to the philosopher's claim about the conditions for the word to have application. This exercise is liable to give us a sense of vertigo: it seems every step of the way was incontestable, and yet we ended up in a place which seems totally alien. We find ourselves marooned in space. As Wittgenstein writes in PI, § 107:

> The more closely we examine actual language, the greater becomes the conflict between it and our requirement. (For the crystalline purity of logic was, of course, not something I had *discovered*: it was a requirement.) The conflict becomes intolerable; the requirement is now in danger of becoming vacuous. – We have got on to slippery ice where there is no friction and so, in a certain sense, the conditions are ideal; but also, just because of that, we are unable to walk. We want to walk: so we need *friction*. Back to the rough ground!

In practice, nothing hangs on the philosopher's claim that we cannot know this or that which, in our customary parlance, we claim to know without hesitation. In struggling with the question whether we really ever do know anything, being reminded of a customary use of the word may refresh us like a cool shower on a muggy day, as in Wittgenstein's response to Moore's claim to know that his hand is a hand:

> Why doesn't Moore produce as one of the things he knows, for example, that in such-and-such a part of England there is a village called so-and-so? In other words: why doesn't he mention a fact that is known to him and not to *every one* of us? (OC, § 462)

If Moore had wanted to remind us of some of the uses we typically make of the word 'know', Wittgenstein seems to be saying, he might, for instance, have brought up an example of someone being in a position to inform his interlocutors of some fact.

4 Conclusion

How we respond to the discovery of the distance between the philosophical claim and our customary ways of speaking will ultimately be up to us. On the one hand we may think that the philosophers' 'subliming' of our use of the word 'knowledge' will have the power of somehow deepening our understanding of human knowledge and our relation to the world. Or on the other hand we may feel that somewhere along the way the philosopher lost track of where he was going. We may then be left wondering where the 'decisive movement in the conjuring trick' was made (cf. PI, § 308) – though in the present case we may feel that the philosopher has tricked himself no less than his audience.

If we are differently minded, however, we welcome the philosopher's radical proposal, and impatiently push the everyday example aside as irrelevant and banal. The notion that we can never know anything for certain may seem to have a romantic appeal. For those on the other side of the debate, this will seem to be an illusion. Knowledge attributions and knowledge claims play a role in every type of human interaction they will argue; rather than instil suspicion of their meaningfulness, we should try to make ourselves aware of their role in human conversation.[12]

References

Austin, J.L. (1970). "A Plea for Excuses". Austin, J.L., *Philosophical Papers*. London: Oxford University Press, 175-204.

Baker, G. (2006). "Wittgenstein on Metaphysical/Everyday Use". Baker, G., *Wittgenstein's Methods: Neglected Aspects*. Edited and introduced by K.J. Morris. Malden, MA: Blackwell Publishing, 92-107.

Baz, A. (2012). *When Words are Called For: A Defense of Ordinary Language Philosophy*. Cambridge, MA: Harvard University Press.

Cockburn, D. (2022). *Wittgenstein, Human Beings and Conversation*. London: Anthem Press.

Hanfling, O. (2000). *Philosophy and Ordinary Language: The Bent and Genius of Our Tongue*. London: Routledge.

Hertzberg, L. (2016). "Review of Avner Baz, *When Words are Called For*". *Philosophical Investigations*, 39(1), 92-5.

Kant, I. (1966). *Critique of Pure Reason*. Transl. by M. Müller. Garden City, NY: Doubleday.

Lehrer, K. (1974). *Knowledge*. Oxford: Clarendon Press.

Levi, D.S. (2000). *In Defense of Informal Logic*. Dordrecht: Kluwer.

Levi, D.S. (2014). "Review of Avner Baz, *When Words are Called For*". *Nordic Wittgenstein Review*, 3(2), 187-90.

Prichard, H.A. (1967). "Knowing and Believing". Griffiths, A.P. (ed.), *Knowledge and Belief*. London: Oxford University Press, 60-8.

Russell, B. (1959). *My Philosophical Development*. London: Unwin.

Unger, P. (1975). *Ignorance: A Case of Scepticism*. Oxford: Clarendon Press.

Weiberg, A. (forthcoming). "'Mancher wird sagen, daß mein Reden über den Begriff des Wissens irrelevant sei' (BPP, II, 289). Wittgensteins Kritik am Ideal der 'Sublimierung'". Pichler, A.; Heinrich, E.; Friedrich Stadler, F. (Hrsgg), *100 Years of Tractatus Logico-Philosophicus – 70 Years after Wittgenstein's Death. A Critical Assessment*. Publications of the Austrian Ludwig Wittgenstein Society. New Series. Berlin: De Gruyter.

Wittgenstein, L. (1969). *On Certainty*. Edited by G.E.M. Anscombe; G.H. Von Wright, transl. by G.E.M. Anscombe; D. Paul. Oxford: Blackwell. [OC]

Wittgenstein, L. (1980). *Remarks on the Philosophy of Psychology*, vol. 2. Edited by G.H. von Wright; H. Nyman; transl. by C.G. Luckhardt; M.A.E. Aue. Oxford: Blackwell. [RPP II]

Wittgenstein, L. (2009). *Philosophical Investigations*. Revised 4th edition. Edited by P.M.S. Hacker; J. Schulte, transl. by G.E.M. Anscombe; P.M.S. Hacker; J. Schulte. Chichester: Wiley-Blackwell. [PI]

The Grammar of the Ordinary

Valérie Aucouturier
UCLouvain Saint-Louis Bruxelles, Belgium

Abstract   In this paper I explore what it means to take ordinary language as the raw material of philosophy. To do so, I contrast what I call 'grammar' or the grammatical approach, which is characteristic of L. Wittgenstein's *Philosophical Investigations* and 'ordinary language philosophy' (OLP) as J.L. Austin understood it. I show that, while 'standard' OLP tends to focus on 'historical situations' understood as virtually plausible stagings of our actual uses of concepts, and thus contrasts 'normal' and 'parasitic' uses of language; the grammarian focuses on the logical possibilities of language through the invention of fictitious language-games. The latter thus extends the 'ordinary' up to the abnormal and the extraordinary and rather contrasts it with mere apparent uses that are no uses of language at all.

Keywords   Wittgenstein. Austin. Grammar. Ordinary Language. Ordinary Language Philosophy. Use.

Summary   1 Wittgenstein and Ordinary Language Philosophy. – 2 Clean Tools. – 3 Actual and Fictitious Uses. – 4 Countless Uses? – 5 Ordinary as Opposed to What? – 6 Historical Situations and Natural History. – 7 Practices.

1   Wittgenstein and Ordinary Language Philosophy

The author of the *Philosophical Investigations* is nowadays commonly presented as an 'ordinary language philosopher'. 'Ordinary language philosophy' (from now on OLP) was, and still sometimes is, a rather pejorative label used by its enemies (Warnock 1998) and was meant to designate a philosophical trend that reached its peak, in the 1930's-1960's, mostly at Oxford University, but also in Cambridge, where Wittgenstein held a chair. Famous figures representing this trend were, for instance, J.L. Austin, G.E.M Anscombe, G. Ryle, N. Malcolm, P.F. Strawson, etc. and are so numerous nowadays that they cannot all be mentioned: A. Baz, S. Cavell, C. Diamond, J. Floyd, S. Laugier, C. Travis, etc. Amongst these people, Avner Baz (2012) is an important explicit advocate of OLP against its critics.

However, as Warnock (1998) and others (Mac Cumhaill, Wiseman 2022, 168-9) rightly noted, although ordinary language philosophers from Oxford (such as Ryle and Austin) were largely influenced by Wittgenstein's ideas, Wittgenstein himself and the Wittgensteinians (among which Elizabeth Anscombe) claimed to be doing something different from OLP:

> Among 'Oxford philosophers' [Wittgenstein] was, well before the publication of *Philosophical Investigations* in 1953, the most esteemed and influential of contemporaries; on the other hand he lived and worked, somewhat reclusively, in Cambridge rather than Oxford, and also (less trivially) himself regarded Oxford as 'a philosophical desert', the meagre fruits of which were to him utterly distasteful. [...] Thus it came about that, while Wittgenstein was always conspicuous among those arraigned as 'ordinary language philosophers', he himself would furiously have disclaimed any kinship with the other targets of that critical fire. (Warnock 1998)

It is indisputable that, in the *Investigations*, Wittgenstein himself extensively refers to the "ordinary" (e.g. PI, I, §§ 60, 93, 98, 132, 156, 412, 600), "ordinary language" (e.g. PI, I, §§ 19, 243, 402, 436, 494), "ordinary sense" (PI, I, §§ 39, 256, 344, 418, 420, 536, 615), and so on. But it is also indisputable that there were and still are various ways of inheriting Wittgenstein and that these ways have an important philosophical relevance. This is somehow illustrated in the way Elizabeth Anscombe deals with the issue of sensation (Anscombe 1981a, 11-14) and defends Wittgensteinian "grammar" against "ordinary language philosophy" illustrated by the work of J.L. Austin (1964). In this paper, I explore some of the philosophical differences that characterise this Oxbridge dispute over what it means to start doing philosophy within ordinary language. What does it mean to take 'ordinary language' as a kind of authority to address or consider philosophical issues?

2 Clean Tools

To understand the difference between the grammarian and the ordinary language philosopher, let us begin with a quote by Austin on ordinary language as a starting point. The "method" of OLP, Austin writes, considers that to "proceed from 'ordinary language'" is to "examin[e] *what we should say when*, and so why and what we should mean by it" (Austin 1979a, 181). Namely, it consists in *imagining* situations (185) and in stating "what words we should use in what situations" (182). I will turn later to the issue of "imagining situations". But first, let us remind a famous quote where Austin advocates why ordinary language should be the "first word":

> First, words are our tools, and, as a minimum, we should use clean tools: we should know what we mean and what we do not, and we must forearm ourselves against the traps that language sets us. Secondly, words are not (except in their own little corner) facts or things: we need therefore to prise them off the world, to hold them apart from and against it, so that we can realize their inadequacies and arbitrariness, and can relook at the world without blinkers. Thirdly, and more hopefully, our common stock of words embodies all the distinctions men have found worth drawing, and the connexions they have found worth making, in the lifetimes of many generations: these surely are likely to be more numerous, more sound, since they have stood up to the long test of the survival of the fittest, and more subtle, at least in all ordinary and reasonably practical matters, than any that you or I are likely to think up in our arm-chairs of an afternoon – the most favoured alternative method. (Austin 1979a, 181-2)

According to this passage, there are three aspects of ordinary language which legitimate the method of OLP. First, we, as competent speakers, ought to be clear about what we mean by the words we use; and we must fight the philosophical tendency to let ourselves fall into the traps of language, for instance by an "artificially induced linguistic uprooting" (Bouveresse 1971, 41). This echoes Wittgenstein: "A philosophical problem has the form: 'I don't know my way about'" (PI, I, § 123). Being lost in language is the first disease of philosophy. This is one of the reasons why philosophy requires conceptual clarification. The first claim suggests that, as competent speakers, we should know and be able to recognise proper uses from uses that are not just improper but that should be revealed as being no uses at all. This recognition, Austin argues, will shed light on "the realities we use the words to talk about" (Austin 1979a, 182).

The second claim, that "words are not facts or things", points to the need to reconsider the way, in philosophy, we tend to articulate

language and the world. It echoes Wittgenstein's remarks about the need to fight our tendency to conflate the words with the things they stand for (PI, I, § 38). OLP is needed to distinguish conceptual issues from empirical issues. Actually, later on in his paper on excuses, Austin points to what he calls "the myth of the verb" and reminds us that doings and actions are not ready-made entities that we simply label (Austin 1979a, 178).

Third, and, as we shall see most importantly, this is precisely because words are not facts or things that they can evolve in time, new uses appear, and old uses disappear. We will address this aspect below.

Anyone familiar with Wittgenstein's 'second' philosophy will admit the close kinship between Austin's and Wittgenstein's approaches to philosophy. But, as Wittgenstein himself recommends, we should draw our attention toward the small yet somewhat crucial differences between them. These differences will reveal crucial because of their consequences regarding the scope of OLP (understood here in the broad sense, as a generic term for both Wittgensteinians, Austinians and others in the same trend of philosophy).

3 Actual and Fictitious Uses

Language is the starting point of philosophy. It is the starting point for disentangling our philosophical perplexities, "to show the fly the way out of the fly-bottle" (PI, I, § 309) or "bring words back from their metaphysical to their everyday usage" (PI, I, § 116). It is the starting point to explore how uses of language, language games exhibit the workings of our ordinary practices, actions and categorisations. Over this matter, there seems to be an agreement between (let us call them respectively) the 'grammarian' and the 'ordinary language philosopher'. The purpose of philosophy is not to propose a theory of language (Anscombe 2011), but to *describe* our uses of language in order to avoid its traps. Description of meaning rather than explanation is the method: "We must do away with all explanation, and *description* alone must take its place" (PI, I, § 109).

But if ordinary language is the object of description, we ought to be clear about what this object is. Now, the first and most obvious discordance between the grammarian and the ordinary language philosopher seems to rest on their understanding of language use as a raw material for philosophy. In the above-quoted passage, Austin explicitly refers to "our common stock of words" as the philosopher's raw material. And he suggests, against the 'armchair philosopher' that this raw material is somewhat 'empirical' or at least the result of some historical evolution of language. Whether or not this is faithful to Austin's philosophy as a whole, in this passage, he seems to be

considering that OLP takes meanings and uses as some special sort of historical data, which have been fixed and will evolve in history.

By contrast, the 'grammarian' will rather consider that, although uses are part of our "natural history" (PI, I, § 25, 415), although remarks on language uses "are really remarks on the natural history of human beings" (PI, I, § 415), grammar is transhistorical and transcultural. Grammar does not reflect a state of language at some time and place, it is rather a mean of comparison (PI, I, § 130) between uses. Grammar does not account for the historicity of language, or for the state of logic and language at a time. Grammar is meant to grasp some aspects of language uses that bound meaning and allow us to point and grasp transhistorical and transcultural differences between uses and language games.

The *Philosophical Investigations* "are anything but a collection of meticulous, detailed observations on how our language actually works" (Bouveresse 1971, 31). This is the reason why it allows for the possibility to invent and grasp new or fictitious uses.

> Whenever we make up "ideal languages" it is not in order to replace our ordinary language by them; but just to remove some trouble caused in someone's mind by thinking that he has got hold of the exact use of a common word. That is also why our method is not merely to enumerate actual usages of words, but rather deliberately to invent new ones, some of them because of their absurd appearance. (BB, 28)

> We are not doing natural science; nor yet natural history – since we can also invent fictitious natural history for our purposes. (PI, II, xii)

Grammar is what gets revealed, by contrast, when we imagine new or limit language games, rather than merely explore "what we should say when" in actual, nevertheless imagined, contexts. In other words, grammar concerns as much old, foreign and non-existing uses of language, as actual uses, provided we can make sense of these uses, i.e. imagine at least a situation or a language game in which they would make sense. Grammar exhibits *logical possibilities* of meaning.

> Although at no point does he wish to deny the influence of linguistic change on the birth and evolution of the philosophical problematic, [Wittgenstein] clearly believes that philosophy, as a therapy, is possible and necessary independently of any history of our language and our forms of life, because what threatens us most seriously, from a philosophical point of view, is not the oblivion of this history, of the history of our linguistic usage, but the oblivion of *current*, familiar usage. (Bouveresse 1971, 58)

This abstraction of grammar from (even present) history of language marks a difference with Austin's interest for realistic (and even real) examples. Our ability to circulate between language games (and possibly between languages) and to make sense of far-reaching conceptual systems is reflected in grammar.

4 Countless Uses?

The job of OLP, according to Austin, is somehow more modest. It is to describe (some uses among) a finite, however numerous, number of admitted uses:

> I think we should not despair too easily and talk, as people are apt to do, about the infinite uses of language. Philosophers will do this when they have listed as many, let us say, as seventeen; but even if there were something like ten thousand uses of language, surely we could list them all in time. (Austin 1979b, 234)

These considerations may suggest one way of reading the modality of "should" in Austin's "what we should say when" (Austin 1979a, 181) – rather than e.g. 'what we *could* say when' – as a rather strong philosophical stance toward the method of OLP: actual possible uses supposedly draw the bounds of sense. There is a normative aspect in this "should" that would bound the domain of what we can do with words.

On the contrary the grammarian allows for indefinitely many uses, which can be the expression of the possibility of indefinitely many forms of life, whether actual or fictional, to the extent that "to imagine a language is to imagine a form of life" (PI, I, § 19):

> There are *countless* kinds; countless different kinds of use of all the things we call "signs", "words", "sentences". And this diversity is not something fixed, given once for all; but new types of language, new language-games, as we may say, come into existence, and others become obsolete and get forgotten. (PI, I, § 23)

The "diversity of the tools of language" is *infinite* for Wittgenstein. Philosophy's job is not to make an inventory of existing uses. Philosophy ought to work with this diversity, to create and explore language games until it bumps against the limits of language (PI, I, § 119).

5   Ordinary as Opposed to What?

Consequently, there appears to be a difference regarding the ways the grammarian and the ordinary language philosopher understand the idea that "ordinary language is all right" (BB, 28) as it stands and can therefore work as a starting point or be the "first word" (Austin 1979a, 185) to provide an understanding of meaning and of human interests and practices. This difference can be exhibited by asking: what is the 'ordinary' (use) both Wittgenstein and Austin appeal to and as opposed to what (to what 'standard' or 'norm') is it 'ordinary'? As opposed to 'extra-ordinary'? 'Abnormal'? 'Non-standard'? 'Stretched uses'?

Austin does not invite philosophers to create or imagine new language games or unnatural situations. He seems rather suspicious about this. However, he invites philosophers to imagine *actual* (realistic) situations where our words are or would be at play and confront these 'normal' and 'ordinary' situations to the misuses of philosophy which are often symptomatic of the philosopher's tendency to focus on abnormal cases and take them as central to the understanding of a concept. This is what happens, for instance, when *sense-data* philosophers (see Ayer 1940) jump from the possibility of using 'see' in an abnormal situation, say of illusion, to the idea that what we do directly perceive in any case is not what is there to be seen but mere *sense-data* (Austin 1964). From this perspective, Austin enquires central, 'normal', 'ordinary uses' of words as opposed to 'parasitic uses' (Austin 1962, 104) or 'stretched uses' (Austin 1964, 15, 91).

The grammatical philosopher, on the other hand, takes any use as a use in its own right, provided "language has not gone on holidays" (PI, § 38), i.e. provided it is actually a determined use we can account for. For Wittgenstein, language is, first of all, a logical space where we can explore and invent language uses. For Austin, language is a finite set of uses that evolves through history and circumstances.

This difference sheds light on the reason why Austin prefers the phrase "linguistic phenomenology" rather than "analysis of language" (Austin 1979a, 182) to characterise his own method. Austin takes ordinary language as some special sort of data for philosophers to investigate the possibilities of meaning. This is the reason why he considers OLP not to be another kind of armchair philosophy but a genuine "field work". Ordinary language, Austin says, is "a good site for *field work* in philosophy" (183). The ordinary language philosopher does not invent new concepts that he thinks would best fit reality without even going outside and looking at the world. The ordinary language philosopher takes a certain state of language as its raw matter and enquires into its uses, tries to disentangle them, thus shedding light on reality. Therefore, Austin does not hesitate to start his enquiry with the dictionary (186-7). According to Austin's

conception of OLP, the current state of language is indeed a result of its evolution through history. Language is good as it stands because it rests on "the inherited experience and acumen of many generations of men" (185).

Of course, to a great extent the grammarian and the ordinary language philosopher are very close. They agree that philosophy should abandon the quest for essences "and get down to the dainty and the dumpy" (183), and "bring words back from their metaphysical to their everyday use", "in the language where it is at home" (PI, I, § 116), "back to the rough ground" (PI, I, § 107). But for the one the ordinary is embodied in a "historical situation", whereas for the other the ordinary goes beyond historicity to characterise a certain 'form of life'.

If one proceeds with the assumption that some substantial difference can be identified between Wittgenstein and Austin in their understanding of the ordinary, one can conclude from this section that, from Austin's perspective, the 'ordinary' is akin to 'normal uses' and their instantiation in virtually possible situations. Whereas the 'ordinary', from Wittgenstein's perspective, can take various aspects. It can amount to the description of rather familiar and central uses of language. But extended or marginal uses are no less part of the grammar (meaning) of our concepts. This is the reason why, according to Wittgenstein, we can imagine improbable or 'abnormal' or 'extraordinary' situations *within* grammar. The default of philosophy, when it bumps its head against the limits of language, is not to invent eccentric uses of language, but to lose track of uses tout court and get lost in its own language.

6 Historical Situations and Natural History

Now that I have sketched these differences, I would like to briefly explore their consequences for my initial question concerning the relevance and scope of ordinary language philosophy (in the broad sense).

Whereas OLP excludes some uses of language, for Wittgenstein any possible use (even one that has never been actualised or that instantiate an alternative 'natural history') is a use, whether it is an actual, past or invented use, as long as we can make sense of it, i.e. as long as we can imagine a situation (even an unlikely situation) where it would make sense.

On the other hand, for Austin, non-actual or unlikely situations have limited authority in OLP. Austin is confident that, at least to some extent, we would say similar things in similar *historical* situations (i.e. whether actual or not but which instantiate some existing use of a phrase), granted that the situation has been described fully enough:

> The more we imagine the situation in detail, with a background of story […] the less we find we disagree about what we should say. (Austin 1979a, 184)

And disagreement over 'what we should say when' does not constitute a counterexample but a further opportunity to clarify our uses:

> A disagreement as to what we should say is not to be shied off, but to be pounced upon: for the explanation of it can hardly fail to be illuminating. (Austin 1979a, 184)

We may sum up this difference by saying that the starting point of OLP according to Austin are actual (or virtually actual) uses of language, considered in a historical situation, whereas for Wittgenstein it is the "form of life", understood as the point wherefrom a logical space is shaped. This logical space is bounded by local language games. To understand a use of language, we need to picture the world or reality where this use makes sense. This is obvious in the following remark:

> I am not saying: if such-and-such facts of nature were different, people would have different concepts (in the sense of a hypothesis). Rather: if anyone believes that certain concepts are absolutely the correct ones, and that having different ones would mean not realizing something that we realize then let him imagine certain very general facts of nature to be different from what we are used to, and the formation of concepts different from the usual ones will become intelligible to him. (PI, II, xii, § 366)

Elizabeth Anscombe (1981b) has commented extensively on this quote in her paper on linguistic idealism. There is a parallel mistake against which both Austin and Wittgenstein are fighting. The mistake consists in thinking that we are bound to chose between saying that either words are (or ought to tend to be) faithful representations of reality, or reality is shaped by words (which would somehow magically create what they are meant to represent). But words are neither things we pick up or discover in our environment, nor are they arbitrary productions of our imagination.

> This dubious companionship between an ethereal extra-linguistic reality, presumed to be stable, and an inconstant linguistic associate, which has its own life and avatars, is in a sense responsible for all philosophical perplexities. (Bouveresse 1971, 58)

Reality exercises empirical constraints on language, but still concepts are made by us for our practical purposes. Of course, we could claim that describing colours properly would imply to provide an

infinite range of words for shades of colours (or even an infinite range of words for shades of red, of blue, etc.), until we reach a virtually un-reachable adequate description of the spectrum of colours. But such a fantasy proves to be forgetful of the fact that our descriptions have contextual and practical purposes: sometimes we need to distinguish some varieties of red, sometimes 'red' is enough. "We can introduce as many new distinctions as we like, but we cannot set ourselves the goal of introducing them *all*" (Bouveresse 1971, 59). That "essence is expressed in grammar" (PI, § 371) only means that "grammar im-poses certain forms of description on us, but we cannot, despite our best efforts, exhibit any 'reality' that would justify them and make all others impossible" (Bouveresse 1971, 48).

Austin and Wittgenstein seem to have distinct understandings of the role of "historical situations" (Austin 1979a, 186) and "natural history" (PI, I, § 25): a "historical situation" being rather a virtual-ly actual situation given the state of our present uses of language and "natural history" being rather an anthropological variable that helps us realise the contingent articulation between our form(s) of life and our linguistic practices. However, both understandings converge toward some sort of what Cora Diamond calls a "realistic spirit" (Diamond 1995), i.e. the idea that the contingency of ordi-nary uses has nothing to do with plain arbitrariness, but is rather constrained by our form(s) of life and our environment.

7 Practices

This distinction between historical situations and imagined situa-tions or logical spaces, between actual uses and forms of life, reveals several difficulties OLP (in the broad sense) may face.

First, a difficulty with Austin's insistence on 'normal' versus 'par-asitic' uses within some actual state of language, is that it may lead to undermine or hierarchise the variety of uses instead of enlight-ening their intertwinements (Anscombe 1981a). Is not there a kind of arbitrariness, in our philosophical remarks, at least to the extent that they are necessarily localised in history and even in a certain social class and language register?

On the other hand, if we consider the a-historical perspective of the grammarian, we may wonder whether she will not fall in the trap that OLP is made to avoid, namely, forget where it speaks from and essentialise grammar and raise philosophical remarks to a transcen-dental level. In other words, the risk is to fall down the parallel mis-take, which would be to overestimate the scope of what we can say in philosophy. Too much localisation of our starting point threatens to lead to triviality, whereas too few localisations of our starting point threatens to lead to exaggerated generality.

I do not think these difficulties are overwhelming or insurmountable, as many advocates of OLP have shown.[1] Although there is no room here to do them justice, let me conclude by suggesting that the key to these difficulties is to be found in a philosophy of action (Aucouturier, forthcoming): something both Wittgenstein and Austin clearly saw.

Indeed, Austin and Wittgenstein both agree that considering ordinary language to be "the first word" soon leads to question the traditional view of language as mainly aiming at truly representing states of affairs. One of the great ideas they have in common is the idea that philosophers should focus more on the various things we *do* with words, the various functions words can have, rather than being obsessed with the question of truth.

> Certainly [...] ordinary language is *not* the last word: in principle it can everywhere be supplemented and improved upon and superseded. Only remember it is the *first* word. [footnote: And forget, for once and for a while, that other curious question 'Is it true?' May we?] (Austin 1979a, 185)

If the main function of language is not to adequately represent the world, but rather to serve our indefinitely various practical purposes in given circumstances, we should not think of the evolutions of language as a succession of attempts to best represent states of affairs.

For Austin, as well as Wittgenstein, language serve our interests and must be understood in relation to what interests us. "Concepts are the expression of our interest and direct our interest" (PI, I, § 570). Therefore, the evolution of language can be understood on the ground of the evolution of practical human interests in accordance with a given situation. If ordinary language is indeed the 'first word', Austin reminds us, it certainly is not the last word:

> If a distinction works well for practical purposes in ordinary life [...], then there is sure to be something in it, it will not mark nothing: yet this is likely enough to be not the best way of arranging things if our interests are more extensive or intellectual than the ordinary. (Austin 1979a, 185)

Indeed, the set of situations we may imagine is not given or closed. For Austin, being confronted to new situations and/or interests, for Wittgenstein, imagining radically new, yet unimagined situations, opens and extends the logical space. Now the mere difference seems to be that, for Austin, it is when confronted to an unheard-of situation

---

1 See e.g. Diamond 1995; Cavell 2000; Baz 2012; Laugier 2013.

that language – understood as some sort of historical data – gets modified; for Wittgenstein, we can still imagine unheard-of situations and new language game. This does not entail these new language game will become actual ones. This just reveals the current workings of language. Historical situations may call for the need of new language games, but philosophy itself is not primarily interested in these language games, simply in what makes them logically possible (and meaningful).

So, although Austin focuses on the actuality of language-uses when Wittgenstein insists on the potentialities of language to draw the bounds of sense from within ordinary language, both philosophers agree on the necessity of thinking language in continuity with what we *do*. Language is part of our practices, and this is the reason why philosophy needs to focus on what we do with word.

This does not call for a theory of language (Anscombe 2011), but for a special attention to the question: how did we come to 'learn' the use of this word or concept? What sort of form of life do we need to share in order to be able to use a word or phrase – i.e. with a shared meaning? And the answer is not to be found in the objectivity of a phenomenal world, but in the regularity of what we do with words in a given context and environment.

References

Anscombe, G.E.M. (1981a). "The Intentionality of Sensation. A Grammatical Feature". Anscombe, G.E.M., *Metaphysics and the Philosophy of Mind: Collected Philosophical Papers*, vol. 2. Oxford: Basil Blackwell, 3-20.

Anscombe, G.E.M. (1981b). "The Question of Linguistic Idealism". Anscombe, G.E.M., *From Parmenide to Wittgenstein: Collected Philosophical Papers*, vol. 1. Oxford: Basil Blackwell, 112-33.

Anscombe, G.E.M. (2011). "A Theory of Language". Geach, M.; Gormally, L. (eds), *From Plato to Wittgenstein*. Exeter: Imprint Academic, 193-203.

Aucouturier, V. (forthcoming). "Wittgenstein, Meaning and Action". Cahill, K.M. (ed.), *Wittgenstein on Practice: Back to the Rough Ground*. Basingstoke: Palgrave Macmillan.

Austin, J.L. (1962). *How to Do Things with Words*. Oxford: Clarendon Press.

Austin, J.L. (1964). *Sense and Sensibilia*. Edited by G.J. Warnock. Oxford: Oxford University Press.

Austin, J.L. (1979a). "A Plea for Excuses". Austin, J.L., *Philosophical Papers*. Oxford: Oxford University Press, 175-204.

Austin, J.L. (1979b). "Performative Utterances". Austin, J.L., *Philosophical Papers*. Oxford: Oxford University Press, 175-204.

Ayer, A.J. (1940). *The Foundations of Empirical Knowledge*. New York: Macmillan.

Baz, A. (2012). *When Words Are Called For*. Cambridge, MA: Harvard University Press.

Bouveresse, J. (1971). "Langage ordinaire et philosophie". *Langages*, 6(21), 35-70. Numéro thématique *Philosophie du langage*.

Cavell, S. (2000). *Must We Mean What We Say?* Cambridge, MA: Cambridge University Press.

Diamond, C. (1995). *The Realistic Spirit. Wittgenstein, Philosophy and the Mind*. Cambridge, MA: The MIT Press.

Flew, A. (1954). "Philosophy and Language". *Philosophical Quarterly*, IV, 21-36.

Gellner, E.A. (1959). *Words and Things: A Critical Account of Linguistic Philosophy and a Study in Ideology*. London: Gollancz.

Laugier, S. (2013). *Why We Need Ordinary Language Philosophy*. Transl. by D. Ginsburg. Chicago: The University of Chicago Press.

Mac Cumhaill, C.; Wiseman, R. (2022). *Metaphysical Animals. How Four Women Brought Philosophy Back to Life*. London: Penguin.

Ryle, G. (1964). "Ordinary Language". Chappell, V.C. (ed.), *Ordinary Language*. Englewood Cliffs: Prentice-Hall, 24-40.

Travis, C. (2001). *The Uses of Sense. Wittgenstein's Philosophy of Language*. Oxford: Oxford University Press.

Warnock, G. (1998). s.v. "Ordinary Language Philosophy, School of". *The Routledge Encyclopedia of Philosophy*. https://www.rep.routledge.com/articles/thematic/ordinary-language-philosophy-school-of/v-1.

Wittgenstein, L. (1965). *The Blue and Brown Books*. Oxford: Blackwell. [BB]

Wittgenstein, L. (2009). *Philosophical Investigations*. Oxford: Wiley-Blackwell. [PI]

Xenakis, J. (1959). "Ordinary Language Philosophy: Language, Logic and Philosophy". *Synthese*, 11(3), 294-306.

The *Philosophical Investigations* in Philosophy of Religion

Thomas Carroll

The Chinese University of Hong Kong (Shenzhen), China

Abstract Despite overlooking religious topics, Wittgenstein's *Philosophical Investigations* [PI] has had a large impact in philosophy of religion. This article surveys that influence and the reasons for it. In what follows, I first describe the reception of certain key concepts from the PI in philosophy of religion. Second, I examine a few scattered remarks on religious topics in the PI. Third, I consider the relevance of the PI for contemporary philosophy of religion. I argue that the dialogical nature of the PI, allowing different generations of readers to engage it with their particular philosophical problems, is key to its long-term influence.

Keywords Philosophy of religion. Religious language. Remarks on religion. Fideism. Metaphilosophy. Historical reception of Wittgenstein.

Summary 1 Introduction. – 2 Themes in the Reception of the PI in Philosophy of Religion. – 2.1 Language-Games. – 2.2 Forms of Life. – 2.3 Family Resemblances. – 2.4 Grammar. – 2.5 Aspect Perception. – 2.6 Metaphilosophy. – 3 References to Aspects of Religions in the PI. – 3.1 Prayer. – 3.2 God. – 3.3 Soul. – 4 Globally Engaged Philosophy of Religion. – 5 Conclusion.

1 Introduction

While Wittgenstein's *Philosophical Investigations* [PI] had an enormous influence on analytic philosophy generally, particularly in the third quarter of the twentieth century, the book also had a large influence on twentieth century philosophy of religion. Its role in Wittgensteinian philosophy of religion was especially substantial as the text was one of the earliest and most authoritative sources available to philosophers of religion who did not know Wittgenstein personally. This is perhaps somewhat ironic since Wittgenstein barely makes reference to religiosities within the book. As Genia Schoenbaumsfeld remarks,

> Wittgenstein published next to nothing on the philosophy of religion and yet his conception of religious belief has been immensely influential. While the concluding, 'mystical' remarks in his early work, the *Tractatus*, are notorious, we find only a single allusion to theology in his *magnum opus*, the *Philosophical Investigations*. (Schoenbaumsfeld 2014, 162)

Schoenbaumsfeld rightly directs her readers' attention to other Wittgenstein sources since published, such as the *Lectures and Conversations on Religious Belief* (1967), the "Remarks on Frazer's *Golden Bough*" (1993), and the miscellaneous collection of remarks known as *Culture and Value* [1977] (1998). Yet, as the first and most polished work in Wittgenstein's corpus dating from his later period of philosophical activity, the PI has long been seen as the most authoritative source for Wittgenstein's mature philosophy. Furthermore, as the earliest publication from Wittgenstein's more mature period – published now seventy years ago – the text has had a long time in which to make its impact felt across the subfields of philosophy, including philosophy of religion (Carroll 2014, 31).

As one of the most important texts of mid-century analytic philosophy, the PI was bound to influence many fields across the discipline, from philosophy of language and mind to aesthetics and even to some extent political philosophy. The text has continued to stimulate topics in philosophy of religion steadily over time, meaning that philosophers have drawn lessons from the book now across multiple generations. Naturally enough, the lessons drawn from the book have been shaped by the philosophical situations of respective eras of philosophers – from the lingering threat of verificationism to the meaningfulness of 'religious language' to the prospects for inter-religious and inter-cultural dialogue.

In this article, I explore the influence of the PI in philosophy of religion in three ways. First, I explore the reception of certain key ideas from the text, such as 'language-games', 'forms of life', and

'family resemblances' by philosophers of religion. Second, I examine the trace references to religiosity in the PI and how these passages are relevant to philosophy of religion. Third, I conclude with some observations on recent developments of philosophy of religion that are influenced by the PI, especially concerning globally engaged philosophy of religion.

2 Themes in the Reception of the PI in Philosophy of Religion

2.1 Language-Games

This notion of a 'language-game' (*Sprachspiel*) could well be the most discussed topic in secondary literature on the PI. Thus, it is not surprising that it would be a major focus of work in philosophy of religion. The idea of a language-game is introduced very early in the PI (I, § 2). In that passage, just following the well-known opening remark on Augustine and his theory of language-learning, Wittgenstein introduces the simple or 'primitive' instance of the language use of a pair of builders and their routinised use of expressions by builder A to order builder B to produce a 'block', 'pillar', etc., in the joint activity of building a structure. In PI, I, § 7, Wittgenstein refers back to this localised instance of language activity as a "language-game".

A key reason for the appeal of this social picture of language in use is how it reframed what it is for language to have meaning. The lingering problem of verificationism persisted in some corners of philosophy well into the second half of the twentieth century. Reductive naturalistic metaphysics continued to prevail when it came to the consideration of language with supernatural and other sorts of unverifiable components: references to gods, spirits, and souls. From A.J. Ayer (1935) to Anthony Flew (1955, 98), scepticism about so-called 'religious' or 'theological' language because of the imperceptibility of its putative referents led many naturalists to suppose that such language was meaningless. The development of Wittgenstein's later philosophy along with ordinary language philosophy allowed for possibilities of understanding meaning in language that went beyond reference. This is not to say that Wittgenstein thought reference was unimportant. It is just that, as Wittgenstein develops through his remarks on Augustine, ostension is not how most language is learned or functions; language instead has a seemingly endless variety of possible uses. While other avenues, such as Alvin Plantinga's burden-shifting common-sense realism about theistic claims (cf. Plantinga 1967), would appear in the next decade,

the evident usefulness of the paired notions of language-games and forms of life in the PI would open new avenues for the interpretation of religions (Malcolm 1960, 56).

Following on the verificationist critique of theological expressions, Wittgenstein's idea of language-games inspired philosophers of religion to consider the roles of 'religious language' within religious practice. An early example relevant to philosophy of religion comes from Peter Winch in *The Idea of a Social Science*:

> A religious mystic, for instance, who says that his aim is union with God, can be understood only by someone who is acquainted with the religious tradition in the context of which this end is sought; a scientist who says that his aim is to split the atom can be understood only by someone who is familiar with modern physics. (Winch 1990, 55)

Appearing a mere five years after the publication of the PI, Winch's book helped inaugurate some core themes of Wittgensteinian philosophy of religion. In conveying the importance of deep attention to social context for the understanding and interpretation of religious language, Winch also highlights the importance of paying attention to the end of the social activity in question. The idea here, eventually commonplace in Wittgensteinian philosophy of religion, is that 'religious language' must be interpreted with an eye to the religious activities in which the language is meaningful. P.F. Bloemendaal observes that Winch's emphasis on the seemingly endless variety of human "modes of social life" and the necessity of interpreting them according to their own criteria set the stage for accusations of the epistemic isolation of instances of social life from one another (Bloemendaal 2006, 112). While Winch's work is most clearly relevant to anthropology of religion, it has also been highly important to the development of Wittgenstein philosophy of religion in general. Indeed, due to the practice-oriented approach of Wittgenstein's philosophy – and the eventual publication of Wittgenstein's "Remarks on Frazer's *Golden Bough*" ([1967] 1993) – links between anthropology and Wittgensteinian philosophy of religion have frequently been made.

Bloemendaal also identifies Norman Malcolm's early contribution to Wittgensteinian philosophy of religion as underlining similar themes drawing on ideas from the PI for philosophy of religion: first, through reference to 'religious language' and second, through conceiving of religions as language-games (Bloemendaal 2006, 199). While 'religious language-games' have often been the focus of Wittgensteinian philosophy of religion, subsequent philosophers – such as Rush Rhees and D.Z. Phillips – have added more nuanced or focused

analyses of the significance of conceiving of aspects of religions in light of the notion of a language-game (cf. Von Der Ruhr 2009, 223).

Wittgenstein's student, friend, and literary executor, Rush Rhees, is another centrally important figure in the early development of Wittgensteinian philosophy of religion. Because Rhees knew Wittgenstein personally, his reception of Wittgenstein's ideas and philosophical methods goes far beyond the PI; however, since Rhees co-edited the book with G.E.M. Anscombe, it is not surprising that one can see themes from the book appear in Rhees's writings. While many of Rhees's writings were not published until much later when they were edited by his former student and colleague D.Z. Phillips (and later literary executor), Rhees had a large influence along with other members of the Swansea School – including Winch and Phillips – on what Wittgensteinian philosophy of religion would become during the early decades after Wittgenstein's death. Notably, Rhees was a sympathetic critic of Wittgenstein's, especially when it came to the interpretation and use of the notion of a language-game (Rhees 1960). Rhees argued that Wittgenstein's remarks in PI lent themselves to the idea that language-games were autonomous smaller instances of language, rather than useful abstractions of actual language use. For this reason, Rhees preferred the notion of "conversation" to language-game when describing the use of language in the flow of life (Von Der Ruhr 2009).

Rhees's student and colleague D.Z. Phillips was also an early interpreter of Wittgenstein with respect to philosophy of religion. In his 1970 essay, "Religious Beliefs and Language-Games", Phillips seeks to defend his interpretation of Wittgenstein's bearing on philosophy of religion from numerous criticisms of an isolationist understanding of language-games (Phillips 1993). A key feature of criticism of isolationist readings of language-games is that they remove the role for religious argumentation (e.g., foundationalist versions of theistic arguments). To some philosophers of religion and Christian apologists, this renders the isolationist language-game view of religion to be absurd. Phillips counters that a language-game reading of religion should not be understood as isolating such language-games from other parts of life. In this respect, he echoes themes argued for by Rhees. In order for religious beliefs to have the importance they clearly have for those who hold them, they would need to be related to many aspects of a person's life. Yet, according to Phillips, a difference in the grammar of religion and those areas of language involved in giving proofs should be observed.

This Phillips essay offers an early example of a particular genre of writing on Wittgenstein and philosophy of religion, the correction of exaggerated or otherwise perceived inaccuracies in interpretation. It is of a piece with the critique of scientism one finds elsewhere in Wittgenstein's writings. Three years before the publication

of Phillips's essay, Kai Nielsen had published his highly influential article, "Wittgensteinian Fideism" (1967), which called into question what Nielsen saw as the isolationist, self-protective – or "fideistic" – character of Wittgensteinian approaches to understanding religious language. 'Fideism' is a term that has been used by philosophers and theologians to refer to a variety of viewpoints on the epistemic standing of religious beliefs. Most, but not all, uses are pejorative, signalling an epistemically defective approach downplaying the role of reason or enquiry in grounding faith (Carroll 2008, 19). In this way, Nielsen's understanding of 'fideism' was in line with that of many secular critics of theistic religious discourses, particularly Christianity; interestingly, this criticism was mirrored by traditional Protestant and Catholic philosophers who sought to maintain the viability of natural theology. The idea is that Wittgensteinian philosophy of religion presented a picture of religious discourse where its intelligibility rested on a commitment that could only be understood by those who held it. To the extent that it provided an accurate depiction of Wittgensteinian philosophy of religion or not, Nielsen's article identified features of an excessively relativistic picture of 'religious language-games', something that both came to frame the philosophical lore about Wittgenstein and philosophy of religion and to provide an example of deficient interpretation of Wittgenstein vis-à-vis religion.

Brian Clack offers a helpful overview of the early history of the development of Wittgensteinian philosophy of religion, from Malcolm to Nielsen in his *An Introduction to Wittgenstein's Philosophy of Religion* (1999). Clack observes that any tendency to think of religions as language-games misunderstands how Wittgenstein introduced and developed the notion:

> Though Wittgenstein never attempted a definition of a 'language-game', the examples he provides of these linguistic phenomena do not suggest that he had in mind anything as large as science or religion, or indeed any practice or institution whatsoever. Language-games seem, rather, to be quite small-scale units of language-usage which occur in various human contexts. (Clack 1999, 87)

This note of interpretive caution reflects the sorts of contributions Wittgenstein scholars would make by way of correction of early extravagances when it comes to the interpretation of religion.

2.2 Forms of Life

The expression "form of life" ("*Lebensform*") appears just a handful of times in Wittgenstein's corpus, and just five times in the PI; yet, the notion has had a quite significant influence in philosophy of religion. While the expression predates Wittgenstein's writings, it was the PI that brought the concept to the attention of a broad audience. While the narrow idea of a language-game conveys the uses of language in a specific social activity, the broad notion of a form of life suggests the comprehensiveness of the social life of a community of people that grounds or puts into motion the meaningfulness of language, including particular instances of language. It is notable that Wittgenstein uses the two expressions to inform each other. Yet, from the context of a few remarks in PI, it is not entirely clear what is to be understood by the expression:

> It is easy to imagine a language consisting only of orders and reports in battle. — Or a language consisting only of questions and expressions for answering Yes and No — and countless other things. —— And to imagine a language means to imagine a form of life. (PI, I, § 19)

Here, we have a relation between the local instances of language use – language-games – and a form of life: "[T]o imagine a language means to imagine a form of life". Juliet Floyd argues that this notion replaced Wittgenstein's earlier embrace of "culture" ("*Kultur*") as capturing what lay behind and informed the use of language (Floyd 2020). What one imagines in imagining language is all of the functions that language performs within the lives of people.

Winch's analysis of interpretation draws on the notion of a form of life to identify distinct discursive practices and traditions. Winch writes:

> [W]hereas the philosophies of science, of art, of history, etc., will have the task of elucidating the peculiar natures of those forms of life called 'science', 'art', etc., epistemology will try to elucidate what is involved in the notion of a form of life as such. (Winch 1990, 41)

That is, in order to interpret what it is to know something in a particular area, one must first attend to the social practices of conceiving and gathering knowledge in that area. Thus, Winch takes sociology and epistemology to be much more closely linked than is commonly thought.

As with language-games, Malcolm links religions with forms of life very closely, and likewise holds that understanding of a form of

life will be closely associated with participation *within* that form of life. Malcolm writes about Anselm and the ontological argument:

> At a deeper level, I suspect that the argument can be thoroughly understood only by one who has a view of that human 'form of life' that gives rise to the idea of an infinitely great being, who views it from the inside not just from the outside and who has, therefore, at least some inclination to partake in that religious form of life. (Malcolm 1960, 62)

While Winch allows that one well acquainted but as yet outside of the way of life being studied could still understand it, Malcolm's view seems to have been that participation is necessary to understanding. As mentioned above, the tendency towards aversion to theistic argumentation can be seen in Rhees's critique of natural theology (cf. Rhees 1969). While Malcolm defends a minimal role for ontological arguments – "it may help to remove some philosophical scruples that stand in the way of faith" – nevertheless, such arguments gain their force within the context of a religious form of life.

In an essay of Rhees's titled "Religion and Language", published in 1969 but written earlier as a philosophical letter, one sees the expression of numerous themes that would appear frequently in philosophical works on Wittgenstein and religion. First, there is the assertion that religious language and religious life are "internally related" (Rhees 1969, 120). In this vein, Rhees compares "religious language" with the "language of love". Rhees writes:

> And people who have tried to understand love – or explain it – by approaching it from biology have got nowhere; and they generally end by ignoring it. If men come to love women, and if men come to love God, this has to do with the life which they lead and in which they take part. (122)

Second, there is the focus on religion in the singular, which can be presented as abstract (potentially applying to all religions) or specifically, which is always synonymous with forms of Christianity. In this respect, Rhees is no different from most of his contemporary peers in philosophy of religion. Third, there is the idea that religious language is different in grammar than other forms of discourse. Rhees writes:

> "God exists" is not a statement of fact. You might say that it is not in the indicative mood. It is a confession – or expression of faith. (131)

This does not mean that "God" does not refer to something, but the reference will be different from ordinary physical objects because the grammar of the two is different. In saying that "God exists" is not in

the indicative mood, Rhees is opening the door to pragmatics in the analysis of religious language. Furthermore, Rhees lays the groundwork for an anti-scientistic argument. Fourth, there is the idea that language about God, spoken by religious people, is more confessional than referential. That is, religious language (understood prototypically as Christian) functions within liturgical contexts as well as moments of exhortation, prayer, and fellowship.

Some critics of Wittgensteinian philosophy of religion have claimed that the emphasis on understanding local contexts of language use amounts to or potentially leads to "protective strategies" (Proudfoot 1987), "fideism" (Nielsen 1967), or "relativism" (Trigg 1983). The metaphor of combat in the first charge is notable, as it evokes the potentially competitive relationship between religious and philosophical forms of language. Admittedly, commentators such as Malcolm, Rhees, Winch, and Phillips emphasise a stark difference between scientific and religious modes of discourse. It is not surprising that these views were interpreted as conveying the incommensurability of religious and scientific discourses (even as a careful reading of the sources reveals more nuance than critics generally register). As Clack would remind us, embracing contextualism in interpretation need not lead to protective strategies; it can lead to atheism:

> This is not an atheism based on denying the existence of super-empirical realities (religion never was about that), nor is it the rebellious atheism of an Ivan Karamazov, nor yet is it the positivistic atheism of denying sense to religious propositions. It is, rather, a despairing, apocalyptic atheism that arises from Wittgenstein's philosophy of religion, the frustrated and bitter recognition that the passionate beauty of the religious life is no longer open to us. (Clack 1999, 129)

The decline in the plausibility of a mode of expression or form of life can happen as one comes to see religions as rooted in instinctual feelings (as Clack interprets the "Remarks on Frazer's *Golden Bough*") and not in a really existing God.

Patrick Sherry and Richard Bell are relatively cautious about the application of ideas such as language-games and forms of life to religions. This is because they both view these ideas as being highly local descriptions of the social activities with language, and the forms of agreement necessary for the social activities to work. Thus, it does not make sense to think of a whole religion – or, indeed, "religion" itself – being a language-game or form of life. Instead, these ideas, if they are to be applied to the interpretation of religions, should be applied to highly specific social aspects of religious activity (e.g., this form of worship in this tradition). Reminding his readers to turn again to Wittgenstein's texts, Sherry cautions against the enthusiasms of important and influential figures like Malcolm.

2.3 Family Resemblances

Wittgenstein introduces the notion of "family resemblances" ("*Familienähnlichkeiten*") a little bit later in the PI, just following his argument against language having an essential feature. The forms of language are varied, as Wittgenstein remarks in § 65:

> Instead of pointing out something common to all that we call language, I'm saying that these phenomena have no one thing in common in virtue of which we use the same word for all — but there are many different kinds of *affinity* between them. And on account of this affinity, or these affinities, we call them all "languages". (PI, I, § 65)

Wittgenstein then lists in § 66 many examples of things that we call games and their lack of a single uniting feature. Concerning 'family resemblances', Wittgenstein writes in § 67:

> The various resemblances between members of a family — build, features, colour of eyes, gait, temperament, and so on and so forth — overlap and criss-cross in the same way. — And I shall say: 'games' form a family. (PI, I, § 67)

Wittgenstein uses the notion of family resemblance to convey the idea that there are similarities across the many uses of language without there being a single common essence across uses.

Where this notion has had its biggest influence in philosophy of religion is with understanding the concept of religion itself. John Hick endorses a family resemblance conception of 'religion' in his *An Interpretation of Religion* (1989). Hick writes of the family resemblance analogy:

> [I]t is, I think, illuminating to see the different traditions, movements and ideologies whose religious character is either generally agreed or responsibly debated, not as exemplifying a common essence, but as forming a complex continuum of resemblances and differences analogous to those found within a family. (Hick 1989, 4)

For Hick and others, 'religion' is thus an open interpretive concept, where its boundaries are contestable and where borderline cases are somewhat common.

Ninian Smart also advanced an approach to thinking about religion drawing on a family resemblance conception (1996). Instead of seeing any one feature of a religion as being a necessary condition (e.g., belief in a supernatural agent), of something being religious, Smart's approach explores a growing number of different

'dimensions' as together tending to express religiosity. And it is not that these dimensions jointly determine the religiosity of something. Rather, in Smart's view, religious worldviews (we might imagine he has something like 'form of life' in mind) generally manifest along these diverse dimensions; keeping these dimensions in mind helps the scholar not to overlook otherwise salient features of religious worldviews. Thus, noting the diversity of forms religions take is a help to noticing their features, to interpreting them.

Timothy Fitzgerald has argued against Wittgensteinian approaches to thinking about 'religion' as a family resemblance concept. Fitzgerald sees in these approaches either a back-door way of entry for a universalised Protestant conception of religiosity as private faith or an unclear and therefore academically inept analytical concept. Fitzgerald writes,

> The idea that English-speaking academics can be free to describe selected practices and institutions of other cultures as 'religions' or as 'religious' if they so choose, as though this can be simply a decision made for convenience of Western academics, seems dangerous when placed in the contemporary context of Anglo-American imperialism. (Fitzgerald 2003, 218)

Due to the danger of reifying designations imposed by powerful agents such as imperial states, Fitzgerald subsequently argues against the family resemblance use of the term in scholarly discourse.

Fitzgerald presents significant problems for proponents of family resemblance approaches to understanding the concept of religion. The danger of imposing from the outside a distorting category on local traditions and practices is real and is moreover a concern very much in line with some of Wittgenstein's philosophical sensibilities ("don't think but look!" (Wittgenstein 2009, 35)). How could a family resemblance conception of something be distorting? While the non-essentialist conception may give the impression of local sensitivity, in drawing connections of putative resemblance, to prototypical religions, cultures in which religions are conventional institutions or ways of life are privileged. Thus, Fitzgerald worries family resemblance approaches could crowd out local vocabulary for making sense of social life. So, if global use of a family resemblance conception of religion to describe ways of life of a certain sort is thus problematic, it should be avoided above all for Wittgensteinian reasons.

Yet, the concept, variously understood, *is used* in contexts around the world and there is, arguably, a family resemblance among these uses (Carroll 2019). There is very good reason to proceed carefully here and to avoid broad generalisations. When local discourses tend to agree that something is or is not religious, this should provide a strong reason to agree. However, histories are rarely so simple, as

the history of classification of Confucianism shows (Yang 2008; Sun 2013); disputation over religion-status may be motivated by a variety of factors, from the local cultural assimilation of Catholicism into Chinese culture (The Rites Controversy) to Marxist critique of Confucian revival following the Cultural Revolution. Furthermore, in some social contexts, ascribing religion-status to Confucianism enables minority groups in particular societies, e.g., Indonesia, to satisfy government requirements that all citizens have a religion.

Notably, this avenue of influence of Wittgenstein in philosophy of religion cuts against the religion-as-form-of-life influence. While the latter tends to reify religions as distinct from non-religions, as distinct from each other, and as ahistorical entities, the former allows for the social construction of what are labelled as religions and builds in internal diversity within the category. Because of concerns raised by Fitzgerald, I agree there is good reason to proceed carefully when using the term in contexts culturally distant from the modern European contexts in which it first formed.

2.4　Grammar

Through the PI and in other works from this period, Wittgenstein uses the term "grammar" ("*Grammatik*") frequently in a specialised sense to refer to the possibilities of meaning for a piece of language. This is a philosophical or metaphorical extension of the term from its ordinary use. For Wittgenstein, clarifying grammar thus becomes the focal point of philosophical clarification. In § 90, he writes:

> We feel as if we had to *see right into* phenomena: yet our investigation is directed not towards *phenomena*, but rather, as one might say, towards the '*possibilities*' of phenomena. What that means is that we call to mind the *kinds of statement* that we make about phenomena [...] Our inquiry is therefore a grammatical one. And this inquiry sheds light on our problem by clearing misunderstandings away. (PI, I, § 90)

Grasping the grammar of an expression enables one to understand it, while confusion about the nature or application of grammar is a key source of philosophical problems. In § 122, Wittgenstein writes:

> A main source of our failure to understand is that we don't have *an overview* of the use of our words. — Our grammar is deficient in surveyability. (PI, I, § 122)

In Wittgenstein's view, there is no ideal metalanguage in which grammar may be definitively expressed. Instead, descriptions of the

possibilities of use of language take place within language. So, clarifications are local rather than global, as Wittgenstein writes in § 97:

> We are under the illusion that what is peculiar, profound and essential to us in our investigation resides in its trying to grasp the incomparable essence of language... Whereas, in fact, if the words "language", "experience", "world" have a use, it must be as humble a one as that of the words "table", "lamp", "door". (PI, I, § 97)

In this way, grammar is sometimes thought of as having a kind of ineffability; the actual possibilities of use may extend beyond what is described in any particular concrete description.

From the conception of grammar in the PI and the related modes of philosophical enquiry that Wittgenstein demonstrates and advises stems a model for philosophical enquiry into religions: grammatical investigations into the possibilities of concepts and practices such as prayer, faith, God, and liturgy. From D.Z. Phillips's contemplation of the possibilities of sense when it comes to prayer (Phillips 1965) to George Lindbeck's comparative study of Christian denominations and their doctrines (Lindbeck 1984), the Wittgensteinian notion of grammar has figured prominently in twentieth century philosophy of religion.

A well-known remark on grammar in the PI (§ 373) links it with theology. Wittgenstein writes:

> Grammar tells what kind of object anything is. (Theology as grammar.) (PI, I, § 373)

If grammar is what established the possibilities of sense within language, then theology would seem to establish the possibilities of sense within a theistic religion (and here especially, Christianity). We might wish Wittgenstein had written more about the topic in the PI (or indeed elsewhere). How is theological clarification similar to philosophical clarification? How much can this idea be generalised beyond Christianity (and Protestantism, at that) to diverse religious traditions? As we saw in connection with the notion of grammar above, the idea that religious beliefs (especially of a foundational or central focus) could play a regulatory role with respect to religious ways of life and their accompanying language has had a considerable influence in philosophy of religion.

Perhaps because Wittgenstein refers at one point in the PI to conceiving of "theology as grammar", Wittgensteinian philosophers of religion of frequently focused their attention on the grammar of "God". William Brenner writes:

> This perspective highlights the fact that many of us first learned a theology in the course of learning the practices of a religion, much as all of us first learned a language in the course of learning how to speak… But 'theology as grammar' (PI, sec. 373): doesn't this comparison trivialize theology? Not if we understand that the grammar in question is for teaching and celebrating a new form of life. (Brenner 1999, 140)

While the remark of Wittgenstein is exceedingly brief, in concert with other remarks elsewhere in the corpus, a developed viewpoint can be reconstructed.

An approach known as Grammatical Thomism also takes inspiration from these remarks, seeing a hybrid Wittgensteinian-Thomist point of view as being intelligible and helpful for elaborating Christian theological commitments using contemporary philosophical parlance. Importantly, while Grammatical Thomists find Wittgensteinian therapy helpful for some unfruitful philosophical questions, they do not refuse metaphysical claims entirely; they are not thoroughgoing non-cognitivists about God-talk. Simon Hewitt writes about Grammatical Thomism:

> The grammatical thomist invites us to consider a way-in to the use of the word 'God' which both secures the sense-making nature of the word and, under very minimal assumptions (the existence of anything whatsoever), the truth of canonical sentences containing it, whilst also placing severe constraints on what we are entitled to assert about God. In Wittgensteinian terms, they supply a way of understanding the grammar of the word 'God', which provides a basis for subsequent philosophical and theological enquiry and which does duty, in a fashion relatively uncommon in the analytic philosophy of religion, to the stress on divine ineffability so often found in living religion. (Hewitt 2021, 35)

In Hewitt's analysis of Grammatical Thomism, the ineffability of grammar meets divine ineffability in a variation on apophatic theology. "God" thus plays a grounding and determinative role within Christian practice while not being an object among objects.

2.5 Aspect Perception

In the second part of the PI – now called by some "Philosophy of Psychology. A Fragment" – another highly influential idea appears. In connection with the famous duck-rabbit diagram, Wittgenstein entertains what it is to see or notice an aspect of a thing. When it comes to the perception of ambiguous objects, the perceiver must introduce a framework to disambiguate the object. In a way, the

framework one applies to the interpretation of the perceptual object is similar to what the language user brings by way of grammatical understanding to a linguistic occasion in order to grasp possibilities of meaning.

While this idea has entered into philosophy of religion in more than one way,[1] an influential approach comes from John Hick. Hick explores the relevance of this idea to religious experience through his related notion of "experiencing-as". Hick refers to the role that faith plays in interpreting the world. The idea is that religious knowledge is a product of experience which is itself framed by a pre-existing interpretation, which according to Hick is what people commonly call "faith". Hick writes,

> To reach the religious case, however, we must expand the notion of "seeing as" into that of "experiencing as", not only visually but through all the modes of perception functioning together. We experience situations as having different kinds of significance and so as rendering appropriate different kinds of practical response. The Old Testament prophets, for example, experienced their historical situation as one in which they were living under the sovereign claim of God. (Hick 1966, 142)

While Wittgenstein was concerned in the PI with a narrow feature of the phenomenology of perception, Hick is interested in the broader picture of religious experience, which can itself be cashed out in a variety of ways. The shifting perspective that frames experience of the world is fundamental to narratives of conversion or spiritual transformation. Indeed, Wittgenstein himself uses similar narratives elsewhere when accounting for the existential character of religious faith (cf. PPO).

Aspect perception also appears periodically in work on Wittgenstein and ethics, especially when it comes to seeing the humanity in another person. In *The Claim of Reason*, Stanley Cavell considers the moral consequences of the failure to see another person as human; he terms this phenomenon, "soul-blindness". Cavell considers the topic of "soul-blindness" in connection with the moral psychological capacity for enslaving others. While this notion perhaps pertains to more directly to ethics than to philosophy of religion, it is relevant to work in religious ethics (an area overlapping with or otherwise adjacent to philosophy of religion) considering the spiritual dynamics involved in the identifying and overcoming racist bias to come to see the humanity in another.

---

[1] For example, Espen Dahl explores the relevance of these passages from the PI for the perception of purported miracles. See Dahl 2018, 106f.

2.6   Metaphilosophy

Another influential theme in the PI concerns Wittgenstein distinguishing his approach to philosophy from more historically influential approaches. The PI contains numerous remarks on the nature of philosophical problems and clarificatory philosophical practices. To some, the PI presents a revolutionary approach to philosophy – recasting the nature of philosophical problems in a way that sets the stage for a completely new way of doing philosophy. In § 123, Wittgenstein writes:

> A philosophical problem has the form: "I don't know my way about". (PI, I, § 123)

When we think about the philosophical problems of philosophy of religion, we may come to see a wide open field rather than a closed set of 'classic' problems in the field (e.g., theistic arguments, the problem of evil, the logic of divine attributes). We may see that philosophical problems concerning religions can appear anywhere the grammar of language having to do with religions has become confused. (cf. Carroll 2014, 2021).

In conceiving of philosophical method as grammatical investigation, it might seem to some readers that the aim would be complete and final clarification. Wittgenstein writes in § 91:

> But now it may come to look as if there were something like a final analysis of our linguistic expressions, and so a single completely analysed form of every expression. (PI, I, § 91)

Yet, some have interpreted this to mean that grammar is ineffable, while others call into question this very idea of any limitation on expression (Floyd 2007). The issue of the expressibility of grammar is pertinent to philosophy of religion insofar as ineffability (broadly construed) is a phenomenon in some traditions of religious philosophy – such as negative or apophatic theology in the Abrahamic traditions, emptiness in Mahayana Buddhism, and the instability of descriptions of the *dao* in the *Daodejing*. The question of the possibility of language expressing all meanings is directly related to the viability of these ineffable traditions of religious philosophy. What I take from this is the idea that clarifications are made in local contexts (i.e., actual) instances of language and not in some meta-language. Wittgenstein continues in § 122 describing his view of philosophical clarification:

> A main source of our failure to understand is that we don't have *an overview* of the use of our words. — Our grammar is deficient in surveyability. A surveyable representation produces precisely that kind of understanding which consists in "seeing connections".

> Hence the importance of finding and inventing *intermediate links*. (PI, I, § 122)

Wittgenstein's picture of clarification does not offer a perspective claiming to be a theory (i.e., a final, factual picture of the grammar) but instead a description, a description using local vocabulary and potential linguistic moves.

Local clarifications find and provide those intermediate links, since grammar is difficult to survey. Moreover, Wittgenstein recognises something in human beings that makes them prone to make blunders. There is a tendency to reach beyond what is available and to offer a theory. Thus philosophy (in Wittgenstein's sense) is a mode of resistance to both human tendencies to go beyond what can be said and philosophical tendencies to develop theories. As Wittgenstein puts it in § 109, achieving clarify requires *striving*:

> Philosophy is a struggle against the bewitchment of our understanding by the resources of our language. (PI, I, § 109)

Where perhaps the metaphilosophical remarks have had the most significant impact is in dialogue between Wittgenstein's philosophy and Buddhism. Chris Gudmansen write:

> For Wittgenstein, getting people to understand is much more than presenting them with the facts. He is prepared to use any means in accordance with what works best. There are no irreducible acts of understanding and therefore no "ultimate explanations" [...] An explanation need not be the "presentation of facts" at all — it could be a gesture or pricking someone with a pin. In different cases, different measures are called for, if liberation is to be achieved. (Gudmunsen 1977, 71f)

In Buddhist philosophy, one often sees that practices aimed at enlightenment are not so much theoretically framed as practically structured. There is not a theory of non-thinking that a Buddhist adept should work towards; for example, in Zen it is through practising enlightenment that one may come to encounter it. Moreover, from the point of view of a teacher, liberatory explanations will be tailored to the particular person (a gesture, a pricking of a pin). In this way, Rupert Read's recent liberatory reading of Wittgenstein – and its implications for overcoming blocks in addressing our climate crisis – draws connections with Buddhist practice and values, especially concerning Mayahana Buddhism's emphasis on interdependence among people, as well as between humanity and nature (Read 2021).

3 References to Aspects of Religions in the PI

While concepts developed in the PI have been influential by way of interpretation in philosophy of religion, matters directly relating to religion are almost entirely absent from the text. Yet, if one looks closely, there are a handful of scattered remarks that seem to show how Wittgenstein would apply the central ideas in the PI to thinking about religiosities. So, these passages are relevant to philosophy of religion inspired by Wittgenstein and have at times been the focus of philosophical commentary. It is my objective in this section to describe them and account for their relevance to future philosophy of religion.

3.1 Prayer

When explaining what a language-game is in remark § 23, Wittgenstein includes as an example of prayer, perhaps to indicate just how varied the interpretive use of "language-games" can be when it comes to human life with language. Wittgenstein writes:

> The word "language-*game*" is used here to emphasise the fact that the *speaking* of language is part of an activity, or of a form of life.

Consider the variety of language-games in the following examples, and in others:

> Giving orders, and acting on them 🗆🗆

> [...] Requesting, thanking, cursing, greeting, praying. (PI, I, § 23)

Wittgenstein describes language as existing within a form of life, as part of an activity. Here, prayer is as much an instance of a human form of life as cracking a joke or forming and testing an hypothesis. Wittgenstein presents prayer here as being on a par with any other instance of language. This inclusion anticipates Wittgenstein's general humanistic attitude towards the wide variety of forms of language use and ways of life human beings may sincerely undertake.

For the philosophers of religion, it is unfortunate that Wittgenstein did not elaborate. Wittgenstein clearly thought much about religious matters, but the PI is a source that is nearly entirely missing explicit reference to religiosities. One more remark involving prayer occurs in the second part of PI. Wittgenstein writes:

> When it is said in a funeral oration "We mourn our...", this is surely supposed to be an expression of mourning; not to communicate

anything to those who are present. But in a prayer at the grave, these words would be a kind of communication. (PI, II, § 81)

In this passage, Wittgenstein contemplates the different meanings that the same expression can have, as one imagines shifting contexts. Different audiences reframe an expression so that it can be used in quite different speech acts.

Prayer can be thought of as an established form of using language to address God. One might think here of the Lord's Prayer or even the Serenity Prayer. Established forms of prayer may be communal or individual. Some are prescribed and/or liturgical, while others are spontaneous. And many forms of prayer do not only address a divine being but also secondarily address the speakers themselves and also fellow congregants (consider here the ways that prayers can function as instances of spiritual teaching, to form and reform modes of engaging God, other people, or oneself). Whether there are many language-games of prayer or if there is something that unites all instances of prayer, or whether forms of prayer will always be indexed to particular religious traditions is up to the analysis of philosophers of religion working in a Wittgensteinian mode. A comparative study of prayer activities both within and across religious traditions could be helpful explored by means of central ideas from the PI.

3.2 God

"God" appears rarely in the PI, and when it does, the word is invoked obliquely. In remark 342, Wittgenstein contends with William James and the idea that thought could be possible without speech. James recounts the story of a Mr. Ballard, a person who only learned to speak as an adult, reported having thoughts about God. Wittgenstein marvels at the notion but arrives at a sort of agnosticism about what it could mean to have such views:

> Are you sure — one would like to ask — that this is the correct translation of your wordless thoughts into words? And why does this question — which otherwise seems not to exist — arise here? Do I want to say that the writer's memory deceives him? — I don't even know if I'd say *that*. These recollections are a strange memory phenomenon — and I don't know what conclusions one can draw from them about the narrator's past! (PI, I, § 342)

Wittgenstein's respectful agnosticism does not mean that he rejects Mr. Ballard's testimony, only that he cannot imagine what it would mean to say such a thing. Thus, experiences and ideas of God are dependent on language and its use. In this way, this remark mirrors the

respectful agnosticism also on display in the "Lectures on Religious Belief". Wittgenstein cannot participate in the framework used by the religious person, but he maintains throughout his life and corpus a respect for the *sincerity* of expressions of religious faith.

Wittgenstein also invokes the idea of God again a few remarks later in the text when considering the law of the excluded middle. The specific scenario being entertained is whether in the expansion of the number π, the group of numbers "7777" should appear. Either it does or it does not, whether or not any human being is able to calculate that far: "That is to say: God sees — but we don't know". (PI, I, § 352). Here, Wittgenstein immediately launches into contemplation of the possibilities of meaning:

> But what does that mean? — We use a picture: the picture of a visible series, the whole of which one person can survey and another can't. Here the law of excluded middle says: it must look either like *this* or like *that*. So really — and this is surely obvious — it says nothing at all, but gives us a picture. And the problem is now supposed to be: does reality accord with the picture or not? And this picture *seems* to determine what we have to do, what to look for, and how — but it does not, precisely because we do not know how it is to be applied. (PI, I, § 352)

In this example, the reference to God is roughly similar to a philosophical concept of God (i.e., a concept used when necessary to make sense of some phenomenon that is otherwise the focus of the philosophical activity). Perhaps such a conception of God is metaphysically useful for stipulating the existence of an answer unknowable to human beings. The question is about the truth or utility of the logical principle rather than God. While the idea is invoked in this example one does not get the sense from Wittgenstein's later writings that he in any way thought of God as philosophically necessary; the only salient concept of God one finds in Wittgenstein's later writings is of God as a devotional focus for existentially engaged forms of religiosity, a concept of God that is quite distant from the "God of philosophy".

3.3 Soul

The word "soul" appears in a handful of remarks. Sometimes, Wittgenstein uses the idea as a commonplace notion indicating personhood rather than as a nonnatural reality to which Wittgenstein is committing himself. Consider this short remark from part two of the PI:

> My attitude towards him is an attitude towards a soul. I am not of the *opinion* that he has a soul. (PI, II, § 22)

We can see immediately how this remark dovetails with the earlier examination of the aspect perception section. This is a matter of perception of another person, as would be relevant to ethics and as may or may not be influenced by religious ideas. Yet elsewhere, a different use appears. Wittgenstein continues:

> Religion teaches that the soul can exist when the body has disintegrated. Now do I understand what it teaches? — Of course I understand it — I can imagine various things in connection with it. After all, pictures of these things have even been painted. And why should such a picture be only an imperfect rendering of the idea expressed? Why should it not do the *same* service as the spoken doctrine? And it is the service that counts. (PI, II, § 23)

The picture here of a religious teaching concerning bodies and souls and personal identity after death does not present a propositional description of the doctrine as being most fundamental. Wittgenstein presents here the *artistic* as not being derivative or an "imperfect" duplication of the *doctrinal*. Wittgenstein instead challenges the idea that a pictorial representation would be inferior to a spoken teaching. In addition to acknowledging any one dimension to religiosity, this remark coheres with Wittgenstein's tendency to downplay any particular description of a grammatical feature of language as being definitive. What is crucial is that which enables people to learn how to play the game.

4 Globally Engaged Philosophy of Religion

In the opening chapter of D.Z. Phillips's *The Concept of Prayer* (1965), Phillips remarks on the diversity within the field of philosophy of religion, comparing it to the Biblical Tower of Babel. Phillips writes:

> To work in the field of philosophy of religion is like working on the Tower of Babel: one cannot take for granted that one's colleagues understand what one is saying. The position, if anything, is worse for the philosophers, since the builders at least were engaged on a common task, they were trying to do the same thing. No such agreement exists among philosophers of religion: the nature and purpose of their subject is itself a philosophical controversy. It becomes essential, therefore, to try to give some indication of what I think philosophy can say about religion. (Phillips 1965, 1)

An interesting thing about this metaphor is that Phillips imagines philosophers of religion continuing to work on the Tower of Babel *after*, one supposes, God has confused the people's language and scattered

them. In this circumstance, any builders remaining would have a difficult time communicating with each other. It is not clear from Genesis, at least, how *extensive* the linguistic confusion is among the people, but if one reads the passage as a polemic against Babylonia, then perhaps the point is not so much that God confuses the languages as that God disperses the univocal Babylonian tower builders. Perhaps Phillips's passage and the myth it invokes simply registers the idea of human beings having diverse projects, languages, and societies (and, of course, worldviews). Yet, however varied the approaches to philosophy of religion were in 1965, they are vastly more varied in today's universities and interconnected world. This is also arguably true for approaches to Wittgensteinian philosophy of religion, where scholars take quite different lessons from the text in addressing philosophical problems.

From the preceding overview of the influence of the PI in philosophy of religion, there is no one definitive way in which philosophers of religion have contended with the text. The PI has tended to inspire, in one way or another, hermeneutically rich interpretations of religiosities; some emphasise the meaningfulness of "religious language" within its social context, others argue for the lack of a common core to all things that are called religions, some readers see in Wittgenstein's remarks resources for understanding differential cognitive responses to the same objects or world, and still others consider the relevance of Wittgenstein's metaphilosophical remarks to philosophy of religion. While there are scant remarks on religiosities in the PI, those that do appear cohere with themes found elsewhere in the PI or in Wittgenstein's corpus (e.g., respect from a distance for sincere belief, a tendency to downplay the importance of intellectual aspects of religions, seeing religious belief as being related to the framework, or grammar, to which one is philosophically committed). Perhaps *because* of the near absence of religious topics in the PI, the text has inspired a wide variety of approaches in philosophy of religion. Thus, the dialectical features of the text stand out, questions and provocations that get to the heart of the assumptions that readers may bring to the text and to their philosophical projects.

The use of Wittgenstein in work aimed at hermeneutically rich encounters between people identifying with different religious and/or cultural traditions also exemplifies recent work on Wittgenstein and the PI. For example, Wittgenstein has been used by scholars interested in interreligious dialogue for many decades (e.g., Lindbeck 1984), but recent years have seen a new generation of scholars develop these resources. In more recent times, Gorazd Andrejč (2016) has explored the resources in Wittgenstein for making sense of religious differences in religiously diverse social contexts. While Andrejč makes use of notions like "grammar" in appraising religious differences, he also draws on Wittgenstein's shifting focus across various works from grammar to instinct to existential concerns as he contemplated religions.

In recent years, there has been a push to diversify philosophy of religion beyond its historical preoccupation with Christianity. Often paired with comparative philosophy, this strand of philosophy of religion seeks to open the field to address philosophical problems concerning a wide diversity of religious and nonreligious philosophical traditions. To some extent, this thread has a long history with work done by Gudmunsen in the 1970s and Hick and Smart in the 1970s, 80s, and 90s. John Clayton beginning in a series of articles published in the 1980s and 90s and culminating in his posthumous book *Religions, Reasons, and Gods* (2006) intentionally drew on Wittgensteinian themes such as family resemblances and forms of life in his contextually-sensitive approach to cross-cultural philosophy of religion (Clayton 2006, 83). Brian Clack has written on atheism in connection with Wittgenstein's philosophy. In more recent times, Mikel Burley has written numerous articles and books on Wittgensteinian philosophy of religion and religious pluralism. His primary focus has been on understanding South Asian religions, but he has also written on indigenous American and African religions. Burley describes his project thusly:

> [I]t aspires to do conceptual justice to the radically plural character of religious phenomena themselves, aiming to deepen understanding of the variegated nature of religious – and indeed nonreligious – forms of life without rushing to evaluate them in terms of some supposedly universal standard of truth or rationality. (Burley 2020, 2)

Even as Burley is looking forward to diversifying the field, we can see readily how this approach is linked with themes we have encountered while surveying the history of Wittgensteinian philosophy of religion.

5 Conclusion

Having been published seventy years ago, the PI is no longer a contemporary work of philosophy. It is through the work of multiple generations of scholars that audiences now encounter the text and/or the ideas within it. While the first wave of influence of the PI happened in the 1960s, since then scholars have had access to so much more of Wittgenstein's corpus, which is especially important for philosophy of religion given the relevance of sources such as the "Remarks on Frazer's *Golden Bough*", the "Lectures on Religious Belief", the "Lecture on Ethics" (2014), and the miscellaneous remarks included in *Culture and Value*, as well as personal writings and memories of conversations. Yet, the PI looms large over all these other sources when it comes to constructing a philosophy of religion inspired by Wittgenstein.

Despite the seventy years of philosophical encounters with the PI, the book will likely continue to exert a significant influence on twenty-first century philosophy of religion – even as *On Certainty* (1969) is having an extended moment (e.g., hinge epistemology and religion). The PI's emphasis on the micro-level of social context for understanding uses of language linked with these things we call religions has not yet really been plumbed to the extent that is needed to understand our culturally and religiously diverse societies and world. With the expansion of many areas of philosophy to approaches and traditions that have been marginalised or otherwise overlooked, this work retains great potential for dialectical engagement as philosophers ask metalevel questions about the nature and boundaries of philosophies, religions, and ways of life.

While there is no one way to do Wittgensteinian philosophy of religion after the PI, several lessons recur that provoke the present author to consider a more sustained integration of different elements from the text into a particular philosophical response. Wittgenstein directs his reader's attention to the uses of language over against pictures of language use that rely entirely on ostention and reference for anchoring meanings of expressions. The text reminds readers that linguistic activity takes place within a form of life. The book problematises attempts to boil it down into simple analyses by introducing, for example, anti-essentialist ideas like family resemblances. The text considers the role of cognitive framing to the interpretation of ambiguous objects and redefines philosophical problems in such a way that the focus of philosophy could radically shift to instances of conceptual confusion rather than some notion of "classic" problems. A synoptic reading of the PI might be out of keeping with the spirit of Wittgenstein's philosophy; he was much more of a reactive philosopher than a system-builder. In that spirit, letting oneself be provoked by Wittgenstein's varied philosophical lessons enables one to develop a hermeneutically rich approach to philosophising about religions that answers to a wide variety of philosophical problems.

References

Andrejč, G. (2016). *Wittgenstein and Interreligious Disagreement: A Philosophical and Theological Perspective*. Basingstoke: Palgrave Macmillan. https://doi.org/10.1057/978–1–137–49823–6.

Ayer, A.J. (1935). *Language Truth and Logic*. New York: Dover Publications.

Bloemendaal, P.F. (2006). *Grammars of Faith: A Critical Evaluation of D.Z. Phillips's Philosophy of Religion*. Leuven: Peeters.

Brenner, W. (1999). *Wittgenstein's Philosophical Investigations*. New York: SUNY Press.

Burley, M. (2020). *A Radical Pluralist Philosophy of Religion: Cross-cultural, Multireligious, Interdisciplinary*. London: Bloomsbury. https://doi.org/10.5040/9781350098343.

Carroll, T.D. (2008). "The Traditions of Fideism". *Religious Studies*, 44(1), 1-22. https://doi.org/10.1017/S0034412507009250.

Carroll, T.D. (2014). *Wittgenstein Within the Philosophy of Religion*. Basingstoke: Palgrave Macmillan. https://doi.org/10.1057/9781137407900.

Carroll, T.D. (2019). "Wittgenstein and Ascriptions of 'Religion'". Andrejč, G.; Weiss, D. (eds), *Interpreting Interreligious Relations with Wittgenstein: Philosophy, Theology and Religious Studies*. Leiden: Brill Publishers, 54-72. https://doi.org/10.1163/9789004408050_004.

Carroll, T.D. (2021). "'Grasping the Difficulty in Its Depth': Wittgenstein and Globally Engaged Philosophy". *Sophia*, 60(1), 1-18. https://doi.org/10.1007/s11841–019–00742–y.

Cavell, S. (1979). *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780195131079.001.0001.

Clack, B. (1999). *An Introduction to Wittgenstein's Philosophy of Religion*. Edinburgh: Edinburgh University Press. https://doi.org/10.1515/9781474465748.

Clayton, J. (2006). *Religions, Reasons, and Gods: Essays in Cross-Cultural Philosophy of Religion*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511488399.

Dahl, E. (2018). "Wittgenstein and Augustine on Seeing Miracles". Doody, J.; Eodice, A.R.; Paffenroth, K. (eds), *Augustine and Wittgenstein*. New York: Lexington Books.

Fitzgerald, T. (2003). "Playing Language Games and Performing Rituals: Religious Studies as an Ideological State Apparatus". *Method & Theory in the Study of Religion*, 15(3), 209-54. https://doi.org/10.1163/157006803322393378.

Flew, A. (1955). "Theology and Falsification". Flew, A.; MacIntyre, A. (eds), *New Essays in Philosophical Theology*. London: SCM Press, 96-9.

Floyd, J. (2007). "Wittgenstein and the Inexpressible". Crary, A. (ed.), *Wittgenstein and the Moral Life: Essays in Honor of Cora Diamond*. Cambridge, MA: MIT Press, 177-234. https://doi.org/10.7551/mitpress/7265.003.0005.

Floyd, J. (2020). "Wittgenstein on Ethics: Working Through Lebensformen". *Philosophy & Social Criticism*, 46(2), 115-30. https://doi.org/10.1177/0191453718810918.

Gudmunsen, C. (1977). *Wittgenstein and Buddhism*. London: Palgrave Macmillan. https://doi.org/10.1007/978–1–349–03128–3.

Hewitt, S. (2021). "Grammatical Thomism". *Religious Studies*, 57(1), 30-48. https://doi.org/10.1017/s0034412518000896.

Hick, J. (1966). *Faith and Knowledge*. 2nd ed. Ithaca, NY: Cornell University Press.

Hick, J. (1989). *An Interpretation of Religion: Human Responses to the Transcendent*. London: Macmillan. https://doi.org/10.1057/9780230371286.

Lindbeck, G. (1984). *The Nature of Doctrine: Religion and Theology in a Postliberal Age*. Louisville, KY: Westminster John Knox Press.

Malcolm, N. (1960). "Anselm's Ontological Arguments". *The Philosophical Review*, 69, 41-60. https://doi.org/10.2307/2182266.

Nielsen, K. (1967). "Wittgensteinian Fideism". *Philosophy*, 42(161), 191-209. https://doi.org/10.1017/s0031819100001285.

Plantinga, A. (1967). *God and Other Minds: A Study in the Rational Justification of Belief in God*. Ithaca, NY: Cornell University Press.

Phillips, D.Z. (1965). *The Concept of Prayer*. London: Routledge. https://doi.org/10.4324/9781315819914.

Phillips, D.Z. (1993). "Religious Beliefs and Language-Games". Phillips, D.Z., *Wittgenstein and Religion*. London: Palgrave, 56-78. https://doi.org/10.1057/9780230377035_5.

Proudfoot, W. (1987). *Religious Experience*. Berkeley; London: University of California Press. https://doi.org/10.1525/9780520908505.

Read, R. (2021). *Wittgenstein's Liberatory Philosophy: Thinking Through His Philosophical Investigations*. New York: Routledge. https://doi.org/10.4324/9781003090977.

Rhees, R. (1960). "Wittgenstein's Builders". *Proceedings of the Aristotelian Society*, 60, 171-86. https://doi.org/10.1093/aristotelian/60.1.171.

Rhees, R. (1969). *Without Answers*. New York: Routledge & Kegan Paul.

Schoenbaumsfeld, G. (2014). "Ludwig Wittgenstein". Oppy, G.; Trakakis, N. (eds), *Twentieth-Century Philosophy of Religion*. Vol. 5, *The History of Western Philosophy of Religion*. New York: Routledge, 161-74. https://doi.org/10.1017/upo9781844654673.014.

Smart, N. (1996). *Dimensions of the Sacred: An Anatomy of the World's Beliefs*. Berkeley: University of California Press.

Sun, A. (2013). *Confucianism as a World Religion: Contested Histories and Contemporary Realities*. Princeton, NJ: Princeton University Press. https://doi.org/10.23943/princeton/9780691155579.001.0001.

Trigg, R. (1983). "Religion and the Threat of Relativism". *Religious Studies*, 19(3), 297-310. https://doi.org/10.1017/s0034412500015250.

Von Der Ruhr, M. (2009). "Rhees, Wittgenstein and the Swansea School". Edelman, J. (ed.), *Sense and Reality: Essays out of Swansea*. Frankfurt a. M.: Ontos Verlag, 219-35. https://doi.org/10.1515/9783110328813.219.

Winch, P. (1990). *The Idea of a Social Science and Its Relation to Philosophy*. London: Routledge. https://doi.org/10.4324/9780203820766.

Wittgenstein, L. (1967). *Lectures and Conversations on Aesthetics, Psychology, and Religious Belief*. Edited by C. Barrett. Berkeley, CA: University of California Press. [LC]

Wittgenstein, L. (1969). *On Certainty*. Edited by G.E.M. Anscombe; G.H. von Wright. Oxford: Blackwell. [OC]

Wittgenstein, L. (1993). "Remarks on Frazer's *Golden Bough*". Klagge, J.C.; Nordmann, A. (eds), *Ludwig Wittgenstein: Philosophical Occasions 1912-1951*. Indianapolis: Hackett, 115-55. [PO]

Wittgenstein, L. (1998). *Culture and Value*. Transl. by P. Winch, edited by G.H. von Wright; H. Nyman, revised edition by A. Pichler. Oxford: Blackwell. [CV]

Wittgenstein, L. (2003). *Ludwig Wittgenstein: Public and Private Occasions*. Edited by J.C. Klagge; A. Nordmann. Lanham, MD: Rowman & Littlefield Publishers. [PPO]

Wittgenstein, L. (2009). *Philosophical Investigations*. 4th ed. Revised edition by P.M.S. Hacker; J. Schulte, transl. by G.E.M. Anscombe; P.M.S. Hacker; J. Schulte. London: Blackwell Publishers. [PI]

Wittgenstein, L. (2014). *A Lecture on Ethics*. Edited by E. Zamuner; E.V. Di Lascio; D. Levy. Oxford: Wiley-Blackwell. [LE]. `https://doi.org/10.1002/9781118887103`.

Yang, X. (2008). "Some Issues in Chinese Philosophy of Religion". *Philosophy Compass*, 3, 551-69. `https://doi.org/10.1111/j.1747–9991.2008.00139.x`.

Answering Sraffa on Religion: Wittgenstein Walking the Tightrope

Mauro Luiz Engelmann
Universidade Federal de Minas Gerais, Brazil

Abstract   The subject of religion, one might think, although discussed in the *Tractatus*, is conspicuously absent from or in Wittgenstein's later works, particularly in the *Philosophical Investigations*. Using Sraffa's comments as a starting point, I *tentatively* deal with the absence of the subject by considering the question whether Wittgenstein's religious views are compatible with his philosophies in the *Tractatus* and in the *Philosophical Investigations*. Answering this question involves examining Wittgenstein's own central concern about his convictions in his later years, namely, to what extent one can honestly be a religious thinker nowadays. Presumably, his philosophy might not allow a 'philosophy of religion', if his own views on religion are not compatible with his philosophy. I tackle these issues beginning with Wittgenstein's conversations with Ludwig Hänsel, then move to his later views and relevant passages in the *Philosophical Investigations*. With this in place, I uncover a religious moment that is not completely apparent in his later book, namely, an admission of errors that is a sort of confession.

Keywords   Wittgenstein. Religion. Religious belief. Confession. Sraffa. Hänsel. Tolstoy.

Summary   1 Sraffa, Wittgenstein and Religion. – 2 On Wittgenstein's *Early* (Christian) Religious Values. – 3 *Later* Views: A Tightrope Walker. – 4 Backdoor Metaphysics? Religion and the Inevitability of a *Weltanschauung*. – 5 *Investigations*: Two Remarks on Religion. – 6 A Religious Point of View Expressed in a Confession. – 7 "Religious Puzzles" and the Critique of Religion.

> An honest religious thinker is like a tightrope walker. He almost looks as though he were walking on nothing but air. His support is the slenderest imaginable. And yet it really is possible to walk on it.
>
> (CV, 73)

> If someone tells me he has bought the outfit of a tightrope walker I am not impressed until I see what is done with it.
>
> (Drury 1984, 88)

1 Sraffa, Wittgenstein and Religion

In October 1941, Wittgenstein gave a copy of the *Blue Book* to Sraffa, who wrote some comments. One of them concerns the following passage in the *Blue Book*:

> When we talk of language as a symbolism used in an exact calculus, that which is in our mind can be found in the sciences and in mathematics. Our ordinary use of language conforms to this standard of exactness only in rare cases. Why then do we in philosophizing constantly compare our use of words with one following exact rules? The answer is that the puzzles which we try to remove *always spring from just this attitude towards language.* (BB, 25-6; emphasis added)

In this passage, Wittgenstein is criticising philosophers who take language (or thought) as structured by a kind of calculus. Of course, although he does not say so, he was himself one of them in the past: he had a "calculus attitude to language" in the *Tractatus* and in the *Big Typescript* (see Engelmann 2013, ch. 3). Contrary to his philosophy in the *Tractatus* and in the *Big Typescript*, the point of the *Blue Book* is to uncover the calculus conception as the *source* of philosophical troubles.

Among Sraffa's comments one finds the question "Metaphysics, Why Not Theology?" referring to the quoted passage of the *Blue Book*. He explains his point to Wittgenstein in the following way:

> Also, why do you deal always with metaphysics and never with theology? Are not their puzzles very similar (e.g., omniscience in god and freewill in man)? But could it be said that theol[ogical] puzzles only arise when people take the calculus' attitude to language? (N.B. I am not suggesting that this is the reason you leave theology alone). (Venturinha 2012, 184)

Sraffa's criticism is expressed in the first and third questions, the latter being ironical. The ground for it is the second question. Indeed,

metaphysical and theological puzzles, as Sraffa points out, if not identical, are at least very similar, for theologians and metaphysicians ask, for example, whether God's omniscience is compatible with the existence of freewill in human beings. However, if they are similar, how can Wittgenstein say that the puzzles he tries to remove *always* spring from the calculus attitude? Giving the similarity between theological and metaphysical puzzles, it seems very strange indeed to say that puzzles always arise from a calculus attitude towards language, for no theologian seems to deal with such conception at all. Thus, one obvious point of Sraffa's critique is Wittgenstein's dogmatic statement that something is always the case.

However, Sraffa's critique is interesting because he is obviously teasing Wittgenstein when he says that Wittgenstein wants *to leave theology alone*. Wittgenstein does not seem to investigate in the *Blue Book*, or in any other work, *how* theological problems/puzzles arise, or which are their sources. Rather, he uses his genetic method in a restricted way by examining only the genesis of puzzles in philosophy (and perhaps in science, depending on how one sees it).[1]

Of course, there is the question of how exactly his criticisms of metaphysics should apply to religion, but one might think the following about his various philosophies.[2] If there are no sentences of ethics or metaphysics, as argued in the *Tractatus*, there are no sentences concerning God either. Pseudo-sentences concerning God must be *merely* nonsense. If metaphysical claims are unverifiable nonsense, "wheels turning idly", as Wittgenstein argues in *Philosophical Remarks* (1930), then theological claims are also unverifiable nonsense, i.e., simply nonsense. If philosophy is full of misleading analogies, as argued in the *Blue Book* and in *Philosophical Investigations*, apparently the same or worse takes place in theology. In this case, one would need to investigate how puzzles in religion arise and how they dissolve with Wittgenstein's method. Thus, all of Wittgenstein's philosophies seem to imply a harsh critique of theological/religious claims.

Therefore, Sraffa's questioning challenges the compatibility of Wittgenstein's understanding of logic, 'grammar', and method in his philosophy with his views on religion (but also, it seems, his views on culture in general). In Sraffa's view, Wittgenstein suspiciously *decided* to leave theology/religion alone.

---

[1] On the origins of the genetic method see chapter 2 of Engelmann 2013 and Engelmann 2012.

[2] The plural (philosophies) means his central views throughout his career in unfinished works where one finds a systematic treatment of philosophical problems: *Philosophical Remarks*, the *Big Typescript*, the *Blue Book*, and the *Brown Book*. I focus on two of Wittgenstein's works in this paper, although I think that understanding those in-between philosophies in themselves is a very serious and urgent matter.

However, has Wittgenstein really left theology alone? Can he, or should he, do it? Some responses to these difficulties should be avoided. I have in mind jargon-answers like "Propositions in religion are nonsense, but they manage to *show* something lying beyond facts in a *mystical* way" or "There are no religious truths, but only rules of a religious grammar". How could such a jargon satisfy us? On the one hand, the 'showing' metaphor is precisely what is strange and what we need to leave alone or explain away in these contexts, for it provides no explanation. The word 'mystic' has the same problem, but it is a little worse, for it reminds one of superstitious obscurity. On the other hand, when dealing with the later Wittgenstein, 'grammar' becomes a suspicious word. Saying something like "religious discourse is part of the language game of religion and follows its own rules of sense; therefore, religious discourse makes sense", is very fishy.[3] Why should we accept those rules of 'grammar' or even the *talk* about 'grammar' in theology? What is the meaning of 'grammar' here and elsewhere? If mathematical equations are 'rules of grammar', for instance, should we think that theology and mathematics are part of a comprehensive 'grammar'? Are 'rules of theology' somehow 'necessary'? Are rules of mathematics and theology the same sort of rules? The word 'grammar' is jargon that has invited jargon abuse.

We need to take a different road. We know that Wittgenstein was interested in religion and respected religious writers such as Weininger, Tolstoy, Dostoevsky, and Kierkegaard. Let us begin with the obvious fact that he was a kind of religious person or, as he preferred to say it, "I am not a religious man, but I cannot help seeing every problem from a religious point of view" (see Malcolm 2002, 24). In the following sections I argue, tentatively, for the compatibility of his personal views with his early and later philosophies.

2 On Wittgenstein's *Early* (Christian) Religious Values

It is a widespread belief that Wittgenstein was a sort of 'mystic' at the time of the *Tractatus*. What sort? In a letter from 1919, Russell told Ottoline Morrell, who was herself a sort of 'mystic', that Wittgenstein "has become a complete mystic" who was reading Silesius and Kierkegaard, but that "all started with William James's *Varieties of Religious Experience*", Tolstoy, and Dostoevsky (Russell 2002, 198-9). However, as already mentioned, the label 'mystic' is not helpful at all. The word suggests a mystery, perhaps something superstitious. For Tolstoy, for instance, the mysterious and mystical was just the opposite

---

[3] Although such a rough view is not explicitly defended in the literature, it fueled, for instance, the classical debate between Philips (2005) and Nielsen (2005).

of the simple, clear, and reasonable teachings of Christ (Tolstoy 1922, 113). Moreover, as observes Tolstoy when criticising traditional religion, "the recommendation to obey the moral law was put in the most obscure, vague, and mystical terms" (81). Indeed, the mystical, the vague, and the obscure are very close. Thus, the word 'mystic' does not give us anything useful and concrete as a Wittgensteinian view.

Fortunately, there are facts that allow us to get concrete here, for Wittgenstein's friend Ludwig Hänsel is a good source. He provides valuable information about the issue in his diaries from the time that he met Wittgenstein in 1919 in the Prison Camp of Monte Casino. He notices that for Wittgenstein the "gospel faith" is "astonishingly certain" (Hänsel 2012, 47). This means that "the gospel is sacrosanct, untouchable, above all talking" (44-5). At the time of the *Tractatus* at least, this relates to his conviction that Tolstoy's presentation of the gospels – "heretic" according to Hänsel – was *accurate*:

> Wittgenstein has unshakable faith in the accuracy (*Genauigkeit*) of Tolstoy's textual work – he prefers to believe in variants unknown to us rather than in arbitrariness. (55)

Considering that Tolstoy might not be the most precise scholar regarding the bible, Wittgenstein's view is astonishing. The motivation behind such faith in Tolstoy's interpretation, however, is as astonishing as interesting: Wittgenstein was really touched by the message of Tolstoy's *Gospel*. Hänsel was impressed by his seriousness, a seriousness that went to the point of *conversion*:

> The depth and seriousness with which Wittgenstein thinks of his conversion, with which he suffers from procrastination. (56)

In which way the conversion could take place, and to what exactly Wittgenstein would convert, we might never know. However, we know through Hänsel that quite apart from the conversion plan, Wittgenstein indeed accepted essential traits of a Tolstoian *Weltanschauung*. In Hänsel's words, Wittgenstein saw the Tolstoian/Christian *Gospel* in the following way:

> Relationship to God and to the Gospels strengthened by Tolstoy's *godless* religiosity. Jesus is God because he is the man in whom there is nothing ethically deficient, because he is good without overcoming. *He does not want to accept that God means something else, namely Creator, Lord of Being*, and that the angels are not God despite their unswerving ethical purity. (51; emphasis added)

This shows that Wittgenstein agreed with the essence of Tolstoy's "heretical" views, particularly with the belief that God is among us

(one finds God in other human beings), that Christ is the example of what is moral, and that an external God is a non-needed fiction (see Tolstoy 1922, 420-1, Recapitulation III). As we will see in what follows, there is also agreement in some relevant details between the *Tractatus* and Tolstoian Christianity.

However, is such an agreement compatible with the *Tractatus*? I think it is, if we do not transform the book into a metaphysical doctrine of nonsensicality about God and ethics. Wittgenstein's *non*-commitment to certain doctrines, as italicised in the quote above, is crucial. First, as Hänsel makes clear, Tolstoy's *Gospel* is in a sense godless. I.e., what is really fundamental is the ethical perfection expressed in Christ, whereas God the Creator is dispensable. Therefore, we need to investigate the dispensability of a creator and the viability of ethics, for there are no sentences of ethics (TLP, 6.42).

We must be quite careful when we interpret sentences like "God does not reveal himself *in* the world" (TLP, 6.432) or when we want to grasp what it means that what makes the world non-accidental "must lie outside the world" (TLP, 6.41). This cannot mean that a Lord of Being created the world with ethical "necessity" and that such Lord and his ethical imperatives are outside the world in the realm of value. This would not agree with Wittgenstein's non-acceptance of a Lord of Being (godless Tolstoianism). Moreover, and this is essential, the philosophy of logic of the *Tractatus* would not allow for such a conclusion anyway. It is crucial that we stick to what the *Tractatus* really demonstrates (its limits) and to what the arguments in the book can answer for honestly.

The point of the mentioned passages really concerns what takes place *in* the world. That God does not reveal himself in the world means that there is nothing like a miracle of God, for all facts (all are contingent, of course) are dealt with by science. Of course, if God's existence is erroneously supposed to be a necessity (obviously, it is not a tautology), then it cannot be derived from the contingency of the world anyway. As Wittgenstein points out in his *Lecture on Ethics*, when we look at the world scientifically, i.e., by considering all true propositions that we know (TLP, 4.11), "everything miraculous has disappeared" (LE, 43). It is *despite that* that God-Christ and ethics are fundamental. Evidently, Wittgenstein (and Tolstoy) did not believe in miracles (Tolstoy 1922, 284).

One *might* see the world differently, considering that the very existence of the world might bring us to a mystical *feeling*. The point here is that there is *no logical compulsion* for any of the alternatives: the scientific or the religious. Logic itself, and all that we know a priori, does not imply a specific worldview (see Engelmann 2016). How one feels about or sees the world might vary, but none of such views is a priori excluded or derivable from what we really know a priori. This is the result of the *Tractatus* and its logical point of view.

While Wittgenstein was a sort of Tolstoian concerning religious ethics, he was also a critic of religion as Tolstoy himself, who did not spare offering *well-argued* critiques of the whole Christendom (see Tolstoj 1922, *Gospel*, preface, and *My Religion*). It is interesting to note that Hänsel immediately understood that this critical aspect of Wittgenstein's views was a *result* of the *Tractatus*, and that it threatened his own Catholic views. In his diaries, Hänsel writes that his own "metaphysical belief" (Hänsel 2012, 72) was made weaker by Wittgenstein and asks if he himself should "remain silent", which *meant*, according to him, "disengagement from the church" (45). He admits, however, that he cannot get rid of the "intellectual search for God, of the metaphysics" (45).

Second, the *Tractatus* is arguing against the idea that one can ground ethics (as at some point Moore and Russell wanted to do) – see chapter 4 of Engelmann 2021a. Note, however, that the fact that ethics or value is ungrounded does not imply that one should not live an ethical (or religious) life and have values. That would be like not playing or listening to Beethoven because his musical principles of harmony are not grounded philosophically. The point is rather: if one wants to live an ethical (or religious) life, one does it because one accepts it (in spite of everything), and not because one makes a philosophically grounded choice, a sort of derivation from more fundamental principles or a priori truths. The "philosophically grounded" in all fundamental philosophical questions is an illusion that the symbolism of the *Tractatus* dissolves (see Engelmann 2021a, ch. 4).

Therefore, in a Dostoevskyan mood, one could say that the acceptance of the ethical might take place despite everything. Later, in a meeting with the Vienna Circle, Wittgenstein says, against Schlick, that the deepest view on ethics is not the philosophical one that says "*p* is right and, therefore, God wants *p*", but the religious one that says "God wants *p*, therefore *p* is right" (WVC, 115). The latter view is deeper, for Wittgenstein, not because he is an 'irrationalist' who asks us to accept absurdities. Quite the opposite. It is deeper because it makes clear that there is no grounding for *p*. One can elucidate ethics, but one cannot ground it logically/philosophically. Without grounding, all one can do is accept *p* along with God, or not. "God wants *p* …" is just another way to say, "I cannot go further than this, I simply acknowledge the limit of justification", for obviously God itself is no explanation or grounding at all.

In the *Tractatus* Wittgenstein makes the point by saying that the "ancients" were at least right for not trying to make it appear as if everything was explained, as supposedly "moderns" do (see Engelmann 2016). The ancients "have a clear and acknowledged terminus" (TLP, 6.372). So, from a logical point of view, they were clearer than the moderns. Note that for Tolstoy there is no grounding for the teachings of Christ either. He understands Christ as saying: "My teaching is not

proved in any way, except that men give themselves up to it, because it alone has the promise of life for men" (Tolostoj 1922, 433).

The acceptance of ethical-religious values might depend on examples that one sees, hears, and reads about. Supposedly, for Christians, Christ is an example to be followed. He is certainly so according to Wittgenstein's Tolstoian view. Personally, for Tolstoy, the Russian peasants that he met had an important role in his conversion. They helped him to change his life and accept the teachings of Christ (Tolstoj 1922, 40-1). The change in Tolstoy's life occurred when he stopped looking for the solution of the problem of life and looked at those who live without that problem (Tolstoj 1922, 48-9). He tells us that he looked at two wrong places before solving his problem. First, he thought that science would teach him. That was not true, for science does not deal with that problem (it deals with the problem of describing the world outside the perspective of the individual who asks such questions). Second, he thought that philosophy could help him, especially Schopenhauer. That quest resulted in a big disappointment. Schopenhauer said that life had no meaning, therefore he certainly did not understand the meaning if there is one. So, if there is a meaning of life, he thought, the best would be to try to find it among those who think that there is a meaning (in his case, the peasants). However, once one grasps the meaning of life, one knows nothing more except that the problem vanishes, and cannot therefore instruct someone else, but only say: "Formerly I did not see the meaning of life; now I see. I know no more" (Tolstoj 1922, 433). Of course, TLP, 6.521 is a quite interesting rephrasing of this point.

There is another important result for the lack of grounding for what has value. If one accepts that one needs to live an ethical life, one will not go on and impose dogmas on other people. If the *Tractatus* is right, dogmatism does not work logically, given the lack of ground for ethics and value (note that this is also true for a grounding of a "scientific worldview" (Engelmann 2016)). A dogmatic person concerning ethics and religion, one might say, is a person that does not understand the logic of our language and thought, which cannot ground *a priori* 'principles'. One might say, therefore, that the *Tractatus* is quite compatible with Tolstoy's attack on dogmas of Christendom (see preface to Tolstoy's *Gospel*).

What the lack of grounding of ethics also shows, logically, is the need for tolerance concerning other forms of religion (those that one does not accept as his own). Indeed, a Tolstoian Christian might admire other kinds of religious lives, as Tolstoy's *Hadji Murat*, a Muslim, makes clear. As we know, Wittgenstein read this book in 1912, right after its posthumous publication and thought that it was "wonderful" (Wittgenstein 2005a, 35).

There is also the question of how one might express one's ethical life. Presumably, one will rather express it in actions. If successful,

one's actions might show what the ethical life is. One might not even need to talk. However, actions might include non-dogmatic talking. This means talking in a personal way, in the first person, as Wittgenstein supposedly did in his *Lecture on Ethics* (see LE, 41). He tells us that that was indeed his intention (WVC, 118). Presumably, a confession telling one's story could work properly here.

Thus, it might be no accident that Tolstoy told us his life story in his *Confession*. Moreover, one cannot underestimate the value of a confession in Wittgenstein's own life. Already in 1919, he talks about it with Hänsel (2012, 52), but apparently does it only in 1936-37. The fundamental value of an honest confession also expresses the view of Dostoevsky, as is made clear in *Brothers Karamazov* (I, I, 5), where the significance of a confession is elaborated in Zosima's teachings and its difficulties presented in the life inside the monastery. Besides, first person report and confessions are the Jamesian doors into religion in *Varieties of Religious Experience* (see Engelmann, Floyd, forthcoming). Thus, confession in a context of honest testimony and willingness to act in life appears as a central aspect of an honest view of religion for Wittgenstein, and this agrees with his favourite religious authors. In what follows, we will see that this is in the background of the *Philosophical Investigations*. Later in section 6 I return to the significance of confessions for Wittgenstein, particularly of his 'hidden' confession in the *Investigations*.

3 *Later* Views: A Tightrope Walker

Independently of the radical changes that his early philosophy went through after the recognition of "grave mistakes" (PI, preface), Wittgenstein always kept the fundamentals of his early ethical/religious views. This is by itself a quite significant fact. What changes is the way that he presents his views by considering some complications derived from them. He adapted his views to new challenges and, arguably, developed quite interesting views on the subject. One of those complications is the variety of religions, the fact that religion comes in very different dressings and cultural backgrounds, as is discussed in his *Remarks on Frazer* (see Engelmann, Floyd, forthcoming; Engelmann 2016).

In what follows, I will not be able to show that his religious views are indeed compatible with his later philosophy. This would be a complex and long task that I cannot fulfil here. Instead, I will suggest that for Wittgenstein *himself* his philosophy is compatible with his views on religion (and perhaps other views) as long as the religious views are completely honest regarding their lack of grounding. I do not intend to show him right or wrong about this. I begin by showing how the early and later views come together.

It is to Drury that Wittgenstein explains in a nutshell, probably in 1930, how his old Tolstoian view is supposed to work. The following passage links the early and the later views:

> But remember that Christianity is not a matter of saying a lot of prayers, in fact we are told not to do that. If you and I are to live religious lives, it mustn't be that we talk a lot about religion, but that our manner of life is different. It is my belief that only if you try to be helpful to other people will you in the end find your way to God. (Drury 1984, 114)

The most important thing in religion is (or should be) living according to it, and not praying and talking. For the early and late Wittgenstein, the real issue is the "ethical relation" with Christ, a human being "who is God" (PPO, 223; MS, 183, 215). As he had already told Hänsel in 1919, Christ is "the perfect one" (see PPO, 221, 223, 227, 241, 243; from 1937). Early and later Wittgenstein, in his tentative Christianity, refuses doctrines or dogmas as part of serious religion. It is the non-theoretical character of religion and its significance for a change in life that really matters:

> I believe that one of the things that Christianity says is that sound doctrines are all useless. That you have to change your *life*. (Or the *direction* of your life.) (CV, 53; from 1946)

However, the later Wittgenstein is more open to some complicating facts concerning how religion is practised; for instance, the fact that strange/miraculous doctrines may be believed (life after death, final judgment, and so on). Evidently, not all Christian thinkers are opposed to such views, like Tolstoy was. Dostoevsky, for instance, had firm belief in immortality and put all his hopes in life after death (see Frank 1988, 296-309).

In 1930-31, Wittgenstein still thought according to a purely Tolstoian perspective when he argued with Schlick that talking was not essential to religion, and that he could imagine a religion in which there is no doctrine, "no talking" (WVC, 117). However, such a claim is suspicious, for how do we determine the 'essential' here? Moreover, it is a fact that the most traditional religions on earth have a lot of doctrinal talking and one might say that it indeed appears to be the case that talking is fundamental considering that people talk all the time about doctrines or presuppose them in their religious practices, in their reports about it, and so on. Is one not even asked to convert other people? How is such an activity to take place if not in talking about religion? One might show how to live religiously in acts, but this will not be enough, for one lives in accordance with one specific religion

and its 'beliefs' (a belief, presumably, might be a statement about God, life, etc., or an attitude towards life, the neighbour, god, etc.).

This is one of the reasons why Wittgenstein tries to get clearer about how he stands in relation to beliefs in doctrines around 1936-37. Another reason is an urge to get clearer about what he himself can believe honestly concerning Christianity:

> Not the letter, only conscience can command me – to believe in resurrection, judgment, etc. To believe not as something probable but in a different sense. (PPO, 157).[4]

The different sense of belief is, of course, faith. The trouble is, as I have argued, that "the Christian solution of the problem of life" seems to require "salvation, resurrection, judgment, heaven, hell" (PPO, 169). Besides, apparently many people *honestly* believed those things (Dostoevsky, for instance). However, the real issue underlying this is that if the example of Christ implies a change of life, as Wittgenstein often emphasises, what happens is that "if one lives differently, one speaks differently", "one learns new language games" (PPO, 169). Thus, one might imagine religion with "no talking", but religious people *do* talk a lot.

Since one must mean what one says, it may seem that the concepts, 'salvation', 'final judgment', 'resurrection', and so on, are needed in Christian "language games" after all, and that one must have a faith grounded in them. However, Wittgenstein did not take this extra step into ordinary religion. Whereas the Christian ethical demand always appeared to him as the correct demand on how one must live one's life, some concepts used in Christianity were difficult to swallow. This, again, is very Tolstoian, for his *Gospel* does not contain the story about Christ's resurrection and other passages that are difficult to swallow for *us*, modern human beings (miracles, for instance). There is a thin line between living a religious life, accepting certain concepts, and living dishonestly. In fact, this was a problem for Tolstoy after his conversion, for he had to struggle against all superstitious thinking of the Orthodox Russian peasants who were the inspiration that brought him back to Christianity. He tells us that when he was ready for conversion, he thought the following:

> I was now ready to accept any faith that did not require of me a direct denial of reason, for that would be a lie… (Tolstoj 1922, 47)

This meant getting rid of superstitions, for "much that was superstitious was mingled with the truths of Christianity" (Tolstoj 1922, 49).

---

4   On resurrection see also CV, 33.

For Wittgenstein, this conflict takes place in a way that is a little different, perhaps because of his admiration for Dostoevsky. On the one hand, one might get the impression that only at a higher stage inside a religious life strange concepts like 'resurrection' can play a real role (see PPO, 155, 181). One needs a very strong faith to go as far as believing in resurrection (Dostoevsky had it). On the other, one can only accept such concepts in religion *honestly*, of course, otherwise religion is a lie. This is why Wittgenstein writes:

> I think I should tell myself: "Don't be servile in your religion!" or try not to be! For that is in the direction of superstition. (MS 183, 198; PPO, 207)

This means the following:

> I *believe* that I should not be superstitious, that is, that I should not perform magic on myself with words I may be reading, that is, that I should and must not talk myself into a sort of faith, of unreason. (PPO, 203)

Wittgenstein, as an honest religious thinker, therefore, expresses his opposition to the uncritical acceptance of strange religious concepts:

> I don't have a belief in a salvation through the death of Christ; or at least not yet. I also don't feel that I am on the way to such a belief, but I consider it possible that one day I will understand something here of which I understand nothing now; which means nothing to me now & that I will then have a belief that I don't have now. (PPO, 201-3)

In order to make compatible his reasoning concerning his beliefs and his abhorrence of superstition, Wittgenstein points out that it is a misunderstanding to consider that 'belief' means the same in ordinary beliefs and in religious beliefs. The latter involves a whole worldview and, thus, is not like a specific belief that we give up if it is an error, a false opinion. If we give up a religious belief it is not because it was a wrong *opinion* that we simply change in light of new facts, but rather because we now see it as a wrong way to look at things (see PPO, 231; LC, 53-9).

One might also say that the later Wittgenstein is more conscious of the difficulties involved in being a religious person in the modern world. The honesty of the religious thinker, thus, comes to the forefront when he is confronted with a so-to-speak unfavourable reality. Wittgenstein's fundamental later view, I take it, is that "an honest religious thinker is like a tightrope walker", for he has no real grounding for his certainty or beliefs, yet he can keep his positions

with a great effort (CV, 73). One cannot pretend that one is moving in a ground where there is a foundational argument to sustain one's position when one talks about religion. One grasps religious and ethical views – arguably, for Wittgenstein the only deep/real ethics is religious – without any support, despite everything.

This kind of honesty was already the issue in his early thought, as seen above, although at that time Wittgenstein did not seem to worry much about the lack of sustainability of such views when they are brought in contact with the real world. One can say, thus, that the later Wittgenstein is more realistic concerning his own religiosity. Simultaneously, he seems more inclined to go deeper into religious concepts that play an important role in certain religious thinkers (for instance, Dostoevsky). However, he does not give in to doctrines after all (one might say that he remains Tolstoian).

4 Backdoor Metaphysics? Religion and the Inevitability of a *Weltanschauung*

It is not only religion that Wittgenstein leaves alone in his works. Religion is part of what one might call *Weltanschauung*, i.e., a world view, a general way to look at things, so that one sees facts as expressing certain rules or tendencies. This *kind of metaphysical view is not criticised by Wittgenstein except when it is used dogmatically or dishonestly*. This kind of metaphysics, if you call it nonsense or not, is simply inevitable.

One interesting example is how one takes history. On reflection, one might see history as the accumulation of knowledge and economical power (capitalistically or socialistically), directed towards a better future of progress. However, someone like Spengler or Wittgenstein might see things differently. For them, development "comes everywhere to an end" so that developing is seen as "a self-containing whole which at some point will be completely present & not a sausage that can run indefinitely" (MS 183, 21; PPO, 29).

The very general traits of a *Weltanschauung* are relevant because they show how one sees our human form of life. Interestingly, perhaps one of the most insightful remarks on Wittgenstein's religious *Weltanschauung* comes from Carnap, who disagreed with him (arguably all members of the Vienna Circle did). In his *Intellectual Autobiography*, Carnap writes:

> Once when Wittgenstein talked about religion, the contrast between his and Schlick's position became strikingly apparent. Both agreed of course in the view that the doctrines of religion in their various forms had no theoretical content. But Wittgenstein rejected Schlick's view that religion belonged to the childhood phase

of humanity and would slowly disappear in the course of cultural development. When Schlick, on another occasion, made a critical remark about a metaphysical statement by a classical philosopher (I think it was Schopenhauer), Wittgenstein surprisingly turned against Schlick and defended the philosopher and his work. (Carnap 1991, 26-7)

We cannot know who the mentioned metaphysical philosopher was and the content of the discussion. However, it might well be the case that Wittgenstein "defended" the philosopher because the issue in question was, perhaps, a *Weltanschauung*. In all of Wittgenstein's religious views there is an obvious dislike of and opposition to the fundamental views of our times, the views behind capitalism and socialism, and of utilitarianism: the ideas of progress and calculus of utility (see CV, 6-8).[5] Note that Schlick, and apparently all *the anti-metaphysical* members of the Circle held the opposite "metaphysical" worldview regarding history. One of the aspects of this idea is, as Schlick assumed, the disappearance of religion as part of "the childhood phase of humanity". In Wittgenstein's *Weltanschauung*, however, religion is a fundamental aspect of what makes us humans. Perhaps only art could have a similar status for him among all the important things that characterise us, human beings. Those are traits of Spengler's views on history, culture, and religion – see Engelmann 2016; 2021b. When Wittgenstein read him in 1930, he wrote that most of the thoughts in Spengler "are completely in touch with what I have often thought myself" (MS 183, 16; PPO, 25).

Wittgenstein's *Weltanschauung* in the *Philosophical Investigations* appears directly at the beginning of the book. The incompatibility of his personal views and our time evidently applies to his philosophy as well, as is suggested by the motto and expressed in the preface with the expression "darkness of our times". He told Drury that "my type of thinking is not wanted in the present age" (Drury 1984, 160).

Possibly, however, Wittgenstein's religious worldview expresses itself in his philosophy as a whole. This possibility needs to be elucidated, but space prevents me from doing this here. Instead, I discuss the sole two occurrences of religion in the text of the *Investigations* in the next section, and then in section 6 I uncover a religious point of view expressed in a confession in the book.

---

[5] On this issue, see Engelmann 2016; 2021b.

5   *Investigations*: Two Remarks on Religion

In Wittgenstein's works after the *Tractatus*, God and religion are not topics of discussion, except for his *Lecture on Ethics*, which is arguably still written in the spirit of the *Tractatus*. Nonetheless, there are two references in the *Investigations*. The first is a parenthetical remark in PI, I, § 373:

> Grammar tells what kind of object anything is. (Theology as grammar.)

As I suggested previously (Engelmann 2013, 262-4), this remark points to the deflation of the notion of 'grammar' rather than to an inflation of the notion of 'theology'. That is, one should see that the notion of 'grammar' does not have the weight of a discipline of the bounds of sense/nonsense, a discipline that tells us about "combinatorial possibilities", for nothing of the sort could find a home in theology. 'Grammar' is not a discipline of sense and nonsense grounded in necessary rules concerning possibilities. In most cases, 'grammar' can be replaced with "use in language" or "descriptions of language use" in Wittgenstein's later philosophy (Engelmann 2013, chs. 3-5).

One might well think that the talk about 'grammar' in religion does not even get to the heart of the matter: it remains at the superficial level. This can be gathered from the context where Wittgenstein indeed discusses this remark from 1937 in MS, 183 (PPO, 211). One learns "the grammar of the word 'God'" by knowing simply what is said about God, for instance, "one kneels & looks up & folds one's hands & speaks, & says one is speaking with God" (PPO, 221). This 'grammar' merely gives us habits of behaviour in certain practices and at best it can be used to teach children how to behave in religious contexts in certain traditions. However, right after talking about the grammar of the word 'God', Wittgenstein writes about what really matters in a coded remark: "A *religious question* is either a question of life or it is (empty) chatter. This language game – one could say – gets played only with questions of life" (MS, 183, 203; PPO, 211). All those rules of 'grammar' are obviously not a "question of life", but minor matters when compared to what really matters.

There is a second appearance of religion in "Part II" of the *Investigations*:

> Religion teaches that the soul can exist when the body has disintegrated. Now do I understand what it teaches? Of course I understand it – I can imagine various things in connection with it. After all, pictures of these things have even been painted. And why should such a picture be only an imperfect rendering of the idea

expressed? Why should it not do the *same* service as the spoken doctrine? And it is the service that counts. (PI, II, iv § 23)

The remark appears in the context of a discussion concerning automata in which Wittgenstein claims that it is nonsense to say, in ordinary circumstances, "I believe that he is suffering, but am certain that he is not an automaton" (PI, II, iv, § 19). The nonsensicality consists in presenting the sentence as if it was an opinion similar to other certainties and uncertainties. However, not taking other human beings as automata is not an opinion, but an attitude. Usually, we do have an "attitude towards a soul" in relation to other human beings, which makes such pseudo-certainty misleading, for such an attitude is not an opinion that can be seen as an error at all (PI, II, iv § 22). So, § 23 elaborates on the notion of 'soul' with religion and painted pictures.

If we take into account what has been said above concerning Wittgenstein's explicit opposition to doctrines in religion, and if we remember that this was already a fundamental point very early in his career, we understand that what he is really saying here is that if we consider that the pictures/paintings of a soul as distinct from a body do as much service as a doctrine, what he means is that both do very little or no service at all. If we consider religion seriously, those aspects, pictures and doctrines, should not be determining factors. Moreover, as Wittgenstein makes clear in several places, imagining something does not mean understanding (see, for instance, PI, I, §§ 393-8). The fact that "I can imagine many things" does not mean that I understand a thing about what is in question.

However, there is a deeper use of religion in the *Investigations*, precisely in a context where 'religion' or 'God' does not appear at all.

6 A Religious Point of View Expressed in a Confession

Although in the *Investigations* nothing is said about the motivation for the systematic critique of the *Tractatus* presented there, the motives behind it are religious in a Wittgensteinian sense. The critique is a confession of sins of a sort, i.e., an admission of errors in order to maintain integrity, honesty, and seriousness. Religion is in the background of Wittgenstein's examined life, in the confession or admission of errors present at the core of his "edifice of pride", namely, the *Tractatus*.

In Engelmann (2013) I argued that Wittgenstein's later philosophy is characterised by traits that were incorporated gradually: what I call the "genetic method" takes centre stage only in the *Blue Book* (Engelmann 2013, chs. 3-4), the anthropological view appears first systematically in the two versions of the *Brown Book* (chapter 4), and finally the systematic critique of the *Tractatus* by means of the

application of the genetic method appears for the first time in the *Philosophical Investigations* (chapters 4 and 5). I suggested in a footnote that there was "an interesting connection of the application of the method to the T[ractatus] and Wittgenstein's private life", the most striking example of "how his life and philosophy are closely related" (Engelmann 2013, 299 fn. 22). I had in mind the contemporaneity of his later critique of the *Tractatus* and his confession of weaknesses and errors (or sins, if one prefers) to several of his friends and family in 1936-37, as is described for instance by Pascal (1984) and attested in Wittgenstein's letters (see, for instance, PPO, 281-91). Wittgenstein's first step in 1936 was to confess to his old friend Hänsel, to whom he first said something about confession already in 1919 (see section 2), that he had lied to him about his family origins when they first met. He then extended his confession(s) to family and friends.

As I argued in chapter 4 of Engelmann (2013), Wittgenstein applies the genetic method to his own early philosophy in the *Investigations*, i.e., he uncovers the false pictures, analogies, and trains of thoughts that led him to the central views and the "grave mistakes" of the *Tractatus* (PI, preface). He does so to exemplify his own method with his own case in many remarks between PI, I, § 1 and § 136. However, this gesture is a lot more than that. For if we see the context in which many of the remarks criticising the *Tractatus* in the *Investigations* were originally written, we can determine that the critique expresses a religious attitude. MS, 157a and MS, 157b, where the most important aspects of the genetic critique of the *Tractatus* (PI, I, §§ 89-136) first appear, follow a time of intense religious/ethical reckoning when Wittgenstein wrote an enormous quantity of remarks on religion, Christ, death, and personal beliefs in MS 183. In fact, the first critical evaluation of the *Tractatus* appears in MS, 183, 152 (PPO, 161), 27 January 1937. On 9 February then, after he had written some remarks about the source of the "sublime" conception of logic in the *Tractatus* in MS 183 (see PPO, 161, 167, 173), Wittgenstein writes extensively about it in MS, 157a.

Part of those remarks on religion were presented in section 3, where we saw how Wittgenstein had to examine his old convictions about religion, for he was certain of one thing: "Let me not shy away from any conclusion, but absolutely also not be superstitious! *I do not want to think uncleanly*!" (MS, 183, 173; PPO, 181). What characterises a (Christian) religious struggle against the unclean is trying to get rid of vanity and pride. This struggle is documented in his notebooks. What happens at the time of his confession in 1936-37 is that his most important object of pride needed to be addressed:

> The *edifice of your pride* has to be dismantled. And that is terrible hard work. (CV, 26; MS, 157a, 57r)

The edifice of his pride was the *Tractatus*, whose fundamental moves are ethically and philosophically examined in 1936-37. In the mentioned MSs Wittgenstein searches for the source of his errors in the *Tractatus* as a religious person searches for the source of her sins in order to plainly confess the deed and what motivated it. One of the sources was the misunderstanding of the ideal of the sublimity of logic, which was taken as the a priori essence of language, thought and world. This led Wittgenstein to grave errors.

Evidently, sometimes one needs to confess that what looked like a good action was in fact motivated by something bad (for instance, one helps a friend out of pride and not out of love). From the religious point of view, as the quote makes clear, the source of the errors of the *Tractatus* was the pride of showing in a sublime logical symbolism nothing less than the essence of everything.

7 'Religious Puzzles' and the Critique of Religion

One of Sraffa's points in his comments on the *Blue Book* was that there is a similarity between metaphysical and religious puzzles (see section 1). He had in mind issues of free-will and the existence of God (presumably their compatibility). Thus, Wittgenstein should not leave religion alone. When religion is puzzling in this way, its destiny should be the same as metaphysics.

We have seen that Wittgenstein himself was critical of religion in several instances. Often in discussions of Wittgenstein on religion one forgets how critical he was, particularly when one intends to use his philosophy to defend or 'understand' religion. However, indeed he did not use his philosophy directly in the *Investigations* or in other later works as a critique of religion. There are a few reasons for this, I think. First, the fact that he discussed aspects of religion grounded in his philosophy in his *Lectures on Religious Belief* in 1938-39. In these lectures, arguably, he uses his philosophy to show how to avoid *misunderstandings* concerning religion. I think that this aspect of his philosophy of religion might be understood as a tentative elucidation of the *possibility of walking the tightrope*. The strategy of avoiding misunderstandings is obviously an important characteristic of the philosophy of the *Investigations* where he aims at "clearing misunderstandings away" (PI, I, § 90). Second, it is also important to notice that in 1941, the year that Sraffa commented on the *Blue Book* and mentioned the free-will puzzle (see section 1), Wittgenstein discussed the problem in his *Lectures on Freedom of the Will* (see Wittgenstein 2017). Third, he might have preferred to avoid a critique of religion in his works because religion is arguably under scientific and philosophical scrutiny or attack very often. Fourth, perhaps he never felt clear *enough* about his own religiosity from the emotional

and from the intellectual point of view. Lack of clarity might lie in the heart of the matter. The tightrope walker might avoid misunderstandings concerning religion, particularly what appears paradoxical, and get rid of "the irritation of the intellect", but the result then must be taken for what it is: "Nothing at all is *intelligible*, it is just not *unintelligible*" (PPO, 247).

Of course, the fact that Wittgenstein thought critically about religion is not incompatible with his profound admiration for it. In it he saw the ultimate source of ethics and the most extraordinary human passion.

References

Carnap, R. (1991). "Intellectual Autobiography". Schilpp, P.A. (ed.), *The Philosophy of Rudolf Carnap*. La Salle: Open Court, 3-86.

Dostoevsky, F. (1992). *The Brothers Karamazov*. Transl. by R. Pevear; L. Volokhonsky. New York: Vintage Classics.

Drury, M. (1984). "Conversations with Wittgenstein". Rhees, R. (ed.), *Recollections of Wittgenstein*. Oxford: Oxford University Press, 76-171.

Engelmann, M.L. (2013). *Wittgenstein's Philosophical Development: Phenomenology, Grammar, Method, and the Anthropological Perspective*. Basingstoke: Palgrave Macmillan.

Engelmann, M.L. (2016). "The Faces of Necessity, Perspicuous Representation, and the Irreligious 'Cult of the Useful': The Spenglerian Background of the First Set of Remarks on Frazer". Albinus, L.; Rothhaupt, J. (eds), *Wittgenstein's Remarks on Frazer: The Text and the Matter*. Berlin: De Gruyter, 129-74.

Engelmann, M.L. (2018). "Phenomenology in Grammar: Explicitation-Verificationism, Arbitrariness, and the Vienna Circle". Kuusela, O.; Ometita, M.; Ucan, T. (eds), *Wittgenstein and Phenomenology*. London: Routledge, 22-46.

Engelmann, M.L. (2021a). *Reading Wittgenstein's Tractatus*. Cambridge: Cambridge University Press.

Engelmann, M.L. (2021b). "La articulación de aspectos de un 'punto de vista religioso' en Wittgenstein y la crítica del 'culto de la utilidad' de la Modernidad (de Spengler a Frazer, por medio de Keynes, Kierkegaard, Hänsel, Tolstoi y Dostoievski)". Velasco, F.S.; Maza, N.A.R; Almanza, K.M.C. (eds), *Perspectivas wittgensteinianas. Lenguaje, significado y acción*. México: Tirant lo Blanche, 305-54.

Engelmann, M.L. (2021c). "O Caso Mauthner: O Tractatus, o Círculo Kraus e a Significatividade da Negação". *Analytica*, 25(2), 67-83.

Engelmann, M.L. (2022). "O Kierkegaard Krausiano de Wittgenstein à Época do *Tractatus*". Dall'Agnol, D.; Peruzzo Júnior, L.; Sattler, J. (eds), *TLP 100: Revisitando a Obra de Wittgenstein*. Curitiba: Pucpress, 265-82.

Engelmann, M.L. (forthcoming). "The Meaning of Silence: Wittgenstein's Kierkegaard and the *Tractatus*". Arnswald, U.; Jareño-Alarcón, J. (eds), *The Nature of Religious Belief in Wittgenstein's Philosophy*.

Engelmann, M.L.; Floyd, J. (forthcoming). "Wittgenstein on Religion as a Form of Life: From a 'Jamesian Type' to Links with Religious Community".

Pritchard, D.; Venturinha, N. (eds), *Wittgenstein and the Epistemology of Religious Belief*. Cambridge: Cambridge University Press.

Frank, J. (1988). *Dostoevsky: The Stir of Liberation, 1860-1865*. Princeton, NJ: Princeton University Press.

Hänsel, L. (2012). *Begegnungen mit Wittgenstein. Ludwig Hänsels Tagebücher 1918/1919 und 1921/1922*. Edited by I. Somavilla. Innsbruck; Vienna: Haymon Verlag.

James, W. (2002). *Varieties of Religious Experience: A Study in Human Nature*. Amherst: Prometheus Books.

Malcolm, N. (2002). *Wittgenstein: A Religious Point of View?* Edited with a response by P. Winch. London: Routledge.

Nielsen, K. (2005). "Wittgenstein and Wittgensteinians on Religion". Nielsen, K.; Philips, D.Z.(eds), *Wittgensteinian Fideism?* London: SCM Press, 225-79.

Pascal, F. (1984). "Wittgenstein: A Personal Memoir". Rhees, R. (ed.), *Recollections of Wittgenstein*. Oxford: Oxford University Press, 12-49.

Philips, D.Z. (2005). "Religion and Obstacles of the Will". Nielsen, K.; Philips, D.Z.(eds), *Wittgensteinian Fideism?* London: SCM Press, 280-304.

Russell, B. (2002). *The Selected Letters of Bertrand Russell: The Public Years 1914-1970*. Edited by N. Griffin. London: Routledge.

Tolstoy, L. (1922). *My Confession, My Religion, and The Gospel in Brief*. New York: Charles Scriber's Sons. Originally circulated in Russia.

Venturinha, N. (2012). "Sraffa's Notes on Wittgenstein's *Blue Book*". *Nordic Wittgenstein Review*, I, 180-92.

Wittgenstein, L. (c. 1937). "Remarks on Frazer's *Golden Bough*". Reprinted in Wittgenstein 1993, 115-55.

Wittgenstein, L. (1960). *The Blue and Brown Books. Preliminary Studies for the "Philosophical Investigations"*. Edited by R. Rhees. Oxford: Blackwell. [BB]

Wittgenstein, L. (1965). "A Lecture on Ethics". *The Philosophical Review*, 74(1), 3-12. Reprinted in Wittgenstein 1993, 36-44, references are to this reprint. [LE]

Wittgenstein, L. (1966). *Lectures and Conversations on Aesthetics, Psychology, and Religious Belief. Compiled from the Notes taken by Yorick Smythies, Rush Rhees and James Taylor*. Edited by C. Barrett. Berkeley: University of California Press. [LC]

Wittgenstein, L. (1975). *Philosophical Remarks*. Edited by R. Rhees. Oxford: Blackwell. [PR]

Wittgenstein, L. (1979). *Wittgenstein and the Vienna Circle. Conversations Recorded by Friedrich Waismann*. Edited by B.F. McGuinness, transl. by J. Schulte. Oxford: Blackwell. [WCV]

Wittgenstein, L. (1981). *Tractatus Logico-Philosophicus*. Transl. by C.K. Ogden. London; New York: Routledge & Kegan Paul. [TLP]

Wittgenstein, L. (1993). *Philosophical Occasions 1912-1951*. Edited by J. Klagge; A. Nordmann. Indianapolis: Hackett Publishing. [PO]

Wittgenstein, L. (1998). *Culture and Value: A Selection from the Posthumous Remains*. Revised 2nd ed. Edited by G.H. von Wright; H. Nyman; A. Pichler, transl. by P. Winch. Oxford: Blackwell. [CV]

Wittgenstein, L. (2003). *Ludwig Wittgenstein: Public and Private Occasions*. Edited by J. Klagge; A. Nordmann. Oxford: Rowman & Littlefield Publishers. [PPO]

Wittgenstein, L. (2005a). *Wittgenstein: Briefwechsel*. Innsbrucke elektronische Ausgabe. Edited by M. Seekircher et al. Online edition, Past Masters. Charlottesville, VA: Intelex Corporation.

Wittgenstein, L. (2005b). *The Big Typsecript. TS 213*. Edited and transl. by C.G. Luckardt; M.A.E. Aue. Oxford: Wiley-Blackwell. [BT]

Wittgenstein, L. (2009). *Philosophical Investigations*. Revised 4th ed. Edited by P.M.S. Hacker; J. Schulte, transl. by G.E.M. Anscombe; P.M.S. Hacker; J. Schulte. Chichester: Wiley-Blackwell. [PI]

Wittgenstein, L. (2017). *Wittgenstein's Whewell's Court Lectures: Cambridge, 1938-1941. From the Notes by Yorick Smities*. Edited by V. Munz; B. Ritter. Oxford: Wiley Blackwell. [WCL]

Wittgenstein, L. (2009). Wittgenstein Source Bergen Nachlass Edition. Edited by A. Pichler. Bergen: Wittgenstein Archives at the University of Bergen. `http://www.wittgensteinsource.org`.

Wittgenstein on Use, Meaning and the Experience of Meaning

Elena Valeri
Università Ca' Foscari Venezia, Italy

Abstract In this paper I discuss the relationship between the notion of experience of meaning, introduced by Wittgenstein in the *Philosophical Investigations*, and the idea of meaning as use, central to much of his thought. In particular, I ask whether the former is to be seen as a development, an integration and a specification of the latter, or whether its emergence in Wittgenstein's work indicates a change in his attitude to meaning. My answer is that the notion of experience of meaning does not bring back some form of psychologism, but rather it develops, integrates and specifies that of use.

Keywords Wittgenstein. Use. Rules. Experience of Meaning. Anti-psychologism.

> Yes, young lady. I cannot give you any rule.
> One must have a feeling for it, and well,
> that's it. But in order to have it, one must
> study, study, and then study some more.
>
> Eugène Ionesco, *The Lesson*[1]

> So does the word 'Beethoven' have a
> Beethoven-feeling?
>
> Ludwig Wittgenstein,
> *Last Writings on the Philosophy
> of Psychology II*[2]

1 Introduction

While Wittgenstein was introducing and discussing – especially, but not exclusively, in the *Philosophical Investigations* – his well-known idea that there is some kind of link between meaning and use, a question gradually arose which led him to new questions and problems. The question can be formulated as follows: Is there something more to meaning than its use – something that escapes 'use' and that 'use' fails to account for? In this paper, I aim to illustrate what might be considered the stages in the gradual emergence of this question and Wittgenstein's treatment of it, and my focus will be on how to interpret the topic which Wittgenstein arrived at: the experience of meaning. Specifically, the point at issue is whether we should regard this reference to the experience of meaning as a kind of re-evaluation of psychologism and, together, a questioning (and downgrading) of the notion of use, or whether it is consistent with his anti-psychologism and his appeal to use. In order to provide some answers, I will first outline how and for what purpose Wittgenstein introduces the notion of use; then, I will examine the context in which the notion of experience of meaning occurs; and finally, I will make an interpretative proposal for how to understand the relationship between (meaning as) use and the experience of meaning.

2 Meaning and Use

Readers of the *Philosophical Investigations* will be well aware of Wittgenstein's constant reference, when he speaks of the meaning of a word, an expression or a sentence, to use (*Gebrauch*, *Verwendung*), employment (*Verwendung*, *Benützung*) and application (*Anwendung*),[3]

---

1 Ionesco 1958, 68-9.

2 LW II, 3.

3 Here I use the 2009 edition of *Philosophical Investigations*, edited by Peter M.S. Hacker and Joachim Schulte, who in their editorial preface state that they "have

and of the central role these notions play in his investigation of meaning.[4] The key section is, of course, section 43, the first paragraph of which reads:

> For a *large* class of cases of the employment (*Benützung*) of the word 'meaning' – though not for *all* – this word can be explained in this way: the meaning of a word is its use (*Gebrauch*) in the language. (PI, § 43; italics in the original, here and in what follows)

The fundamental question presented to interpreters by this paragraph is whether or not it contains a definition of what for Wittgenstein is (what he takes to be) meaning. That is, whether (a) Wittgenstein undertakes to argue that meaning is (coincides with or is identified with) linguistic use, or (b) his aim is different and, perhaps, entirely different. This is also tantamount to asking whether or not there exists in the 'later' Wittgenstein a theory of meaning as use (a use-theory of meaning) that is to be regarded as primarily different from (or even opposite and antithetical to) that of the 'early' Wittgenstein (the author of the *Tractatus Logico-Philosophicus*).[5]

The affirmative or negative answer to this question defines two opposite readings and two corresponding approaches: (1) the reading and approach of those who believe that in Wittgenstein there is, like it or not, a theory which, in competition and opposition with other

translated *Gebrauch* by 'use', *Verwendung* by 'use' or 'employment', and *Anwendung* by 'application'. 'Use' also does service for *benützen*"; "[i]n general, however, [they] have not allowed [them]selves to be hidebound by the multiple occurrence of the same German word or phrase in different contexts" (Hacker, Schulte 2009, xiv; italics in the original).

4   The central role of use is certainly evident in the case of the *Philosophical Investigations* and the texts coeval with their composition. I would point out, however, that the centrality which this notion assumes in Wittgenstein's later texts is anticipated by its significant presence in the *Tractatus*. Indeed, in the *Tractatus*, meaning is connected to the usefulness (or uselessness) of the sign (cf. TLP, 3.328), and 'usefulness' means that the sign can "determine a logical form" only if "taken together with its logico-syntactical employment" (TLP, 3.327); as we also read, one can "recognize a symbol by its sign" only by observing "how it is used with a sense" (TLP, 3.326). Although Wittgenstein speaks of "logico-syntactical employment" (TLP, 3.327), there is already a tension here between a notion of use linked to logic and a broader notion of use, which anticipates a certain view of use, found in the *Blue Book* and the *Philosophical Investigations*, as that which 'gives life' to signs, which would otherwise be 'dead' and 'inert': as Wittgenstein suggests, "if we had to name anything which is the life of the sign, we should have to say that it was its *use*" (BB, 4). Consider also this remark: "Every sign *by itself* seems dead. *What* gives it life? – In use it *lives*. Is it there that it has living breath within it? – Or is the *use* its breath?" (PI, § 432). For a more detailed discussion of the relationship between sign and use in Wittgenstein, see Perissinotto 2009.

5   Of course, this presupposes something that is neither obvious nor taken for granted, namely that there is such a thing as a theory of meaning in the early Wittgenstein. See, for example, the position of the Neowittgensteinians (Crary, Read 2000), who radically deny that there is any theory of meaning in the *Tractatus*.

and different theories (e.g. in competition and opposition with the picture-theory of the *Tractatus*, or with referentialist or ideational theories), leads back (reduces) meaning to use,[6] and (2) the reading and approach of those who recognise, on the contrary, that the emphasis on use is nothing but a step – perhaps the most important and decisive step – in the Wittgensteinian philosophical method. According to the latter reading and approach, in section 43 (and similar sections) Wittgenstein is suggesting that looking at use is a way "to cure you of the temptation", so widespread in philosophising, "to look about you for some object which you might call 'the meaning'" (BB, 1) and of the "mental cramp" (BB, 1) that this temptation produces.[7] In the view of (2), what is contained in section 43 of the *Philosophical Investigations* would not be, as in the view of (1), a theoretical definition of meaning, as if use were something definable that accompanies words, expressions and sentences, but rather a methodological indication that invites one to follow such a maxim: "Don't look for the meaning of a word (an expression, a sentence), look for its use".[8]

Even if one accepts and adheres to this methodological reading – and I think there are good reasons for doing so[9] – one is left with a question that Wittgenstein himself poses more than once: Where should I look if it is the use which I am looking for? That is, what is the use for which I am supposed to look? Such a question arises precisely because 'use' (as well as 'employment' and 'application') is a rather vague term, far from unambiguous. Even with 'use', as it were, we need to ask (and know) how it is used. Moreover, not only does the emphasis on use seem to be 'operationally' unhelpful (where should I look? What do I say, when I am asked about meaning?), but also, as Paul Snowdon (2018, 29) observes, the fact "that the term 'use' is very indefinite" is itself the reason why "[i]t is very difficult to assess" Wittgenstein's proposal. In short, whether we read it as a theoretical definition, i.e. as "meaning is use", "meaning = use", or

---

6   The problem with this reading and approach, which evaluates Wittgenstein's proposal as a theoretical hypothesis (which can then be said to be correct or not), is expressed by Paul Horwich in the following way: "Moreover, no matter how these matters are decided, his proposal surely isn't going to be *obviously* correct; but in propounding a controversial hypothesis, is he not guilty of contravening his own anti-theoretical meta-philosophy?" (Horwich 2008, 134).

7   In particular, questions such as "What is meaning?" produce this impasse: "We feel that we can't point to anything in reply to them and yet ought to point to something" (BB, 1).

8   In a late note from January 1948, Wittgenstein makes this point as follows: "Nicht nach der Begleitung des Wortes ist zu suchen, sondern nach dem Gebrauch" (It is not the accompaniment of the word that is to be sought, but its use) (MS 136, 64b).

9   It should be noted that the expression "For a *large* class of cases" (PI, § 43) already shows that Wittgenstein had no theoretical intention, since generality or universality is unanimously a hallmark of the theoretical.

methodologically as "Don't look for the meaning, look for the use", we need to know what use is and where, so to speak, it is to be found. In order to attempt some sort of answer, we can begin by taking the negative route, that is, by pointing out what use is certainly is not, and how its relation to meaning is not to be understood.

2.1 What (Linguistic) Use Is Not

First of all, (1) there is a sense of 'use' which obviously leads us to exclude that 'use' could have the meaning we are looking for: there is a clear difference between a very generally understood use and a use which has to do with linguistic meaning. 'Use' does not always convey the linguistic meaning, since there are clearly uses which have nothing to do with this kind of meaning (I mean linguistic meaning), either in the sense that there are words (expressions, sentences) which have a use, but of which we would not say that they have a (linguistic) meaning (e.g. a 'lalala' that we repeat for our own amusement, or magic words such as 'abracadabra' and 'bibbidi-bobbidi-boo'), or in the sense that there are words (expressions, sentences) which have meaning without any connection to the particular use we make of them (e.g. the word 'cat' used as a password or as a decorative motif on some wallpaper). In such cases, the use (of a word, an expression or a sentence) does not seem to be that "use in the language" (PI, § 43) of which Wittgenstein speaks in section 43, but rather a use *of* the language, so to speak.

(2) But 'use' should not be understood, even trivially, as the use I make of a word, an expression or a sentence to mean something: what I use a word (an expression, a sentence) to mean. For Wittgenstein, the point is not that I can *use* a word (an expression, a sentence) *to mean* something – assuming that it can be established "what using an expression [a word, a sentence] to mean something actually amounts to, or, comes down to" (Snowdon 2018, 30) – but that what I mean with a word (an expression, a sentence) results from (or is in) the use I make of that word (expression, sentence) in different circumstances.

(3) Even what would be the most obvious thing to do, namely, to turn to Wittgenstein's examples in order to find the characteristics of use according to him, does not seem to be decisive. An examination of the examples with which Wittgenstein begins his *Philosophical Investigations* (such as the examples of red apples in section 1, building stones in section 2, or numbers in section 8; see PI, §§ 1, 2, 8) – together with the numerous others scattered throughout his later writings – suggests that all these uses refer to "something interpersonal and social" (Snowdon 2018, 30), which can be traced back to forms of training and acquired habits (see PI, § 199). Wittgenstein's examples are admittedly, very simple, and probably deliberately simplified,

but it is true that they highlight mostly, or rather exclusively, a single aspect of use: they are imperatives, where use only "amounts to a speaker getting a hearer to do something" (Snowdon 2018, 30). So much so that it is easy to see that they leave out much of the meaning – for instance, what we might call 'descriptive' uses or meanings. One must then ask whether they really represent the 'locus' of meaning and whether they serve to shed full light on what is meant by 'use'.[10] In short, while Wittgenstein's examples give us some pointers – above all, the idea that "use in the language" (PI, § 43) is not to be understood in an intralinguistic sense – they leave the question of where to look for meaning open and undecided in many respects.

Hence, it seems far from easy to determine what this "use in the language" (PI, § 43) is that is supposed to give us the meaning. After all, it seems safe to say that Wittgenstein, who, as we shall see, tries to give us some hints as to how to understand the notion of use, was fully aware of these difficulties. In the pages of the *Philosophical Investigations*, as well as in many other pages of his manuscripts and typescripts, we find Wittgenstein often dissatisfied and constantly struggling to come to terms with the mental cramps, confusions and misunderstandings which his own repeated emphasis on use risks producing; as we might also say, and as he knew well, in philosophising it can sometimes happen that what is presented as the solution turns out to be the problem, or at least part of the problem. As we have seen, it is possible to appeal to use while remaining fully within the theoretical stance that Wittgenstein unfailingly questions.

2.2 The Rest of the Task: Gains and Losses

Clarifying what use is would only be a part of the task. Even once we have established a non-extrinsic link between use and meaning, the problem remains if, by looking for the meaning of a word (an expression, a sentence) in its use, we do not see many things about meaning that we would do much better to see: assuming that there are gains in looking for meaning in use, are we sure that they compensate for any losses? And what, if anything, would these losses be? These are questions that recur insistently in the writings of the 'later' Wittgenstein, though not always in this form, and so explicitly. It is from these questions, and Wittgenstein's (almost obsessive) engagement with them, that topics such as the experience of meaning emerge. In

---

10   As Snowdon puts it: if "in thinking of imperatives, interpersonal responses seem a reasonable aspect to bring in […] it is far harder to make this seem plausible as a model of what we might call descriptive meaning", so that "if 'use' means something like interpersonal responses there is no obvious application for the slogan [meaning is use] to large central parts of language" (Snowdon 2018, 30).

order to better understand the significance of this, however, it will be necessary to say something more about the problems that Wittgenstein's emphasis on use, so important as to be considered a hallmark of his later philosophy, can give rise to.

3 Three Problems with Use

There are (at least) three aspects that, according to Wittgenstein himself, are problematic in indicating the locus of meaning in use. It should be noted that these are not fictitious problems to which Wittgenstein already has, or thinks he already has, the answer; while his questions are sometimes rhetorical, they are not always so, and not even, I would say, in most cases. The question and answer between two or more interlocutors in which Wittgenstein assumes multiple roles and positions, so typical of the *Philosophical Investigations*, is real and not merely a dramatized staging of already established and, so to speak, archived results. Let us see in detail what these three aspects are and what problems they raise.

3.1 Use and Calculus

When asked what the use in which meaning is to be sought actually is, Wittgenstein initially (i.e. in the years of his return to philosophy, 1929-30) did not hesitate to answer that (linguistic) use is to be understood as a calculus defined, as in the paradigmatic case of arithmetic calculus, by precise and rigorous rules. At this stage, he was even convinced that the calculus was something more than a simile: as he "deliberately" pointed out to his interlocutors in the Vienna Circle, "there is not a mere analogy" (WVC, 168) between (linguistic) use and calculus; one could even say that the concept of calculus encompasses that of (linguistic) use.[11] That is to say, the use of words (expressions, sentences) is not like a calculus, but is a real calculus, because:

> [w]hat I am doing with the words of a language in understanding them is exactly the same thing I do with a sign in the calculus: I operate with them. (WVC, 169-70)

However, the certainty with which Wittgenstein expresses this identity between (linguistic) use and calculus is gradually lost. In the

---

[11] In Wittgenstein's words: "I can actually construe the concept of a calculus in such a way that the use of words will fall under it" (WVC, 168).

Philosophical Investigations, as already in the *Blue Book*, Wittgenstein casts doubt on his previous conviction by pointing out the error that may lie behind the fact "that in philosophy we often *compare* the use of words with [...] calculi with fixed" and "definite rules" (PI, § 81; see BB, 24): a dogmatic identification of use and calculus. Indeed, it is one thing to treat calculus as a good analogy for illuminating language, and quite another to claim that our language, despite its apparent imperfections, is a rigorous calculus.[12]

Should we, then, to prefer a more attenuated, less dogmatic version of the idea of calculus, treating it only as a term of comparison, as a model? But why keep it and not get rid of it altogether? As is almost always the case with Wittgenstein, the problem is not the word 'calculus' as if it were in itself misleading. Indeed, having freed the calculus from those 'logicising' implications we have seen, we can preserve it and transform it methodologically into a term of comparison for clarification purposes. Wittgenstein's answer to our question is then clear: the comparison between (linguistic) use and calculus should be preserved because it is helpful, and it helps us precisely insofar as it sheds light on (clarifies), by means of analogies and differences, the confusions that can arise when 'use' is dogmatically identified with 'calculus'. After all, as epistemologists have always emphasised, this is the function that a good model must fulfil: to highlight analogies and differences.

Even if we give this methodological value to the notion of calculus, what we might call 'the problem of the rule and of following (applying) a rule', which the identification of language with a calculus had helped to bring to the fore, does not disappear. This problem occupies a substantial part of the *Philosophical Investigations*, but also of the *Remarks on the Foundations of Mathematics*. More precisely, Wittgenstein is led to ask himself two questions, the first of which can be formulated as follows: apart from the fact that use can be fully identified with a calculus, in what sense can we say that use (like calculus) is limited by rules? The second question is: how do we know how to follow or apply a rule? What is it that allows us to say that someone who answers "It's 14" to the question "What is 8 plus 6?", or who brings a chair after being ordered "Bring me a chair!", is correctly applying the rule of addition, or the rule for using the word 'chair', unlike someone who answers "It's 19", or brings a hammer?

Through a series of examples and comparisons, Wittgenstein repeatedly invites us to see that the use of a word (an expression, a

---

12   In this claim, Wittgenstein recognises a form of that dogmatism, against which his whole philosophy seeks to fight, which, as he very effectively explains, consists in predicating "of the thing what lies in the mode of representation", i.e. in taking "the possibility of comparison, which impresses us, as the perception of a highly general state of affairs" (PI, § 104). On Wittgenstein's dogmatism, see Kuusela 2008.

sentence) is "not everywhere bounded by rules" (PI, § 68).[13] There is, however, a persistent tendency in philosophy to affirm that a use which is not entirely bounded by rules, or even without rules, is (and remains) an 'inexact' use, since it is open to hesitation and doubt, and that what one should aspire to is a use "that is everywhere bounded by rules", i.e. "whose rules never let a doubt creep in, but stop up all the gaps where it might" (PI, § 84). Against this aspiration (shared by both Descartes and Frege), Wittgenstein suggests that we compare a rule to a signpost. Indeed, "[a] rule stands there like a signpost", and a signpost "sometimes leaves room for doubt, and sometimes not" (PI, § 85). Sometimes, and usually, we follow it without even thinking about it; sometimes, because of the way it is placed, or for other reasons, we may hesitate and wonder exactly which way it is pointing. Why should the fact that there are times when we doubt lead us to conclude that we should always doubt? Or that we should never follow the signpost without first stopping and thinking? As this simile of the signpost shows, hesitation, doubt and uncertainty are *part* of the rule, not its negation or dissolution. Certainly, 'doubtful' or 'inexact' "does not mean 'unusable'" (PI, § 88).[14]

Acknowledging all this, however, does not settle the question of what it might mean to 'follow (or apply) the rule' (whatever it is or however it works). The problem that Wittgenstein faces in some of the most famous passages of the *Philosophical Investigations* is basically this: there are rules, all right, and these rules may be more or less 'exact', but what does it mean exactly to follow (or apply) a rule? Consider, for example, the rule "Add 2". Wittgenstein asks: how can this "rule teach me what I have to do at *this* point" (PI, § 198)? How can it teach me that, having arrived at 1004, what I have to say is exactly "1006", and that any other number would be wrong?

The question posed in section 198 could be answered – in a Platonist tone – by remarking that the rule teaches me what I am to do because it already contains its applications within itself: the rule, as it were, "traces the lines along which it is to be followed through

---

[13] For instance, the same applies to both rules of use and rules of play. Indeed, as Wittgenstein points out using one of his favourite examples, there are certainly rules, even codified rules, that distinguish the game of tennis from other kinds of games: "tennis is a game [...], and has rules too", even if there are "no [...] rules for how high one may throw the ball in tennis, or how hard" (PI, § 68).

[14] Likewise, 'undoubtful' or 'exact' does not mean 'usable'. An order such as "Stay roughly here!" is inexact when compared, say, to the order: "Stay right here!". Nevertheless – Wittgenstein ask rhetorically – "[i]f I tell someone 'Stay roughly here' – may this explanation not work perfectly? And may not any other one fail too?" (PI, § 88). On the contrary, just to follow the "ideal of exactness", should we think that the measurement of the width of the table we give to a joiner must be exact "to the nearest thousandth of a millimetre?" (PI, § 88) Wouldn't that get in the way of his work? Would he really understand what we are telling him and asking him to do?

the whole of space" (PI, § 219), like a groove or "a visible section of rails invisibly laid to infinity" (PI, § 218). According to this Platonist-sounding answer, every subsequent step is already ideally envisaged in the rule.[15] I cannot develop this point here, except to say that Wittgenstein's key statement (unlike that of the Platonist, who thinks that the applications are already in the rule, and also unlike that of the anti-Platonist, who, on the contrary, separates the rule from its applications) is that there is not the rule *and then* its applications, but that the rule *is* its applications. In short, following (or applying) a rule is one and the same thing as recognising it as a rule.[16] Still, I have dwelt on this for a moment because Platonism anticipates certain questions and problems – questions and problems very similar to those which, as we shall see, Wittgenstein will address when he introduces the notion of experience of meaning.

3.2 Sudden Understanding and Use in Time

Wittgenstein again clashes with the Platonist attitude to which, as we have seen, he is opposed in his various remarks on rule-following when he turns his attention to a phenomenon which seems to cast some shadow on the idea that it is in the use that meaning is to be sought: sudden understanding.

It is not uncommon for us to say, about a rule (an arithmetic rule or not) or the meaning of a word (an expression, a sentence), something like: "*Now* I have understood how I should proceed!"; "At *this* precise moment, the meaning has become clear to me", "*Suddenly* I have understood what it means", etc. Now, how does this sudden understanding (of the rule, the meaning of a word, etc.) fit in with the fact that the use of a word (or the application of a rule) unfolds over time, so to speak? When we suddenly understand or grasp the meaning of a word, what exactly is it that we have understood or

---

15   The Platonist answer is not the only one Wittgenstein considers. He dwells at length on the stance of those who assume that between the rule and its applications there must be 'something' that, from time to time, establishes that 'this', and not 'that', is the step to be taken: to the Platonist answer, one might counter, in an anti-Platonist spirit, by asking what guarantees I have that the actual step I take is precisely what the rule ideally envisages. In particular, Wittgenstein considers the answer of those who maintain that there must always be an interpretation between the rule and its applications, and for whom, therefore, applying a rule is always equivalent to interpreting it. As is well known, this interpretationism gives rise to the famous 'paradox' of the section 201 of the *Philosophical Investigations*, which shows that this reading of the relationship between a rule and its application leads to the dissolution of the rule itself (see PI, § 201).

16   This lies in the background of Wittgenstein's statement that "'following a rule' is a practice" (PI, § 202). The focus of Wittgenstein's investigation has completely shifted – to put it in a formula – from the rules of use to the use of rules.

grasped? These questions seem complicated to answer. But one thing seems certain, namely that "[w]hat we grasp in this way is surely something different from the 'use' which is extended in time" (PI, § 138). So here 'meaning' seems to be something different from 'use': while the former can be grasped 'suddenly', the latter cannot, precisely because it extends 'in time'. And then a problem arises: are we forced to conclude that meaning is not to be found in use, since it can be grasped 'suddenly', i.e. before any use? How can we escape from this trap?

Wittgenstein's way out is, once again, to invite us to change our perspective or point of view and to look more closely at the various circumstances in which we happen to say things like "Now (suddenly) I understand what this is (or what this word means)!". Let's imagine, for example, a person who is at first puzzled by certain explanations, and at some point exclaims: "Now (suddenly) I understand what an Allen key is (or what 'Allen key' means)! An Allen key is...". In the face of this exclamation, it is of little use to ask where (in the mind?) and what this 'thing' is that the person has suddenly understood, and which has led them to exclaim "Now I understand...". "Now I understand..." can mean many different things, depending on the circumstances in which it is said (see PI, § 154). It can mean that one no longer needs explanations, something like: "Now I can do it myself!", or "That wasn't so complicated!", or "Try me!". The individual in question seems to have understood what 'Allen key' means if they know how to use the word in the appropriate way and circumstances. For example, we can say that someone has understood and knows 'Allen key' if, when they need to loosen or tighten the screws on the handlebars of their bicycle, they ask a neighbour or friend: "Do you happen to have an Allen key I can borrow?". Against the idea that when I grasp the meaning there is 'something' that I grasp, Wittgenstein observes that "[i]f something has to stand 'behind the utterance of the [rule]', it is *particular circumstances*", that is to say, those circumstances that "warrant my saying that I can go on" (PI, § 154), and that now I know e.g. how to use 'Allen key'.[17] This is why Wittgenstein emphasises that the grammar of the words 'understand' is "closely related" to the grammar of the words 'know', 'can' and 'is able to', and that the family to which they all belong is that of "[t]o have 'mastered' a technique" (PI, § 150).

---

17   As Wittgenstein also says, "[t]he use of the word in practice is its meaning" (BB, 39).

3.3 Meaning and the Experience of Meaning

As we have seen, the phenomenon of sudden understanding can lead us to think that meaning is 'something' that we grasp and understand. This temptation is even stronger when our attention is drawn to another phenomenon which Wittgenstein, with explicit reference to William James,[18] calls the 'experience of meaning' (*Bedeutungserlebnis*) or the 'feeling of meaning' (*Bedeutungsgefühl*) and the related (by negation) phenomenon which he calls 'blindness to meaning' (*Bedeutungsblindheit*).[19] What prompts Wittgenstein to carefully consider these other two phenomena are, once again, the perplexities, uncertainties and doubts that his maxim "The meaning is (in) the use" might raise. What Wittgenstein is wondering here, as in the aforementioned case, but in a stronger and more obvious way, is whether the emphasis on use (as well as on understanding as the mastery of a technique) is somehow limiting or reductive, and, in particular, whether it risks disregarding the fact that not everything in meaning is use, or of neglecting, by relegating it to the background, certain experiences which seem to be fundamental to every speaker and to the actual use of language. Such experiences include, for example, that feeling of 'familiarity' which sometimes seems to accompany the words we use and that feeling of 'fusion' between a word and what it signifies, which gives the impression that there is a kind of close, and not at all arbitrary or conventional, intimacy or consonance between, say, our name and ourselves.[20]

The experiences of meaning illustrated and discussed by Wittgenstein also include: (a) those connected with proper names, such as the

---

[18] For example, James 1950, 472. But Wittgenstein also has in mind a passage by George Moore which appears in chapter three "Propositions" of his *Some Main Problems of Philosophy*: "It is quite plain, I think, that when we understand the meaning of a sentence, something else does happen in our minds *besides* the mere hearing of the words of which the sentence is composed. You can easily satisfy yourselves of this by contrasting what happens when you hear a sentence, which you *do* understand, from what happens when you hear a sentence which you do *not* understand: for instance, when you hear words spoken in a foreign language, which you do not understand at all. Certainly in the first case, there occurs, beside the mere hearing of the words, another act of consciousness—an apprehension of their meaning, which is absent in the second case. And it is no less plain that the apprehension of the meaning of one sentence with one meaning, differs in some respect from the apprehension of another sentence with a different meaning" (Moore 1953, 58-9).

[19] On the experience of meaning and its related phenomena, see also Goldstein 2004.

[20] Wittgenstein describes this impression of us by saying that the words of our language are like faces, whose particular and peculiar expressions are familiar to us: "Meaning – a physiognomy (PI, § 569). He also writes: "The familiar face of a word, the feeling that it has assimilated its meaning into itself, that it is a likeness of its meaning – there could be human beings to whom all this was alien. (They would not have an attachment to their words). – And how are these feelings manifested among us? – By the way we choose and value words" (PPF, XI § 294).

one described in the remark "Goethe's signature intimates something Goethean to me" (RPP I, § 336), when we feel that a name perfectly suits its bearer (his personality, physicality, etc.) "as if the name were an adjective" (LW I, § 69); (b) so-called 'synaesthetic experiences', such as those with coloured vowels (see LW I, § 59),[21] which he links and almost assimilates with the experiences of meaning; (c) experiences that enable us to understand orders or requests such as the following: "Pronounce the word 'till' and understand it as a verb, and not as a conjunction" (see PPF, XI § 261),[22] "Repeat the word 'March' to yourself and understand it now as an imperative now as the name of a month (see PPF, XI § 271)[23] or, again, "Read the word 'rank' as a verb and not as an adjective" (see LPP, 342).

The fact that Wittgenstein considers all these different experiences with words shows that, while insisting on use and gradually clarifying the sense and scope of his insistent appeal, he also asks himself, perhaps with no less insistence, whether it is indifferent to use one word instead of another, that is, whether meaning has a dimension which cannot be limited or reduced to use, because it is, so to speak, prior to use and independent from it. As Wittgenstein acknowledges, there often seems to be 'something' in our words, a sort of character or soul which we feel and experience, and which makes us inclined to use a specific word because it seems to us to be the most, or even the only, suitable and convenient word for our purposes and intentions. Wittgenstein, who devotes a great deal of space to the discussion of the phenomenon of the experience of meaning, attempts to explain precisely this component which he calls, with two pithy metaphors, the 'aroma' or 'atmosphere' of words (see, for example, PI, §§ 594, 596, 610; PPF, VI §§ 35, 42, 50; RPP I, § 243), and at the same time wonders whether it is not lost if we look too emphatically and exclusively at use. To insist on this, considering experiences such as those mentioned raises the question of whether an overemphatic and exclusive focus on use ultimately loses the experiential, even aesthetic, dimension or component of meaning, or, to put it differently, whether taking these experiences seriously also means admitting the limits and shortcomings of the appeal to use, however methodologically circumscribed. What Wittgenstein needs to do is to

---

[21] On this and other cases of synaesthesia in Wittgenstein, see ter Hark 2009. This is an interest and concern that Wittgenstein shares with the phenomenological tradition, among other, and that refers back to many aesthetic-artistic experiences of great significance. Think, for example, of nineteenth-century Symbolist poetry, in particular Rimbaud and Mallarmé.

[22] The German word used by Wittgenstein in his example in PPF, XI § 261 is 'sondern', which means 'to separate' as a verb and 'but' as a conjunction.

[23] Wittgenstein's example in German is with the word 'weiche' which means 'soft' as an adjective, 'side' as a noun and 'move away' as a verb.

come to terms with an objection which the maxim "The meaning is (in) the use" can easily raise, and to which he is by no means insensitive: by strictly adhering to this maxim, does one not end up treating every word as indifferent and interchangeable with every other, and thus failing to see or disregarding (what seems to be) the undeniable aesthetic-experiential dimension or component of meaning?

We might think that Wittgenstein's tendency is to also apply to the case of the experience of meaning the same critical strategy that he uses on several occasions with regard to those (mental) images and feelings that may accompany words and their use. As we know, although Wittgenstein would never dream of denying that they exist and that they often accompany the use of words, he does not hesitate to declare that these images and feelings have nothing to do with meaning, i.e. with our use of words and our knowledge of how to use them (in different circumstances) words. But is this really so? The answer is neither simple nor obvious, as evidenced by the complexity of this passage of Wittgenstein's, which shows more doubts than certainties:

> It is as if the word I understand has a specific slight aroma, which corresponds to its being understood. It is as if two words well known to me were distinguished not only by their sound, or their appearance, but, even if I do not associate any representation with them (*nichts bei ihnen vorstelle*), by their certain atmosphere. (RPP I, § 243)

As is quite clear, in the quoted passage, Wittgenstein distinguishes the case of *Vorstellungen* (mental images or representations), from that of the aroma or atmosphere of a word, suggesting that the former do not serve the same function as the latter, since (1) aroma or atmosphere does not depend on *Vorstellungen* and that (2) unlike the latter, aroma or atmosphere has to do with understanding. In short, as we might also say, it is one thing to recognise that words have an aroma or atmosphere and that to understanding them is also, so to speak, to feel or experience this aroma or atmosphere; it is quite another to identify, as mentalists of the most varied schools do, the meaning of a word with the mental images or representations which are associated with it or accompany it. In this respect, Wittgenstein is an anti-mentalist.

Having ascertained this, the question remains as to whether, and if so to what extent, the attention paid to the experience of meaning compels Wittgenstein (or us) to reconsider or reformulate what has always been regarded as the guiding maxim of his research, namely, the maxim with which I began, which states that "the meaning of a word is its use in the language" (PI, § 43). Here is the question in brief: what is the ultimate relationship between 'meaning' and 'use' if, as all the phenomena mentioned and similar ones suggest, the use

of a word is often, if not always, connected with our experience of its meaning? Wittgenstein explicitly asks this question in a remark on 'reading expressively'. After observing that "[w]hen I pronounce this word while reading expressively (*ausdrucksvollen*), it is completely filled with its meaning", he asks himself (or makes himself ask), I think not rhetorically: "'How can this be, if meaning is the use of the word?" (PPF, XI § 265). Indeed, if meaning is the use of the word, then 'this', i.e. a word completely filled with its meaning, appears to be nonsense. Suffice it to observe that if, in the expression "This word is completely filled with its meaning", we replace 'meaning' with 'use', we are faced with the nonsense of a word filled with its use; yet, the experience of meaning is given, and then, unless we argue that it is only an illusory appearance, we must conclude, as Wittgenstein himself seems to do, that meaning is not, or is not always, in the use of a word. Phenomena such as those on which Wittgenstein dwells seem to show that "more to meaning than the use of the word" (Zemach 1995, 490).[24] We can, however, interpret this conclusion in two different ways. (1) On the one hand, it can be argued that in realising that there is such a thing as the experience of meaning, Wittgenstein finally came to recognise the serious limitations of the maxim "The meaning is (in) the use". (2) On the other hand, it can be argued that this phase of his thought is only one part of the process that led him progressively to free himself from the image of language as calculus. This is the question that Wittgenstein then asks himself, and which effectively leaves him without an answer:

> How about this: you can set up *certain* rules, but only a few, which are of such kind that the person usually learns them through experience anyway – but what if, what is left, the most important part, is *imponderable*?? (LW I, § 921)[25]

---

[24]   A key passage in this regard is the following: "When I supposed the case of a 'meaning-blind' man, this was because the experience of meaning seems to have no importance in the *use* of language. And so because it looks as if the meaning-blind could not lose much. But it conflicts with this, that we sometimes say that some word in a communication meant one thing to us until we saw that it meant something else. First, however, we don't feel in this case that the experience of the meaning took place while we *were hearing the word*. Secondly, here one might speak of an experience rather of the sense of the sentence, than of the meaning of a word" (RPP I, § 202).

[25]   Here by 'imponderable' – elsewhere by 'imponderable evidence' (see PPF, XI §§ 358-60) – Wittgenstein seems to refer to all those circumstances in which the choice of one word over another makes a great difference, e.g. the difference between a good and a bad poem, even if the difference between the two words belongs to what Wittgenstein calls 'subtle' difference (see PPF, XI § 297). On the significance of Wittgenstein's appeal to imponderable evidence, see Putnam 1992, 39-46; Boncompagni 2018. It should also be emphasised here that in this context Wittgenstein recovers the value of 'experience' that resonates in the word '*Erfahrung*', for example, when we say of someone (think of the Homeric Ulysses) that they have seen many things and had many experiences, or

Besides, we should not forget, as we are often inclined to do, that the famous section 43 of the *Philosophical Investigations* excludes that 'meaning' and 'use' are always interchangeable; 'experience of meaning' seems to apply to those cases which do *not* belong to the "*large* class of cases*" (PI, § 43) referred to in the section. For, to repeat, while the expression 'experience of meaning' can give us a headache, 'experience of use' is utterly nonsensical.

But even if we leave aside how the specification of section 43 should be interpreted, the question remains: "'What would someone be missing if he did not *experience* the meaning of a word?'" (PPF, xi § 261).[26]

4 A Modest Proposal

Ideally, there are two almost antithetical ways of interpreting the relationship between (meaning as) use and the experience of meaning.

According to a first interpretation (see Bouveresse 2007),[27] which focuses above all on the fact that the experience of meaning is an

---

when we acknowledge that "we learn certain things only through long experience (*Erfahrung*)", not simply "from a course in school" (LW I, § 925). It's through experience that we "develop a feeling for the rules" (LA, 5), so to speak. It is in this way, for example, that we form those tastes and aesthetic sensibilities that characterise the person who is usually called a 'connoisseur'. To form "the eye of a connoisseur" requires "[a] great deal of *experience*": one does not learn to evaluate a painting "in the *same* way as one learns to calculate", but, say, by looking at and comparing, with the help of a teacher, "a large number of pictures by various masters again and again" (LW I, § 925). Therefore, although "[i]n most cases" the connoisseur is "able to list reasons for his judgement" (e.g. for the judgement: "This picture could not have been painted by this master"), we have to admit that "generally it wasn't *they* that were convincing" (LW I, § 925). Indeed, even in cases where this man is not able to give good reasons for his judgment, what makes us accept it as evidence is, precisely, that he is a connoisseur, i.e. that he has long and extensive experience of painting, and "this is more or less the *only* way of weighing such evidence" (Monk 2005, 104).

26   Alongside the descriptions that can be found in various of Wittgenstein's writings (largely in RPP I), an effective description of what this person would be missing is provided by fiction; consider this excerpt from Mark Haddon's famous novel *The Curious Incident of the Dog in Night-Time*, whose young protagonist describes his 'blindness to meaning' in this way: "This will not be a funny book. I cannot tell jokes because I do not understand them. Here is a joke, as an example. It is one of Father's. **His face was drawn but the curtains were real**. I know why this is meant to be funny. I asked. It is because *drawn* has three meanings, and they are **1)** drawn with a pencil, **2)** exhausted, and **3)** pulled across a window, and meaning **1** refers to both the face and the curtains, meaning **2** refers only to the face, and meaning **3** refers only to the curtains. If I try to say the joke to myself, making the word mean the three different things at the same time, it is like hearing three different pieces of music at the same time which is uncomfortable and confusing and not nice like white noise. It is like three people trying to talk to you at the same time about different things. And that is why there are no jokes in this book" (Haddon 2003, 10; bold in the original).

27   Michel ter Hark (2011) seems to be going in a similar direction. See e.g. what he writes in presenting the point of his reading: "In this chapter, I will show otherwise

experience, Wittgenstein introduces this notion, much as Frege introduced the notion of representation in *Sense and Reference*, in order to declare it irrelevant to the question of meaning. The experience of meaning would be something like the crown on the head of the chess king of which Wittgenstein speaks of in this passage from the *Blue Book*:

> — I want to play chess, and a man gives the white king a paper crown, leaving the use of the piece unaltered, but telling me that the crown has a meaning to him in the game, which he can't express by rules. I say: "as long as it doesn't alter the use of the piece, it hasn't what I call a meaning". (BB, 65)

Applying to this case an image that has been used in the case of Frege, we could say that the notion of experience of meaning ends up in Wittgenstein's 'wastebasket' (see Bar-Hillel 1971). In short, according to this first interpretation, the way in which Wittgenstein handles this notion would only be a confirmation of his radical anti-psychologism.

According to a second possible interpretation,[28] the introduction of the notion of experience of meaning corresponds to Wittgenstein's recognition that meaning cannot be entirely or totally in the use and that, indeed, the presence of something like an experience of meaning (and here the emphasis is on 'of meaning') would entail at least a partial return to the idea that meaning is something we can experience here and now. This kind of interpretation may perhaps explain why, as seen above, several scholars have expressed suspicion and distrust of any attempt to give weight to the topic of the experience of meaning. Wittgenstein's rejection of the idea that meaning is 'something' is so strong and repeated that any attempt to revalue it seems hardly in keeping with the spirit of his thinking, even in his later years.

With respect to these two almost mirror-image interpretations, I would like to conclude by suggesting a third interpretation, which obviously requires further development and investigation. My idea is

---

and establish, exegetically and argumentatively, that the discussion of the experience of meaning is not supplementary to the earlier account of meaning and understanding. It is not the case that Wittgenstein gradually came to see that the earlier account had left something out, i.e. the familiar feel of words to which Moore and James refer. Rather, the point of the discussion is to determine what it is that philosophers think that is left out in an account of language which emphasizes 'only' the use of signs. Put otherwise, what is under investigation here is the very concept of experience of which not only Moore and James but also Wittgenstein's commentators say that it has to be included in any account of language distinctive of human beings" (ter Hark 2011, 501). See also ter Hark 2013.

28 Probably no one has supported this interpretation in its most explicit and strongest form, although it clearly serves as a critical lens for the first interpretation.

that we can apply to our topic (the relationship between use and the experience of meaning) part of what Wittgenstein says when, in the *Philosophical Investigations*, he addresses the relationship between our concepts and some "very general facts of nature" (PPF, XII § 365). According to Wittgenstein, it must be acknowledged that there is a "correspondence" (PPF, XII § 365) between these facts and our concepts. To acknowledge this, however, is not to acknowledge that our concepts have their basis in these facts. Rather, taking up some suggestions from *On Certainty*, we might say that these facts are "incorporated into" (OC, § 61) our concepts and related language games. Something similar could also be said about the relationship between use and the experience of meaning. Far from being the basis of use, this experience is incorporated into use and modulates it in certain ways. Let us take two examples. (1) Proper nouns are such because they are used in certain ways and circumstances (to call someone, to roll call at school, to sign a document, etc.). However, this use also involves (incorporates) the experience of 'fusion' with one's own name and 'attachment' to it (see PPF, XI § 294), without which proper names would have a different place and role in our lives. Our name, for example, could be changed without any suffering on our part. (2) If the experience of meaning were not incorporated into the use of words, that spasmodic attention to the choice of each individual word which is characteristic of poetry would not be there, or would be very different. From a certain point of view, we could say that without the experience of meaning we would only have 'unpoetic' uses of words.

On closer inspection, many of Wittgenstein's remarks about blindness to meaning have an analogous (methodological) function to that of "imagin[ing] certain very general facts of nature to be different from what we are used to" (PPF, XII § 366). In short, Wittgenstein never says that without the experience of meaning there would be no meaning, but he does say that without the experience of meaning our uses would be different, and perhaps more like the calculus he had in mind in the intermediate phase of his thought. Some support for this reading of mine can be found in this passage from *On Certainty*, at least if we assume (as it is reasonable to do) that the facts of which Wittgenstein speaks here also include what we might call 'psychological facts', such as, for instance, the fact that we cherish our name and consider it part of who we are:

> If we imagine the facts otherwise than as they are, certain language-games lose some of their importance, while others become important. And in this way there is an alteration—A gradual one—in the use of the vocabulary of a language. (OC, § 63)

References

Bar-Hillel, Y. (1971). "Out of the Pragmatic Wastebasket". *Linguistic Inquiry,* 2(3), 401-7.

Boncompagni, A. (2018). "Immediacy and Experience in Wittgenstein's Notion of 'Imponderable Evidence'". *Pragmatism Today*, 2(9), 94-106.

Bouveresse, J. (2007). "Wittgenstein on 'Experiencing Meaning'". Moyal-Sharrock, D. (ed.). *Perspicuous Presentations: Essays on Wittgenstein's Philosophy of Psychology*. Houndsmills, Basingstoke; New York: Palgrave Macmillan, 75-94.

Crary, A., Read, R. (2000) (eds). *The New Wittgenstein*. London-New York: Routledge.

Hacker, P.M.S., Schulte, J. (2009). "Editorial Preface to the Fourth Edition". Wittgenstein 2009, i-xvii.

Goldstein, L. (2004). "What Does 'Experiencing Meaning' Mean?". Moyal-Sharrock, D. (ed.). *The Third Wittgenstein: The Post-Investigations Works*. Aldershot: Ashgate, 107-23.

Haddon, M. (2003). *The Curious Incident of the Dog in Night-Time*. London: Jonathan Cape.

Hark ter, M. (2009). "Coloured Vowels: Wittgenstein on Synaesthesia and Secondary Meaning". *Philosophia*, 37, 589-604.

Hark ter, M. (2011). "Wittgenstein on the Experience of Meaning and Secondary Use". Kuusela, O.; McGinn, M. (eds). *The Oxford Handbook of Wittgenstein*. Oxford: Oxford University Press, 479-521.

Hark ter, M. (2013). "Wittgenstein on the Experience of Meaning: Historical and Contemporary Perspectives". *Philosophy Study*, 3(10), 974-81.

Horwich, P. (2008). "Wittgenstein's Definition of 'Meaning' as 'Use'". *Annals of the Japan Association for Philosophy of Science*, 16(1-2), 133-41.

Ionesco, E. (1958). *Four Play*: *The Bald Soprano/The Lesson/Jack or the Submission/The Chair*. Translated by D.M. Allen. New York: Grove Press, Inc.

James, W. (1950). *The Principles of Psychology (1890)*. Vol. 1. New York: Dover Publications.

Kuusela, O. (2008). *The Struggle against Dogmatism: Wittgenstein and the Concept of Philosophy*. Cambridge MA: Harvard University Press.

Monk, R. (2005). *How to Read Wittgenstein*. London: Granta Books.

Moore, G.E. (1953). *Some Main Problems of Philosophy*. London: George Allen & Unwin Ltd; New York: The Macmillan Company.

Perissinotto, L. (2009). "Il respiro del segno. Significato e interpretazione in Wittgenstein". Chiurco, C.; Sciuto I. (eds), *Verità, fede, interpretazione*. Padova: Il Poligrafo, 303-14.

Putnam, H. (1995). *Pragmatism: An Open Question.* Oxford: Blackwell.

Snowdon, P.F. (2018). "Wittgenstein and Naturalism". Cahill, K.M.; Raleigh, T. (eds), *Wittgenstein and Naturalism*. New York-Abingdon; Oxon: Routledge, 15-32.

Wittgenstein, L. (1958). *The Blue and the Brown Books*. Edited by R. Rhees. Oxford: Blackwell. [BB]

Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus*. With an introduction by Bertrand Russell. Translated by D.F. Pears and B.F. McGuinness. London-New York: Routledge. [TLP]

Wittgenstein, L. (1966). "Lectures on Aesthetics", in *Lectures and Conversations on Aesthetics, Psychology and Religious Belief*. Edited by C. Barrett. Oxford: Basil Blackwell, 1-40. [LA]

Wittgenstein, L. (1969). *On Certainty*. Edited by G.E.M. Anscombe and G.H. von Wright; translated by D. Paul and G.E.M. Anscombe. Oxford: Blackwell. [OC]

Wittgenstein, L. (1979). *Wittgenstein and the Vienna Circle. Conversations recorded by Friedrich Waismann*. Edited by B. McGuinness; translated by J. Schulte and B. McGuinness. Oxford: Basil Blackwell. [WVC]

Wittgenstein, L. (1980). *Remarks on the Philosophy of Psychology*. Vol. 1. Edited by G.E.M. Anscombe and G.H. von Wright; translated by G.E.M. Anscombe. Oxford: Basil Blackwell. [RPP I]

Wittgenstein, L. (1982). *Last Writings on the Philosophy of Psychology*: *Preliminary Studies for Part II of* Philosophical Investigations. Vol. 1. Edited by G.H. von Wright and H. Nyman; translated by C.G. Luckhardt and M.A.E. Aue. Oxford: Blackwell. [LW I]

Wittgenstein, L. (1988). *Lectures on Philosophical Psychology 1946-1947. Notes by P.T. Geach, K.J. Shah, A.C. Jackson*. Edited by P.T. Geach. Hemel Hempstead: Harvester-Wheatsheaf. [LPP]

Wittgenstein, L. (1992). *Last Writings on the Philosophy of Psychology*: *The Inner and the Outer 1949-1951.* Vol. II. Edited by G.H. von Wright and H. Nyman; translated by C.G. Luckhardt and M.A.E. Aue. Oxford: Blackwell. [LW II]

Wittgenstein, L. (2000). *Wittgenstein's Nachlass. The Bergen Electronic Edition*. [MS]

Wittgenstein, L. (2009). *Philosophical Investigations*. Revised 4th edition by P. M. S. Hacker and J. Schulte; translated by G.E.M. Anscombe, P.M.S. Hacker and J. Schulte. Chichester: Wiley-Blackwell. [PI = *Philosophical Investigations*, 1-181; PPF = *Philosophy of Psychology – A Fragment*, 182-244]

Zemach, E.M. (1995). "Meaning, the Experience of Meaning and the Meaning-Blind in Wittgenstein's Late Philosophy". *The Monist*, 78(4), 480-95.

"Following According to a Rule Is FUNDAMENTAL to Our Language-Game"
Rules and Meaning in Wittgenstein

William Child
University of Oxford, United Kingdom

Abstract It is commonly thought that, according to the later Wittgenstein, the meanings of words are determined by rules and using language involves following those rules. Against that standard interpretation, Kathrin Glüer and Åsa Wikforss have argued that, though Wittgenstein did hold these views in his middle period, he decisively rejected them in his mature work. And, they think, he was right to do so. The paper defends the standard interpretation on textual and philosophical grounds: both as an account of Wittgenstein's later views and as a philosophical position in its own right.

Keywords Rule-Following. Wittgenstein. Meaning. Philosophical Investigations. On Certainty.

1 Introduction

What is the connection between linguistic meaning and rules? In *Philosophical Investigations*, and in Wittgenstein's later philosophy more generally, the discussion of meaning is intertwined with the discussion of rules and rule-following. And despite the continuing controversy about how exactly to understand his views about rules and rule-following, there is widespread agreement about how he sees the relation between meaning and rules. According to that general consensus, what a word means is a matter of the rules for its use. To grasp the meaning of a word is to grasp the rules for its use. And using the word with that meaning is a matter of following those rules. As Wittgenstein's student and literary executor, Rush Rhees, puts it, when "I have learned what [an expression] means [...] I have learned a rule" (Rhees 1954, 77); and "using [expressions] in their meanings is what we call following a rule" (88).

Kathrin Glüer and Åsa Wikforss have challenged the "received view [...] that the later Wittgenstein subscribes to [...] the thesis [that] speaking a language is a rule-guided activity" (Glüer, Wikforss 2010, 148).[1] They argue that the later Wittgenstein does not think that meaning is determined by rules. Instead, he thinks that the meanings of words are determined by use: by the practice of applying them. Though he rejects the received view, they argue, he does hold that there is a fruitful *analogy* between meaning and rules. For instance, following a rule is a custom, a usage, an institution; so is using language to make a report, to give an order, and so on (PI, § 199). An action is correct or incorrect in the light of rule; similarly, an application of a word is correct or incorrect given its meaning. And so on. The reason why the discussion of rules and rule-following in Wittgenstein's later philosophy is so closely related to his discussion of meaning is that "he is exploring the analogy between meaning and rules" (Glüer, Wikforss 2010, 150). But, Glüer and Wikforss insist, it is only an analogy. In *Philosophical Investigations*, Wittgenstein does not conceive of speaking a language as a matter of following meaning-determining rules.

Though they reject the 'received view' of Wittgenstein's position in *Philosophical Investigations*, Glüer and Wikforss acknowledge – indeed, insist – that Wittgenstein's middle-period writings of the early 1930s do conceive of meaning as constituted by rules and of language as a rule-guided activity. Thus, for instance, he wrote in 1931 that an ostensive definition of a colour word is a rule:

---

[1] I shall use the expression "received view" sometimes to refer to a view about Wittgenstein's philosophy (the view that Wittgenstein thinks of speaking a language as a rule-governed activity) and sometimes to refer to a philosophical view (the view that speaking a language is a rule-governed activity). The context should always make clear which is meant.

the ostensive explanation "*That* is 'red'" [...] is one of the symbolic rules for the use of the word 'red'. (Ms 110, 213[7], 24 June 1931. See also Ts-213,176r[5])[2]

And he held that the meaning of a word is given by the rules for its use:

There can be no debate about whether these or other rules are the right ones for the word 'not' (I mean, whether they accord with its meaning). For without these rules, the word has as yet no meaning; and if we change the rules, it now has another meaning (or none), and in that case we may just as well change the word too. (Ms 110, 133[3], 3 March 1931)[3]

Now if we are to understand language as a rule-guided activity, we need some account of what it is for a speaker to follow or be guided by linguistic rules, as opposed to merely acting in accordance with them. And, according to Glüer and Wikforss, having struggled to develop an account of linguistic rule-following in his middle-period writings, Wittgenstein came to see that no satisfactory account could be given. In his later writings, therefore, he abandoned the idea that understanding language is a matter of following rules. On their interpretation, *Philosophical Investigations*, §§ 198-202, which is often seen as a statement of the received view, actually argues against the association of meaning with rules. And, they suggest, Wittgenstein's rejection of the received view emerges particularly clearly in his final notebooks, published as *On Certainty*, which "leaves no room for doubt" that he thinks only that there is an *analogy* between meaning and rules (Glüer, Wikforss 2010, 150).

The target of Glüer and Wikforss's critique is, as we have seen, the "received view" that "speaking a language is a rule-guided activity". That formulation of the view combines two elements: there is the idea that the meaning of a word is *constituted* or *determined* by rules for using it; and there is the idea that using a word involves *following* or *being guided by* those rules. Glüer and Wikforss's discussion focuses mainly on the second element. In a fuller treatment of the topic, it would be worth reflecting on the relation between the two elements. For instance, would it be coherent to hold that the meanings of words are constituted by rules for using them but that someone can

---

2 References in this form are to items from Wittgenstein's Nachlass, using the versions available at Wittgenstein Source http://www.wittgensteinsource.org.

3 The translation is taken from PI, § 549. The *Big Typescript* version of the remark continues: "Thus these rules are arbitrary, because it is the rules that first give meaning to the sign" (BT, 234-5).

grasp the meanings of those words, and use them with those meanings, without following or being guided by those rules? For present purposes, however, I leave those questions aside.

I shall defend the "received interpretation" of Wittgenstein's later views on rules and meaning against Glüer and Wikforss's interpretative case against it. And I shall defend the claim that speaking a language involves following rules against their substantive philosophical attack.

2 Some Textual Evidence for the "Received View"

Before considering Glüer and Wikforss's case for rejecting the "received view" of Wittgenstein on rules and meaning, I will point to some passages that strongly support the received interpretation. Of course there may be evidence on both sides. But, at a minimum, a defence of Glüer and Wikforss's interpretation needs to explain how it is consistent with the passages I shall cite.

2.1 Following a Rule "Characterizes Description"

In *Remarks on Foundations of Mathematics*, Wittgenstein writes this:

> Following according to the rule is FUNDAMENTAL to our language-game. It characterizes what we call description. (RFM, VI, § 28 (Ms-164,81[2], 1941.01.01?-1944.12.31?))

On the face of it, that passage says that when we use words to describe something we are following a rule for the use of those words. Someone might point out that Wittgenstein thinks that not all language-use involves *describing*; so even if we agree that following a rule characterises *description*, it doesn't follow that *every* use of language involves following rules. Maybe so. But many uses of language *do* involve describing. And if we accept that describing involves following rules for the words we employ in our description, there seems just as much reason to accept that giving an order, say, or asking a question, involves following rules.

It is worth quoting the context in which Wittgenstein makes this remark. He writes:

> Someone asks me: What is the colour of this flower? I answer: "red".—Are you absolutely sure? Yes, absolutely sure! But may I not have been deceived and called the wrong colour "red"? No. The certainty with which I call the colour "red" is the rigidity of

my measuring-rod, it is the rigidity from which I start. When I give descriptions, *that* is not to be brought into doubt. This simply characterizes what we call describing.

(I may of course even here assume a slip of the tongue, but nothing else.) (RFM, VI, § 28)

Then comes the claim that following according to a rule is FUNDA-MENTAL to our language-game. Applying the claim to the example that precedes it, Wittgenstein seems absolutely clear that applying the word 'red' to a flower involves following a rule.

2.2 Using a Word as "Following Certain Rules"

In *Lectures on the Foundations of Mathematics* Wittgenstein says this:

> If you have learned a technique of language, and I point to this coat and say to you, "The tailors now call this colour 'Boo'" then you will buy me a coat of this colour, fetch one, etc. The point is that one only has to point to something and say, "This is so-and-so", and everyone who has been through a certain preliminary training will react in the same way. We could imagine this not to happen. If I just say, "This is called 'Boo'" you might not know what I mean; but in fact you would all of you automatically follow certain rules.

> Ought we to say that you would follow the *right* rules?—that you would know *the* meaning of "boo"? No, clearly not. For which meaning? Are there not 10,000 meanings which "boo" might now have? [...] To know its meaning is to use it *in the same way* as other people do. "In the right way" means nothing. (LFM, 182-3)

In that passage, Wittgenstein takes it for granted that using the word 'Boo' with a given meaning involves following certain rules for the use of 'Boo'. His view is not that there is an *analogy* between using a word and following rules. He is saying that using a word *is* following rules.

The passage comes from lectures Wittgenstein gave in 1939. So someone might suggest that the views expressed belong to his middle period, when he did think of language as a rule-guided activity, and do not threaten Glüer and Wikforss's account of his position in *Philosophical Investigations* and beyond. But they themselves suggest that Wittgenstein's mature views about rules and meaning are already starting to be visible in the *Brown Book*, which was dictated

in 1934-35.[4] So it would be surprising for them to argue that the views expressed in the 1939 lectures belong with his middle period rather than his later view of rules and meaning.

2.3 Rules and the Meanings of Logical Constants

In the Introduction, I quoted a passage from Ms110, which was composed in March 1931:

> There can be no debate about whether these or other rules are the right ones for the word 'not' (I mean, whether they accord with its meaning). For without these rules, the word has as yet no meaning; and if we change the rules, it now has another meaning (or none), and in that case we may just as well change the word too.

That remark, which dates from Wittgenstein's middle period, is an explicit statement of the view that the meaning of a word is a matter of the rules for its use. But the passage does not just appear in Ms110; it also occurs in *Philosophical Investigations*, as paragraph (b) in the boxed comment following § 549. Of course its appearance in that context does not show that Wittgenstein still endorsed this view at the point when he attached this comment to the typescript of *Philosophical Investigations*. Hacker and Schulte say that the boxed comments in *Philosophical Investigations* were "probably meant to be taken into account in future revisions of the text" (PI, xxi). But who can say what such a revision would have involved? Maybe Wittgenstein would have used this remark as an example of a view that is tempting but should ultimately be rejected.

However, there is good reason to think that Wittgenstein did not come to reject that view, and that when he attached this remark to the typescript of *Philosophical Investigations* he still held the view it expresses. For a passage from RFM, composed in March 1944, offers essentially the same account of the connection between meaning and rules as the Ms110 remark from 1931:

> Is logical inference correct when it has been made according to rules; or when it is made according to *correct* rules? Would it be wrong, for example, if it were said that p should always be inferred from $\neg p$? But why should one not rather say: such a rule would not give the signs '$\neg p$' and 'p' their usual meaning?

---

4 "In the *Brown Book*", they write, "Wittgenstein suggests that rules cannot play the fundamental role in our linguistic practices that they had earlier been ascribed" (Glüer, Wikforss 2010, 155).

We can conceive the rules of inference—I want to say—as giving the signs their meaning, because they are rules for the use of these signs. So that the rules of inference are involved in the determination of the meaning of the signs. In this sense rules of inference cannot be right or wrong. (RFM, VII, § 30, Ms 124,113[2], 9th March 1944)

All the indications are that Wittgenstein wrote that remark in 1944 as an expression of what he thought at the time. There is no reason to treat it simply as a record of a view he had held more than ten years earlier and had now given up.

3 Glüer and Wikforss's Textual Evidence Against the "Received View"

3.1 *Philosophical Investigations*

Glüer and Wikforss argue that the text of *Philosophical Investigations* – and specifically the key discussion of rule-following leading up to §§ 201-2 – supports their contention that Wittgenstein came to reject the received view. They write:

> For a rule to guide a speaker, Wittgenstein holds, an expression of the rule has to be involved in the speaker's use of terms. However, any expression can be variously interpreted; consequently, the idea that meaning is determined by rules leads to a regress of interpretations: "'But how can the rule show me what I have to do at *this* point? Whatever I do is, on some interpretation, in accord with the rule'. – That is not what we ought to say, but rather: any interpretation still hangs in the air along with what it interprets, and cannot give [it] any support. Interpretations by themselves do not determine meaning" (PI, § 198). Thus, Wittgenstein is here rejecting his own earlier idea that meaning is determined by rules that guide our use – instead, he suggests, meaning is determined by this use itself, by the *practice* of applying the sign. (Glüer, Wikforss 2010, 155)

But Glüer and Wikforss's reading seems to me to mistake the significance of § 198.

In the first place, the topic of § 198 is not specifically how a *linguistic* rule can show me what I have to do at a particular point; the discussion concerns rules in general. Indeed, the only example of a rule that Wittgenstein gives in this section involves a signpost. A signpost, he says, is an expression of a rule: as we might say, an expression of

the rule *go this way*. His question is, how *can* the signpost show me that I have to go *this* way? And the lesson of Wittgenstein's discussion is a general one: the rule does determine what I have to do at this point; but its determining what I have to do does not depend on its being supplemented by an interpretation. That is the message of § 198 and of the closely-related § 201: "[T]here is a way of grasping a rule which is *not* an interpretation".

Glüer and Wikforss argue that § 198 rules out the "idea that meaning is determined by rules". If the idea that meaning is determined by rules depended on the idea that meaning is determined by interpretations, it would indeed be a non-starter. That is the point of § 198: if a rule cannot determine anything unless it is supplemented by an interpretation then, by the same token, an interpretation cannot determine anything unless it is supplemented by another interpretation, and so on; if we go down that path, the whole idea of anything being determined by a rule collapses. But there is no reason to think that the idea that meaning is determined by rules *does* depend on the idea that meaning is determined by interpretations. And as far as I can see, § 198 says nothing at all against the idea that the meaning of a word is a matter of rules for its use.

Glüer and Wikforss make a further interpretative point against the received view. They remind us that, for the later Wittgenstein, meaning is determined by use. But that view, they suggest, is inconsistent with the idea that meaning is rule-determined; the earlier idea that meaning is determined by rules, they think, is *replaced* in Wittgenstein's later work by the idea that meaning is determined by use.[5]

Contrary to what Glüer and Wikforss say, however, there is no tension between the idea that the meaning of a word depends on the rules for its use and the idea that the meaning of a word is determined by use. Consider the analogy between language and games. Chess is the game it is in virtue of having the rules it does. But chess, with the rules that define it, did not appear in the world by magic. We might have used the same pieces to play a different game, or none at all. The game of Chess exists, and has the rules it does, because we play it according to those rules: because we 'use' the pieces in the way we do. Similarly for linguistic meaning. Wittgenstein says that the meaning of the word 'not', say, is determined by the rules for its use. But what determines that those *are* the rules for the use of that word is the way that we use it: specifically, our using the word 'not' according to those rules. Had we used the word 'not' in a different way, observing different rules, it would have had a different meaning. In short, the idea that the meaning of a word is determined by our use of the word is not in competition with the idea that meaning

---

5   For this argument, see Glüer, Wikforss (2010, 156).

is determined by rules. On the contrary, it is an essential accompaniment to it.

3.2   *On Certainty*

According to Glüer and Wikforss, it is in Wittgenstein's latest writings – the notebooks published as *On Certainty* – that we see the clearest and most explicit rejection of the idea that using words is a matter of following rules. They highlight two passages in particular. But neither passage, I shall argue, gives compelling support to their reading of Wittgenstein.

The first passage is *On Certainty*, § 46. In German:

> Das Wichtigste aber ist: Es braucht die Regel nicht. Es geht uns nicht ab.

And in the published translation:

> But the most important thing is: The rule is not needed. Nothing is lacking.

Taken in isolation, that remark might be thought to imply that language has no need for rules and that speaking a language is not a matter of following rules. But when we look at the context in which it occurs, we can see that that is not what Wittgenstein is saying at all.

A preliminary point is this. The passage that Glüer and Wikforss quote from OC, § 46 continues like this:

> We do calculate according to a rule, and that is enough.

So Wittgenstein is talking not about language-use in general but about a case of calculating according to a rule. He says, of that case, that our calculating according to a rule "is enough". Whatever he means when he says that "the rule is not needed", then, he is not denying that calculation is a rule-governed activity or that, when we calculate, we are acting according to a rule.

What is the point of the passage? OC, § 46 is part of Wittgenstein's discussion of a question that is raised some twenty remarks earlier:

> One may be wrong even about "there being a hand here". Only in particular circumstances is it impossible.—"Even in a calculation one can be wrong—only in certain circumstances one can't."

> But can it be seen from a *rule* what circumstances logically exclude a mistake in the employment of rules of calculation?

> What use is a rule to us here? Mightn't we (in turn) go wrong in applying it?
>
> If, however, one wanted to give something like a rule here, then it would contain the expression "in normal circumstances". And we recognize normal circumstances but cannot precisely describe them. At most, we can describe a range of abnormal ones. (OC, §§ 25-7)

Wittgenstein is interested in the kind of certainty that attaches to such Moorean propositions as "there is a hand here" or "I have two hands". In some circumstances, he thinks, I could be wrong in thinking that I have two hands; consider how things might be after an accident or a medical procedure, say. But in normal circumstances, according to Wittgenstein, the proposition "I have two hands" is a basic certainty: I cannot give grounds for believing it; I couldn't be making a mistake about it; and so on. Similarly for mathematical calculations. In some circumstances, he thinks, it makes good sense to suppose that we have made a mistake when we perform some calculation: when I calculate the product of two ten-digit numbers, for instance, it is easy to see that my answer could be mistaken. In other circumstances, however, one cannot be wrong in a calculation: he insists, for instance, that we couldn't all be making a mistake in thinking that $12 \times 12 = 144$; in such a case, a mistake is "logically excluded". Now the question Wittgenstein presses in the quoted passage from OC, §§ 25-7 is this: what distinguishes the case where a mistake in applying the rules of calculation is logically excluded from the case where such a mistake is perfectly possible? He suggests that there is no general rule for distinguishing between the two kinds of case; we can recognise the difference, case by case, but we cannot give a precise rule for doing so.

Wittgenstein returns to this question in the sections leading up to OC, § 46:

> What sort of proposition is this: "We *cannot* have miscalculated in $12 \times 12 = 144$"? It must surely be a proposition of logic.—But now, is it not the same, or doesn't it come to the same, as the statement $12 \times 12 = 144$?
>
> If you demand a rule from which it follows that there can't have been a miscalculation here, the answer is that we did not learn this through a rule, but by learning to calculate.
>
> We got to know the *nature* of calculating by learning to calculate.
>
> But then can't it be described how we satisfy ourselves of the reliability of a calculation? O yes! Yet no rule emerges when we do

so.—But the most important thing is: The rule is not needed. Nothing is lacking. We do calculate according to a rule, and that is enough. (OC, §§ 43-6)[6]

The message of that discussion is this. When we calculate according to a rule, we do not need *another* rule to tell us whether we could be making a mistake in our application of the first rule. We learn the difference between cases where miscalculation is possible and cases where it is not by learning to calculate, not by learning a rule for distinguishing the two cases. But none of that takes away from the fact that learning to calculate is learning to follow rules: "We do calculate according to a rule".

Understood in the context in which it appears, then, Wittgenstein's remark *"Es braucht die Regel nicht"* does nothing to challenge the idea that grasping the meaning of a term involves grasping rules, or that applying the term is a matter of following rules.

The second passage that Glüer and Wikforss quote from *On Certainty* is OC, §§ 61-2, which, they say, "leaves no room for doubt" that Wittgenstein's view is simply that there is an *analogy* between meaning and rules:

A meaning of a word is a kind of employment of it.
For it is what we learn when the word is incorporated into our language.

That is why there is an analogy between the concepts "meaning" and "rule". (OC, §§ 61-2)

The final sentence of that passage is given in Glüer and Wikforss's own translation. The printed translation is different:

That is why there exists a correspondence between the concepts 'rule' and 'meaning'.

And Wittgenstein's German is this:

Darum besteht eine Entsprechung zwischen den Begriffen 'Bedeutung' und 'Regel'.

Glüer and Wikforss are plainly right to correct the published translation by putting the words "meaning" and "rule" in the same order as

---

6 I have quoted the published translation. But the sense of the last paragraph would in my view be better captured by translating *"Es braucht die Regel nicht"* as *"A rule is not needed"*, rather than *"The rule is not needed"*.

their German equivalents. But are they right to translate "*eine Ent-sprechung*" as "an analogy" rather than "a correspondence"? The suggestion that the concepts "meaning" and rule are *analogous* seems to rule out the idea that what you grasp when you grasp the meaning of a word *is* a rule or a set of rules. By contrast, that idea is not ruled out by the suggestion that there is a *correspondence* between the concepts "meaning" and "rule". I am in no position to pass judgement on this question of translation. That said, it does seem plausible that the English "analogy" is a narrower or more specific notion than the German "*Entsprechung*"; after all, German has the word "*Analogie*" to express the narrower notion. At the very least, it is not clear that the passage that Glüer and Wikforss quote from OC, § 62 bears the weight that they put on it, as establishing that Wittgenstein came to think that using a word with a given meaning is definitely not a matter of following rules for its use.[7]

4   Following Rules and Conforming to Rules

So far, I have focused on the textual grounds for accepting or rejecting Glüer and Wikforss's contention that, in *Philosophical Investigations* and his later work, Wittgenstein rejected the "received view" that speaking a language is a rule-guided activity. But Glüer and Wikforss also argue that the "received view" is unacceptable in its own right. They write:

> The received view stands [or] falls with its ability to supply us with a plausible account of what it is to follow, or be guided by, a rule – in contradistinction to merely acting in accordance with one. (Glüer, Wikforss 2010, 156)

And if we try to conceive of using a word as being a matter of following a rule, they argue, we face a choice between three unacceptable options. The first is to accept commitments that Wittgenstein explicitly rejects.[8] The second is to collapse the distinction between following a rule and acting in accord with a rule, so that every sort of regular behaviour is construed as an instance of rule-following. The third is to endorse a kind of quietism or anti-reductionism that helps itself to the distinction between following a rule and merely

---

[7]   In a fuller treatment, it would be interesting to examine Wittgenstein's use of "*Ent-sprechung*" and its cognates in other contexts for the light they cast on this question of translation. My sense is that that would not provide support for translating "*Entsprechung*" as "analogy".

[8]   That will only be unacceptable, of course, if we are aiming to give an account of Wittgenstein's views; it might be an acceptable view in its own right.

acting in accord with a rule without giving any informative account of that distinction.

For reasons of space, I cannot consider all the details of Glüer and Wikforss's case for their view. But I shall argue that Wittgenstein shows us a way to understand language-use as a form of rule-following behaviour that is consistent with his other commitments and maintains the distinction between following a rule and merely acting in accord with a rule. His account of that distinction is an anti-reductionist one; there is no prospect of giving an account of what it is to follow a rule that is entirely non-circular. But there is nothing philosophically unsatisfactory about that.

Wittgenstein writes:

> Are the propositions of mathematics anthropological propositions saying how we men infer and calculate?—Is a statute book a work of anthropology telling how the people of this nation deal with a thief etc.?—Could it be said: "The judge looks up a book about anthropology and thereupon sentences the thief to a term of imprisonment?" Well, the judge does not USE the statute book as a manual of anthropology. (RFM, III, § 65)

As Wittgenstein says, there is a difference between an anthropological work that records regularities in people's behaviour and a statue book that sets down rules they follow. But how should we characterise the difference? Central to Wittgenstein's account of the distinction is the idea of *using* something *as a rule*. The judge uses the statutes in the statute book as rules for sentencing criminals. And more generally, following a rule involves recognising or using it as a rule. But we should not over-intellectualise what that requires.

Here is an example. English has the saying: "Cometh the hour, cometh the man". That saying is sometimes adapted to fit other contexts. I once came across this instance: "Cometh the hour, cometh the caring people of Chicago". My immediate reaction was that that was wrong; you cannot say "cometh the caring people Chicago". I could not articulate exactly why it was wrong; but I knew that it was. Later, I worked out why it is wrong. "Cometh" is the (archaic) third-person singular of "come": I come, thou comest, he/she/it cometh. The third-person plural is "come". So you can say "Cometh the hour, come the caring people of Chicago"; you can not say "Cometh the hour, cometh the caring people of Chicago". But even before I could explicitly articulate the rule for "cometh", I had grasped that rule and was following it. I was not just acting in a regular way. On the contrary; I treated "cometh" as grammatically *correct* in the third-person singular and *incorrect* in the third-person plural.

Similarly, when someone plays chess, she follows the rules of chess. She may not be able to *state* the rules accurately – or even at

all. But she must be able to make judgements like these: you *can't* move the bishop like that; you're only *allowed* to move it like this; you *have to* move the pawn to the last square before you *can* have a Queen; if you move your pawn like that, I'm *allowed* to take it like this. Such a player is not merely moving the pieces on the board in a regular way: a way that accords with the rules. She *treats* or *uses* the rules as rules. And, on Wittgenstein's view, that is enough for her to be *following* those rules.

4.1 "Our Grammar is Lacking in […] Perspicuity" (PI, § 122)

On the view just sketched, following a rule involves treating it as a rule. But you can only treat something as a rule if you know that it is a rule. So following rules, as opposed to merely conforming to them, requires knowing the rules you are following. Glüer and Wikforss object that such a view of linguistic rules is incompatible with Wittgenstein's other commitments. Their reasoning is this. If using a language involves treating its rules as rules, we must know the rules of our language; otherwise we could not treat them as rules. But Wittgenstein says repeatedly that the grammar of our language is not perspicuous. And to say that is to say that we *do not* know the grammatical rules that govern our language. So the current view of rule-following conflicts with Wittgenstein's insistence that we often misunderstand the grammar of our own language.[9]

However, there is no tension here – provided we avoid over-intellectualising what it takes to be following linguistic rules. A central insight in Wittgenstein's later work is that even though we have a practical grasp of the use of our language, we often have no reflective understanding of that use. For instance, we have a practical grasp of our language for talking about time and of the procedures for measuring time. But we lack a reflective, philosophical understanding of the grammar of that language: that is why we are easily puzzled by the question, 'What is time?'; and it is why we can get into the position of wondering how it is so much as possible to measure time.[10] Now what does it take to have a practical grasp of our language? It is not enough that we merely apply words in regular ways: ways that conform to the grammatical rules of our language. Having a practical grasp of our language also includes being able to recognise what does and does not make sense; to identify this use as *right* and that as *wrong*; to recognise that you *can say* this and *cannot say* that. Someone who can do

---

9 For this argument, see Glüer, Wikforss (2010, 157-9).

10 See Wittgenstein's comments about time at PI, §§ 89-90 and about the measurement of time at BB, 26.

all that is not just conforming to the rules that govern their language; they are following the rules. But following the rules, understood in that way, is perfectly compatible with being unable to give a reflective account of those rules. That is Wittgenstein's point.

4.2   Crispin Wright and Basic Rule-Following

Finally, we should consider Glüer and Wikforss's discussion of the account of rule-following developed in Crispin Wright's later work on that topic. Glüer and Wikforss think that Wright's account obliterates the distinction between following a rule and merely conforming to a rule. But Wright highlights a feature of rule-following that is clearly important in Wittgenstein's treatment. Is there a problem, here, for the "received view"?

Wright draws attention to passages like PI, § 219:

When I follow the rule, I do not choose. I follow the rule *blindly*.

At the basic level, as Wright puts it, we can give no reason for following a rule in the way we do. And the message of Wittgenstein's discussion of rule-following, he thinks, is that

All rule-following involves basic rule-following. And basic – 'blind' – rule-following, properly understood, is rule-following without reason. (Wright 2007, 497)

Glüer and Wikforss argue that, if we accept that account of rule-following, we lose the distinction between following a rule and merely conforming to the rule. Intuitively, they think (and I agree), following a rule involves treating the rule as a reason for acting as one does. But on Wright's account, we have no reason at the basic level for following any rule in the way we do. Applying that to the case of language gives the view that we use words in regular ways but, at the basic level, have no reason for using them as we do. So, Glüer and Wikforss conclude, if we accept the view of rules that Wright derives from Wittgenstein, we must give up the idea that using language involves following rules.

As before, I do not think this is a telling criticism of the "received view" that using language involves following rules. The point about basic rule-following that Wright takes from Wittgenstein needs handling with care. Properly understood, I shall argue, there is no conflict between Wittgenstein's observation that the application of a familiar rule is "blind" and the idea that, when someone is following a rule as opposed to merely acting in accord with the rule, the rule is involved in her reasons for acting as she does.

We have already quoted PI, § 219. Here are two other relevant passages from *Philosophical Investigations*:

> "No matter how you instruct him in continuing the ornamental pattern, how can he *know* how he is to continue it by himself?" – 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. (PI, § 211)

> "How am I able to follow a rule?" — If this is not a question about causes, then it is about the justification for my acting in *this* way in complying with the rule.
>
> Once 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". (PI, § 217)

Now consider a familiar case. I am writing down a series of numbers, following the rule 'add 2 each time'. I write down "996, 998, 1000, 1002". A conversation ensues:

Q: What reason do you have for writing "1002" after "1000"?

A: I'm following the 'add 2' rule and the rule requires me to put "1002" at this point.

Q: But what reason do you have for thinking that "1002" is what the 'add 2' rule requires you to put at this point?

A: Well, following the 'add 2' rule requires you to put "2, 4, 6, 8, 10…" and to go on doing the same thing at each successive step. Putting "1002" after "1000" is doing the same thing as that.

Q: But what reason do you have for thinking that putting "1002" after "1000" *is* doing the same thing as that?

A: It just is. Putting "1002" after "1000" just is what counts as doing the same thing as before.

What should we say about my reasons in this case? The position is this. In the first place, I did have a reason for continuing the series in the way I did: my reason was that the 'add 2' rule requires putting "1002" after "1000". Furthermore, I could give some reasons for thinking that that is what the 'add 2' rule requires. Those reasons 'soon gave out'. At that point, I wrote "1002" without having any further reasons *for thinking that that's what the add 2 rule requires at that point*. In that sense, I acted "without reasons". But that does not mean that, in writing "1002", I had no reasons *for doing what I did*. On the contrary, I did have a reason for writing "1002"; namely, that "1002" was what the 'add 2' rule requires one to put after "1000".

Glüer and Wikforss worry that, if we accept that basic rule-following is "blind", we lose the distinction between following a rule and merely acting in accord with a rule. But the points just made give us an answer to that worry. We can imagine a parrot or a machine

making noises that conform to the rule 'add 2' without understanding what it is doing. It makes the sounds "2, 4, 6, 8…. 996, 998, 1000, 1002" and so on. But it has no sense that "1002" is the right way to continue the series and that "1004" would be wrong: it has no reason for putting "1002" after "1000"; it is not trying to follow the rule 'add 2'; indeed, it has no idea that there is such a thing as the rule 'add 2', or that there are such things as rules at all. In short, the parrot or the machine is making sounds that conform to the rule 'add 2'; but it is not following the rule. Contrast the parrot or the machine with me. When I write "1002" after "1000", I am trying to follow the rule 'add 2' and, as we have seen, I do have a reason for putting "1002": namely, that that is what the rule requires at this point. That is the difference between me and the parrot or the machine. And it is entirely consistent with Wittgenstein's point that my reasons for thinking that the 'add 2' rule requires acting in *this* way at this point 'soon give out'.

References

Works by or originating from Wittgenstein

Wittgenstein, L. (1969). *The Blue and Brown Books*. 2nd edition. Oxford: Blackwell. [BB]

Wittgenstein, L. (2005). *The Big Typescript*. Edited and transl. by C.G. Luckhardt; M.A.E. Aue. Oxford: Blackwell. [BT]

Wittgenstein, L. (1989). *Wittgenstein's Lectures on the Foundations of Mathematics, Cambridge, 1939*. Edited by C. Diamond. Chicago: Chicago University Press. [LFM]

Wittgenstein, L. (1969). *On Certainty*. Edited by G.E.M. Anscombe; G.H. von Wright, transl. by D. Paul; G.E.M. Anscombe. Oxford: Blackwell. [OC]. http://dx.doi.org/10.1007/bf02379321.

Wittgenstein, L. (2009). *Philosophical Investigations*. 4th edition. Edited by P.M.S. Hacker; J. Schulte, transl. by G.E.M. Anscombe; P.M.S. Hacker; J. Schulte. Oxford: Wiley-Blackwell. [PI]

Wittgenstein, L. (1978). *Remarks on the Foundations of Mathematics*. 3rd edition. Edited by G.H. von Wright, R. Rhees; G.E.M. Anscombe, transl. by G.E.M. Anscombe. Oxford: Blackwell. [RFM]. http://dx.doi.org/10.2307/3609432.

Other works

Glüer, K.; Wikforss, Å. (2010). "Es braucht die Regel nicht: Wittgenstein on Rules and Meaning". Whiting, D. (ed.), *The Later Wittgenstein on Language*. Basingstoke: Palgrave Macmillan, 148-66.

Rhees, R. (1954). "Can There Be a Private Language". *Aristotelian Society Supplementary*, 28, 77-94. https://doi.org/10.1093/aristoteliansupp/28.1.63.

Wright, C. (2007). "Rule-Following Without Reasons: Wittgenstein's Quiet-
ism and the Constitutive Question". *Ratio*, 20(4), 481-502. `https://doi.`
`org/10.1111/j.1467–9329.2007.00379.x.`

Frege and Wittgenstein's Later Philosophy

Marie McGinn
University of York, United Kingdom

Abstract Interpreters have tended to focus on the relation between Frege and the early Wittgenstein, but Frege also posed problems for the later Wittgenstein. Wittgenstein's concept of a language-game was inspired by Hilbert's alleged formalism, a view criticised by Frege, and it points to an important dialogue that Wittgenstein engages in with Frege. Wittgenstein expresses formalist views and invokes Frege's critique of formalism at the beginning of the *Big Typescript* and *The Blue Book*. He engages more deeply with the problems posed by Frege and formalism in the remarks collected in *Philosophical Grammar*, where the issues raised set the agenda for the first §§ 242 of the *Philosophical Investigations*. The radical transformation in our understanding of meaning and understanding that takes place enables Wittgenstein to escape the problems which P.T. Geach believes were posed for him by Frege's paper "Thought", concerning first-person thoughts about sensations.

Keywords Wittgenstein. Frege. Formalism. Meaning. Language-games.

Summary I. Introduction: Wittgenstein and Frege's "Thought" – II. Kuusela on the Continuities between Later Wittgenstein and Russell and Frege: Language-Games as a Method of Logic – III. Hacker on Later Wittgenstein and Frege on Meaning: The "Disastrous Effect the Preoccupation with the 'Sense' of a Proposition… Has Had" – IV. Wittgenstein's Response to Frege's Attack on the Formalists: Getting Rid of Intermediaries – V. The Paradoxes of PI § 95 and PI § 201: Recognising the Patterns in our Life with Language.

> Something in itself not perceptible by sense, the
> thought, is presented to the reader – and I must
> be content with that – wrapped up in a perceptible
> linguistic form.
> (Frege 1997, 334 fn. D)

> Sense is not the soul of a proposition. So far as
> we are interested in it, it must be completely
> measurable, must disclose itself completely in
> signs.
> (BT, 210)

> I am compelled to occupy myself with language,
> although it is not my proper concern here.
> (Frege 1997, 334 fn. D)

> Everything is carried out *in language*.
> (BT, 283, 286)

1 Introduction: Wittgenstein and Frege's "Thought"[1]

It is hard to escape a sense of dialogue in these paired remarks. Wittgenstein appears, at least in part, to be responding to ideas expressed by Frege. Wittgenstein famously lists Frege as one of the thinkers who influenced him (CV, 16), although it is clearly a matter of interpretation how this influence is to be understood. Interpreters have tended to focus on the relation between Frege and the early Wittgenstein, but the appearance of dialogue above suggests that Frege posed problems that also stimulated the thought of the later Wittgenstein.

In her paper on the Frege-Wittgenstein correspondence, Juliet Floyd records an anecdote related by P.T. Geach concerning Wittgenstein's estimate of "Der Gedanke" ("Thought"), a copy of which he received from Frege when he returned to Vienna at the end of the war, in 1919. Wittgenstein, Geach reported, considered the paper an inferior work – the attack on idealism a particular focus for his criticism – and he persuaded Geach and Max Black not to include it in their collection of translations of Frege's works. However, Geach went on to say that "in spite of Wittgenstein's unfavourable view of 'Der Gedanke', his later thought may have been influenced by it" (Floyd 2011, 99). Floyd quotes Geach's description of one of the influences he believes Frege's paper had:

> Frege affirms (1) that any thought is by its nature communicable, (2) that thoughts about private sensations and sense-qualities and

---

about the Cartesian *I* are by their nature incommunicable. It is an immediate consequence that there can be no such thoughts. Frege never drew this conclusion, of course [...] Wittgenstein was to draw it. (Floyd 2011, 102)

Here Geach sees Frege as posing a problem for Wittgenstein: to clarify how our psychological concepts, and the first-person thoughts in general, function, in such a way that the following pictures no longer tempt us:

[I]t [is] necessary to recognise an inner world distinct from the outer world, a world of sense impressions, of creations of his imagination, of sensations, of feelings and moods, a world of inclinations, wishes and decisions. (Frege 1997, 334)

Now everyone is presented to himself in a special and primitive way, in which he is presented to no one else... And only [he] himself can grasp thoughts specified in this way. (Frege 1997, 333)

The question is whether other elements in Frege's way of thinking about thought and language had to shift before Wittgenstein could arrive at the destination Geach identifies for him. And if so, what is the nature of the shift that takes place? Is it, as Peter Hacker suggests, that Wittgenstein's later philosophy "is propounded to a very large extent in opposition to Frege's. They can no more be mixed than oil and water" (Hacker 2001, 219)? Or should the dialogue between the two philosophers be understood in a less oppositional, more constructive way? In *Wittgenstein on Logic as the Method of Philosophy*, Oskari Kuusela argues for a much more positive view of the relationship between Frege and the later Wittgenstein. These are the questions I want to look at in this paper.

2　Kuusela on the Continuities between Later Wittgenstein and Russell and Frege: Language-Games as a Method of Logic

According to Kuusela, the break with his early philosophy begins with Wittgenstein's disappointment with the limited capacity for calculus-based approaches to the task of logical clarification to capture the complex and fluctuating uses of the expressions of natural language. He came to see that the major obstacle to progress in philosophy is the assumption, shared by Frege, Russell and the early Wittgenstein, that behind the messy, surface phenomena of natural language there is an ideal, abstract system of propositions. It is this conception that allows philosophers to conceive of logic as the laws

of thought, as what is common, or essential, to everything that can be called thought or language. The idea of logic as a precise calculus or system of rules and the conception of propositions as ideal entities of which linguistic expressions are only the impure manifestation go together. This is the picture Frege expresses in "Thought"; it is how Kuusela understands Wittgenstein's shift away from it that we are interested in.

The root of the problem, according to Kuusela, lies in the idea that logic requires us to speak of language in a purified or idealised sense. We are driven to this by the conflict between logic's aspiration for exactness and the actual vagueness of everyday language: everyday language is not, on its surface, a calculus operated according to precise rules. Since everyday language does not appear to meet the ideal, it must be met at an underlying level: we are led to reify the ideal. This constitutes a fundamental misunderstanding of the role of the ideal. And for Kuusela it follows that to understand the shift Wittgenstein makes means understanding how the role of the ideal is reconceptualised.

The shift, as Kuusela understands it, is fundamentally a methodological one. The ideal calculi which logicians construct with the aim of clarifying how expressions function are no longer to be considered as something with which reality must correspond. There is no single system of propositions and one cannot assume that the same logical laws apply irrespective of the objects of thought. However, we can treat these precise calculi as objects of comparison, which may be useful for shedding light on a particular aspect of how an expression of natural language functions, with the aim of clearing up particular misunderstandings. The ultimate aim is to clear away misunderstandings, by describing aspects of the complex, fluid, dynamic uses of linguistic expressions. But there should be no expectation that these descriptions will cover all the varied cases in which we use an expression, or that they are in any way definitive. Putting the ideal in its proper place means we can acknowledge without falsification the complexity and diversity of the uses of the expressions of everyday language: our ideal descriptions are merely approximate descriptions of reality, which we construct for a particular purpose.

The break with Frege is not, on this understanding, an outright rejection of his conception of logic, but a repositioning of it. This is how Kuusela understands Wittgenstein's idea that "[t]he *preconception* of [the] crystalline purity [of logic] can only be removed by turning our whole enquiry around" (PI, § 108). It means putting the ideal in its proper place, as an object of comparison, and at the same time reorientating our attention towards the actual use of expressions within our everyday lives: towards "the spatial and temporal phenomenon of language" (PI, § 108). However, it is crucial that Wittgenstein's philosophical aims remain unchanged: "[T]he inquiry must be

turned around, but on the pivot of our real need" (PI, §108), namely, the logical clarification of the functioning of expressions as a means to resolve philosophical problems. It is within the context of this understanding of important continuities between Frege, Russell and the early Wittgenstein that Kuusela understands Wittgenstein's introduction of the idea of language-games. Kuusela sees the concept of a language-game primarily as a method for describing the use of expressions in a way which extends the capacity of calculus-based methods and overcomes their limitations.

Kuusela sees Wittgenstein's development of the method of language-games as amounting to a Kuhnian paradigm-shift, in the sense that, while it can handle the cases that calculus-based methods (which it absorbs as a special case) can handle, it vastly extends the possibilities for describing the uses of the expressions of natural language. At the heart of the method is the idea that it is in the use of expressions as it is interwoven with human activities that their specific roles are revealed. The method of language-games – a method for describing the scene of language-use – is devised as a means for studying the functioning of expressions within the context of the activities of the life into which their use is interwoven. If in logic we are trying to clarify the use or logical function of words, and their use is embedded in our life, then it is our life with words and the different circumstances of their use that reveals their function, and which we need to describe.

This indicates the way in which Wittgenstein's method is connected with a particular conception of language. Wittgenstein has clearly rejected the picture of language as a mental phenomenon hidden away in our minds. When the investigation of language takes the form of an investigation into how human beings operate with signs within their everyday, active lives, then we are regarding language in a particular way, as constituted by a form of life in which speakers employ expressions in ways that are governed by certain rules. Kuusela raises the question whether this means that Wittgenstein's method is based on a conception of the nature of language and whether this is compatible with his claim that he is not putting forward any philosophical theses. Kuusela argues that, properly understood, Wittgenstein

> is not committed to such theses [...] the method of language-games eschews commitment to philosophical theses about language, including the thesis of language use as embedded in actions or language as a form of life. (Kuusela 2019, 169)

Kuusela argues that Wittgenstein's method only depends upon "*comparing* language with a game according to rules, or regarding it as or describing [it] in the form of such a game" (Kuusela 2019, 170).

Kuusela's emphasis here is on the idea that the use of a word may be something constantly fluctuating, yet for purposes of clarification, and for a specific purpose, we may find it useful to envisage its use as a game with fixed rules. We can capture an aspect of its use by means of an ideal use regulated by a definite rule which we set alongside the actual, fluctuating use. His point is that there is "no claim that such a description captures language use in all its actual complexity" (Kuusela 2019, 171) and hence, by implication, no claim that language *is in its nature* a language-game played according to precise rules. Kuusela's focus here is solely on the issue of whether the use of an expression is essentially governed by rules that can be made fully perspicuous. He wishes to acquit Wittgenstein of dogmatism in this respect and in that he is surely entirely correct. However, the foundational conception of logic was dependent on a conception of propositions which conceived of the sense of a proposition as something that was instantaneously grasped by the mind. This is the position Frege expresses in "Thought":

> The grasp of a thought presupposes someone who grasps it, who thinks. He is the owner of the thinking, not of the thought. Although the thought does not belong with the contents of the thinker's consciousness, there must be something in his consciousness that is aimed at the thought. (Frege 1997, 342)

Kuusela has said very little about the nature of Wittgenstein's shift away from this position, although the shift to a method that is open to the dynamic and fluctuating nature of the use of the expressions of natural language has also, clearly, made a shift away from that conception of sense.

The claim that "nothing is hidden" (PI, § 435), that "everything lies open to view" (PI, § 126), that we are concerned entirely with "the spatial and temporal phenomenon of language" (PI, § 108), essentially amounts to the claim that we are concerned with *signs* and their use. This marks a major shift away from Russell, Frege and the early Wittgenstein, insofar as it abandons the idea of the instantaneously meaningful symbol, and accepts that all there is are signs whose use is extended in time. Kuusela notes that Wittgenstein's concept of a language-game was inspired by Hilbert's alleged formalism – the idea that syntax can be conceived as a system of rules for a game – a view criticised by Frege. Clearly, this might be seen as an indication of another important dialogue going on between Frege and the later Wittgenstein. Here the question is not merely whether the rules governing the expressions of natural language are determinate and can be made fully perspicuous, but whether logic can be preserved without the concept of meanings, understood as something distinct from the sign and graspable by the mind in an instant. Kuusela does not directly address

the issue of signs and their use versus meanings as a focus for understanding the transformation from the early to the later philosophy. Yet it is one of the questions that preoccupies Wittgenstein in the first 242 paragraphs of the *Philosophical Investigations* and it is central, I should argue, to the shift away from the ideas of Frege in "Thought".

3 Hacker on Later Wittgenstein and Frege on Meaning: The "Disastrous Effect the Preoccupation with the 'Sense' of a Proposition… Has Had"

Peter Hacker sees Wittgenstein's revisionary thoughts concerning the concepts of thinking, meaning something and understanding as the main pivot of the transformation of his philosophy in the early 1930s. He writes:

> It is no coincidence that the opening chapters of the *Big Typescript* are concerned with the investigation of understanding, meaning, and explanation, for it is this that signals the transformation in Wittgenstein's conception of language and representation. (Hacker 2001, 229)

However, Hacker does not focus on Wittgenstein's engagement with Frege's critique of formalism, but sees the dispute between the two conceptions of sense as one that Wittgenstein settles by appeal to the tribunal of ordinary language:

> For the thought that a speaker might know or understand what an expression that he uses correctly means, but be altogether incapable of saying what he means by it, is incoherent. (Hacker 2001, 229)

With this, and a series of other observations about what it makes sense to say, ordinary usage is taken to settle the matter:

> The meanings of words are not entities correlated with the words by 'a method of projection' (as had been argued in the *Tractatus*) or by the abstract machinery of 'senses' (modes of presentation of a meaning – as Frege had argued). To know what a word means is not to 'grasp' an abstract entity, a sense, which is associated with the word, nor to know what entity the word stands for, but rather to know its use. The meaning of an expression is best conceived as its use – that is, the manner in which it is to be, and normally is, used. (Hacker 2001, 229)

Wittgenstein invokes Frege's critique of formalism in the opening paragraphs of both the *Big Typescript* and *The Blue Book*, both of which

originated at about the same time, in 1933-34. In the *Big Typescript*, the reference to Frege is prefaced by a remark in which Wittgenstein appears to express his commitment to a version of formalism:

> It can also be put this way: If one always expresses oneself in a system of language and so uses only propositions of this system to explain what a proposition means, then in the end meaning drops out of language completely, and thus out of consideration; what remains is language, the only thing we can consider. (BT, 3)

He then goes on to make his fundamental objection to Frege's attack on formalism:

> When Frege argues against a formal conception of arithmetic he is saying, as it were: These pedantic explanations of symbols are idle if we understand the symbols. And understanding is like seeing a picture from which all the rules follow (and by means of which they become understandable). But Frege doesn't see that this picture is in turn nothing but a sign, or a calculus, that explains the written calculus to us. (BT, 3)

He makes the same point against Frege in the opening pages of *The Blue Book*:

> Frege ridiculed the formalist conception of mathematics by saying that the formalists confused the unimportant thing, the sign, with the important thing, the meaning. Surely, one wishes to say, mathematics does not treat of dashes on a bit of paper. Frege's idea could be expressed thus: the propositions of mathematics, if they were just complexes of dashes, would be dead and utterly uninteresting, whereas they obviously have a kind of life. And the same, of course, could be said of any proposition: Without a sense, or without the thought, a proposition would be an utterly dead and trivial thing. And further it seems clear that no adding of inorganic signs can make the proposition live. And the conclusion which one draws from this is that what must be added to the dead signs in order to make a live proposition is something immaterial, with properties different from all mere signs.
> [O]ne is tempted to imagine that which gives the sentence life as something in an occult sphere, accompanying the sentence. But whatever accompanied it would just be another sign. (BB, 4-5)

Despite the sureness of his response to Frege, there is a clear sense that Frege's belief that the domain of language is not on its own enough to understand the nature of the proposition posed a problem for Wittgenstein, a problem about which he was led to think very

deeply. It is the issue that sets the agenda for many of the first 242 paragraphs of the *Philosophical Investigations*, where Wittgenstein's engagement with it is the occasion for a transformation in our understanding of the nature of meaning and understanding that is more radical than Hacker can allow. And it is one in which the idea that "the speaking of language is part of an activity, or of a form of life" (PI, § 23) serves as part of a picture that is intended to guide us from error to truth. The nature of Wittgenstein's engagement with Frege here is complex. For while he wants us to "appreciate what a disastrous effect the preoccupation with the 'sense' of a proposition, with the 'thought' that it expresses, has had" (BT, 210), what he ultimately wants to show is that "Frege's basic idea in his theory of sense and meaning [is]: that the *meaning* of a proposition, in Frege's sense, is its *use*" (BT, 210). The deep engagement with Frege is, in the end, intended to bring about a solution to the problem Frege posed – "What makes a sign a proposition?" – and in such a way that Frege himself would have recognised it as a solution.

The extent of Wittgenstein's engagement with Frege is very clear in the *Big Typescript*. He raises the issue of our use of signs – to give orders, answer questions, etc. – and the question of whether something needs to be added to them in order for the mere signs to become *a command*, or *an answer*, again and again. He is constantly working against the "often held view", expressed in the quotation from "Thought" at the beginning of this paper

> that one can show one's understanding only incompletely, as it were. That one can only point to it from afar, as it were, can get closer to it, but can never grab it with one's hand. And that finally what matters must always remain unsaid. (BT,10)

It is this inexpressible thought, we are tempted to think, that fills the gap between an order and its execution, between a wish and its fulfilment. And against this in the dialectic, Wittgenstein over and over again makes the case for his more formalist approach:

> [L]et's not talk about "meaning something" as an indefinite process that we don't know very well, but about the (actual), "practical" use of the word, about the actions we carry out with it. (BT, 157)

Later he acknowledges that his approach is exactly the one that Frege ruled out:

> Here I am touching on the way of explaining signs that Frege ridiculed so much. For one could explain the words "knight", "bishop", etc by citing the rules that apply to these pieces. (BT, 206)

4 Wittgenstein's Response to Frege's Attack on the Formalists: Getting Rid of Intermediaries

The *Philosophical Grammar*, based on manuscripts that were also written in the early 1930s, also begins by invoking Frege's attack on the formalist conception of arithmetic, and Wittgenstein gives the same objection to it: "Frege does not seem to see that such a picture would itself be another sign, or a calculus to explain the written one to us" (PG, 40). And again, the same dialectic ensues, in which Wittgenstein repeatedly affirms the formalist picture:

> I want to say the place of a word in grammar is its meaning.
>
> But I might also say: the meaning of a word is what the explanation of its meaning explains.
>
> The use of a word in the language is its meaning.
>
> The meaning is the role of the word in the calculus. (PG, 59-60)

But then he raises a question:

> But it might be asked: Do I *understand* the word just by describing its application? Do I understand its point? Haven't I deluded myself about something important?
>
> At present, say, I know only how men use this word. But it might be a game, or a form of etiquette. I don't know why they behave in this way, how *language* meshes with their life.
>
> Is meaning then really only the use of a word? Isn't it the way this use meshes with our life?
>
> But isn't its use a part of our life? (PG, 65)

I think that we should see this as Wittgenstein being pushed, through his engagement with the problem posed by Frege's attack on formalism, to notice something about what is involved in our grasp of the use of the expressions of natural language, which leads him to a much deeper understanding of what it is that he is describing. It dramatically shifts the focus away from an impersonal conception of language as a calculus, which can be described by means of a rule for the use of a word, and acknowledges the central importance of the role of the speaker as an agent, whose active participation in a life with language is essential to our understanding of what language is. Wittgenstein immediately goes on to make the point explicit:

> Do I understand the word "fine" when I know how and on what occasions people use it? Is that enough to enable me to use it myself? I mean, so to say, use it with conviction.
>
> Wouldn't it be possible for me to know the use of the word and yet follow it without understanding? (As, in a sense, we follow the singing of birds). So isn't it something else that constitutes understanding – the feeling "in one's breast", the living experience of the expressions? – They must mesh with *my own* life.
>
> Well, language does connect with my own life. And what is called "language" is something made up of heterogeneous elements and the way it meshes with life is infinitely various. (PG, 65-6)

It is important that we should not see this as an attempt on Wittgenstein's part to *explain* what meaning consists in. He is careful to say in a remark that follows closely on the ones I have just quoted that he is "only *describing* language, not *explaining* anything" (PG, 66). There is a danger, in placing the emphasis on the speaker as agent and on our life with language, that it could appear that Wittgenstein is claiming that it is the human agent who uses language who breathes life into the words he utters: "[A]s if one must be doing the meaning of it oneself in order to understand it as meaning", with the result that one would no longer be

> considering it as a phenomenon or fact but as something intentional which has a direction given to it. [And] what this direction is we do not know; it is absent from the phenomenon as such. (PG, 143)

This is a view fundamentally at odds with the idea that "nothing is hidden" (PI, I, § 435), "everything lies open to view" (PI, § 126), and it is not one that Wittgenstein intends to embrace. Our being alive is not used to explain what gives life to language; the concepts of living and the capacity to use language are on the same level; the capacity to use language is one of the criteria of being a living thing. However, these issues lead him into a much deeper engagement with the problem posed by Frege's attack on the formalists and a much more expansive treatment of the dialectic between the opposing views: the question of whether what comes before my mind when I hear and understand a word is the meaning of the word or just the word itself. It leads him to develop an increasingly naturalistic approach to the description of our linguistic practices, as the significance of viewing our practices from *within* is made clear.

Wittgenstein had responded to Frege's attack on formalism by pointing out that anything added to a sign would be just another sign. And he suggested that the way out of the difficulty is

to recognise that the sense of a proposition – what gives life to a sign – is its use in a calculus. But now he observes a difficulty with his own solution:

> I imagine the expression of a wish as the act of wishing, the problem appears solved, because the system of language seems to provide me with a medium in which the proposition is no longer dead.
>
> But now someone will say: even if the *expression* of the wish is the wish, still the whole language isn't present during this expression, yet surely the wish is!
>
> So how does the language help? (PG, 149)

Once again, we may feel forced into thinking of the wish as a shadow of its fulfilment, which will admit of no interpretation. The use is something extended in time, yet the wish is surely all there at the moment I have it. Once again, we will be faced with the question of how a wish can prefigure its fulfilment. For whatever is before my mind, can it not be interpreted in many different ways? But then how can I know what it is that I wish?

Wittgenstein responds to these worries as follows:

> I said that it is the *system* of language that makes the sentence a thought and makes it a thought *for us*.
>
> That doesn't mean that it is while we are using a sentence that the system of language makes it into a thought for us, because the system isn't present then and there isn't any need for anything to make the sentence alive for us, since the question of being alive doesn't arise. (PG, 153)

The reason that it does not arise is that the language we are investigating is *my* language, the language I understand and within which I am at home. It is not that in using language I breathe life into dead signs, but in mastering the techniques for employing the expressions of my language, in the way this use meshes with my life in infinitely various ways, those signs are alive for me:

> But if we ask: "[W]hy doesn't a sentence strike us as isolated and dead when we are reflecting on its essence, its sense, the thought etc" it can be said that we are continuing to move in the system of language. (PG, 153)

What becomes clear is that Wittgenstein finds himself drawn more deeply into a dialogue with Frege. What removes the temptation to

look for intermediaries between a sign and its application – for something which cannot be interpreted – is our natural way of responding to the expressions of our language:

> If I see the thought symbol "from outside", I become conscious that it *could* be interpreted thus or thus; if it is a step in the course of my thoughts, then it is a stopping-place that is natural to me and its further interpretability does not occupy (or trouble) me. (PG, 147)

His dynamic solution to the problem of the sense of a proposition that Frege had posed prompts him to reflect, not only on the way in which the different functions of expressions are revealed in their use within our everyday lives, but on how to describe what goes on when I use the expressions of my language and understand them. Can he show, as it were to Frege's satisfaction, that understanding can be understood without recourse to meanings? That turns out to be a question requiring a more protracted treatment than perhaps it seemed at first sight.

5 The Paradoxes of PI § 95 and PI § 201: Recognising the Patterns in our Life with Language

When, in *The Blue Book*, Wittgenstein describes, as "a simple case of operating with words", the case in which "I give someone the order: '[F]etch me six apples from the grocer'", he uses it to present ways of using signs which are simpler than ours. Here, he suggests, "[w]e see activities, reactions, which are clear-cut and transparent", but which we recognise are "not separated by a break from our more complicated ones" (BB, 16-17). This is the aspect of Wittgenstein's method of language-games that Kuusela focuses on. However, when the example appears in § 1 of the *Philosophical Investigations*, a further thought has been added:

> "But how does he know where and how he is to look up the word 'red' and what he is to do with the word 'five'?" – Well, I assume that he *acts* as I have described. Explanations come to an end somewhere, – But what about the meaning of the word "five"? – No such thing was in question here, only how the word "five" is used. (PI, § 1)

It is not only that the way a speaker operates with a sign makes clear what he means by it – whether, for example, he means a colour, a shape, or a number, by a sign he ostensively defines – but that, in the end, a speaker *acts* without guidance from anything we might call the *meaning* of the sign in applying the expressions of his language

in the way he has learned to use them. It is this idea that Wittgenstein returns to in PI, § 138, which marks the beginning of a discussion that culminates in the remarks on the paradox of interpretation in PI, I, § 201 ff. What now seems clear is that the discussion can be seen as having its roots, at least in part, in Wittgenstein's response to Frege's attack on the formalists.

We should see the paradox of both PI, § 95 and PI, § 201 as something Wittgenstein evolves as a way of demonstrating something he has long been committed to: that anything added to a sign is just another sign. The idea of an intermediary between a sign and its application, which settles what counts as a correct application of it – either a shadow (a rule of projection, Frege's sense of a proposition) or an interpretation – is an illusion. As he said in his original objection to Frege's view: any addition "would itself be another sign, or a calculus to explain the written one to us" (PG, 40). The aim of his dynamic conception of meaning was to put an end to temptation to think of meaning as occurring in a peculiar medium, independent of the act of expressing our thoughts. The point of PI, § 1 is that "everything lies open to view" (§ 126) in how the speaker operates with signs. However, PI, § 138 appears to acknowledge that the pressure to introduce intermediaries is not easily removed. The dynamic conception of meaning can seem to exert a pressure of its own to introduce intermediaries and Wittgenstein has to do more work to show, on the one hand, that that idea is an illusion, and on the other, that everything we need to understand language and linguistic mastery lies open to view in how speakers operate with words in the context of their everyday lives.

In the remarks which follow PI, § 138, Wittgenstein uses his interlocutor to pose a series of challenges to his dynamic conception of meaning. How can I know that I mean one series rather than another? How can I know that I have understood the principle of a series when I say, "Now I understand"? How can I say that I meant an order to develop a series in a particular way at the time I gave it? How do I know what I am to do at *this* point, if whatever I do can, on some interpretation, be made compatible with the rule? How am I able to follow a rule if the rule itself does not tell me which way I am to go? What is my justification for my applying a rule in the way that I do? Does his dynamic conception of meaning mean that human agreement decides what is true and what is false?

Wittgenstein's response to all these questions is to describe aspects of our life with signs in a naturalistic manner. We saw him introduce a performative element into his conception of language in PI, § 1: "I assume he *acts* as I have described". The challenges the interlocutor makes to his dynamic conception of meaning provide the occasion for Wittgenstein to explore this performative aspect more fully. In his investigation of "Now I understand", "Now I can go on", Wittgenstein

does not give anything that could be considered even a partial definition of these words. He compares their use with a "signal" (PI, § 180), "an instinctive sound, a glad start" (PI, § 323), but clearly does not intend to claim that the words, even on a particular occasion, *mean* a signal, an instinctive sound or a glad start. The words are, rather, to be *seen as* a sign, an instinctive sound or a glad start: expressive of the speaker's confidence that he will go on correctly, if the occasion arises. It is a description given from within our practice, from the perspective of a practioner, and it depends on the person reading it also being a practioner and recognising the description as apt. It is very far from a calculus-based method: there is an investigation which is intended to elucidate how these words are used, but it depends on the evocation in the reader of one's own life with signs – a way of operating with them – that is found to be recognisable.

Whether the words "Now I understand" are justified or used correctly on a particular occasion is another question. Here Wittgenstein points to the role of the *context* – or the *circumstances* – in which a speaker says these words, in determining whether they are correctly used. The tacit conventions by which we judge whether what a speaker claims is correct are immensely complex, touching on the speaker's past history and training, his established abilities, our confidence in his capacities, and so on. Our third-person criteria are complex and involved, but what forms the background to their employment is an existing linguistic practice and a speaker's manifest possession of abilities to participate in it. Wittgenstein overcomes the idea that "Now I understand" must describe a mental state that makes its appearance in an instant by showing *a pattern in our use of words* when we speak of coming to understand. Recognising the pattern turns, on the one hand, on seeing the way in which our employment of expressions displays the first-person/third-person asymmetry that is distinctive of agency, and on the other, on seeing the way in which our criteria are responsive to what is revealed over time, to the circumstances in which things are said and done. This alternative conception is not merely being made the methodological basis for a novel model for how to describe the use of the expressions of natural language, it is Wittgenstein's working out of a modified version of formalism that meets the objections of his interlocutor.

One of the central questions of the remarks on rule-following is the one we saw anticipated in *Philosophical Grammar*:

> But that is just what is remarkable about *intention*, about the mental process, that the existence of a custom, of a technique, is not necessary to it […].

> But isn't chess defined by its rules? And how are these rules present in the mind of someone who intends to play chess? (PI, § 205)

And again, Wittgenstein uses his naturalistic method to get us to see that it is not a question of what is "present in the mind". As he says at PI, § 199:

> To understand a sentence means to understand a language. To understand a language means to have mastered a technique.

We come back to the idea of a certain pattern in the life of a speaker, understood as an autonomous agent possessing a manifest ability to participate in a practice which provides the context for what he says and does. What is in question here is that the words "I want to play chess", "I meant he should say '1002' after '1000'", are spoken by a speaker who has mastered certain techniques, in the context of a practice which he points to if he wants to specify which game he intends to play, which rule he meant by 'add 2'. The practice depends upon agents' acquiring the ability to act confidently and independently and autonomously – without further guidance – in ways that agree. We resist Frege's temptation to think that ordinary signs need supplementing with super-signs which cannot be interpreted, by recognising ourselves as active participants in a linguistic community in which we are bound together in agreed, regular, stable and established ways of acting with signs that constitute our "form of life" (PI, § 241).

What Wittgenstein has tried to make clear is that the formalist is right: it is not anything that accompanies an act of following a rule that makes it an event that we can, for example, describe as a move in chess, adding 57 and 68, or developing the series +2. It just is a fact about us that, after a certain sort of training, we do for the most part go on independently in a way that sustains our practices. We may, in certain circumstances, give justifications for how we apply a particular rule, but in the end, as Wittgenstein observed in PI, I, § 1, "[e]xplanations come to an end". We come back to the actions of an autonomous agent who applies the techniques he has been trained to use, without guidance, in ways which count as "following the rule".

This shift in how we see language and linguistic mastery is key to Wittgenstein's achieving the solution to the problem Geach held was posed for him by "Der Gedanke". The effect of training in the use of the psychological expressions 'think', 'imagine', 'expect', 'wish', 'intend', etc., is to initiate a speaker into the complex form of human life, whose distinctive patterns are laid-down in the language-games of thinking, inferring, calculating, measuring, imagining, expecting, intending, and so on. As a speaker acquires the capacities of an autonomous agent who operates with words in ways that are characteristic of our complicated form of life, he gradually takes on the form of life distinctive of a minded human being. The ideas of private

objects, an inner realm and of introspection are seen to have no role to play: "everything lies open to view" (PI, § 126) in a speaker's form of life with signs. In the same way, 'I' in sentences in which a speaker gives expression to what he feels, affirms his intention, gives voice to what he believes, expects, wishes, etc., does not function as a name. As Geach says, there are no thoughts of the kind Frege held were incommunicable; it is a matter of describing the distinctive use of first-person present indicative sentences. But seeing this depends on our making a radical adjustment in our conception of the nature of language and recognising that there is nothing to meaning over and above a sign and its use.

References

Floyd, J. (2011). "The Frege-Wittgenstein Correspondence: Interpretive Themes". De Pellegrin, E. (ed.), *Interactive Wittgenstein*. New York, NY: Springer, 75-107.

Frege, G. (1997). "Thought". Beaney, M. (ed.), *The Frege Reader*. Oxford: Wiley-Blackwell, 325-45.

Hacker, P.M.S. (2001). *Wittgenstein: Connections and Controversies*. Oxford: Oxford University Press.

Kuusela, O. (2019). *Wittgenstein on Logic as the Method of Philosophy: Re-Examining the Roots and Development of Analytic Philosophy*. Oxford: Oxford University Press.

Wittgenstein, L. (1958). *The Blue and Brown Books*. Oxford: Blackwell. [BB]

Wittgenstein, L. (1974). *Philosophical Grammar*. Edited by R. Rhees, transl. by A. Kenny. Oxford: Blackwell. [PG]

Wittgenstein, L. (1998), *Culture and Value*. Edited by G.H. von Wright; H. Nyman; A. Pichler, transl. by P. Winch. Oxford: Blackwell. [CV]

Wittgenstein, L. (2005). *The Big Typescript: TS 213*. Edited and transl. by C.G. Luckhardt; M.A.E. Aue. Oxford: Wiley-Blackwell. [BT]

Wittgenstein, L. (2009). *Philosophical Investigations*. Edited by P.M.S. Hacker; J. Schulte, transl. by G.E.M. Anscombe; P.M.S. Hacker; J. Schulte. Oxford: Wiley-Blackwell. [PI]

From the Referential to the Relational: Duchamp and Wittgensteinian Family Resemblance

Marjorie Perloff
Stanford University, USA

Abstract This essay identifies and discusses certain affinities between Wittgenstein's philosophy and Duchamp's artistic work. While acknowledging the great differences between the one and the other, it cannot fail to strike one that the two take similar attitudes on certain issues concerning the way they look at mathematics and numbers, the importance given to the relational over the referential, and the peculiar use made of the notion of context.

Keywords Duchamp. Wittgenstein. Family Resemblance. Context. Relation.

Summary 1 Introduction. – 2 Context and Contact. – 3 All in the Family.

1 Introduction

In 1913, Marcel Duchamp cut a length of white thread exactly one metre long, stretched it at a distance of one metre above a rectangular canvas painted Prussian blue, and let it fall. He did the same thing with two more threads, each one to fall onto a separate canvas, and then to be glued down with varnish in whatever shape it had assumed **[fig. 1]**. Calling the piece *Three Standard Stoppages* (*Trois Stoppages Etalon*), Duchamp was amused to note that the supposedly 'fixed' metre assumed three slightly different shapes when it fell to the ground (see Cabanne 1968, 46-7). Duchamp called it "canned meter" or "canned chance": "pure chance" he tells Pierre Cabanne, "interested me as a way of going against logical reality". Or, to give this thread the twist we find in Wittgenstein's *Philosophical Investigations*:

> There is *one* thing of which one can state neither that it is 1 meter long, nor that it is not 1 meter long, and that is the standard meter in Paris. — But this is, of course, not to ascribe any remarkable property to it, but only to mark its peculiar role in the game of measuring with a meter-rule. (PI, I, § 55)

In later life, Duchamp remarked that *Three Standard Stoppages* was his most important work:

> That was really when I tapped the mainspring of my future. In itself it was not an important work of art, but for me it opened the way—the way to escape from those traditional methods of expression long associated with art [...] For me the Three Stoppages was a first gesture liberating me from the past. (Kuh 1962, 81)

The three glued threads were permanently affixed to glass plate strips, which served as imprints for the preparation of three wood templates. The entire assembly was then enclosed in a wooden croquet box **[fig. 2]**, and it is in the context of this box that most viewers know the work. What Duchamp liked is that his curved threads questioned the authority of *metre* as a standard unit of measure. The work reminds us, as Francis Naumann notes, that metre is itself "a unit of length generated through approximation: the straightening out, as it were, of a curved meridian" (Naumann 1989, 30). Duchamp thus parodies our faith in scientific authority, our trust in causality.

At around the same time that Duchamp was playing with "canned chance", Wittgenstein, who was serving in the Austrian army on the Eastern Front during World War I, wrote in his notebook:
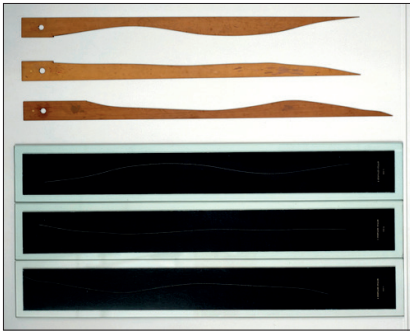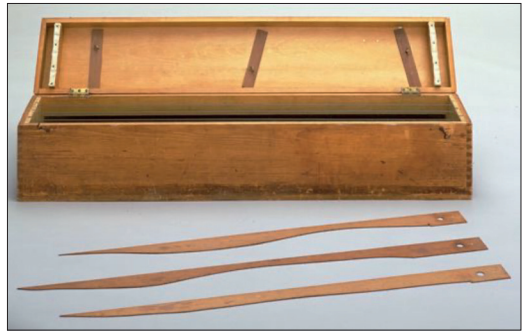
> In essence, the whole modern conception of the world is based on the illusion that the so-called laws of nature are explanations of natural phenomena.
> So they stop short at the "laws of nature" treating them as something untouchable, just as their ancestors did with God and Fate. And in fact both are right and both are wrong. The Ancients were actually clearer, in that they acknowledged a clear-cut limit, while with the new system, it is supposed to look as if everything can be explained. (PN, 6.5.1916, 171)

In slightly different form, these lines reappear in the 1922 *Tractatus Logico-Philosophicus*, expressing Wittgenstein's repeated caution that the "so-called laws of nature" are not to be trusted as explanations of natural phenomena (TLP, Prop. 6.371). And in the lectures delivered at Cambridge between 1930-32 – lectures that first introduce many of the key issues taken up in the *Philosophical Investigations* – we find an uncanny echo of the experiment behind *The Three Standard Stoppages*:

> What does it mean to hold that there are *a priori* concepts? If we pull a piece of cotton very tight, then to say that it is straight is to refer to what is manifest to our senses […] But we know perfectly well that if we look through a magnifying glass we shall see that what was apparently straight actually is not so. (WLC, 77-8)

Then again the "uncanny" echo may not be so surprising. For although Marcel Duchamp (1887-1968) and Ludwig Wittgenstein (1889-1951) could hardly have been more different – indeed oppositional – in their

tastes, habits, and values[1] – they shared a particular interest in mathematics as a science with 'poetic' possibilities. The two came of age at the moment in history when geometry, traditionally the dominant branch of mathematics, was giving way to a new understanding of *number*. As Andrea K. Henderson has argued in an important essay on numerical abstraction in Victorian literature, this was the period when logicians first concerned themselves with the seemingly simple reality that, while "7 inches is a concrete reality, 7-ness is not" (Henderson 2024). The shift was from a world in which mathematics was still grounded in spatial intuitions (geometry) to one that turned to the temporal, mathematics concerning itself with sets of objects to be enumerated.

> What was enumerated, moreover, was not things in themselves but the differences between them. Thus mathematicians came to conceive their work not as a referential science, but as a science of *relationships*.[2]

Readers of the *Philosophical Investigations* will recognise this view of *relatedness* as central to Wittgenstein's own thinking. We routinely refer, he remarks early in the *Investigations*, to 5 apples or 3 slabs, but how do we define the number two?

> The definition of the number two, "That is called 'two'"—pointing to two nuts—is perfectly exact.—But how can the number two be defined like that? The person one gives the definition to doesn't know *what* it is that one wants to call "two"; he will suppose that "two" is the name given to *this* group of nuts!—He *may* suppose this; but perhaps he does not. He might make the opposite mistake: when I want to assign a name to this group of nuts, he might take it to be the name of a number. (PI, I, § 28)

> Perhaps someone will say "two" can be ostensively defined only in *this* way: "This *number* is called 'two'". For the word "number" here shows what *place* in language, in grammar, we assign to the word. But this means that the word "number" must be explained before that ostensive definition can be understood.
> Whether the word "number" is necessary in an ostensive definition of "two" depends on whether without this word the other person takes the definition otherwise than I wish. And that will

---

[1] See my "Introduction" to PN, passim.

[2] Cf. the logician William Stanley Jevons (1874), as cited by Henderson 2024: "Number is but another name for diversity. Exact identity is unity, and with difference arises plurality".

depend on the circumstances under which it is given, and on the
person I give it to.
And how he "takes" the explanation shows itself in how he uses
the word explained. (PI; I, § 29)

Indeed, it follows, numbers can be understood only in relation to one
another. Seven is one more than six and one less than eight of what-
ever the items in question. Or again, seven is two times three plus
one. And further (PI, I, § 552-3), the meaning of a given number will
also depend on context:

> What if I were to ask: does it become evident, while we are utter-
> ing the sentences "This rod is 1 metre long" and "Here is 1 sol-
> dier", that we mean different things by "1", that "1" has different
> meanings?—It does not become evident at all. —Say, for example,
> such a sentence as "1 metre is occupied by 1 soldier, and so 2 me-
> tres are occupied by 2 soldiers." Asked, "Do you mean the same
> by both 'ones'?" one would perhaps answer, "Of course I mean the
> same: *one!*" [Perhaps raising one finger.] (PI, I, § 552)

> Now has "1" a different meaning when it stands for a measure and
> when it stands for a number? If the question is framed in *this* way,
> one will answer affirmatively. (PI, I, § 553)

Thus a given number – say, 3 – as in Duchamp's *Three Standard Stop-
pages* takes on different meanings according to its context and use.
The profound shift from the *referential* to the *relational* is at the
core of one of the concepts central to the *Philosophical Investigations*:
namely, *family resemblance*:

> Consider, for example, the activities that we call "games". I mean
> board-games, card-games, ball-games, athletic games, and so on.
> What is common to them all? […] if you look at them, you won't see
> something that is common to *all*, but similarities, affinities, and a
> whole series of them at that. […] Look, for example, at board-games
> with their various affinities. Now pass to card-games; here you
> find many correspondences with the first group, but many common
> features drop out, and others appear. When we pass next to ball-
> games, much that is common is retained, but much is lost. […] Or is
> there always winning and losing, or competition between players?
> Think of patience. In ball-games, there is winning and losing; but
> when a child throws his ball at the wall and catches it again, this
> feature has disappeared. Look at the parts played by skill and luck,
> and at the difference between skill in chess and skill in tennis.

Figure 3
Marcel Duchamp. *The Bride Stripped Bare by Her Bachelors*, *Even (The Large Glass)*, 1915-23. Oil, varnish, lead foil, lead wire, and dust on two glass panels (cracked), each mounted between two glass panels, with five glass strips, aluminum foil, and a wood and steel frame, 109 1/4 × 69 1/4 inches (277.5 × 175.9 cm). © ARS, NY. Bequest of Katherine S. Dreier, 1952. Philadelphia Museum of Art. Photo: The Philadelphia Museum of Art / Art Resource, NY

And the upshot of these considerations is: we see a *complicated network of similarities overlapping and criss-crossing*: similarities in the large and in the small. (PI, I, § 66)

I can think of no better expression to characterize these similarities than "*family resemblances*"; for the various resemblances between members of a family—build, features, colour of eyes, gait, temperament, and so on—overlap and criss-cross in the same way.— And I shall say: 'games' form a family.
And likewise the kinds of number, for example, form a family.... we extend our concept of number, as in spinning a thread we twist fibre on fibre. And the strength of the thread resides not in the fact that some one fibre runs through its whole length, but in the *overlapping of many fibres*. (PI, I, § 67, emphasis added)

This account of family resemblances is nowhere better exemplified than in the world of Duchampian figuration, especially in the famous *Large Glass*, otherwise known as *The Bride Stripped Bare by Her Bachelors, Even* (*La Mariée mise à nu par ses célibataires, même*) **[fig. 3]**. There is, in this "delay" in glass, as Duchamp playfully called it, only one bride: the enigmatic tube work hanging from the "Milky Way" in the upper half of the *Glass*, but in the lower half, there are *seven* conelike shapes, known as the Sieves or Parasols, *three* Oculist Witnesses (circular diagrams used by oculists to test people's eyesight), *three* roller-drums that support the Chocolate Grinder, which stands on a circular platform, supported by *three L*ouis XV-style legs, and – most prominently of all – the figures called *Nine Malic Moulds*
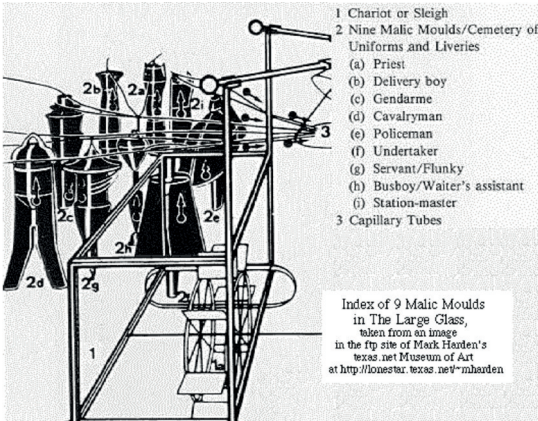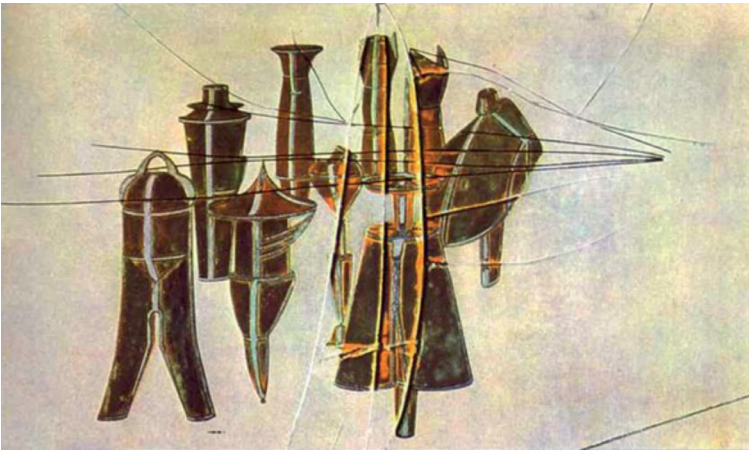
at the left centre-rear of the *Glass*. These are the "bachelors" of the title **[figs 4]**; the group was also known as 'the Cemetery of Uniforms and Liveries' or, because of the paint used, 'red fellows'. Here lead wire is used to 'draw' the forms, which are painted on glass, sealed with lead foil, and presumably, so Duchamp remarks in his notes, filled with "gaz d'éclairage" (illuminating gas) (Sanouillet, Peterson 1975, 51).

Glass proved to be just the right medium for Duchamp's spatial structures. In a note about the *Large Glass*'s composition, he wrote:

> Make a painting on glass so that it has neither front nor back; neither top, nor bottom. (to serve probably as a *three-dimensional physical medium in a 4-dimensional perspective*.) (Duchamp 1983, 67)

Linda Henderson, in her important study of the scientific sources, adds:

> Glass allowed Duchamp to suspend the Bride and her Top Inscription in an indefinite space without clear orientation and without the earthbound quality of the Bachelors below. (Henderson 1998, 81)

The Nine Malic Moulds are remarkable for their equivocal sameness and difference. Made from the same materials in the same way and grouped together, they are a clear-cut unit, a 'family'. Further, all nine 'Moulds' are semi-abstract forms, suggestive without any clear designation, whether of gender, age, or physical appearance. But there are also specifications. In a diagram of the components of the *Large Glass* **[fig. 5]**, Duchamp playfully ascribes the following names to the Nine Malic Moulds (going from left to right): Priest, Delivery Boy, Gendarme, Curassier (cavalryman), Policeman, Undertaker, Flunky (liveried servant), Busboy, Stationmaster. This catalogue is designedly absurd, none of the 'moulds' resembling their given titles. The first on the left, for example, exhibits two legs in trousers, perhaps with a sleeveless vest on top, but the figure also looks like a dress designer's dummy. In either case, no. 1 is far from priestly. Gendarme (no. 3) and Policeman (no. 5) are synonymous characters, but Duchamp's two figures do not look alike: no. 3 has a lantern shape, no. 5 a flag or trophy, whereas no. 4, the Curassier, resembles a bowling pin. Not only do the names fail to define the forms in question; the designations are in no way parallel or in any sort of rational sequence: "Priest" (no. 1) is a vocation: priests may serve in various professions. "Flunky" (no. 7) is a derogatory social designation rather than a profession, and the Curassier (no. 4) has no military colleagues. As for employment status, how does Undertaker (no. 6) relate to Stationmaster (no. 9)?

It is all very arbitrary and yet the group has certain common characteristics; all are 'malic' – male-ish – rather than fully male, which allows Duchamp to create figures like Undertaker (no. 6) and Busboy (no. 8) which could be said to be wearing dresses. None have faces or arms and hands, giving them the look of machine parts or bullets. Further, as Duchamp explains it, "Each of the [...] malic forms is *built* above and below a common horizontal plane, the plane of sex cutting them at the pnt. of sex" (Sanouillet, Peterson 1975, 51). This remark must be taken as tongue-in-cheek because in fact we see no such line of demarcation in the *Large Glass* itself. Rather, the big 'cut' is between the Bride panel and the Bachelors panel, the nine Bachelors being unable to reach the tubing, much less the Milky Way of the Bride up above them. Their family status is thus assured, each figure depending somehow on the others for

completion. One malic mould would be nothing at all; nine make a significant grouping.

When Duchamp later reproduces *Nine Malic Moulds* in miniature for their appearances in his *boîtes en valise*, we recognise them as if they are old friends. Their identity depends upon number as well as form: 9 is 3 x 3 and there are, as I mentioned above, 3 Oculist Witnesses, three Parasols, and three Wheels of the Chocolate Grinder's drum. In astrology, 9 is associated with Mars, the planet of ambition, passion, and aggression, the irony here being that the Malic Moulds are not aggressive or passionate at all; indeed, they are curiously passive. The nine are closer to the Tarot pack of cards, in which the number 9 is that of the Hermit. Since the 'uniforms or hollow liveries' in this 'cemetery' have no arms or hands to touch with, they can only 'hold' the illuminating gas up to the 'planes of flow' above them.

However we interpret Duchamp's composition, the *Nine Malic Moulds* are a perfect example of Wittgensteinian family resemblances. And, as in the case of Wittgenstein, the notion of these resemblances allows Duchamp to play with the concept of *difference* rather than with the similarity between items or with the features of a single isolated work like a geometric figure. A single liquid poured into a number of identical moulds will exhibit minute but significant variations. And even identical twins, Duchamp reminds us, are not entirely alike, thus echoing Wittgenstein's query in the *Investigations* (PI, § 215): "But isn't the same at least *the same*?". "Then are two things the same when they are what *one* thing is? And how am I to apply what the one thing shows me to the case of two things?"

2 Context and Contact

The answer to these pressing questions, as both Wittgenstein and Duchamp understood, had to do with context. Consider Wittgenstein's discussion of the way we use the colour word blue:

> "Is this blue the same as the blue over there? Do you see any difference?"—
> You are mixing paints and you say, "It's hard to get the blue of this sky.
> "It's turning fine, you can already see blue sky again."
> "Note how different these two blues look."
> "Do you see the blue book over there? Bring it here."
> "This blue light means…"
> "What's this blue called?— Is it 'indigo'?" (PI, I, § 33)

And note that these shades of meaning are merely variations at the *denotative* level; if we added the connotations of blue, as in "Am I blue?" or "He's a blue-blood", the list would be much longer. What Wittgenstein is trying to show us is that a single word may have so many possible meanings that we must contextualise and delimit our words as fully as possible. "The meaning of a word is its use in the language" (PI, I, § 43).

The poet, in this scheme of things, is one who understands that the same is never *the same*, and that hence every word, every morpheme and phoneme, and every rhythmic form chosen by the poet makes a difference. To be a poet or artist, in other words, is to draw on the verbal or visual pool we all share but to *choose* one's words and phrases with an eye to unexpected relationships – verbal, visual, sonic – that create a new construct and context – relationships that create what Duchamp termed *inframince* (infrathin) possibilities (see Perloff 2022, esp. ch. 1). When Wittgenstein famously declared that "*Philosophie dürfte man eigentlich nur dichten*" ("Philosophy should actually be written only as a form of poetry") (CV, 28), what he means, I think, is that it is poetry that makes us aware of what language can do and what a difference a single word or phoneme or number can make. Accordingly, the attentive reader must be attuned to difference. Wittgenstein once remarked:

> Hegel seems to me to be always wanting to say that things which look different are really the same. Whereas my interest is on showing that things which look the same are really different. (Drury 1978, 171)

For Duchamp, difference became the basis of the readymade, with its astonishing visual puns. The "assisted readymade" *Fresh Widow* [fig. 6], for example, is a miniature french window, its frame painted an ugly blue-green like that of beach furniture, and its windows' eight glass panes covered with sheets of black leather. By erasing a single letter, **n**, from each word in "fre**n**ch wi**n**dow" the object becomes a *Fresh Widow* – perhaps a recent widow or war widow, but also 'fresh' in the sense of bold, not easy to repress or squelch. What is this widow thinking? We do not know because the leather panes are impenetrable: we cannot see what is behind them. Then, too, the window is closed, and yet those little knobs on the wood 'open' the door, suggesting that perhaps one could see inside![3]

---

3 There is the further joke that no two *Fresh Widows* are quite the same: the leather varies. In the version found at the Chicago Art Institute, there are the outlines of breasts on some of the black leather panes, and so on.
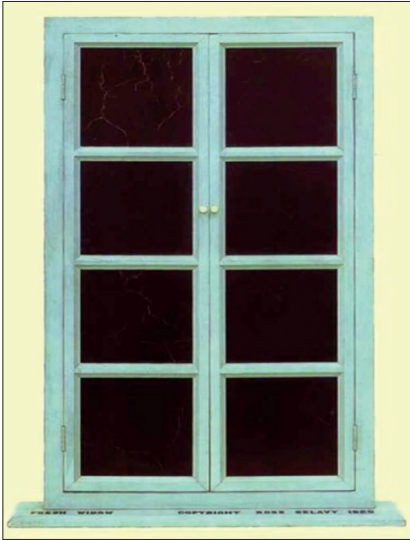
Figure 6
Marcel Duchamp, *Fresh Widow*. 1920.
Miniature window: wood painted blue
and eight rectangles of polished leather.
77.5 × 45 cm on a wooden board,
1.9 × 63.3 ×10.2 cm. Museum of Modern Art:
Bequest of Catharine S. Dreier

Not comparison or generalisation but difference: this, as both Duchamp and Wittgenstein foresaw, from their very different perspectives, would be what is required in the age of social media, where "our craving for generality", "our tendency to look for something in common to all the entities which we commonly subsume under a general term" (BB, 17), dominate the scene. The emphasis on the *infrathin* helps us to look more exactly at what is before us; it allows us to recontextualise the *ordinary*, the everyday. And here again Wittgenstein and Duchamp see eye to eye.

3　All in the Family

> When philosophers use a word — "knowledge", "being", "object", "I", "proposition/sentence", "name"—and try to grasp the *essence* of the thing, one must always ask oneself: is the word ever actually used in this way in the language in which it is at home?—
> What we do is to bring words back from their metaphysical to their everyday use. (PI, I, § 116)

And in line with this distinction:

> The aspects of things that are most important for us are hidden because of their simplicity and familiarity. (One is unable to notice something —because it is always before one's eyes.) (PI, I, § 129)

> Our clear and simple language-games are not preliminary studies
> for a future regimentation of language. [...] Rather, the language
> games stand there as *objects of comparison which*, through sim-
> ilarities and dissimilarities, are meant to throw light on features
> of our language. (PI, I, § 130)

Again, Wittgenstein might be describing the avant-garde ready-
mades of the Marcel Duchamp he never knew – those ordinary ob-
jects brought back, so to speak, from their metaphysical to their
everyday use, and the language games in which these objects
participate.

The readymades – *Bottle Dryer*, *Bicycle Wheel*, *Dog Comb*, *Tzanck
Check*, and of course the famous urinal called *Fountain* **[fig. 7]** – are
often characterised as arbitrarily selected objects regarded as works
of 'art' because Duchamp declared that they were.[4] But the fact is
that the readymades exhibit strong family resemblances: all refer to
manmade industrial products and all relate somehow to the erotic:
think of the bicycle wheel with the rod of the single wheel inside the
hole in the stool beneath it.

When I teach a class on Duchamp and hold up, say, a sock as po-
tential readymade, the students immediately and intuitively insist
that "no, that's not a readymade!". At least not one that belongs to
the Duchamp family.

Revealing family resemblance often means taking the object in
question out of its actual context and putting it in a new one Consider
Duchamp's first American readymade *In Advance of the Broken Arm*
[fig. 8], a snow shovel, with a flat, galvanised iron blade and a wood-
en handle, which Duchamp bought in a hardware store on Columbus
Ave in New York in 1915. As Calvin Tomkins notes:

> There were thousands just like it in hardware stores all over Amer-
> ica, stacked up in advance of the winter storms, or, as Duchamp
> would say in the title that he inscribed on the metal reinforcing
> plate across the business end, *In Advance of the Broken Arm*. Why
> did he choose this particular item? He [...] had never seen a snow
> shovel before, he explained some years later—*they did not make
> such things in France*. [...] Duchamp, after taking it home and sign-
> ing it "[from] Marcel Duchamp 1915" (to show that it was not 'by'
> but simply 'from' the artist), tied a wire to the handle and hung it
> from the ceiling" (Tomkins 2014, 157-8, italics added)

---

4 Perhaps the most authoritative case for this position is that of Thierry de Duve in
his many seminal studies, culminating in de Duve 2023.

Describing his newest readymade in a letter to his sister Suzanne, Duchamp remarked: "Don't try too hard to understand it in the Romantic or Impressionistic or Cubist sense—that has nothing to do with it" (cited in Tomkins 2014, 157).

No doubt the idea of shovel made to remove the snow (and possibly break the arm of the shoveller) was one that Duchamp, newly arrived from France in 1915, found intriguing, and its family resemblance to bottle dryer, urinal, or Chocolate Grinder, must have pleased him. But as in *Fountain*, there is also the parodying of the original context for the object in question. Right about the time, he made *In Advance of the Broken Arm*, Duchamp was organising the Salon of the Independents, held in New York in 1917, on the eve of World War I. This was the famous exhibition where anyone could submit up to two art works for the fee of $ 6 plus a membership fee of $ 1. One of the paintings shown was Henrik Hillblom's *The Making of an American* **[fig. 9]**, which, as it happens, has recently been advertised on E-Bay on a site called Fantasia Antiques. In the ad, the painting was described as follows:

This wonderful oil on canvas painting is ca 1910 and was painted during the first world war. It shows a standing liberty figure, a man, a woman, baby, child, eagle and cornucopia and much more. Note the patriotic influence of Impressionist Childe Hassam [also in the Independents Exhibition], one of Hillbom's compatriots

at The Old Lyme Colony who also painted in the patriotic impressionist style.
This painting measures 30¨H x 24¨W. The colors are marvelous. It has been brought back to its original vibrancy by Page Conservation in Washington DC, restores for the National Gallery of Art. […] It is a real treasure.[5]

And the website copy goes on to tell us about Henrik Hillblom (1863-1948), who was born in Sweden and studied in Paris with Benjamin Constant and Jules Lefebvre. Hillblom

was a member of the Old Lyme Colony School of Artists, gaining its name due to the large number of painters then living in Old Lyme, Connecticut, which became the first major art colony in America to encourage Impressionism. Old Lyme was accessible to its New York City-based painters by excellent rail service.

Duchamp would have relished this delicious description of the Old Lyme School, especially the misdating of this "World War I" painting as belonging to 1910! The tradition of *The Making of an American* is that of Edwin Markham's classically sentimental American poem, *The Man with a Hoe* (1899), which begins:

---

5 http://www.fantasia-antiques.com/Fantasia/hillbom.html/ [URL available until July 2008].

> Bowed by the weight of centuries he leans
> Upon his hoe and gazes on the ground
> The emptiness of ages in his face
> And on his back the burden of the world. (Markham 1899)

The painting of the famed Liberty figure – a secular goddess – silhouetted against the American flag, proffering a huge shovel to the eager man, who is flanked on one side by a young boy, no doubt his son, and on the other by his wife, holding a baby in her shawl, is the quintessential patriotic image of the welcoming of immigrants to the American soil, where a cornucopia of fruits and vegetables (lower right) greets the new worker-to-be. And the title immediately brings to mind Gertrude Stein's *The Making of Americans*, written in Paris between 1906-08.

In this context, Duchamp's *In Advance* can be construed as his own "Making of an American": the snow shovel, rendered useless, providing his own line of work as a new immigrant in the US. Romanticism – Impressionism – more specifically Patriotic Impressionism as it is called on the Fantasia Antique site – the very core of the Independents' 1917 Exhibition – was thus turned inside out.

But there is more at play here than parody. Family resemblance, as Wittgenstein has taught us in his discussion of numbers and language games, is not duplication – a congerie, in this case, of shovels – but rather that resemblance which does not elide the crucial differences within it. As in the case of the *Nine Malic Moulds*, relatedness is not repetition.

It is this concept that Wittgenstein understood so profoundly and made central to his discussion of language games and numbers in the *Investigations*. The meaning of *shovel* is its use in the language. Just as those basic words like *blue* and *read* and *pain* must be understood contextually, so, Duchamp suggests, his own shovel, hanging from the ceiling like a mobile, takes on a very different aura from that in Hendrick Hillblom's *Making of an American*. Its real "family" includes, not hoes or spades or hammers, but the *Bicycle Wheel*, the *Bottle Dryer*, and the *Three Standard Stoppages* – all those pataphysical children of measurement and industry that bear the unique stamp of Duchamp's inventiveness and wit. They are members of the Duchamp family even as Wittgenstein's propositions are part of his. And as witnesses to a Modernist ethos now almost a century old, we readers / viewers can begin to see family resemblances between artists and thinkers who, until recently, were judged to have absolutely nothing in common.

References

Cabanne, P. (1968). *Dialogues with Marcel Duchamp*. Transl. by R. Padgett. New York: Viking.

de Duve, T. (2023). *Duchamp's Telegram*. London: Reaktion.

Drury, M. (1978). "Conversations with Wittgenstein". Rhees, R. (ed.), *Ludwig Wittgenstein: Personal Recollections*. Totowa, NJ: Rowan and Littlefield, 112-90.

Duchamp, M. (1983). *Notes*. Transl. by P. Matisse. Boston: G.K. Hall.

Henderson, A.K. (forthcoming). "Victorian Equations". *Critical Inquiry*.

Henderson, L.D. (1998). *Duchamp in Context: Science and Technology in the "Large Glass" and Related Works*. Princeton: Princeton University Press.

Kuh, K. (1962). "Marcel Duchamp". Kuh, K., *The Artist's Voice: Talks with Seventeen Artists*. New York: Harper & Row.

Markham, E. (1899). *The Man with the Hoe and Other Poems*. New York: Doubleday & McClure (The Project Gutenberg ebook).

Naumann, F.M. (1989). "Marcel Duchamp: A Reconciliation of Opposites". Kuenzli, R.E.; Naumann, F.M. (eds), *Marcel Duchamp: A Reconciliation of Opposites*. Cambridge (MA): MIT Press.

Perloff, M. (2022). *Infrathin: An Experiment in Micropoetics*. Chicago: University of Chicago Press.

Sanouillet, M.; Peterson, E. (eds) (1975). *The Essential Writings of Marcel Duchamp: Marchand du Sel/ Salt Seller*. London: Thames and Hudson.

Tomkins, C. (2014). *Duchamp: A Biography*. New York: MoMa.

Wittgenstein, L. (1960). *The Blue and Brown Books: Preliminary Studies for the "Philosophical Investigations"*. Transl. by R. Rhees. New York: Harper. [BB]

Wittgenstein, L. (1980). *Wittgenstein's Lectures: Cambridge 1930-32*. From the notes of J. King and D. Lee. Chicago: University of Chicago Press. [WLC]

Wittgenstein, L. (1992). *Tractatus Logico-Philosophicus*. Transl. by B.F. McGuinness; D.F. Pears. London; New York: Routledge. [TLP]

Wittgenstein, L. (1998). *Culture and Value*. Edited by G.H. von Wright; H. Nyman, revised ed. by A. Pichler, transl. by P. Winch. Oxford: Blackwell. [CV]

Wittgenstein, L. (2022). *Private Notebooks 1914-16*. Transl. by M. Perloff. New York: Liveright. [PN]

The Bridge from Language to Mind: PI, §§ 240-56

Meredith Williams
Johns Hopkins University

Abstract This article is an examination of a remarkable set of 16 passages that mediate Wittgenstein's treatment of language in the *Philosophical Investigations* prior to these remarks and his turn to mind, beginning with his two private language arguments of PI, § 258 and PI, § 293. The intervening remarks, PI, §§ 240-6, function to identify six central lessons concerning the nature of language that are essential to mind as well. These six stages are linked by what Wittgenstein considers the two fundamental philosophical problems for language, and now for mind: the Problem of Reference (or Identity) (PI, § 239) and the Problem of the Criterion of Identity. The general conclusion is two-fold: ordinary language is necessary to the human mind; and neither are reducible nor eliminable in favour of literally inner events.

Keywords Wittgenstein. Language. Mind. Reference. Criterion of Identity.

Summary 1 Introduction 2 PI, §§ 240-1. Framework of Language: Truth. – 3 PI, § 242. Form of Language: Measurement. – 4 PI, § 243. Interpretation: Three Voices. – 5 PI, § 244. Problem of Reference: Private Naming. – 6 PI, §§ 245-9. Problem of Identity: Possibility and Necessity. – 7 PI, §§ 253-7. Criterion of Identity: Logical Grammar. – 8 Concluding Remarks.

1   Introduction

This article is an examination of a remarkable set of 16 passages that mediate Wittgenstein's treatment of language in the *Philosophical Investigations* prior to these remarks and his turn to mind, beginning with his two private language arguments of PI, § 258 and PI, § 293. The intervening remarks, PI, §§ 240-6, function to identify six central lessons concerning the nature of language that are essential to mind as well. These six stages are linked by what Wittgenstein considers the two fundamental philosophical problems for language, and now for mind: the Problem of Reference (or Identity) (PI, § 239) and the Problem of the Criterion of Identity. The general conclusion is twofold: Ordinary language is necessary to the human mind; and neither are reducible nor eliminable in favour of literally inner events.

It is well known that Wittgenstein binds the case for his positive views with the philosophical theories he seeks to eliminate. This is certainly true of his treatment of mind. The target here is the picture of the individual mind as the inner arena of objects and events that are private and knowable directly only by that individual mind or self. Their metaphysical and epistemological properties mark them off from ordinary physical objects and events. These include sensations, perception, imagination, intentions, belief, thought among other mental states. We owe the classical account of this picture to Descartes' theory of mind. Though there are many who hold that the inner mental arena is the brain, this is typically described as the mind-brain in an effort to forestall the problems that arise with attempts to identify mental states either directly or indirectly with brain states or functional neural roles. For others, the computer is the arena of mental activity, taken as systems of representations manipulated in accordance with algorithm and/or other formal structure. But it is our ordinary ways of attributing and explaining our actions and mental states are the indispensable housing for both the mind-brain and computer that creates the illusion that our sensations, intentions, imaginings and so on are actually 'in there'.

Wittgenstein does not repudiate the relevance of neural activity to the functioning of mind and body, but it is not mental activity. He would acknowledge what a computer contributes to our world, but he would deny that its inner Turing machine is a mind. What he would hold, in this philosophical context, both of the brain and the computer is that our ordinary ways of talking and acting are the cocoon within which each is thought to house the mind. The cocoon is the necessary projection of our ordinary ways of attributing and explaining our actions and mental states onto the brain and the computer that creates the illusion that our sensations, intentions, imaginings and so on are actually 'in there'.

The point of this paper is not to take on these reductionist accounts of mind, but to look carefully at a remarkable set of 16 passages that mediate Wittgenstein's treatment of language in the *Philosophical Investigations* prior to these remarks and his turn to mind, beginning with the two private language arguments of PI, § 258 and PI, § 293. His overall position is that language, ordinary mental language, is integral to our mental life. The intervening passages, PI, §§ 240-56, function to identify the central lessons concerning the nature of language that Wittgenstein has already defended. These identify six major features that belong not only to language, but to mind as well. It is for this reason I call the set of 16 passages 'The Bridge Passages'. They take us from the lessons of language to a defence of Wittgenstein's conception of mind.

The Bridge is bookended by PI, § 239 and PI, § 257. Together these introduce what Wittgenstein takes as the two fundamental philosophical problems for language, and now for mind: the Problem of Reference (or identity) and the Problem of the Criterion of Identity, or, as I prefer to express it, the Problem of Normative Similarity. Wittgenstein marks the beginning of the Bridge by asking what a colour word means to an individual person:

> How does he know which colour he is to pick out when he hears "red"?—Quite simple: he is to take the colour whose image occurs to him when he hears the word.—But how is he to know which colour it is 'whose image occurs to him?' Is a further criterion needed for that?... "'Red' means the colour that occurs to me when I hear the word 'red'—would be a *definition*. Not an explanation of *what it is* to use a word as a name". (PI, § 239)

For the cartesian, what 'red' means to me would be given by an ostensive definition of the word. But that definition cannot by itself determine whether the colour I experience answers to the word 'red', whether the colour I experience is the colour to call 'red'. What is needed is an explanation of '*what it is* to use a word as a name'. The core problem of any version of mind as interior or inner is that it cannot explain the meaning or meaningfulness of mental states, e.g., that this state is red. I shall call any such inner theory of mind a 'cartesian theory'. Any cartesian theory hypothesises that mind is an interior system of episodes and events. The price of this achievement is the elimination of meaning or content from the interior system. The major contemporary cartesian accounts are formal logical theories, syntactic theory, computational theories, and mind-brain theories. Any of these might be important contributors to our understanding of human life, but they are not theories of mental states. They use mental concepts for constructing their hypotheses, but in doing this, they do not replace mind with their hypothesised systems.

The second bookend addresses the problem of the *criterion* for identity, or normative similarity. An answer to this philosophical question must provide 'an explanation of *what it is to use* a word':

> [...] But what does it mean to say that he has 'named his pain'?—How has he done this naming of pain?! And whatever he did, what was its purpose?—When one says "He gave a name to his sensation" one forgets that a great deal of stage-setting in the language is presupposed if the mere act of naming is to make sense. And when we speak of someone's having given a name to pain, what is presupposed is the existence of the grammar of the word "pain"; it shews the post where the new word is stationed. (PI, § 257)

Certain philosophical tropes have come into play that support the notion that ostensive definition or naming can be used, not only to provide a definition for a word, but also an explanation of how the word is to be used. The latter is possible because we know, individually and privately, the colour-image we have or the pain we experience. Knowing one's conscious states provides the criteria for use. Wittgenstein's idea of a logical grammar is thus unnecessary. It is at best, a critic might hold, the emergence of what is secondary to the primary foundations of individual knowledge and meaning. Wittgenstein's explanatory reversal of the picture of the cartesian mind and that of logical grammar is the primary target of the six stages of the Bridge. As he says in PI, § 257 just before introducing the private diary argument, giving a name to pain presupposes the grammar of the word 'pain'. What the Bridge provides is a summary of the arguments that take us from inner naming of objects and events as the fundamental semantic capacity of the mind to the logical grammar of mental concepts, realised through the use or role of mental terms within language games.

Wittgenstein's aim is thus to show that the cartesian theory, though a revolutionary picture of the mind, linked to the seventeenth century Scientific Revolution, persists into present-day conceptions and theories of mind. And yet the cartesian theory cannot replace our ordinary concepts of mental activity either methodologically or explanatorily. Wittgenstein's method is to link powerful criticism of philosophical theories of mind as an inner arena to his presentation of an alternative conception of mind, which is the one we all work with ordinarily.[1] The kinds of criticism Wittgenstein develops open the way themselves to Wittgenstein's logical grammar picture of mental activity. The critique points to the human mind as systemically informed by *ordinary* language.

---

[1] For a fuller discussion of Wittgenstein's method used in the *Philosophical Investigations*, see Williams 2010, ch. 1.

Wittgenstein's conception of language as having a logic is not a formal logic, but a logical grammar. Mind too is thus informed by logical grammar. The Bridge passages provide reminders of the elements of logical grammar that are as essential to mind as they are constitutive of language. The idea that language has the structure of a formal logic like the propositional calculus, though intended to explain the systematic features of speech and rational activity, *eliminates* what Wittgenstein calls our form of life. In fact, our form of life is the indispensable background to the use of language of any kind –

> to imagine a language means to imagine a form of life. (PI, § 19)

> the speaking of language is part of an activity, of a form of life. (PI, § 23)

– and so too our human minds:

> Can only those hope who can talk? Only those who have mastered the use of a language. That is to say, the manifestations of hope are modifications of this complicated form of life. (PPF, § 1)

> "Grief" describes a pattern which recurs, with different variations, in the tapestry of life.[2] (PPF, § 2)

The aim of Wittgenstein's later work is to recall and restore that human form of life to our understanding of mind. To do this, he develops his own method of argumentation and persuasion. To find ultimately the certainties that bring that form of life into focus once more, Wittgenstein constructs simple language games that involve our form of life even as they are used to diagnose and criticise the philosophical theories other philosophers endorse. The language games are simple pictures that identify the essence of misleading philosophical theories though without the typical use of technical vocabulary. Reference and the criterion of identity are the two great problems that Wittgenstein seeks to understand. The opening passage of the *Philosophical Investigations* gives us what we need to understand these two problems. The first is the grocers' language-game (PI, § 1). A orders five red apples from the grocer by saying: "I want five red apples".[3] Wittgenstein shows what the grocer is to *do* in order to comply with this request. He must proceed in different ways in order to

---

[2]   Passages like these are especially emphasised by Cavell 1979.

[3]   In conversation, Michael Williams has repeatedly urged that the opening grocer's language-game has all the key features that Wittgenstein aims to establish in the *Investigations*.

respect each element of the order. Apples are to be found in a bin la-
belled "apples"; "red" apples are to be compared with a colour chart;
and "five" apples require counting "1-2-3-4-5 stop" red apples. The
uses of the words occurring in the request are what give the words
meaning. This requires underlying mastery of techniques that con-
sist not only in the utterance of these words narrowly construed, but
in the relation between the buyer and the grocer and the actions that
are appropriate within a greengrocer's store. Language games are an
integral part of his method for both displaying his picture of language
and using it as a critical tool. The Bridge passages are no exception.

The Bridge passages consist of six stages, listed here with in-
dications of where similar discussions take place earlier in the
Investigations:

- PI, §§ 240-1. Framework of Language: Truth [PI, §§ 136-7]
- PI, § 242. Form of Language: Measurement [PI, §§ 49-50]
- PI, § 243. Interpretation: Three Voices [PI, §§ 201-2]
- PI, § 244. Problem of Reference: Private Naming [PI, § 2 ff.]
- PI, §§ 245-9. Problem of Identity: Possibilities and Necessity
  [PI, §§ 90-104]
- PI, §§ 250-6. Criterion of Identity: Logical Grammar [PI, § 253,
  § 288, §§ 370-3]

These six stages have been discussed in connection with language in
earlier passages of the *Investigations*. The point here is not to *estab-
lish* these points again, but to examine mental concepts and words
in the light of these earlier arguments. The reason for this repeti-
tion is to inform the reader that the six stages are a preparation for
Wittgenstein's picture of the *human* mind and its episodes, function-
ing and activities, all of which are informed by our mastery of lan-
guage. It does not provide an empirical theory of mind and its capac-
ities and functioning. But it does provide the logical grammar that is
indispensable for the human mind. Mental states and mental func-
tioning are thus not eliminable in favour of physical states nor are
they reducible to physical states or functions nor are they identical
with such states. Mental states have their own logical structuring
and functioning, but only as states and events of individual human
beings as members of a community. Never to have had a communi-
ty is never to have had a full human mind. To lose one's community
is to live in perpetual grief, a virtual retaining of that lost commu-
nity. Again, 'logical structure' is not that of a formal logic, say that
of Frege or the *Tractatus*, nor of a computational system. The rele-
vant logical structure is that of logical grammar, which will be dis-
cussed more fully below.

2   PI, §§ 240-1. Framework of Language: Truth

Wittgenstein begins with a few remarks that identify essential features of language: agreement, truth, framework, and form of life.[4] As we have already seen, the need for the bridge is introduced by the classic philosophical problem, how does an individual know *what* colour is referred to by a particular colour word in a language (cf. PI, § 239). Wittgenstein presents this problem as a problem of reference (or ostensive definition) when it must be solved by the individual alone: "How is [an individual] to know what colour he is to pick out when he hears 'red'?". Left to the individual alone there is no way for that individual to pick out knowingly the correct colour, namely, that the colour he experiences now is called 'red' in English. PI, § 240 begins by identifying agreement among speakers as essential to following the rules of language. The agreement that is essential to us all as language users is not that of our opinions or hypotheses or speculations. Where disagreement is common to political debate or whether chocolate tastes better than raspberry, Wittgenstein contrasts these areas of speech with that of mathematicians. Mathematics is an exemplar of agreement precisely because the rules of mathematics, of *how to do* mathematics, do not tolerate disagreement. The rules of counting require and receive complete agreement. The application of mathematical procedures is a necessary part of mathematics itself. Agreement among participants following a rule is essential because it, the agreement, belongs to the framework/scaffolding of language that fixes the application of the rule that underwrites its necessity. Whether we are teaching a novice the colour palette or building a bridge or discussing a film, agreement in the tacit rules of each project is taken for granted by all discussants. Agreement in application belongs to the framework/scaffolding of language and so stands independent of the opinions or hypotheses entertained by the participants of the game or project. Language must have rules to which we are blind in our shared behavioural respect for these rules or norms. We are unaware for the most of that which secures our agreement. We may not even see ourselves as in agreement with our interlocutors. This raises the question, what secures the scaffolding? Is it the tacit agreement? Or is it something else?

The importance Wittgenstein assigns agreement can be met with the objection that, if agreement is a necessary part of the framework

---

4   In *On Certainty*, Wittgenstein modifies the notion of a framework or scaffolding for language in his discussion of the structure of belief. There he says that "[t]he truth of certain empirical propositions belongs to our frame of reference" (OC, § 83). The expression "frame of reference" is the translation of the German word *Bezugssystem*. But, as I shall discuss later, this German word is better translated as 'coordinate system' or 'axial system'. It gives us a far better understanding of what Wittgenstein has in mind.

for language, it looks as though "[…] human agreement decides what is true and what is false" (PI, § 241), and this cannot be acceptable. What is true or false must be a matter of reality, of the way things are, whether or not we agree. Wittgenstein responds to this objection by holding that "[w]hat is true or false is what human beings *say*" (PI, § 241). By this he means, once again, that our opinions or our hypotheses, our assertions that we make are true or false, and are subject to epistemic principles of evaluation. But even when engaging in disagreement and discord in the opinions we hold, we nonetheless agree in the language we use. 'Language', as Wittgenstein uses it here, is agreement in form of life. This distinction between what can be true or false and truth, although it evolves in his writing, most particularly in *On Certainty*, is one that he never gives up. An important passage in *On Certainty* is the following:

> The truth of certain empirical propositions belongs to our frame of reference. (OC, § 83)

Two important points are made here: first, certain *apparently* empirical propositions belong to our frame of reference, our form of life; and second, these propositions are not true or false like ordinary empirical propositions; they belong to truth. Here Wittgenstein distinguishes between that which is true or false, and so subject to possible dispute, and truth, which is our frame of reference, our logical grammar and form of life. Putting those two points aside for the moment, we can see that Wittgenstein is committed to the distinction between the *uses* of language as applied in different situations or language games; and the logical grammar of language games in terms of how they are used given our form of life. Agreement must exist at the level of logical grammar; otherwise agreement as to what is true and what is false is unreachable. What is language at the level of logical grammar? This is the level of truth, agreement, and form of life: the framework that is necessary for the functioning of our language games, whether political, mathematical, culinary, or any other language game.

3 PI, § 242. Form of Language: Measurement

Given the necessity of agreement within "the framework on which the working of our language is based", it may seem "to abolish logic". Not only must we agree on definitions of words (like names), we also must agree on how to use words. A form of life is a way of acting and engaging with others in relation to the world. The human form of life is a non-reflective or blind way of acting through and with language in relation to others and to the world. This involves judging

with others, counting as others do, and virtually any other use of language. The grocer must choose the fruit from the 'apple' bin, compare colours with the colour of the apple, and he must count out '5' apples, 1, 2, 3, 4, 5 STOP. The procedures used to apply to the individual words of the order are distinct. Merely naming the objects indicated would not enable the grocer to fulfil the order. He must not only know the object each of the three words denotes, he must know how the words are to be used. The former associations might be called 'definitions', but that could be determined only through the proper use of those words in making judgments. If we cannot construct assertions out of those words, we do abolish logic from language. That is to say, we abolish language as assertion or judgment in favour of stimulus-response couplets. So, how are use and rules of use to be added to names, which otherwise are mere words in a list?

Agreement must extend not only to definitions but to judgments as well. We need to be able to say not only 'red' in the presence of red, but also 'roses are red'. In other words, behind the problem of naming is the problem of the unity of the proposition or judgment. Just as the mere association of the word 'red' with a colour is problematic, so a particular string of words need not establish the connection of predication of a subject. What these problems clearly indicate is the need for a very different model of how names and use of names relate. This is what Wittgenstein provides in PI, § 242, using the metaphor of measurement:

> It is one thing to describe methods of measurement, and another to obtain and state results of measurement. But what we call "measuring" is partly determined by a certain constancy in the results of measurement. (PI, § 242)

At this stage of the crossing, we have established that agreement is essential to language on the grounds that without agreement, communication and truth are impossible. This requires that language must have two levels, the first is our ordinary linguistic interactions, opinions, hypotheses and the like: the use of language is volatile, in dispute, and changeable. The second is language as a foundational framework (or scaffolding) which is the base for 'the workings of our language', for which agreement is essential, and there is no room for falsehood. All judgments of the second level are part of truth. Truth is a necessary part of the foundation for falsehood, error, and disagreement as well as what is subject to being true or false.[5] Language in this sense is agreement in form of life. The two levels of language

---

[5]   I shall turn to this two-fold level of language when I focus on the structure of belief, which will focus on Parts 2-3 of *On Certainty*.

are separated for analytic purposes. They 'live' inseparably in our lives. The distinction does not answer the two problems we face: naming and unity of judgment and thought.

It is with these problems in mind that Wittgenstein introduces one of his important metaphors for language and so for the idea of language being embedded in a form of life. Language is a form of measurement (PI, § 242). To provide a clearer picture of this idea, we will first consider a far simpler case, namely, that of the standard meter stick stored in a vault in Paris (PI, § 50).[6] The meter stick is the rule or standard for the meter length. The stick is neither one meter long nor not one meter long: it is the standard that fixes the one meter length. There are three components of such a standard, and all are necessary. The first is a physical element. Something physical is necessary; moreover, it must be rigid and immune to easy deterioration. The bar or stick situated in the vault is made of steel, which is rigid and protected from material processes that might undermine the integrity of the stick. Secondly, though the stick is hard and rigid, it must be amenable to calibration, that physical marks can be engraved on the surface of the stick that regulate sublengths. It must be accepted socially with a shared understanding of the procedures for using it. The meter stick is a hard and rigid bar that can be calibrated for the shared use as the standard for one meter. Once calibrated and designated as 'one meter', the metal bar fixes, in a permanent and unchangeable way, the length that is one meter.[7] This standard or norm is protected in its role as one meter by placing the calibrated metal bar into the sealed vault.

The metal bar designated 'one meter' is thus a means of representation; it is not an object itself being represented. Using the metal bar to measure meter lengths in the world is to judge the world in relation to this standard. Using this standard successfully requires a rich physical and cultural domain in which to act. A similar strategy can be used for fixing shades of colour. Such fixed colour shades are, for Wittgenstein, also a means of representation and not that which is represented (PI, § 48, §§ 50-1). Saying the 'standard meter stick' is a means of representation is a special way of saying that it is a name. The difference is, as the representation of one meter, it acquired that 'name' through its having been assigned its role to play. Names are static, simply attached to some object. Good enough for saying of that object that it is a such-and-such, and nothing more.

---

6   The item that is used to measure meter length has undergone several changes since the first introduction of the meter stick which was placed in a vault in Paris. But this fact is irrelevant, as the reader shall see, to the point that Wittgenstein is making in the *Investigations*.

7   It is a mistake to think that the replacement of the Parisian standard meter with other devises in any way makes a difference to the philosophical point being made here.

An exception to this is the name of a person, one who can claim the name for him or herself.

Language too is a measuring device to be used for certain purposes. If we carry out the analogy with the standard meter stick, we must ask what counts as the calibration and use of language *qua* measurement. "Agreement in judgments [...] is required for communication by means of language" (PI, § 242) just as the rigidity of the metal stick is required of the standard for one meter. Judgments calibrate language. As Wittgenstein says elsewhere "judgments can be standards for judging". But this then also "seems to abolish logic, but does not do so" (PI, § 242). Judgments as standards for judging seem to abolish logic since they do not have the *kind* of objectivity that is assigned assertoric sentences or propositions. Propositions are the form of assertoric sentences that enable them to stand in formal relations to other propositions, without regard to their meaning: assertion, negation, conjunction, and the conditional. Judgments are evaluated for nuance, sensitivity, appropriateness, correctness, wisdom, and other normative properties. One cannot identify a judgment in purely formal terms. It is critical to their use that are meaningful. Yet Wittgenstein tells us that judgment does not abolish logic. What can this mean? There is an ambiguity here. It can mean that the sentences that comprise judgments, like the metal bar of the standard meter stick, can be assigned logical properties – calibrations – that fix the propositions into a system of formal relations of same as, shorter than, longer that. These are identified solely in terms of the calibrations being identified with marks along the length of the stick. This creates a kind of formal calculus. Calibrations of length, like propositions, belong to a calculus. The calculus provides the procedures of use of the propositions or calibrations. Judgments as standards are not rigid in these ways. They do require a 'certain constancy' in their use in the world. So judgments as standards are embedded in the world as are meter sticks. Judgments are calibrated much in the way that the assertoric sentences that express propositions are calibrated, by way of the words used in constituting a sentence. But the use of this calibration is, however, not identical. The meanings of sentences that belong to the propositional calculus, that is, propositions, are given in terms of truth conditions. The meanings of judgments, on the contrary, are embedded in the social world as the procedures by which we act correctly or rationally or wisely. The use provides the meanings or 'methods of measurement' which, when so used, enable us to 'obtain and state results of measurement'.

This then is Wittgenstein's summation of those features of language, discussed and defended in detail earlier in the *Investigations*, that have significance for the discussion of mind. The method of measurement shows the difference between the picture of proposition and the picture of judgment or thought. The analysis of propositions

presents them as having an assertoric form of the subject-predicate sort whose semantics corresponds to truth conditions as fact-stating conditions. The picture of such assertions, or propositions, shows their place in a logical web of formal relationships among propositions. According to the Tractarian interpretation, the assertoric sentence 'the dress is blue' shows that the object named by 'the dress' and the object named by 'blue' relate as the fact that the dress is blue. A central difficulty for this account is how the two objects relate in a single fact. This problem must be solved before we can understand the relations among truth-conditions that accord with the truth-preserving operators of the propositional calculus, namely, negation, conjunction, disjunction, conditionals. The world is understood in a fact-stating way, as having a linguistic or assertoric structure itself.[8] This is better understood as the world having a factual structure that harmonises with the assertoric structure of the calibration of language. Agreement is possible, then, so long as we attempt to state facts about the world. Nothing else can be seen through language though the world may be able to affect us in ways without any awareness on our part.

Wittgenstein's picture of judgment, on the other hand, removes the dominant role assigned names and naming; and replaces the strictures of the propositional calculus with rules of judging. There may be judging that is not fact-stating as when we make normative ethical, aesthetic, or intentional judgments by which to measure human actions and reactions. Agreement underwrites our capacity to understand the world and human minds. On this picture, it might seem that there is no way that a solitary mind could be understood. It is at this point (PI, § 243) that Wittgenstein refines the notion of agreement by considering three cases of possible monologists.

4 PI, § 243. Interpretation: Three Voices

There are three voices[9] that can be heard in this passage, not counting Wittgenstein's voice as the moderator for the other three. Each one imagines "human beings who spoke only in monologue; who accompanied their activities by talking to themselves". His aim is to identify

---

[8] This is, of course, a highly truncated account of Wittgenstein's picture theory of meaning. It is meant only to highlight the fact that meaningful language to the factual structure of reality. Meaningful language is given only in fact stating sentences 'propositions'; and truth only obtains with object-related states of affairs. See TLP, sections 1-3 and 4.2-4.5.

[9] I am following David Stern's (2004) recommendation to identify the distinct voices (or philosophical positions) that 'name' positions in the discussion of a passage (or set of related passages). It is an excellent device for following the dialectic movement within the passage(s).

three distinct cases of monologic speech of which the cartesian monologist is the third. This longer discussion enables us to get a clear picture of the relation each such form of language has to the world and other people. I shall represent each monologic language with the name of a philosopher placed in scare quotes. The first is 'Wittgenstein' speaking for the conception of ordinary language he has introduced at the beginning of the Bridge. The second is 'Quine'[10] speaking as an explorer who aims to translate a monologist who speaks a radically foreign language; and finally 'Descartes' speaks for a private language of sensation.[11] The point of drawing these distinctions is to isolate just what is clearly unique to a cartesian monologist. The Bridge lays out the issues that must be discussed in detail, thus preparing the way for Wittgenstein's own conception of the mind.[12]

All three voices aim to understand those who speak only in a monologue. Yet at the beginning of the Bridge, Wittgenstein emphasised that background agreement is a condition of meaningful language. Is not all monologistic speech therefore meaningless? In restricting the use of language to self-talk, how is agreement achieved? First, what would it be for one of us to speak to himself or herself? asked

---

10   Quine (1960) introduces his own 'explorer' who seeks to translate the language of a wholly alien language. The point is to identify the elements that are necessary to such translation. In this case the native says: "Gavagai". Quine's simple language game requires repeated use of 'gavagai', some common or similar object within the environment, and the capacity to say 'yes' and 'no' in gavagai-language.

11   See Stephen Mulhall's discussion of this passage in his interesting book *Wittgenstein's Private Language*. There he argues that passage 243 is more complicated than I have presented here. In particular, he holds that there is a debate among the monologists over whether the correct interpretation is substantive or resolute. Indeed his chapter ends with Mulhall asking: "Can we explorers of [Wittgenstein's] texts ever really succeed in translating his language into ours?" (Mulhall 2006, 22). The implication of this question is that Wittgenstein is himself a private monologist. Yet Wittgenstein says of one who holds that only he knows his sensations that "in one way this is false. And in another this is nonsense" (PI, § 246). Certainly treating Wittgenstein' words as Mulhall recommends will guarantee that he is speaking nonsense.

12   Mulhall (2006) and I have both recognised the importance of PI passages 243 to 293 with a new special emphasis on § 243 to § 255. Mulhall's excellent treatment of these passages was published before my own and there is some overlap in how they are interpreted, I would say necessarily so. One of the most important differences in the two interpretations is the structure of these passages and their relation to the private language arguments. Where Mulhall treats them as flatly continuous, I show that there is important structure in the occurrence of individual passages and their relation to the private language arguments. They are not meant to be continuous. First there is a parallel only hinted at here to the order in which the passages 240 to 253 occurs and Descartes' application of the method of doubt. Second, three monologists introduce the problem of reference and how it can be solved by Wittgenstein's conception of language. The problem of the criterion of identity follows that of reference as it has occurred in earlier discussions of reference and identity in the Bridge. A treatment of illness is called for and that takes us to the private languages arguments and a preparation for a full analysis of mind. This is why this is a treatment of language bridged to an understanding of mind.

'Wittgenstein'. There is no difficulty here, according to 'Wittgenstein'. An individual can "encourage himself, give himself orders, obey, blame and punish himself". But these acts do not require the individual to relinquish agreement in our form of life and the way that the self-talker uses words in his self-talk: that is the same. Self-talk does not entail that he means something different by 'order' or 'blame and punish'. On the contrary, it is important that these words retain their ordinary meaning when he applies them to himself alone. Were the subject to speak openly about these matters, there would be no difficulty for another, who spoke the same language, to understand him. Otherwise, if he could not be understood by another, he would be muttering nonsense to himself. Ordinary self-talk carries no interesting philosophical consequences. It is just a person talking to himself.

But what would it be for someone to speak a very different language from our own such that we could not hear the vocalisations as meaning anything? And though his actions seem to be related in regular ways to his environment, there is no community visible of like speakers of the alien language. Wittgenstein argues that not even in these more extreme situations is the intelligibility of the speaker impossible to understand. 'Quine' develops a distinctive thought-experiment that is comparable to Wittgenstein's language-game of the explorer who comes upon a person, one who is alone and speaks only to himself (see Quine 1960, ch. 2). The explorer does not recognise the vocables and so must find some technique for rendering the vocables intelligible by relating them to salient objects in the environment, objects assumed to be visible to both native and the explorer. This is part of the form of life we human beings share. 'Quine' must go much closer to the native at which point he can hear the speaker calling out 'gavagai' from time to time. With this he has a concrete problem to solve: how to translate *gavagai* into some English word or phrase. How is this to be done? 'Quine' adopts the simplest way available. He makes himself known to the speaker, and then, pointing to a salient object, asks: "Gavagai?". This strategy requires that 'Quine' not only knows the apparent word 'gavagai' but also the words for 'yes' and 'no'. In this case, he has discovered somehow that 'yes' is *evok* and 'no' is *yok*. As a methodological principle, he assumes that certain kinds of salient objects are most likely to be named by the native. 'Quine' therefore names a living moving animal. He picks out a rabbit hopping by and calls out: "Gavagai!". The native speaker calls back: "Evok!". 'Quine' treats this as supporting the hypothesis that *gavagai* means 'rabbit'. In other words, he must make a large number of tacit assumptions that strike him as obvious objects for any human being to see. If a living moving animal, it is also taken to be obvious how it acts in the environment. This and much more is taken by 'Quine', the native speaker, and most other human beings as obvious, so obvious it need not be spoken.

Of course, this is not what the real Quine concludes on the basis of this thought experiment. He rather thinks that it shows that meaning is indeterminate, that meaning is not obvious. There are many other words and phrases that are extensionally equivalent to the word 'rabbit'. For example, 'undetached rabbit parts'. This phrase is extensionally the same as for the word 'rabbit', but the two terms do not mean the same thing. But for the purposes of this thought experiment, there exist a methodology for translating the words of an unknown language into one that is known. Even an unknown language apparently spoken by a single individual can be translated. This is because the native is a human being who shares our susceptibilities, basic desires and interest in certain living things. It is our shared human form of life that enables the explorer to form hypotheses that have a high probability of being relevant to the native's interests as well as to our own. It is at this level that Wittgenstein finds the agreement that is necessary for translation and meaningful language.

Quine's own thought that the native might mean 'undetached rabbit parts' by 'gavagai' tries to make use of extension as the basic principle of identity. But this clearly will fail in this context. First, the phrase 'undetached rabbit parts' can only be taken to be identical with 'rabbit' and 'gavagai' if the elements of the English phrase mean the same in English in which case there can be no identity of meaning. Quine accepts this point: indeed it is crucial to his argument for the indeterminacy of *meaning*. Second, the identity of the object that is secured through the use of the phrase and words is the living hopping animal. 'Living hopping animal' is not to be understood as yet another descriptive phrase, but as the animal itself. The animal itself is the extension of the phrases and words. Unlike the phrases and words, the animal itself does not mean anything. It just exists. But as an existent object, it can be referred to. The argument uses the two semantic values of words and phrases, which Frege calls sense and reference and which we are calling meaning and reference (see Frege 1997). Translation exploits objects as referents of meaningful words. To do this requires constraints on which objects are relevant to reference in the situation; and the existence of regularities in the linguistic activity of the human being under observation. The explorer must recognise repeated vocalisations co-occurring with the presence of an object in the environment. The key presumption is that objects as possible referents are shared though the names need not be shared; and that the vocable of the native shares reference with the word of the explorer. This means that referent and sense are separable, and must be for Quine's argument to work. What are needed then are hypotheses that link referent, sense, and object, all separable, together under a single banner-word; or in the case of translation under two organising banner-words, 'rabbit' and 'gavagai'.

Having considered these first two cases of monologism, Wittgenstein is enabled to identify what is special and different about the language of the cartesian thinking self. Passages 240-2 emphasised two important points about language. First, agreement as part of shared social background is necessary for language; and, second, language as "methods of measurement" requires "a certain constancy in results of measurement". In short, language use by human beings requires basic agreement that is a function of being human living in the world and displaying regularity in what they do, in their effects on the physical world and other speakers and vice versa. Meanings are not objects that can be separated from the human form of life. There is no separable meaning per se to be grasped. These requirements, it might be thought, should show that monologues could not be meaningful. The real Quine, as noted above, argued that these different words and phrases do not mean the same thing. Both claims would be a mistake. We saw of ordinary self-talk and self-talk that is unrecognisable fall foul of these claims by looking in the wrong place. Once we understand what 'Wittgenstein' and 'Quine' are saying, there is no conflict between the general requirements of language and the monological speech of the single speaker. Yet Wittgenstein's explorer cannot have access to the native's 'yes' and 'no'. The whole language belongs to the native alone. Suppose that 'Quine' is confronted with *gavagai evok*. Can it be translated? *Evok* must be 'yes' if it is to be translated. 'Quine' translates *gavagai* = 'rabbit' where there has occurred many rabbits coordinated with many tokens of 'gavagai'. But if a duck came by and the native said to himself 'gavagai evok', 'Quine' would have to say "yok" meaning 'no gavagai'. Whereas *evok* is tied to concrete cases that are similar, *yok* is tied to possible cases of no similarity. So the native would have to solve the problem of the criterion of similarity. In his monological language, the native must be able to recognise that many hopping rabbits are all similar and so all *gavagai* whereas the passing duck, badger, and are all dissimilar. The 'evok' group and the 'yok' group are dissimilar, so what exactly is the criterion for the sameness of identity? How does the single native solve that problem? If we, having a shared language, solve the problem of the criterion of similarity and possibility, there seems to be no special problem for the native who speaks only to himself provided what he says is translatable by another. Quine's requirements for translation would satisfy Wittgenstein's hypothesis that such a monologue could be shared and understood. The translator would take himself to be the arbiter of similarity, a position he acquired only by being raised in a language speaking community. To impose extension as the determinant of meaning remains to be discussed, though the pressure is great already on its direct relevance to the meanings of our language.

But are there any languages like this that support a single speaker whose words cannot be translated? 'Descartes' presses this problem further by asking if such a self-directed language could be used for our inner sensations and feelings. We do this using our ordinary language without difficulty. But 'Descartes' replies that this is not what he means when he speaks of a monologue. He means a language 'that only the speaker can know [...]. So another person cannot understand the language'. Now we have the philosophical problem identified. This is a language that only one person can use; that cannot be translated; that cannot put ordinary language to a private use. This is a private language. The meaning of this language is given by what the words refer to, and what they refer to are 'immediate private sensations'.

The privacy of this language is inherited by the privacy of our inner sensations. Surely, it is held that no one can know the private inner sensations of another. What strategy could Wittgenstein use to reject such a language of inner sensations first, and inner thoughts secondly? One might think that the Quinean method of translation might enable our explorer to grasp the inner experience of the monologist. The structure of shared linguistic reference is triangulation (see Davidson 2001). To take a paradigm of this, consider an event of a child just learning the word 'table'. The child and adult form two corners of an abstract horizontal bottom line. At the apex above this horizontal line is a third object, namely, the object whose name is being learned. The adult looking at the child says 'table' and looks at the table. The child looks at the table and utters 'tab' and looks back at the adult. And so it can continue. Triangulation thus requires two human beings and an object. Translation also can be understood as involving triangulation: Explorer looking at rabbit and saying 'gavagai?' to native; native looking at rabbit and then explorer, saying 'evok' to explorer; both looking at rabbit.

Let us see how triangulation would work with Descartes' private language of sensation. Call the private language user 'Adult1' and the third party 'Adult2'. Adult1 cries out and squeezes his hand. Adult2 looks at Adult1 and says: "What hurts?". Adult1 says: "This hurts". In a standard triangulation, 'this' would point to the object at the apex, but the object at the apex is nowhere to be seen.

The problem with this is that the most important element of the triangulation whereby words and objects converge is simply left out of consideration altogether. To ensure that it remains the most important element requires, at a minimum, that it is left out of the picture of the conscious mind. There is nothing to be seen. The explorer is faced with a mystery. There is just a blank spot where the apex-object should be. What is needed is reference. Just how is either the cartesian self or the explorer to *refer* to the mystery apex-object? Indeed what are the self and the explorer supposed to look

at together? Triangulation is impossible. Singular 'triangulation' is not possible. There is no way to introduce reference to a common objective object. So there must be something that is private referring, understood only by the monologist.

5   PI, § 244. Problem of Reference: Private Naming

Wittgenstein opens this passage with the question: "How do words refer to sensations?". This question can be expressed in a more pragmatic way as "how is the connection between the name and the thing named set up?". The place to look to get the clearest picture of just that is the initiate language learning situation. Here the child is just learning individual words. In the section above we described initiate learning in terms of triangulation. It applies to sensation words, but with a twist since it must respect the asymmetry between the child in pain and the adult reacting to the child. This asymmetry must be respected within the structure of triangulation if shared reference to the same thing is realised within initiate learning. How is that to be done? Child, who has cut himself, cries and looks at his hand; Adult looks at Child, and says: "Booboo"; Child looks up to Adult. This completes the horizontal line of triangulation. Child looks at his bleeding hand, and says: "Booboo"; and Adult looks at Child's bleeding hand and says: "Booboo". Now we have the complete triangulation: the apex-object referred to is the hurt bleeding hand. The Child *has* the pain of a cut on his hand while the Adult *sees* the pain in the bleeding hand and tears of the child. As Wittgenstein says, "the verbal expression of pain replaces crying and does not describe it". Without Wittgenstein's reply, we are left with the illusion that the Child has a way of privately referring to the propertyless pain as apex-object while Adult can only refer to publicly available physical and behavioural coordinates. So the two lines of referring do not pick the same object, and thereby do not refer to anything. The Child refers to pain without regard to his own behavioural reactions or to the environment that contains many dangers for causing pains. A new problem arises. What then are the identity conditions for sensation? Again, the behavioural squeezing of the hand and crying as well as the bloodied knife on the table are irrelevant to the identity of the sensation. The sensation must be identified first before one can hypothesise correlations between the sensation (careful here or one is importing our ordinary ways of identifying sensations) and behavioural and environmental phenomena. So what is the way that pains can be identified as such.

The response that most find intuitive and obvious is that "[…] only I can know whether I am really in pain; another person can only

surmise it" (PI, § 246). This epistemological reply to the problem results in an epistemic gulf between any given individual and any other human being. We are always inevitably kept at a distance from each other. But this epistemic solution to the problem is, according to Wittgenstein, "[i]n one way [...] wrong, and in another nonsense" (PI, § 246). In the first way, it is false that other people cannot know when I am in pain. That is a human reality. But why is it nonsense? "It can't be said of me at all [...] that I *know* I am in pain. What is it supposed to mean—except perhaps that I *am* in pain?" (PI, § 246). If we make a false claim, the words we use are meaningful, but if we are using words nonsensically, then the words we use are either meaningless, a kind of gibberish, or do not mean what they ordinarily mean. What is most important to notice in this reply from Wittgenstein is that meaning is prior to any other question or use that we can put to language. The problem of identity, even of pain, cannot be solved prior to the meaningfulness of the language that is used to solve it. The classic cartesian reply is a non-starter. Either the language used to refer to pain is meaningful, in which case there is no problem of identity; or it is just a kind of gibberish used to hide the fact that there is no solution to the problem of the identity of sensations under the assumption of radical privacy.

The problem of reference is fundamental to this emerging picture of the language of sensation since there is no sharing with others, no inherited way of going on correctly, and apparently, no background form of life. There is just a bare sensation in the void. This is the same set of issues dominating the opening passages of the *Philosophical Investigations*. There Wittgenstein introduced the builders' language-game (PI, § 2) as his tool for investigating the tripartite relation among a finite set of words, corresponding objects, and human builders. In the Bridge (PI, § 244) he reshapes the problem for inner sensations, and holds that it is most illuminatingly solved by the initiate learning situation, when the child is just acquiring language. How sensation words refer to sensations can be shown by how a child first acquires sensation words, like, 'pain'.

Consider how a child can be taught the word 'pain': "[W]ords are connected to primitive, natural expressions of sensation and used in their place". Suppose the child touches a hot stove. His natural spontaneous reaction is to scream and cry. These are the primitive, natural expressions of burning pain. As such natural expressions, we share them with all other human beings. The parent, who knows pain, tries to soothe the child by, in effect, teaching it new words, that is new pain-behaviour: 'ouch!' 'booboo', 'pain'. On the next occasion on which he is in pain, the child will come to utter 'booboo' or 'ouch' or even 'pain'. When he does so, he replaces his cries with words, and thereby refers to, the experience of pain, through his natural pain behaviour. That is the genesis of the word's referring to the experience.

The sensation word subsequently gains autonomy from the natural pain behaviour in its meaning pain. The expression 'It hurts' replaces the crying. When the expression 'it hurts' is mastered, it need not require the presence of the natural expression of the inner sensation in order to meaningfully refer to that sensation. The natural pain behaviour directly expresses the sensation. In replacing that naturally expressive behaviour, the learned verbal behaviour refers to the sensation *without* the mediation of thought. Expressive behaviour does not describe the link between the sensation and the word. According to Wittgenstein, primitive expressive behaviour does not refer in and of itself. It does not have semantic properties. Reference comes within acquisition of a socially shared language-game. Acquisition of reference requires the two-party linguistic learning of words. Precisely what the cartesian conception of language cannot provide. The child cannot teach itself what the word for pain is. He does not even know what pain is: he experiences pain.

Wittgenstein will use this expressive conception of sensation as a model for many more states of mind, for feelings, emotion, sensations and the like. Descartes by contrast can only teach himself, which means that he must wait, in his own words, until he is fully mature and can form the appropriate mental concepts. The consequence of this view is that much of a person's intellectual and conceptual mental life awaits inner innate development. Such concepts must be independent of ordinary public language, primitive behavioural expressions, and so are not embedded in a shared form of life, subject to a logical grammar. As we have seen the meaningfulness of the words of a language is not acquired through brute acts of naming. Words are meaningful as instruments of measurement within logically structured systems of activity. The reason that the problem of reference is stymied is because it is thought that meaningfulness derives from reference but in fact, as Wittgenstein has argued, successful reference depends upon its being situated within a meaningful language game. Identity questions require logical grammar.

Wittgenstein challenges the cartesian approach by asking: "How can I even attempt to interpose language between the natural expression of pain and the pain?" (PI, § 245) When the child uses 'booboo' or learns 'booboo', he cannot reserve such use to the pain independently of his tears. I cut my hand when cooking. In what sense is that pain separable from my hand, 'it hurts'? Wittgenstein asks, in light of these considerations, "in what sense are my sensations private?" (PI, § 246). To which Descartes answers by appeal to his epistemological goal: "[W]ell, only I can know whether I am really in pain" (PI, § 246). I know my pains I do not just express them. But how I state what I know is left obscure. How do I even come to know that I have pain as opposed to merely having sensation or a tongue? The full development of what is nonsensical about this epistemological strategy is developed

in the private language arguments.[13] But however the question of identity can be answered, pure reference achieved through an act of naming, cannot provide the answer.

6 PI, §§ 245-9. Problem of Identity: Possibility and Necessity

From having focused on the relation between individual private sensations and words, Wittgenstein turns to the role that imagination might play in our understanding of sensations and mental privacy. If the epistemological account of privacy is false or nonsensical, what sort of account can be given? It is perhaps a grammatical proposition, like: "[T]he sentence 'Sensations are private' is comparable to 'One plays patience by oneself'" (PI, § 348). There is a temptation to construe the second sentence as an analytic proposition, true in virtue of the meanings of the words. But this will not capture Wittgenstein's notion of a grammatical proposition. He gives his readers two hints as to how to understand this obscure notion. The first is that it "belongs to the scaffolding from which our language operates" (PI, §

---

13 Here we can briefly review how Wittgenstein uses this double-barreled argumentative method – false or nonsense – to successfully critique philosophical theories of language and mind that he rejects. The first of the pair of arguments is a conflation argument in which the philosophical theory under scrutiny conflates the means by which the theory represents its subject matter with that subject matter itself (cf. PI, § 246). The second argument shows that the theory is self-defeating. It is a paradox because the theory eliminates the very phenomena to be explained. Wittgenstein makes three such arguments in Part I of the *Investigations*. The first theory is informed by Frege's idea that formal logic actually structures natural language. Wittgenstein argues that logic is a means of representing natural language, but logic is not what gets represented. In other words, the advocate of Frege's idea conflates formal logic and natural language. The stronger argument is the paradox argument. Taking an individual thought to mean: this – is – so creates a paradox. In meaning this, a thought cannot be of what is not the case (PI, § 95). But the point of a thought is precisely to be meaningful in a way that is independent of whether it is true or false. Imagination for example would be eliminated on this account of thought. Imagination is meaningful and yet is typically false. The source of the problem lies with Frege's idea, his picture of the relation of logic to reality (see Williams 2010, ch. 4). The second philosophical theory Wittgenstein examines in this way is a theory of rule following. This is a variation of the idea of the question of the identity over time. In this case, causal determination over time is conflated with idealised logical continuity. A metaphor Wittgenstein uses is that of the causal action of a machine, like a watch, being conflated with an idealised machine-as-symbol which is conceived as determining all possible continuations (PI, § 194). There is no way out of this mistake while attempting to preserve the theory that interpretation determines the continuation of a rule: "[I]f every course of action can be brought into accord with the rule, then it can also be brought into conflict with it. And so there would be neither accord nor conflict here" (PI, § 201). The very thing we look to the rule to establish, what is correct and what is wrong, would be eliminated (see Williams 2010, ch. 5). The third pair of conflation-paradox arguments is Wittgenstein's close examination of sensation and consciousness (see Williams 2010, ch. 8).

240); and secondly, briefly introducing the word 'intention' as it occurs in the sentence "Only you can know if you had that intention", a word quite removed from 'sensation', and yet it too seems to name something private like sensation. Wittgenstein offers in place of that claim the claim that 'intention' "means: *that* is how we use it" (PI, § 247). In other words, he offers a reminder that words are like measurements. It is essential to them that we know how to use words just as we know how to use a measuring stick. Whereas analyticity is understood in terms of truth conditions, not human action, grammar is understood in terms of the word's role or use within a living language. That involves that we share the procedures by which we use words in action. The procedure for private reference is empty of meaningful use. It is nonexistent. It might be thought that triangulation cannot be used for private reference, as has been shown already. It might be thought that we should look to the modalities – possibility or necessity – to find the link between word and private object. What would this mean?

The critiques[14] Wittgenstein has already applied against Frege's idea of a propositional logic and the interpretationist theory of rule-following, as providing the scaffolding of language, open space for seeing Wittgenstein's logical grammar. The logical grammar of a language or language-game imposes constraints on how expressions are to be used in providing the background foundational procedures for engaging in our ordinary social and normative use of language which is intertwined with our social and environmental activities. They are procedures of use much in the same way rules of ordinary games open possibilities through constraint. Is there some way that judgments of privacy of object or privacy of reference can be identified through patterns of modality. For this task, it might be thought that imagination, in its liberality, might impose the relevant divide between what is possible (or thinkable) and what is not merely possible but necessary, the region in which a proposition cannot but be true. Wittgenstein rejects this crude account of how the modalities are to be understood. The cartesian private inner arena has been imagined as real for centuries. Does this not make it possible and further that we cannot but imagine it to be so? Taking this view seriously, possibility reaches as far as the imagination takes us. But that is way too far to impose the constraints that are necessary for the limits of language or for truth. Even Descartes repudiates imagination as the source for fixing identity over time.

---

14 Here briefly are Wittgenstein's arguments against assigning the formal propositional calculus the role of scaffolding in the human form of life and his arguments against treating interpretation theory as powerful enough to impose the procedures that govern rule following.

Descartes uses as his example for showing that imagination cannot be the means for determining the identity of an object whose range of possible properties can vary over time, in other words just about anything that exists. Descartes shows the inadequacy of this through a simple thought experiment. Consider a piece of wax taken freshly from the bee hive. It smells and tastes sweet; it is hard and cold; it is sticky to the touch. At a later time, it is melted: it is black and acrid smelling and tasting; it is hot to the touch; it is an irregularly shaped puddle. All of its properties of that earlier time are gone as a result of being melted. We cannot determine its particular identity by its surface properties. But could we determine its identity through the imagination, it is suggested. Imagining the changes along the way to mark its slow transformation from the former condition to the present puddle of black liquid. It is impossible to imagine that the cold hard piece of wax had become this very hot and blackened puddle of an irregular shape. It could have had, just as well, any number of other shapes. How does imagination sort out the path that leads to this shape when it could have been easily imagined to have been different. Imagination cannot be the faculty for deciding this.

Consider now some additional examples of what is possible if imagination were our guide. Imagination, as we shall see, is not freedom to go in any direction it can take us, and so it is irrelevant to the problem of identity of pain that faces us now. And if Grammar imposes restrictions on imagination, then again imagination is irrelevant to the problem of identity. It presupposes identity, it does not determine it. We cannot take the smile of a baby to be pretence (PI, § 249) nor can a dog simulate pain (PI, § 250). The baby must learn to lie before it can pretend. And the dog cannot simulate pain because it lacks "the right surroundings for this behaviour to be real simulation". It needs motivation for simulation and the right sort of audience to witness it and be taken in by this behaviour. This does not mean that babies cannot smile or that dogs cannot feel pain. But we cannot imagine a baby to be capable of pretence or lying. To imagine this would be to attribute sophisticated cognitive capacities and motivations to the infant. Though we can imagine a dog being in pain in many different situations, we cannot imagine him to be motivated to simulate the complex behaviour of actually being in pain. Another way to put this point is that we may observe babies speaking a sophisticated adult language with the facial expressions to go with it, but they are computer generated babies of the imagination that everyone knows are not real. Cartoons are filled with dogs and other creatures who speak, pretend, lie, simulate pain and many other sophisticated acts. But these are cartoon characters. No one, not even a child, would take these as real living dogs. What is required of both babies to pretend and dogs to simulate are second-order linguistic capabilities which require full command of first-order speech.

Philosophically, as Michael Williams has repeatedly argued, episte-mological thought-experiments, like Descartes' suggestion that each of us might *just as well* be dreaming our lives away or that any one of us might be an isolated brain in a vat *so far as we know*, are equally fairy tales of the imagination (see Williams 1995).

Giving up the idea that a range of imagined possibilities might fix private reference and so the identity of the private object, Wittgenstein goes on to consider the idea that necessity, not possibility, might fix the relation between reference and the private object. We use a special form of words to appeal to such necessity, namely, 'I cannot imagine the opposite'. Wittgenstein's concern with this way of trying to express the necessity of certain states of affairs through the lens of imagination is that the phrase 'I cannot imagine the opposite' provides, not a standard for necessity, but "a defense against something whose form makes it look like an empirical proposition, but which is really a grammatical one" (PI, § 251). To avoid confusion over this misidentification of the cognitive limitations of babies and dogs, and so the grammatical limits of imagination, Wittgenstein introduces a very simple example of a grammatical proposition: "Every rod has a length" (PI, § 251). This has the look of a general empirical proposi-tion, like 'every squid squirts ink'. The difference between the two propositions is that there could be squids that do not squirt ink while no rod can fail to have a length. Furthermore, and this is emphasised by Wittgenstein, there is nothing that we would call 'the length of a sphere'. These inferences, especially those constituting a negative inference, are part of a 'picture' belonging to the grammatical prop-osition in question. We cannot picture a rod without a length and we cannot picture a sphere with a length. There is an inferential struc-ture associated with the grammatical proposition. This is a holistic structure that has no room for atomistic elements that are nonethe-less meaningful.

The cartesian view, on the other hand, does permit or even require an atomistic treatment of particular propositions and individual as well as general versions of the same proposition. Wittgenstein's ex-ample of a cartesian claim that is atomistic as well as individual is saying "This body has extension" (PI, § 252). We could respond by say-ing "Nonsense!" but typically we do not. To understand what a gram-matical proposition is for Wittgenstein, we need to understand the rationale for both responses. The generalised version of this propo-sition is 'Every body has an extension', a proposition that is neces-sary. The use of 'every' or 'each' does not render it necessary. Rath-er it is the use of 'body' that requires 'every' or 'each'. This is what it is to specify the identity of 'body'. Though it looks like an empirical proposition, it is not. It plays an a priori role in numerous language games; that role can be foundational just because it has an inferential structure that cannot be broken without rendering the proposition or

proposition fragment in which it occurs nonsensical. So what changes with the individual version "This body has extension"? 'This body' implies only that this one has extension, but that is ludicrous. If this were the fundamental notion of 'body', it would become impossible to express the general proposition 'Every body has an extension' or 'All bodies have extension'. To get the general notion of body, we would have to identify each particular body and conjoin them to make a general proposition.[15] But we have no idea how to construct this proposition nor how to use it. An indefinitely long conjunction would have to take the place of the universal grammatical proposition. It would have to become irrelevant to our language games, and so, it could not play the role of giving the identity of 'body'.

This same diagnosis applies to sensations and other mental states. Mental states also occur in language games as part of our form of life, and for which there are foundational grammatical propositions. We have looked at one of them: 'crying expresses pain'. This is not an empirical proposition though the cartesian treats it like a contingent empirical one. Pain is treated as separable from crying or any other expressive feature of being in pain. This separation is what enables the cartesian to raise the question of the identity of pain as independent of crying or being cut with a knife or any other pain-behaviour. Once the connection is severed, the identity of the private sensation is lost. Wittgenstein's point is that the problem of the criterion of identity cannot be resolved unless it is, in a sense, an *a priori* one (as understood by Wittgenstein),[16] a grammatical proposition. It is foundational to our language game of sensations. It is important to note that it is an empirical matter for adults whether their pain is actually accompanied by crying. Nonetheless it remains the case that crying necessarily expresses pain. What is needed now is an understanding of what the criterion of identity for pain is. Neither the imagination's possibilities nor reason's necessity can identify what pain is or what 'pain' means. As we shall see, the criterion for the identity of pain is entwined with the meaning of 'pain'.

---

15 This is the method that Wittgenstein describes in the *Tractatus Logico-Philosophicus*. Everything that exists is particular and atomistic in its essence. General propositions can only be constructed by conjoining individual propositions *ad infinitum*. This is, of course, a hopeless project.

16 The primary point in putting the proposition of identity as *a priori* is to underscore that it is not an empirical proposition that may be true or false even though how the child learns the word 'pain' is contingent.

7　PI, §§ 253-7. Criterion of Identity: Logical Grammar

Wittgenstein raises the problem of the criterion of identity for pains, and by extension, for other mental states. He does so by reminding the reader that the statement "This body has extension" is one that we are inclined to reply with "Of course!" as though it might not have had extension (PI 252). The very suggestion that this body might not have had extension makes nonsense of the emphatic agreement. Particular bodies have extension, with the same implications that are carried by the general form "Bodies have extension". Insofar as Descartes is interested in such propositions, the general form would express a metaphysical necessity. Extension is the essence of particular bodies (Descartes 1996, IV), God be willing, since body as extended is the creation of God who thereby fixes the criterion of identity for body, its extension. Wittgenstein's reply to this theistic solution to the problem of the criterion of identity for body would be 'Nonsense!'. 'Body is extended' is necessary in the same way that 'patience is played by oneself' is necessary. It is one of the fundamental rules of the game played with 'body'. It provides the criterion of identity for bodies. The relation between 'body' and 'extension' cannot be broken without destroying bodies. An image of a body in one's imagination or dream is not a body. It is an image of a body.

Now the question of the criterion of identity is raised for pain, and in a particular way: "'Another person can't have my pains.'— Which are my pains? What counts as a criterion of identity here?" (PI, § 253). A familiar suggestion is to strike oneself while saying "But surely another person can't have THIS pain". In other words, what I alone am able to feel or know is the criterion for the identification of pain for me, and likewise for anyone else. But Wittgenstein thinks that to take such a feeling – THIS – as the criterion of identity for, what else THIS, can only be a way of reminding ourselves of what the criterion of identity for pain really is. So, what is the criterion of identity? Wittgenstein presents a dilemma for the individual who is trying to use words to stand for his sensations. First, he may use words for his sensations as we ordinarily do, but then "my language is not a 'private' one. Someone else might understand it as well as I" (PI, § 256). Using words for sensations as we ordinarily do involves recognising natural expressions for sensations in which case the fact that another cannot have my pains does not entail that another cannot recognise when a person is in pain. Such recognition is an ordinary and essential part of our grasp of the *criterion* of identity for pain.

If we take away our ordinary criterion of identity of pain, especially as it is tied up with the natural expressions of pain and we have only the sensation itself, all that I as an individual could do is to "associate

names with sensations and use the names in descriptions". I would have to do something like the following: Whenever I felt something 'inner' as it were, I would associate it with a name (maybe, 'pain' or 'Fred' or 'sensation') and use that name in description ('Fred' came today). To develop the second half of the dilemma, Wittgenstein invites us to consider a world in which "human beings shewed no outward signs of pain" (PI, § 257). How then could a child be taught the word for 'pain'? Wittgenstein's sarcastic response is that perhaps the child is a genius and invents a name for pain. No matter what the cognitive strengths of the child-genius might be or might become, he cannot discover sensation language for himself and he certainly cannot fix the criterion for identity. It cannot consist in the association of a name with a sensation. How would the child reidentify the sensation? By what criterion? Perhaps he thinks that a pain in the leg is not the same as a pain in the hand or in the tooth. Nothing constrains his choices of names nor the principles of reapplication of those names. Another might see no regularities in the names uttered; the possibility of communication is nonexistent. So where do we look for the criterion of identity?

> And when we speak of someone's having given a name to pain, what is presupposed is the existence of the grammar of the word 'pain'; it shews the post where the new word us stationed. (PI, § 257)

The logical grammar of our words provides the foundational rules or structure in relation to which words are meaningful tools of measurement. They are not mere elements of association. The use of words must involve more sophisticated procedures that enable us to live, create, and maintain the human form of life we occupy.

This completes the bridge to PI, § 258: the private diary argument. At the end of the Bridge, Wittgenstein tells us where we can find the criterion of identity for pain and other sensations: it is the logical grammar of our language games. The arguments to come are his most powerful arguments against the cartesian theory of sensations. Methodologically, they use imagined scenarios that provide a context in which a private diary can be written (PI, § 258), and a context in which private objects can be located (PI, § 293). In providing such contexts, the scenarios give the illusion of supplying the necessary logical grammar that is foundational for the use of sensation-words. But Wittgenstein's arguments overwhelm this illusion and show it for what it is. The private diary rests on a conflation of the means of representation with the object of representation. The beetle in the box is enmeshed in a paradox. There is no further place for the cartesian theory to go.

8   Concluding Remarks

This is what any bridge must be: a truncated route to a new region. There is much that is drawn upon that requires further explanation and development. The most important are logical grammar and the human form of life.[17] Both are involved in what is foundational to language (a framework of logical grammar) and mind (that intersection of our causal situatedness, expressive behaviour, natural activity, and, of paramount importance in making the human mind unique, language through-and-through (our human form of life). Wittgenstein hopes to secure two important philosophical points before he begins his careful examination of mind and mental concepts. The first is the primacy of meaning over any epistemological or metaphysical principle or mode of explanation. If the words used in articulating and defending a philosophical view or theory without an understanding of how words mean and maintain their meaning, the risk is confusion. Secondly, it is equally important to recognise the import of grasping the problem of the criterion of identity for objects. Objects cannot have the requisite identity needed for learning and using unless they already involve a conceptualised identity. The 'this' inside me does not naturally have the label of 'pain' attached to it. It is terribly misleading for Cartesians to simply describe objects of reference in their ordinary English or French terms as though this were neutral in characterising, for example, what the toddler already knows when given an ostensive definition, or any of us know when introduced to a new object. Objects are not conceptualised on their own, not even as 'objects'. Language makes them recognisable. Now that we no longer live in a theistic philosophical world, there is no other way to identify the objects of our interest.

---

[17]   I recommend two excellent Cambridge elements that are directly pertinent to enhancing understanding of logical grammar and the human form of life. These are: Bronzo (2022) and Boncompagni (2022).

References

Boncompagni, A. (2022). *Wittgenstein on Forms of Life*. Cambridge: Cambridge University Press. `https://doi.org/10.1017/9781108946513`.

Bronzo, S. (2022). *Wittgenstein on Sense and Grammar*. Cambridge: Cambridge University Press. `https://doi.org/10.1017/9781108973359`.

Cavell, S. (1979). *The Claim of Reason*. Oxford: Oxford University Press. `https://doi.org/10.1093/oso/9780195131079.001.0001`.

Davidson, D. (2001). "The Second Person". Davidson, D., *Subjective, Intersubjective, and Objective*. Oxford: Clarendon Press, 107-21. `https://doi.org/10.1093/0198237537.003.0008`.

Descartes, R. (1996). *Meditations on First Philosophy*. Edited by J. Cottingham. Cambridge: Cambridge University Press.

Frege, G. (1997). "Sense and Reference". Beaney, M. (ed.), *The Frege Reader*, Oxford: Blackwell, 151-71.

Mulhall, S. (2006). *Wittgenstein's Private Language: Grammar, Nonsense and Imagination in Philosophical Investigations, §§ 243-315*. Oxford: Oxford University Press. `https://doi.org/10.1093/acprof:oso/9780199208548.001.0001`.

Quine, W.V. (1960). *Word and Object*. Cambridge, MS: The MIT Press. `https://doi.org/10.7551/mitpress/9636.001.0001`.

Stern, D. (2004). *Wittgenstein's Philosophical Investigation: An Introduction*. Cambridge: Cambridge University Press.

Williams, M. (2010). *Blind Obedience*. London; New York: Routledge. `https://doi.org/10.4324/9780203870815`.

Williams, M. (1995). *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism*. Princeton, NJ: Princeton University Press.

Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus*. Transl. by B. McGuinniesse; D. Pears. London: Routledge & Kegan Paul. `http://dx.doi.org/10.1093/owc/9780198861379.003.0001`. [TLP]

Wittgenstein, L. (1969). *On Certainty*. Edited by G.E.M. Anscombe; G.H. von Wright, transl. by D. Paul; G.E.M. Anscombe. Oxford: Blackwell. `http://dx.doi.org/10.1007/bf02379321`. [OC]

Wittgenstein, L. (2009). *Philosophical Investigations*. 4th edition. Revised by P.M.S. Hacker; J. Schulte, transl. by G.E.M. Anscombe; P.M.S. Hacker; J. Schulte. London: Blackwell Publishers. [PI for Part 1 of the first edition; PPF for the so-called *Philosophy of Psychology – A Fragment*, formerly known as "Part 2"]

Understanding Others, Conceptual Know-How and Social World

Rémi Clot-Goudard
Université Grenoble Alpes, France

Abstract In contemporary philosophy of mind, understanding others is often presented as an activity of attributing mental states to agents or mindreading – the central question being then how to access their minds. The paper argues that this pervasive approach should be rejected, in favour of the view along which identifying an action comes from exercising conceptual skills acquired through being inserted into shared practices characterising a social world. Examining the conditions of their acquisition then sheds new light on the semantics of psychological concepts as well as on the roots of misunderstanding.

Keywords Wittgenstein. Understanding. Know-How. Intention. Anscombe. Common-sense Psychology.

Summary 1 Introduction. – 2 Does Understanding Others Require Access to their Minds? – 3 Understanding Others as a Practical-conceptual Achievement. – 4 Tuning in with Others. – 5 Conclusion.

> Merely recognizing the philosophical problem as a
> logical one is progress. The proper attitude and the
> method accompany it. (LWI, § 256)

1 Introduction

In contemporary philosophy of mind, understanding others is often presented as the result of a complex activity of interpretation, consisting in the attribution to an agent of mental states that enable us to make sense of his or her behaviour.[1] In a restaurant, your neighbour gets up and walks slowly towards the door: what exactly is he doing? Does he want to leave without paying, thinking that the boss, who is busy elsewhere, will not see him? Or does he want to surprise someone sitting near the entrance, whom he believes to be an acquaintance, by arriving silently behind him? Asking and answering such questions presupposes mastery of a whole range of psychological concepts (such as intending, wanting, desiring, believing, the various concepts of emotion, etc.) that make up the paraphernalia of our so-called 'commonsense (or folk) psychology'. From a philosophical point of view, then, the central question seems to be how to account for our ability to apply such concepts, as well as for their epistemic status.

The debates surrounding these questions (of which we will give a rough idea below) are still lively. But many authors seem to agree on one point: psychological concepts are used to refer to something that is 'in the head' of the agent; and the main problem is how we can gain access to it. This is why it is now common to refer to the interpretive activity that enables us to understand others as 'mindreading' (see for instance Spaulding 2018; 2020). Such a label might seem trivial: is it not obvious what motivates the metaphor? Understanding the meaning of a behaviour is analogous to understanding the meaning of a text. And don't we sometimes say that someone 'reads another's mind like an open book'? Now, far from being just a convenient label, the metaphor actually betrays a presupposition that Constantine Sandis (2019, 241) states as follows: understanding another person implies "obtaining and decoding the information stored in their mind" (see also Hacker 2018, 380). But is the meaning of the agent's behaviour really given by what is 'in his mind'?

In what follows, I would like to show that these widespread semantic presuppositions regarding the use of psychological concepts are

---

doubly erroneous. Firstly, it succumbs to a picture of the mind as a container or interiority, long criticised by Wittgenstein, which feeds the idea that understanding others ultimately results from knowledge of 'mental contents' – thus focusing debates on the question to know how we can have 'access to' the minds of others. Secondly, it has the effect of withdrawing our attention from a more fundamental point. Understanding what someone does is crucially manifested in our ability to *describe* their conduct, that is to say, primarily to classify it under concepts that *are not* psychological concepts, but various concepts of activity, in relation to instituted practices. In fact, the logic of psychological concepts can only be fully understood by starting from a better view of what describing someone's action does imply. It is only by elucidating this through an examination of the conditions under which concepts of activity are learned that we become able to get the full meaning of this Wittgensteinian point: understanding others is the achievement of a sustained interaction that presupposes participation in the same background of life, in a tangle of linguistic and non-linguistic practices that characterise a social world.

2 Does Understanding Others Require Access to their Minds?

Donald Davidson's early work (Davidson 1963) did much to spread the view that understanding others consists in identifying their reasons for acting, which in turn can be analysed as a combination of mental states, namely a desire, giving a general characterisation of the desired thing, and an instrumental belief, specifying a particular means to obtain what is desired. To understand an action is then to be able to rationalise it, i.e. to see it as the conclusion of a practical reasoning of which such desire and belief are the premises. But this requires to get knowledge of the agent's beliefs and desires. Discussions about the nature and epistemic status of folk psychology have therefore tended to focus on the question of what enables us to get such a knowledge: how does one come to have access to the content of other minds? What are the cognitive or non-cognitive capacities, or even the underlying mechanisms, through which the relevant mental content is identified?

Among the answers, two main options stand out. The first, known as the "theory theory", asserts that the ability to understand others ultimately depends on the possession of a theory of mind (e.g. Fodor 1987). According to this approach, 'intention', 'desire' or 'belief' are theoretical terms, designating unobservable internal mental representations, postulated by the theory as the rational causes of observable behaviour. The connection between mind and action is based on

theoretical hypotheses or laws (roughly of the form 'if X is in mental state M in circumstances C, then, all other things being equal, he will perform action A'), by means of which behaviour can be predicted and explained. We attribute mental states to others by a kind of inference to the best explanation of their behaviour. While the details of the analysis are hotly disputed,[2] all those who embrace this approach nevertheless share the idea that understanding others is based on inferential knowledge of mental content.

The plausibility of this first approach has been vigorously challenged by proponents of the second, known as "simulation theory" (see Davies, Stone 1995). According to the latter, understanding someone consists in putting oneself in their shoes, i.e. adopting their perspective on the world in order to imagine or simulate what our own mental states would be in such a case, before projecting them onto the other person in order to predict or explain their actions. The notion of 'simulation', borrowed from the field of artificial intelligence, suggests that understanding others is conceived as an internal psychological modelling process, rather than as reasoning informed by a theory. This approach has thus helped to revive the old notion of empathy that Lipps, following the psychologist Karl Groos, had defined at the beginning of the twentieth century as "internal imitation" (Stueber 2018) and to which many analytical philosophers have recently turned their attention (2006). Neurological discoveries concerning "mirror neurons" have also been interpreted by some as providing a neurobiological basis for the capacity for empathy (Rizzolatti, Sinigaglia 2008; Coplan, Goldie 2011).

However diverse and conflicting these approaches may be, they nonetheless subscribe to the spontaneous image according to which thoughts or intentions are processes that take place *in* the mind of the agent and remain hidden from us, constituting the internal counterparts of behaviour that give it its meaning. Ludwig Wittgenstein, in his work on the philosophy of psychology, often drew attention to the distortions in the account of the logic of psychological concepts that arise from the philosophical use of this picture. For example:

> The intention *with which* one acts does not 'accompany' the action any more than a thought 'accompanies' speech. Thought and intention are neither 'articulated' nor 'non-articulated'; to be compared neither to a single note which sounds during the acting or speaking, nor to a melody. (PI, II, § 280)

---

2 Some proponents of the idea that commonsense psychology is a theory nevertheless believe that it is obsolete and doomed to give way to a robust theory, formulated in sheer neurophysiological terms (e.g. Churchland 1981). For a recent overview, see Hutto, Ravenscroft 2021.

Let us say I am sitting at my desk and I get up to fetch the dictionary. As I stand up, I may say to myself: "I'm going to check the spelling of this word". And it would be correct to say that these words express my intention. It would only be hidden from you insofar as I kept this to myself. But what if I said those words out loud? You would then know the 'contents of my mind'; but you would still not know exactly what I wanted to do, since you would still ignore the word the spelling of which I wanted to check.[3] However, I know what the word is and I could tell you if you asked me. Does that mean that I have already said it to myself? No. In fact, it is quite possible that I have not said anything to myself at all: in a moment of doubt, I stop writing, my eyes stare at a word on the screen for a moment, then I get up to fetch the dictionary. Or maybe the only thing 'on my mind' is a haunting melody that has been playing over and over since I heard it on the radio. Commenting on this, Elizabeth Anscombe writes thus:

> An intention after all needn't be [an occurring] thought, for one can intend what one is not thinking of, as when one intends over a whole period to make a certain journey, but in fact seldom thinks of it, and when one even thinks of it, one's thoughts aren't to the effect that one is going to make that journey. [...] We tend to think it out of a prejudice that an intention *must* be a mental phenomenon, i.e. an event in the mind. (Anscombe 1963, 59)

While it is true to say that intentions are a kind of thought, we cannot equate thought or *cogitatio* with something that presents itself to consciousness at a given moment, on the model of conscious experience, as Descartes did (Anscombe 1963, 60-1; see also Descombes 2004, 190-8). This kind of actualism let aside a logical difference between what we call the content of an intention (or the content of a belief) and the content of an experience. To report the content of my intention is to describe what I am going to do, but not to describe what is happening in my mind at the moment. Similarly, expressing the content of a belief is saying something about the world, not about my experience. On the other hand, to describe the content of an experience is to say how things appear to me at a given moment, what the (visual, auditive, etc.) appearances are. Let us suppose that when I get up to reach to the dictionary, I think I hear my phone vibrating, though it is in fact the neighbour's intercom ringing. Realising my mistake, I could describe my experience by saying "it sounded like the faint noise my phone makes when it vibrates". But having an intention does not imply at all that something presents itself to me in

---

3 Cf. PI, II, § 284: "If God had looked into our minds, he would not have been able to see there whom we were speaking of".

this sense, nor in the way of an inner speech or a melody heard in the mind. All 'mental content' cannot be modelled on the content of an experience. The confusion arises as soon as we start talking about 'mental content' without further examination, and make a theoretical use of the metaphor of the mind as a place or container.

Why is the view that understanding others involves having access to the inner contents of their minds so attractive and why does it have such a powerful influence on us? It is partly because of well-known facts (Anscombe 2000, 9, § 4): for example, the fact that we can hide our thoughts from others or lie to them about our intentions; that we can have an intention but not carry it out; or that the intention with which an agent does something cannot be seen in what he does: in this case, we have to question him – but he does not have to make any particular observation or inference to be able to answer us, which feeds the idea that what an agent thinks or wants he knows it directly, whereas we need to manage ourselves access to his interiority and can have only indirect knowledge of it.

But seeing the problem of understanding along these lines actually leads to miss the point. First of all, it misses the point that thoughts can be expressed and that, from then on, there is nothing hidden about them. On the other hand, it is true that understanding what someone else is doing may depend on knowing more about them – knowledge that I can obtain by questioning them, or by paying attention to other features of the circumstances, or through testimonies, and so on. However, we cannot equate understanding of others with knowledge of anything, and especially not of a 'mental content'. After all, one may have knowledge of this content and still fail to understand it. As Wittgenstein puts it:

> Even if someone were to express everything that is 'within him', we wouldn't necessarily understand him. (LWI, § 191)

I might still be unable to understand the reasons given to me by the agent, even though I am certain that such are his reasons for acting (because he has told me and I have no reason to doubt his sincerity) – that is, I might still be unable to understand the agent himself. Anscombe (2000, 71, § 37) notes that the agent's mere statement of his will is not enough to make me understand what he wants. If someone says "I want a saucer of mud", I will certainly know the object of his desire, but his conduct and his discourse will still remain obscure to me, unless I understand what is the point of wanting a saucer of mud. The answer to this question, Anscombe explains, would consist in a "desirability characterisation", i.e. a specification of the aspect under which the thing desired is good in the agent's eyes and makes it desirable. Now, to understand what the agent might say here requires that I myself be able to recognise the good he is pursuing. But this

presupposes much more (and much else) than the possession of knowledge or information about the agent or what is going on 'in his mind'.

3 Understanding Others as a Practical-Conceptual Achievement

The confusions associated with the use of the metaphors of content and access should encourage us to account for understanding others from another starting point. To this effect, we should reconsider the fact that most of the time, the behaviour of those around us is immediately intelligible to us.

As Dan Hutto (2004) points out, following Shaun Gallagher and other proponents of embodied cognition, those whose believe that understanding others is based primarily on the use of commonsense psychology in order to ascribe mental states to others tend to think of the problem as arising from the third person, from the point of view of a more or less detached spectator. In so doing, they do not pay enough attention to interaction situations, or to the basic abilities that enable us to attune naturally with the expressive responses of others (such as facial and motor imitation, the phenomena of emotional contagion, etc.). This leads them to give an over-intellectualised account of understanding. Against this tendency, writes Hutto:

> I promote the idea that in the basic cases we are able to 'read' others reliably and *vice versa* and that when we are in our historically normal environments this is no accident. For, like all creatures, due to long periods of tinkering and adjustment, we have been shaped precisely to respond to such environments, be they biological or social. Taking this idea to heart makes the alternative claim that our basic social interactions are made possible by means of the tacit predictions and explanations of commonsense psychology deeply suspect. (2004, 554)

Elucidating understanding requires clarifying the nature of this adjustment. But Hutto goes further and also argues that in most cases, and not just "basic cases", understanding others does not depend *at all* on an attribution of reasons in the third person – such an activity being, at best, "peripheral" (558). For the intelligibility of actions derives from the fact that they conform to common norms of conduct. That we are legible to each other in our ordinary interactions is not the result of a specific interpretation activity, explicit or implicit, but results from the fact that we share the same set of "norms and routines that structure these interactions" (558-9). This shared practical background is what our common sense does consist in, on the basis of which others' behaviours are identifiable.

Hutto (as well as those who agree with him) is certainly right to insist that this is the starting point for elucidating the ordinariness of mutual understanding. But we must guard at the same time against the temptation to conceive of this "embodied practice" (550) as a set of adjustment mechanisms divorced from any conceptual or symbolic dimension. In what follows, I would like to argue that understanding means exercising a kind of conceptual know-how – which insists on the fact that a concept is something more akin to the possession of a technique than a product of representational mental activity.

In her masterpiece *Intention*, Anscombe asks how we go about telling someone's intentions: what kind of true statements might we give about someone's intentions and how do we know that they are true? Having suggested that it would be enough to state "what he actually did or is doing", she adds:

> I'm referring to the sort of things you would say in a law court if you were a witness and were asked what a man was doing when you saw him. [...] [I]n a very large number of cases, your selection from the immense variety of true statements about him which you might make would coincide with what he could say he was doing [...]. I am sitting in a chair writing, and anyone grown to the age of reason in the same world would know this as soon as he saw me, and in general it would be his first account of what I was doing; if this were something he arrived at with difficulty, and what he knew straight off were precisely how I was affecting the acoustic properties of the room (to me a very recondite piece of information), then communication between us would be rather severely impaired. (Anscombe 2000, 8, § 4)

It is indeed a remarkable feat that even a fairly young child entering a room can usually give a description such as "she is sitting and writing" with ease, description which identifies an action performed intentionally. Of course, such a description is very rough and may raise a number of questions: what is she writing, to whom, and what for? But it is already a correct answer to the question "what is she doing?". The questions designed to enrich the scenario thus sketched out could not be asked if the child did not first recognise that the person is writing. His ability to correctly describe what the other is doing expresses his understanding of that action; but, like his understanding, it depends on whether or not he possesses some concepts, such as 'writing'.

What does it mean to possess a concept? For a whole tradition born of modernity, to possess a concept of something is to be able to form a representation or idea of it (in the Cartesian or Lockean sense) having a general or archetypal character. But in Wittgenstein's perspective, "a concept is the technique of using a word" (LPP, 50). To learn

such a technique is to be initiated into a kind of know-how – which certainly concerns itself with words, but is also intertwined with a whole range of other activities. Explaining this perspective, Anscombe writes thus:

> The competent use of language is *a* criterion for the possession of the concepts symbolized in it, and so we are at liberty to say: to have such-and-such linguistic practices is to have such-and-such concepts. "Linguistic practice" here does not mean merely the production of words properly arranged into sentences on occasions which we vaguely call 'suitable'. It is important that it includes activities *other* than the production of language, into which a use of language is interwoven. For example, activities of measuring, of weighing, of giving and receiving and putting into special places, of moving about in a huge variety of ways, of consulting tables and calendars and signs and acting in a way which is connected with that consultation. It is plausible to say that we would have no concept of *length* apart from some activity of measuring, and no concept of precise comparative length of distant objects if the activity of measuring had not a quite elaborate use of words interwoven into it. (Anscombe 1976, 117)

In order to know what "writing" means and use the term correctly in describing someone else's action or your own, you need to have been introduced to handling certain objects (pencil and paper, chalk and slate), to have learnt to imitate letter shapes and name them, to form words and read them, etc., and then to have been introduced to the uses of writing (making a list, signing, writing a postcard…). Writing does not simply mean drawing, nor leaving a trace on a surface, even if it is with an ink pen, nor simply tracing shapes that look like letters. To understand what it means to "write", you need to have been admitted into a whole tangle of shared practices – a social world – that form a way of life in which writing occupies a certain place and is of some interest for people.

Generally speaking, by being educated in a human form of life, we learn to identify and name various activities, their characteristic ends and results, and the elements of the world necessary for their accomplishment: baking bread, cooking, driving a bus, taking a tram, thanking or greeting someone, nursing someone, buying and selling… At the same time, we learn to identify the role or status of the agents who perform those activities (the baker, the driver, the doctor, the shop assistant, and so on) and to recognise the patterns of actions and reactions that fit together in them. But we also learn to act on our own in accordance with some of these roles and motives. The ordinary intelligibility of actions comes not from something in the mind of the agent, a kind of mental (inner) supplement to his conduct, but

from the fact that these actions implement instituted ways of doing things, i.e. ways that are both received and authoritative, which we learn to recognise and apply ourselves as agents (Descombes 2014, 295-313). This helps us to understand why the active search for an agent's reasons for acting is not central to ordinary situations, but 'peripheral': most of the time, other people's reasons are obvious to me. If the baker opens his till after I have handed her a note, it is to give me the change for the bread I am buying; if the waiter at the restaurant hands me a menu, it is for me to choose my dish because I am coming for lunch; and so on. The lack of understanding and the need for explanations arise when an incident interrupts the normal course of events. For example, a man suddenly gets up from the table in the middle of lunch and leaves the restaurant; we naturally wonder what has bitten him, but we don't wonder why the other diners stay eating at their table.

How then can we understand the role and use of psychological concepts such as 'believe', 'want' or 'intend', if understanding others does not necessarily require their projection? A complete answer to this question actually involves a whole philosophy of psychology, of the kind Wittgenstein developed in his later writings. It is beyond the scope of this article to set out all the details. But I can at least indicate a few elements here.

First of all, let us emphasise once again that the content of an intention or a belief cannot but refer to the world of the agent: they can only have a content that the agent is able to think or express, through his language or his conduct, because he participates in this normative practical background made up of institutions and customs (PI, I, § 337). As it has been said, it requires both practical and conceptual training. This suggests an important point: psychological concepts like 'intending' or 'believing' are logically dependent in their use on those by which we identify things, facts, activities and events.

How do attributions of intention work indeed, and what purpose do they serve? To find this out, we need to retrace the language game and its roots. A child gradually learns to say what he is doing – and this, because the adults around him talk to him, telling him what he is doing, asking him things, encouraging him, teaching him thereby how is called what he is doing. In this way, he becomes able to answer questions about his current activity: "I'm playing", "I'm drawing a little man", etc., as well as to use the question "What are you doing?" himself. In his answers, he indicates the point of his current activity, possibly associated with a criterion of its achievement. A further stage consists in being able to say what he is about to do: "I'm going to ride my bike", "I'm writing a postcard to Grandma", as well as being able to describe what others are up to. He also progressively learns to articulate the complexity of what he is doing ("I'm writing to thank her for her present") while learning to answer the question

"Why?", which he also applies to others. Following these lines, the language of attributing intention to others ("She intends to do A") can be seen as an extension of the possibilities for describing other people's actions, logically based on their possible or actual first-person expression. Attributing an intention involves mastering a particular linguistic technique, that of indirect style discourse, which makes it possible to report to an addressee the words by which an agent expressing himself, as the child of my former example does, could declare what he is doing or intends to do (see Descombes 2004, 38). Of course, it is not necessary for the agent to have uttered the words, either to someone or for himself. On the other hand, as we have seen, he must have the necessary conceptual resources to be attributed the corresponding intention.

By using descriptions that the agent himself might give of his action ("I am doing A"), this technique makes it possible to identify a complex action from its end, and to articulate the observer's and the agent's points of view when they diverge. We generally identify an action by its intended result, in the same way that we describe processes by reference to their end (Anscombe 2000, 39, § 2). But sometimes the agent's intended result is not the one that actually takes place and which the others are able to observe directly. Suppose someone takes some eggs from a box but clumsily drops them on the floor; it is true to say that he has broken some eggs, but false to say that he has made an omelette, even though 'making an omelette' is the description under which he intended to act, an action which then appears to have failed. The technique of indirect discourse opens up the possibility of distinguishing, in relation to the same action, between a description under which the agent thinks what he is doing or going to do, and a description of what he is doing that is not linked to what he could have said about his own action – a distinction that opens up the possibility of talking about the degree of accomplishment of the action and its failure (see Thompson 2008, 122-8).

The preceding remarks, without exhausting the topic, should suffice here to make plausible the idea that the functioning and the descriptive use of psychological concepts like 'intention' or 'belief' must be grasped, not from the picture of the mind as interiority, interpreted literally, but in relation to situations of interaction between agents and to phenomena of first-person expression. Whatever the full elucidation of their logic, understanding others depends above all on the fact that we have a common conceptual repertoire, both practical and linguistic. It is this conceptual know-how that enables us to identify the actions of others (according to their degree of achievement) and to interact with them.

4   Tuning in with Others

A question arises, however: how is it that, while having the same historical background, both conceptual and practical, is not always enough to understand others? Why is there misunderstanding? And what does it show us about what understanding is?

There are a number of different situations that must be taken into account here. Firstly, 'misunderstanding' can refer to the lack of understanding coming from the ignorance of some important elements of context; for example, if a man sitting at a table in a restaurant suddenly gets up and leaves, we will not understand his behaviour until we know more about the circumstances and his state of mind (has he just remembered an important appointment? Or has someone insulted him? Etc.). The word can also refer to the simple fact of being mistaken about what someone is doing or saying. A mistake or misunderstanding is a kind of hitch in the interaction, a failure, which can nevertheless be repaired. But misunderstanding might be of a more radical nature and mark the failure of the interaction or even its impossibility; this is the situation Wittgenstein refers to when he writes:

> It is important for our approach, that someone may feel concerning certain people, that he will never know what goes on inside them. He will never understand them. (Englishwomen for Europeans.) (CV, 84)[4]

> We also say of a person that he is transparent to us. It is, however, important as regards our considerations that one human being can be a complete enigma to another. One learns this when one comes into a strange country with entirely strange traditions; and, what is more, even though one has mastered the country's language. One does not *understand* the people. (And not because of not knowing what they are saying to themselves.) We can't find our feet with them. (PI, II, § 325)

Our reaction does not stem from the fact that something is hidden from us, even if this picture sometimes comes to us spontaneously to express our confusion. It stems from the fact that we can't really relate to someone. Such an experience, however, is not specific to finding ourselves in a foreign society, whose concepts and ways of life are at odds with our own; it also occurs within our own society. As Peter Winch (1997, 202) has pointed out, we can feel completely alienated by our contemporaries' interest in football, say; Winch also mentions the British philosopher Robin Collingwood who, in his autobiography,

---

[4]   For a complete elucidation of this remark, see Schroeder 2019.

describes the feeling of being out of step with the kind of philosophy practised by his Oxford colleagues. The boundary between what is and what is not 'alien' is actually rather fluid.

The fact is that, as Winch says, a culture in the anthropological sense of the word is not a "seamless web" (1997, 198). In the course of our education, all of us are not introduced exactly to the same ways of living. We are exposed to different facets of the same culture and, in so doing, we are led to take different things for fundamental and important.[5] Hence, I may not understand the distress of people who are unable to attend a football tournament, say. Their expressions of despair seem completely incongruous to me, because I don't understand their relation to football, the place it occupies in their way of life taken as a whole, or, as Winch puts it, the "point" of their passion for football.[6] Winch however suggests that it is sometimes possible to overcome this misunderstanding: to that effect, I need to find connections between their way of life and mine, by means of which I can find my own an analogue of the interest they attach to this activity and the role it plays in theirs. If I play a sport myself, I might have an idea of the passion it can arouse – but the analogy might not be enough to understand the importance of attending matches in person, getting together with others to talk about it, or even the feeling that one's own life might be deprived of value if one's favourite team lost the tournament. Someone who, on the other hand, doesn't particularly like football but passionately follows his basketball team's championships would probably have a better understanding of these aspects than I do. Generally speaking, the possibility of understanding others will depend on the way in which our lifestyles overlap and lend themselves to the building of enlightening analogies. We are far from a theoretical inference or from an effort at simulation.

However, the divergence of lifestyles is not the only cause at stake. This is sometimes overlooked by sociologists or anthropologists who tend to describe acculturation as a simple process of "internalising norms" (Winch 1997, 198). Here, Winch's thoughts echo those of Wittgenstein on the importance of individual spontaneity and "primitive reactions" in learning to follow a rule.[7] For instance, imagine two people, *A* and *B*, such that *A* teaches *B* to write a sequence of signs in a given order, such as the sequence of natural numbers. *A* writes

---

[5] See also Z, §§ 387-8.

[6] On the use of this expression, see Le Du 2013. On the example of football, see also Lyas 1999, 74-5.

[7] Cf. CV, 36: "The origin and the primitive form of the language game is a reaction; *only* from this can the more complicated forms grow. Language – I want to say – is a refinement, 'in the beginning was the deed'".

the sequence from 0 to 9 for *B*'s attention, and *B* has to copy it down. Wittgenstein declares:

> At first, perhaps, we guide his hand in writing out the series 0 to 9; but then the *possibility of communication* [*die Möglichkeit der Verständigung*] will depend on his going on to write it down by himself. (PI, I, § 143)

And the same applies throughout the learning process: "[T]he effect of any further *explanation* depends on his *reaction*" (PI, I, § 145).

Sharing common concepts, mastering a technique for using words, generally presupposes a certain regularity in reactions to learning. But this also applies to their projection into new uses and their application to new situations. The rules for using a word are not like rails on which we would set off once and for all, and which would determine all its possible projections; the possibilities of meaning change and expand with our practice itself. The intelligibility of a new projection will therefore also depend on the similarity of people's reactions. So it is when we introduce a new metaphor, a witty remark or a line of humour: they will be intelligible to others only if they are able to see what the person uttering them sees in them which gives this use of words its "point", i.e. both its meaning and its value. And this variety of possible individual reactions extends its effects to existing cultural forms: some will elicit no significant response from us, and we will therefore be in great difficulty to find any meaning in them. What is more, the divergence of our reactions can lead us into conflict – a conflict, says Winch (1997, 198), which is even characteristic of certain areas of life: morality, politics, religion.

As Severin Schroeder (2019, 183-4) points out, there is therefore a non-intellectual dimension to understanding others which is rooted in individual spontaneity; understanding others is also a matter of affinities, of sharing dispositions that are both moral and aesthetic, i.e. that concern what is valuable and what is not. (It should be noted, however, that understanding does not presuppose agreement or unison: we can very well get along in a conflictual mode, like those couples who share a taste for quarrelling, for example, according to an eroticised perception of confrontation.) The emphasis placed on the diversity of individual agreements, so to speak, allows us to see that misunderstanding cannot be apprehended solely as a case of failure of our cognitive capacities or of missing knowledge, but that it is an irreducible possibility, immanent to human relations, the flip side of the plasticity and indefinite nature of our practices and concepts (see also Hacker 2023, 96-8). For all that, incomprehension can be overcome – sometimes, at least, when we are able to find the right analogies and if we are also inclined to show goodwill. But there is no guarantee that it will be, nor even that it can be.

5 Conclusion

When the meaning of an agent's conduct eludes us, we are prone to think that we might find it 'in his mind'. This inclination still more or less implicitly governs much philosophical thinking about understanding others, by focusing attention on the idea that we should be able to account for 'the access' we have to the mind of another. To understand someone, it is assumed, is, first and foremost to be able to rationalise his behaviour, which implies discovering the content of his desires and beliefs. But by what process? Some believe it is an inference based on the possession of a theory, others a form of simulation through which we find these contents within ourselves before projecting them onto others. Yet interpersonal understanding is not the result of access to content, as if it were a matter of discovering something fundamentally hidden. What's more, this way of looking at the problem reduces understanding to a mere question of knowing someone else's reasons, what they want and what they believe; but this overlooks the fact that reasons for action, even if they are explicit and therefore known, may not be understood.

In a sense, the emphasis in the debates on identifying the content of an agent's reasons and the operations that make this possible, has contributed to obscure a more fundamental point: understanding others is not a specific cognitive achievement, but the manifestation of a shared know-how. In his masterwork *The Concept of Mind*, Ryle (2000, 53) had already taken a step towards an elucidation of this kind:

> Understanding is a part of knowing *how*. The knowledge that is required for understanding intelligent performances of a specific kind is some degree of competence in performances of that kind. The competent critic of prose-style, experimental technique, or embroidery, must at least know how to write, experiment or sew. [...] Of course, to execute an operation intelligently is not exactly the same thing as to follow its execution intelligently. The agent is originating, the spectator is only contemplating. But the rules which the agent observes and the criteria which he applies are one with those which govern the spectator's applause and jeers.

In his formulation, Ryle put the emphasis on technical operations: if I can see myself in what someone is doing, it is because I myself know how to do part of what he is doing. But does all behaviour boil down to the application of a technique? In this article, I have tried to show that Wittgenstein's philosophy provides the means to give a proper formulation to this intuition in a much broader way: understanding others depends in the first place on conceptual know-how, on the possession of concepts that enable the agent to think about and describe

his own action – but they enable him to do so because they structure the action itself, because they are nothing other than the rational order the agent is able to give to his conduct by virtue of his insertion into some social world, made up of norms, roles, rules and practices. To be able to describe your own action by saying "I'm writing", you need to know how to write: mastery of language is thus interwoven with non-linguistic activities, in a huge diversity of ways. And such an action is intelligible to another (interacting with the agent or in the position of an observer) insofar as this other participates in the same social world and is himself, as a result, familiar with its practices.

Such a perspective leads us to re-consider the logical functioning of concepts (such as intention, desire or belief), the mastery of which is at the heart of commonsense psychology. From this point of view, the psychological concepts used to articulate an agent's attitude towards what he holds to be reasons (the end he pursues, the things he holds to be true and on the ground of which he acts) play an auxiliary role in extending descriptions: they make it possible to enrich the minimal scenario suggested by the description that an observer is immediately able to give about what is going on ("she writes", "he takes the tram"…) by crediting the agent with thoughts that he could himself express if questioned; but this in no way implies that psychological concepts are intended to designate 'mental contents', in the sense of objects of knowledge inaccessible for the observer. The logic of psychological concepts does, of course, call for careful study; but to consider that the understanding of others can be elucidated on the basis of their use alone simply misses the point.

However, participation in the same social world is not enough to bring about understanding of another. There is a non-intellectual dimension to understanding which is rooted in individual spontaneity, that can get in the way when we are not sensitive to the same things. But above all, as Winch emphasised in the wake of Wittgenstein, incomprehension remains an irreducible possibility, immanent in human relations; for participation in a common world presupposes a sufficient convergence of reactions in the learning and subsequent application of words and concepts, which nothing can absolutely guarantee.

References

Anscombe, G.E.M. (1963). "Events in the Mind". Anscombe, G.E.M., *Metaphysics and the Philosophy of Mind. Collected Papers*, vol. 2. Oxford: Blackwell, 1981, 57-63.

Anscombe, G.E.M. (1976). "The Question of Linguistic Idealism". Anscombe, G.E.M., *From Parmenides to Wittgenstein. Collected Papers*, vol. 1. Oxford: Blackwell, 112-33.

Anscombe, G.E.M. (2000). *Intention*. 2nd edition. Cambridge, MA: Harvard University Press.

Churchland, P. (1981). "Eliminative Materialism and the Propositional Attitudes". *Journal of Philosophy*, 78(2), 67-90. https://doi.org/10.5840/jphil198178268.

Coplan, A.; Goldie, P. (eds) (2011). *Empathy. Philosophical and Psychological Perspectives*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199539956.001.0001.

Davidson, D. (1963). "Actions, Reasons and Causes". Davidson, D., *Essays on Actions and Events*. Oxford: Oxford University Press, 12-24. https://doi.org/10.1093/0199246270.003.0001.

Davies, M.; Stone, T. (eds) (1995). *Folk Psychology. The Theory of Mind Debate*. Cambridge, MA: Blackwell.

Descombes, V. (2004). *Le Complément de sujet*. Paris: Gallimard. https://doi.org/10.14375/np.9782072765506.

Descombes, V. (2014). *The Institutions of Meaning. A Defense of Anthropological Holism*. Transl. by S.A. Schwartz. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/harvard.9780674419971.

Fodor, J. (1987). *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.

Hacker, P. (2018). *The Passions. A Study of Human Nature*. Hoboken: Wiley Blackwell. https://doi.org/10.1002/9781118951866.

Hacker, P. (2023). "Other Minds, Other People, and Human Opacity". *Ratio*, 36(2), 87-98. https://doi.org/10.1111/rati.12362.

Hutto, D. (2004). "The Limits of Spectatorial Folk Psychology". *Mind & Language*, 19(5), 548-73. https://doi.org/10.1111/j.0268-1064.2004.00272.x.

Hutto, D.; Ravenscroft, I. (2021). "Folk Psychology as a Theory". *Stanford Encyclopedia of Philosophy*. Edited by E. Zalta. https://plato.stanford.edu/archives/fall2021/entries/folkpsych-theory/.

Le Du, M. (2013). "Anthropologie et relations internes". Ambroise, B.; Chauviré, C. (éds), *Le Mental et le Social*. Paris: Editions de l'EHESS, 275-310. https://doi.org/10.4000/books.editionsehess.12069.

Lyas, C. (1999). *Peter Winch*. Teddington: Acumen. https://doi.org/10.1017/upo9781844653201.

Rizzolatti, G.; Sinigaglia, C. (2008). *Mirrors in the Brain. How our Minds Share Actions and Emotions*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780199217984.001.0001.

Ryle, G. (2000). *The Concept of Mind*. Prefaced by D. Dennett. London: Penguin Classics. https://doi.org/10.7208/chicago/9780226922652.001.0001.

Sandis, C. (2019). "Making Ourselves Understood. Wittgenstein and Moral Epistemology". *Wittgenstein-Studien*, 10(1), 241-59. https://doi.org/10.1515/witt-2019-0015.

Schroeder, S. (2019). "God, Lions, and Englishwomen". Padilla Galvez, J.; Gaffal, M. (eds), *Human Understanding as Problem*. Berlin: De Gruyter, 171-84. https://doi.org/10.1515/9783110613384-011.

Spaulding, S. (2018). *How We Understand Others. Philosophy and Social Cognition*. Abingdon: Routledge. https://doi.org/10.4324/9781315396064.

Spaulding, S. (2020). "What is Mindreading?". *WIREs Cognitive Science*, 11(3), 15-23. https://doi.org/10.1002/wcs.1523.

Stueber, K. (2006). *Rediscovering Empathy. Agency, Folk Psychology, and the Human Sciences*. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/5833.001.0001.

Stueber, K. (2018). "Empathy". *Stanford Encyclopedia of Philosophy*. Edited by E. Zalta. https://plato.stanford.edu/archives/spr2018/entries/empathy.

Thompson, M. (2008). *Life and Action. Elementary Structures of Practice and Practical Thought*. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/9780674033962.

Winch, P. (1997). "Can We Understand Ourselves?". *Philosophical Investigations*, 30(3), 193-204. https://doi.org/10.1111/1467-9205.00038.

Wittgenstein, L. (1982). *Last Writings on the Philosophy of Psychology*, vol. 1. Edited by G.E.M. Anscombe; G.H. von Wright; H. Nyman, transl. by C.G. Luckhardt; M.A.E. Aue. Oxford: Blackwell. [LWI]

Wittgenstein, L. (1988). *Lectures on Philosophical Psychology 1946-1947*. Edited by P. Geach. Chicago: University of Chicago Press. [LPP]

Wittgenstein, L. (1998). *Culture and Value*. Edited by G.H. von Wright; H. Nyman, revised by A. Pichler, transl. by P. Winch. Oxford: Blackwell. [CV]

Wittgenstein, L. (2007). *Zettel*. Edited by G.E.M. Anscombe; G.H. von Wright, transl. by G.E.M. Anscombe. Berkeley: University of California Press. [Z]

Wittgenstein, L. (2009). *Philosophical Investigations*. Revised 4th edition. Edited by P.M.S. Hacker; J. Schulte, transl. by G.E.M. Anscombe; P.M.S. Hacker; J. Schulte. Chichester: Wiley-Blackwell. [PI]

Wittgenstein's Methodology of Gestalt Psychology

Michel Ter Hark

Vrije Universiteit Amsterdam, The Netherlands

Abstract This article proposes a methodological understanding of Wittgenstein's remarks about Gestaltpsychology. Wittgenstein is not so much concerned with Gestaltpsychology proper but rather with its understanding of the nature of the problem of seeing (and seeing-as) as dealt with by British empiricism. Gestaltpsychology offers a more sophisticated physiological explanation of seeing and seeing-as than empiricism has done, yet also this explanation bypasses the (conceptual) problem. Physiological explanations are not eschewed by Wittgenstein, he even gives himself interesting physiological hypotheses. The problem with physiological explanations is that they focus on particular items in the brain as underpinning our use of concepts like seeing and seeing as, whereas they are constituted by our reactions and responses to what we see. Such reactions are embedded in language games and acquire their meaning by our 'forms of life' rather than the human brain. This interpretation is finally applied to Wittgenstein's and Köhler's explanation of social understanding'

Keywords Physiological explanation. Transitive use. Intransitive use. Organization. Social understanding.

Summary 1. Psychology and Physiology. – 2. Wittgenstein. – *3. Variety of Aspects*. – 4. Wittgenstein's Critique of Köhler's Explanation of Aspect Seeing. – 5. Wittgenstein and Köhler on Social Understanding. – 5. Conclusion.

Part 2 of Wittgenstein's *Philosophical Investigations* aims to understand the logic of psychological concepts. The main source material from which Part 2 has been drawn is to be found in the *Remarks on the Philosophy of Psychology* and the *Last Writings on the Philosophy of Psychology* which were published in the 1980s.

Wittgenstein's philosophy of psychology has not drawn the same attention as Part 1 of the *Philosophical Investigations*. Yet some topics in Part 2 have received more attention than others. In particular the notion of seeing-as or aspect seeing can count on numerous interpretations. The section concerned with aspect seeing, however, includes also related but clearly different psychological concepts and which have largely failed to draw the attention they deserve. One of them is the concept of seeing, another and related to this, seeing pictures (of human beings), such as portraits but also photographs. And finally, seeing the emotions, feelings and thoughts of other people.

In this article I will focus on Wittgenstein's analysis of the concepts of seeing and seeing-as, and how they apply to 'social understanding'. My approach will be largely historical in that I discuss these topics in the context of Wittgenstein's reading of the work of the Gestalt psychologist Wolfgang Köhler. In my *Beyond the Inner and the Outer* (1990), I presented the earliest detailed interpretation of Wittgenstein's philosophy of psychology against their background in empirical psychology. In my view, Wittgenstein's work is closer to philosophy of science, or methodology, than to what is now conceived as philosophy of mind which is predominantly metaphysical. In ter Hark 1995 I further strengthened my 'methodological' reading of Wittgenstein by interpreting his notorious remarks about mind and brain (cf. RPP I, § 90 3 ff.) in the light of Köhler's theory of isomorphism of mental states and brain states. In Wittgenstein's methodology there is no focus on how science can best proceed, as with e.g. Popper, in order to get better explanations and predictions. Rather the emphasis is on a clarification of concepts in the light of their natural history. Psychological concepts have their habitat in a natural history which relates their meaning to our physiognomy, our gestures, our ways of responding to other people, our use of instruments and samples in explaining and teaching language, etc. When these concepts are transferred to a scientific context, such as the psychological lab, much of their natural history disappears from view, but it is still what gives them their meaning. Notably the concept of seeing (or perceiving) is not a concept which has been coined by psychologists for purely scientific purposes, as is common in the physical sciences. Yet in the 1920s, especially during the rise of Gestalt psychology, the concept of seeing got increasingly used in a physiological context. Indeed, Köhler remarked that not until the physiological underpinnings of psychological processes were discovered,

psychology would remain a preparatory science at best. In Hausen and ter Hark (2013), the methodological understanding of Wittgenstein's remarks about seeing and all aspect seeing has been further deepened. There we employed the conceptual distinction between 'intransitive' and 'transitive' uses of words, which Wittgenstein outlines in the *Brown Book*, to throw light both on his critique of and alignment with Köhler. In this essay, therefore, I will continue the work by Hausen and ter Hark (2013), and elaborate on the misleading effect physiology has (had) on the concepts of seeing and seeing-as in the context of early twentieth century science and philosophy. In addition, I add a new interpretation of Wittgenstein's often overlooked remarks on facial expression and social understanding which also have Gestalt psychology as their background.

1 Psychology and Physiology

According to the British Empiricists, Locke and Berkeley, the senses are only capable of registering form and colour, the intellect being the instrument for all other visual aspects. In nineteenth century psychology, this dichotomy was developed into a division between sensation and perception. Köhler has called this approach the 'Meaning Theory'. An assumption of this theory is that sensation reveals simple and neutral sense data. Sure enough, as adults we do not have such virginal impressions in everyday life, but the demarcation of the adult's visual field into segregated units is the result of learning. By contrast, the original visual field is a mere mosaic of sensations. To explain visual percepts, the empiricist invokes associations or previous knowledge.

When we look at the desk in front of us, we thus perceive a grey object because our previous interactions with objects (and in particular, our interactions with desks) impart meaning to the grey patch of colour that we would *see*. That we do not seem to see simply a grey patch of colour is due to the effects of learning.

Köhler critically observes that "little is left that would be called a true sensory fact by the Introspectionist" (1947, 83). His alternative theory of seeing and perceiving is that our visual field has an 'organisation' and this organisation is a sensory (specifically, a visual) fact, just like colour and shape. According to Köhler, it is in virtue of organisation that "the contents of particular areas [in the visual field] belong together as circumscribed units from which their surroundings are excluded" (1947, 137, 139). Köhler maintains that the segregated wholes or Gestalten are given first as visual facts, and *then* we associate meaning with them (1947, 138-9). He stresses also that when sources outside the organism stimulate the retina, the resulting 'mosaic' on the retina is not itself already organised into Gestalten. Instead, the nervous system responds to the retinal stimulation,

and various Gestalten in the visual field can thereby result (1947, 160-2). Sensory organisation, therefore, constitutes a characteristic achievement of the nervous system.

Thus, for Köhler 'organisation' or 'form' is the primitive unit of perception. Form or organisation is *seen* as much as colour and shape are seen. Organisation takes its properties from 'electric brain fields', or underlying configured brain processes.

Consider some of his examples in more detail. Köhler refers to maps of countries, or to charts of ships captains. On these maps the sea tends to the appearance of which the land has on ordinary maps. The contour of the land on maritime maps is the same as it is on a map we use when touring through the countryside, which means that the geometrical line which separates land and water is normally projected on the retina.

> None the less, when looking at such a map, says, of the Mediterranean, we may completely fail to see Italy. Instead we may see a strange figure, corresponding to the area of the Adriatic, and so forth, which is new to us, but which happens to have shape under the circumstances. (Köhler, 181)

He concludes that to have shape is a peculiarity which distinguishes certain areas of the visual field from others which have no shape in this sense. So long as the Mediterranean has shape, the area corresponding to Italy has no shape.

The retinal stimuli constitute a mere mosaic, in which no particular areas are functionally aggregated and shaped. These stimuli as such do not tell us which organisation of the visual field will be prominent and which will fade into the background. Only when we take into account brain fields and their principles of organisation can we predict which particular organisation will result.

A further example is the figure of two different shapes, either that of a cross consisting of four slender arms, or that of another cross which consists of the four large sectors. So long as the former shape is before us, the area of the latter is absorbed into the background, and its visual shape is non-existent. When the latter shape emerges, the former disappears. Köhler concludes that

> in both cases, the oblique lines are boundaries of the shapes which are seen at the time. They belong to the slender cross in the first case, and to the large cross in the second. (1947, 183)

And in an earlier treatment of a similar figure he says:

> Now the lines which in the first object belong together as boundaries of a narrow sector are separated; they have become

boundaries of large sectors. Clearly, the organization of the pattern has changed… (1947, 171-2)

Thus, to Köhler the change of organisation which occurs when we report a change of our visual impression upon looking at the figure of the double cross is a change of visual reality as we also experience such a change when we are facing a change of shape or a change of colour. It is a change of visual reality in the sense that we (or our brain) *group* the lines in different ways.

Despite Wittgenstein's greater sympathy for Gestalt psychology than for empiricism, their common physiological way of explaining problems concerning seeing is rejected by him. Indeed, the problem situation created can hardly have satisfied him. On the one hand there is empiricism which claims by appealing to the physiology of the retinal image that colour and shape are the only items of perception and that psychological states like emotions are a matter of interpretation. On the other hand there is Gestaltpsychology which claims also by appealing to physiology (of brain processes), that we do see emotions because we see organisation as much as we see colour and shape. To be sure empiricism notices a difference, a difference between seeing colours and shapes and seeing emotions, depth and other phenomena. The question, however, is whether they grasp the nature of the difference?

And Köhler may be right when he observes that empiricism is in conflict with the common, or as he puts it, the naïve view of seeing. However, what does he understand by the naïve view of seeing? It seems as if he assumes that built into common sense is a theory of perception. This is what Wittgenstein explicitly rejects. There is no general theory of perception built into common sense. There are only concepts. It is to the study of these concepts that one has to turn in order to solve the problems of perception.

2 Wittgenstein

It is in particular Köhler's understanding of the problem the empiricist psychologists wrestled with that is the target of Wittgenstein in the first volume of *Remarks on the Philosophy of Psychology*. And because, according to Wittgenstein, Köhler fails to understand the nature of the problems concerning seeing as tackled by empiricism, his own alternative approach fails.

I first discuss Wittgenstein's comments on empiricistic theories of perception.

Wittgenstein concedes that the psychologist has identified an important meaning of the verb 'to see', namely, what is seen is what can be inferred from the retinal image. What is seen, is that of which

one can make an ideal and precise representation. At many occasions Wittgenstein remarks that our gaze continually wanders when we look at objects or scenes, for instance, streaming water (cf. RPP I, § 1080). The point of these remarks can be made clear when considering our perception of people's faces. Especially when we look at another person's face our gaze wanders and our impression of the facial features and the contours of the face consists largely out of edges and subtle transitions of colours and their shadings. A drawing of our impression of the other person's face would not contain the above-mentioned fuzzy areas. Hence, what is called an exact representation of what is seen would always leave out aspects that are truly characteristic of what we see. What then is the use of the ideal of an exact representation?

And how about the use of the concept of interpretation? To be sure there are clear cases of seeing something and interpreting it. Wittgenstein gives the example of a blueprint of a triangle. One may give someone such a blueprint asks the person to hang a triangular shape on the wall with an apex as the upper part. Here the person is not seeing the blueprint as a triangle but he interprets it. When we interpret, Wittgenstein would say, we make a conjecture, we express a hypothesis, which may subsequently turn out false (ter Hark 1990, 179). But in the cases discussed by Köhler in his chapter on sensory organisation, and other chapters, what is called by him "seeing" (and by Wittgenstein "seeing-as") there is not only no fitting together of pieces, but there is no hypothesising either, no verifying, nor falsifying. When we look at the figure of the sea chart our experience of the switch of aspect, i.e. the visual emergence of the Mediterranean and the disappearance of Italy, and vice versa, our experience has 'genuine duration'. And this is one reason why it is legitimate to speak of 'seeing', as Köhler does, rather than seeing plus interpreting as Introspectionism does.

Despite this commonality Wittgenstein and Köhler approach aspect in distinctly different ways. In what respect different is hinted at in this remark:

> "When you get away from your physiological prejudices, you'll find nothing in the fact that the glance of the eye can be seen." Certainly I too say that I see the glance that you throw someone else. And if someone wanted to correct me and say I don't really *see* it, I should hold this to be a piece of stupidity. (RPP I, § 1101)

Köhler thinks that overcoming the empiricist prejudice that real perception remains true to the patterns of the retina is the way to explaining as well as describing real perception, including the role of wholes in real perception. It is only by according the organisation of the visual field a role that real perception can be explained. Changes

in one's visual field, like those occurring during aspect seeing, can be accounted for only by the physiological underpinnings of processes of organisation.

Although Wittgenstein has no problems with causal explanations of mental phenomena in terms of association, he believes that the problem hand – e.g. can we really see a human gaze or do we just see form and colour – is to be solved in a different way. A causal explanation is as much in need of a (conceptual) clarification as is the phenomenon of seeing and seeing-as itself. He has two more specific reasons for this insight. For one, he notices that Köhler wants to treat everything in a uniform way and explain all the Gestalten that we see by the notion of an organised visual field that is on a par with colour and shape. For another, Wittgenstein notices an ambiguity in Köhler's notion of organisation. Following ter Hark (2011) and Hausen and ter Hark (2013), I will explain this ambiguity by means of the distinction between transitive and intransitive use of words that Wittgenstein makes in the *Brown Book*.

3 *Variety of Aspects*

I start with a brief overview of the kinds of aspects Wittgenstein distinguishes in his writings and the lecture notes by his students between 1945-47. In (ter Hark 1990), I distinguished between optical aspects and conceptual aspects. Optical aspects can switch automatically, almost like after-images. Conceptual aspects require the use of words in order to convey that and how one experiences them. In the *Philosophical Investigations*, Wittgenstein speaks of organisational aspects, thereby referring to what I have termed optical aspects. An example of the optical aspects is the picture of a series of points or dots at equal distances from each other: ….. The one who perceives the figure is asked to see them grouped as 2,1,2, or as 1,2,3,4, and then yet another one. In such cases to describe one's changed visual impression in terms of a change of organisation or of grouping is quite apt. Closely related to this case is the figure the 'double cross', of a black cross on a white background and of a white cross on a black background. In this case the aspects can be reported

> simply by pointing alternately to an isolated white and an isolated black cross. One could quite well imagine this as a primitive reaction in a child even before it could talk. (PI II, 217)

Clearly, the black and the white cross need not, and typically do not, switch automatically, hence they are not straightforward optical aspects with their characteristics of after-images. But they do not need the help of concepts in order to be experienced. Even prelinguistic

children might be pointed out the switch of the black cross and the white cross

Things are different with the famous duck/rabbit figure. The ability to see the ambiguous figure as a duck, or as a rabbit, does not come off the ground simply by pointing as in the preceding case. One needs already to be "conservant with the shapes of these animals" before one can say that one sees it so or so (cf. PI II, 217).

I believe there is here a connection with another dimension of aspect seeing that Wittgenstein observes but Köhler does not. Aspect seeing, Wittgenstein notes, are subject to the will. It makes sense to give the command "See this as a rabbit" but it makes no sense to say "now see this leaf green" (cf. PI II, 213). Wittgenstein does not explain his reason for bringing this difference between aspect seeing and seeing under our attention. But let us counter factually suppose that it would make sense to command someone to see this leaf green. In that counterfactual situation we could not learn the meaning of colour words by ostensible definition, i.e. by making a pointing gesture at a green leaf and saying "that is what we call 'green'". In the case of the duck/rabbit, though, one could point to colours and shapes but not to a rabbit, or to a duck. To see it as a rabbit is not a matter of what but of *how* we see what we see. And how one sees it is to make a comparison, a comparison with e.g. pictures of rabbits. Clearly the ability to make comparisons is voluntary.

The voluntariness of aspect seeing is not to be seen as a direct refutation of Köhler. He simply has failed to take account of the role of language in aspect perception and instead proceeds from association and brain processes. Yet there is a difference between association, which is involuntary, and the role of language in aspect perception. It may be true that eye movements are involved in noting aspects, and it may be true that association works in the background. But the point is that one may see a certain aspect, e.g. a duck in the duck/rabbit figure, just by *saying* or pronouncing the word 'duck'. Hence, language and therefore language games, have a role to play, even at this transitional point where physiology seems to take over psychology.

This role of language is even more prominent in a number of different examples of aspect seeing. Wittgenstein gives the example of the figure of a triangle and the question to see the triangle as if it is hanging from its apex or as if it is standing on its base (cf. PI II, 200). The person who receives this command normally understands it immediately and also asks for no other explanation of what is meant by 'hanging' or by 'standing'. The situation is not be explained in the following way. There you see a real physical hanging object and what you see on the picture resembles it so that is makes sense to say that it is hanging. These words are meant in their ordinary sense, as possible states of a physical object. In particular they do not seem to be an indirect description of what one sees as when we speak of the colour of

blood rather than of red. The use of standing or base are *essential* here and hence being able to apply them in other situations is an essential condition for underhand applying them in the case of aspect seeing.

Again the case of the triangular hanging figure is different from both double cross and the duck/rabbit figure. In the latter two cases it is possible that someone fails to note the ambiguity and takes, e.g. the duck/rabbit for a rabbit, but it is not possible to take the bare triangular figure for the picture of an object that has fallen over: "To see this aspect of the triangle demands *imagination*" (PI, 207).

4 Wittgenstein's Critique of Köhler's Explanation of Aspect Seeing

Köhler's physiological departure to the question of the nature of seeing is especially dominant in his treatment of aspect seeing. His description of our seeing the duck/rabbit figure would be as follows. When we look at the ambiguous figure and see first the rabbit and then the duck we first of all experience that a real rabbit looks like X and then that a real duck looks like Y, and that the ambiguous figure switches between X and Y, and back again from Y to X. Our visual experience thus changes, from X to Y and conversely from Y to X. To explain this change in our visual experience Köhler invokes the concept of organisation. When we see the figure as X our visual experience is organised differently from the situation in which we see the ambiguous figure as Y.

Wittgenstein rephrases Köhler thus: an example of organisation aspects is when I suddenly see the solution of a puzzle picture. Before there were only branches and twigs, now there is suddenly a human shape. "My visual impression has changed and now I recognize that it has not only shape and color but also a quite particular 'organization'" (PI II, 196).

Wittgenstein's initial response to this explanation in terms of organisation is that it makes no sense to say this. This 'scientific' explanation is not better than the naïve view that the ambiguous figure would *move* when we undergo an experience of aspect seeing. Although Köhler does not hypothesise inner mental objects that change, he still falls victim to a 'category mistake' for it seems that the only (hypothetical) change that may occur when we experience a change of aspect is a physiological change. But it is also obvious that a (hypothesised) physiological change is not what we see when we say that we see something as. Put otherwise, what we see does not change and what does (possibly) is not what we see.

The hypothesised physiological change therefore has no bearing on the solution of the conceptual problems concerning seeing and seeing as. As Wittgenstein puts it nicely:

> You have now introduced a new, a physiological, criterion for see-
> ing. And this can screen the old problem from view, but not solve
> it. (PI II, 212)

A fairly standard view of Wittgenstein's comments on the notion of
organisation in his *Philosophical Investigations* goes as follows. Were
we to represent our experience of a change of aspect by means of two
drawings, one of the situation in which we note the rabbit and one
of the situation in which we note the duck, the drawings would show
no differences at all, they would be exactly the same. It is also fair-
ly standard in the secondary literature to continue as follows there-
by drawing on some remarks Wittgenstein subsequently makes. If
someone, notably Köhler, concedes that change of organisation is not
the same as change of colour or shape then change of organisation
becomes an object which is vulnerable to Wittgenstein's arguments
against private inner objects.

Commentaries who argue in this way fail to see that Wittgenstein's
discussion of Köhler has not only a negative side but also a positive
one. The negative and the positive side, however, are interdepend-
ent, hence I once more have to discuss the negative side in order to
explain what is positive about Köhler's use of organisation. I will do
this by drawing on earlier work by Hausen and ter Hark (2013). I will
show how Wittgenstein's negative and positive arguments rely on a
subtle and rarely discussed distinction that Wittgenstein makes be-
tween the transitive and intransitive use of terms.

Wittgenstein develops the transitive/intransitive distinction in the
Brown Book, immediately prior to a discussion of aspect seeing. His
example is the word particular, as when we say, "The face has a par-
ticular expression". On the one hand we may mean 'particular' in a
transitive sense, as when we say "This face gives me a particular
impression which I cannot describe". We also may mean it intransi-
tively as when we say: "This face gives me a strong impression" (cf.
BBB, 158). So, in the transitive case, the word 'particular' is used as
a precursor to a further specification. To the question 'Peculiar in
what way?', an answer can be given that explains this way in differ-
ent words. In the intransitive case, however, the word 'particular'
is used for emphasis, hence there is no further specification or com-
parison to be made.

Transitive and intransitive uses of words are not always easy to tell
apart, however. This is especially true when the sentences in ques-
tion involve what Wittgenstein calls a "reflexive construction" (BBB,
159-61). The use of words in a reflexive construction is intransitive
yet appears to be a special case of a transitive use (namely, the re-
flexive constructions appear to be comparing something with itself
or describing something by appealing to the thing itself). The impor-
tant feature of reflexive constructions is that the sentences can be, as

Wittgenstein says, 'straightened out'. What he means by this is that the sentences seem to involve a comparison or description that loops from an object back to itself. But when the sentences are straightened out, we see that there is no loop. Rather, the sentences involve only an intransitive use; that is, they involve emphasis, not comparison or description. For instance, Wittgenstein says that "That's that" is a reflexive expression. Although "That's that" appears to compare a thing to itself, it can be straightened out as "That's settled" and in fact is used to emphasise the finality of the situation.

Wittgenstein's objective in discussing these distinctions is to point out that confusion can arise if intransitive uses are not properly distinguished. Hausen and ter Hark have argued that Köhler's notion of organisation falls into the transitive/ intransitive trap. Specifically, it looks as if Köhler is using the term organisation transitively when he speaks about the organisation of the visual field. But actually what is involved is an intransitive use.

As explained earlier, Köhler defines the organisation of the visual field as a sensory fact in addition to colour. So, when we experience a change in aspect of (for example) the pie figure, there is a change in the sensory facts, namely, the organisation of our visual field changes.

Wittgenstein wonders whether the change of our visual impression can be attributed to change of organisation as Köhler would have it. He seems to take the use of 'change of organisation' from other linguistic situations. As Wittgenstein notes:

> "The organization of the visual image changes" has not the same kind of application as: "The organization of this company is changing". *Here* I can describe *how it is*, if the organization of our company changes. (RPP I, § 536)

That is, a company's organisation may be described by a flowchart that shows the company's hierarchy and structure. It makes sense to ask, "How did the organization change?", and the response could involve pointing to changes in the flowchart. But there is no comparable way to describe the organisation of the visual field (cf. ter Hark 1990; Hausen and ter Hark 2013, 98). We might, as Wittgenstein suggests, represent our visual impression by means of drawings. Such drawings would reflect a change in colour. Yet, these drawings will show no change when there is a change of aspect, "they will be the same before and after the theorized change in organization takes place" (LW I, § 439).

Now the sentence "The organization of my visual field has changed" seems similar to a sentence "The color of the sky has changed", yet in answer to the question "How has your visual field changed" one can say no more than "Like *this*", thereby pointing to inner (mental)

objects. But this response is not an informative further specification, and the use of 'organisation' in "The organisation of my visual field has changed" is not transitive. Rather the sentence involves a reflexive construction, and the use of 'organisation' is intransitive.

In making this argument, Wittgenstein's comments in RPP I, § 1118 (which immediately follow a remark about Köhler and the pie figure) also are relevant to discuss. Wittgenstein notes:

> Indeed, you may well see what belongs to the description of what you see of your visual. Impression is not merely want to copy shows, but also the claim for example to see this solid this other as intervening space hear it all depends on what we want to know when we ask someone what he sees. (RPP I, § 1118)

A central idea in Wittgenstein's analysis of aspect seeing is that in everyday contexts, the change in what we see is adequately described by, for instance, pointing to part of the pie and saying "I used to see this part of the figure as intervening space, and now I see it as solid". For example, if Wittgenstein was looking at the pie figure and wanted to describe a change in what he sees, he could say, 'I now see the narrow sectors as solid'. The situation is different in Köhler's case, however. Suppose that Köhler would suggest that "I now see the narrow sectors as solid" describes a change in organisation of the viewers visual field. That is, suppose that Köhler were to suggest that a (transitive) answer to "How has the organisation of your visual field changed?" is "I now see the narrow sectors as solid". In this case, the answer is not sufficient.

> The reason why it is insufficient is because Köhler needs the answer to provide more than just a description of the change in what is seen, for he intends to *explain* change in what is seen by appeal to a change in organization of the visual field. Yet, saying that I now see the narrow sectors as solid (this is how the organization has changed) does not explain why I now see the narrow sectors as solid (this is what I now see). In other words, Köhler would be claiming, in effect, "I now see the narrow sectors as solid because I now see the narrow sectors as solid", which clearly does not provide an informative explanation. (Hausen; ter Hark 2013, 101)

Viewing Köhler's situation from the transitive/intransitive perspective hence exposes yet another way that his notion of organisation is unilluminating and mystifying.

> When introducing organization as a sensory fact, he apparently assumes that his notion will have a transitive use similar to that our concept of (ordinary) organization and to that of color and shape.

But, upon inspection, we see that his notion lacks any transitive meaning at all. (Hausen; ter Hark 2013, 101)

On that note, let us turn to Wittgenstein's alternative, positive description of the use of 'organisation' with respect to aspect seeing. Wittgenstein's idea here is that a sentence like "The organization of my visual field has changed" can be useful even if it is not used as a precursor of *how* the organisation has changed. In particular, it can be used to express and thereby emphasise an experience that one has had. Like the sentence "That's that" which can be straightened out as "It is settled", the sentence "The organization of my visual field has changed" can be straightened out as "I am having an experienced that I want to express by saying 'The organization of my visual field has changed'". Moreover, this experience need not be further explicated in order for the sentence to be meaningful. As Wittgenstein says, regarding the feeling of everything being unreal "[a]nd how do I know that another has felt what I have? Because he uses the same words as I find appropriate" (RPP I, § 125). The other person knows what I am talking about not on epistemological grounds, but because we are in tune with the very verbal expression. A continuation of this use of words might be "Yes, I should like to say what you say", but not a statement to the effect that there is something which we both describe by means of the same words. "Accordingly, the inclination to say such and such is not simply a reaction but is itself the psychological phenomenon that matters" (ter Hark 2011, 516). To conclude, by speaking of the organisation of our visual field, we are intransitively emphasising an experience rather than transitively describing the visual field.

5 Wittgenstein and Köhler on Social Understanding

Köhler distinguishes between two notions of behaviour: behaviour in the physical sense and behaviour as perceived. Behaviour in the physical sense is the domain of behaviourism and physiology. Köhler admits that behaviour qua physical has nothing in common with mental processes. To the extent that the philosophical argument by analogy for the existence of other minds proceeds from the notion of physical behaviour, it can be dispensed with. The body and the behaviour of other living human beings, Köhler emphasises, is given to us "only as percepts and changes of percepts" (Köhler 1947, 221-2). Emotions, he continues, tend to express themselves in the behaviour of people as we see them. The question now is, if these expressions resemble what is being expressed? If so, he argues, the main reason for strictly indirect interpretation of social understanding would obviously be removed (223).

Behaviour as perceived by others provides us with all sorts of sensory experiences. Is there a similarity between these sensory experiences and the mental life of other creatures, Köhler asks? He proceeds cautiously by first investigating similarities between different senses that have nothing to do with mental, subjective experiences. Brightness and darkness, he argues, are attributes of both auditory and visual experiences. Again, ik fiets nu weg an object which we touch appears cool, its coolness somehow resembles visual brightness. Even words, in particular their sound, may resemble qualities of different senses. He quotes a line from the German poet Morgenstern:

> Die Möwen sehen all aus, als ob sie Emma hiessen (All seagulls look as though their name were Emma).

"The sound of 'Emma' as a name and the visual appearance of the bird appear to me similar" (Köhler 1947, 224). Köhler rejects the view that these and other synesthetic linkages are mere analogies from which nothing can be inferred about underlying facts. On the contrary, he defends the view that the analogies are all grounded in resemblances that exist between different realms of sense-experience.

From these and other examples he concludes that certain experiences of the inner and the perceptual worlds resemble each other.

As I have shown elsewhere (cf. ter Hark 2011), synesthetic experiences and their relation to language are also discussed by Wittgenstein in detail. For now it suffices to consider his comments on Köhler's reading of Morgenstern. But there is no similarity between the sound of the name Emma and the appearance of seagulls. What could the resemblance be here? It is obvious that the experience might be due to a childhood association between seeing seagulls walking lamely, and the stiffness of women called Emma. Perhaps there is even an association between a particular Emma limping out of the house at the seaside and the gait impeded by stiffness seagulls. But such associations are a far cry from noticing a resemblance between the sound of a name and a certain visual appearance. Indeed, there is no more similarity between Emma and the appearance of seagulls than between the name Beethoven and the Ninth Symphony. Hence Köhler mistakenly believes that giving an associative explanation also amounts to having described this typical use of words.

In a series of three remarks in his *Philosophical Investigations*, Wittgenstein turns to Köhler's similarity thesis concerning the inner and the outer. He begins by commenting on the question typically raised by empiricism: "For how could I see that this posture was hesitant before I know that it was a posture and not the anatomy of the animal?" (PI II, 209; cf. LW I, § 736). We know by now that this is not the epistemological problem the empiricist takes it to be but

a conceptual problem. The question amounts to a refusal to use the concepts of e.g. 'mean' or 'frightened' in describing an object of sight because these concepts do not serve exclusively for the description of what is visual. And if it is 'just' a question of choosing the 'right' concepts, why one does not take recourse to a purely visual concept for describing a mean facial expression? How easy is this? Or how difficult? In the next remark Wittgenstein suggests that theirs might be purely visual descriptions of a mean expression in the way the concepts of major and minor may provide purely auditory descriptions of music, including the emotional value of music. In the next remark he observes that psychological concepts, e.g. 'sad', can be applied to an 'outline face', such as an emoji, in the way major and minor can be used to describe music: "The epithet 'sad', as applied, for example, to the outline face, characterizes the groupings of line in a circle (Major, minor.)" (PI II, 209).

What is the difference between an emoji and a real human face as far as their expression is concerned? A picture face can be described by purely visual concepts. For instance, one can describe a nose as acute-angled, thereby giving the face a certain expression. But in the case of a human being there is no such equivalent to major and minor. And this is not because we haven't defined our concepts sufficiently sharp in order to meet the varieties of the sense experience of a human facial. The reason rather is that the concepts we use for describing a human facial expression have a different *use*. When Wittgenstein earlier said that they have not merely a visual descriptive use, this is not to be understood as if they are defective or vague, but to remind us that our *attitude* to facial expressions is part of their meaning. This is why he says: "We *react* to a hesitant facial expression differently from someone who does not recognize it as hesitant (in the *full* sense of the word)" (LW I, § 746). When we 'sense' the impact of an expression we will often imitate it with our own (747).

In the third remark, Wittgenstein warns us for not overlooking the 'field' of expression. Köhler's preoccupation with visual reality precludes our eyes for this field, or these other dimensions of facial expressions: "Think of this too: I can only see, not hear, red and green, but sadness I can hear as much as I can see it" (PI II, 209). We do not see a person's plaintive cries, we hear them, but especially: we react to them.

That the concept of seeing here reflects also our reactions to what we see is illustrated by yet another striking example that is discussed by both Köhler and Wittgenstein. In his *Dynamics in Psychology* (1940), Köhler tries to explain what happens when we look at the picture of a human face which is turned upside down. "They change so much that what we call facial expression disappears almost entirely in the abnormal orientation" (Köhler 1940, 25). His explanation is that it is not abnormal orientation in perceptual space, but

inversion with regard to retinal coordinates which alters the characteristics of our visual percepts and thus makes it difficult to recognise these percepts.

Wittgenstein approaches this striking phenomenon from an entirely different perspective. Unlike Köhler, Wittgenstein is not interested in the causal question as to whether the radical change of one's visual impression when the photograph is turned upside down is due to a change of perceptual orientation or of retinal orientation. Instead he focuses on a remark that Köhler makes almost in passing, namely that we fail to recognise the face and its smiling expression in upside down position. For Wittgenstein the case of the upside down face demonstrates a deep difference between language games, or between different descriptions of what is seen. Or what comes to the same, it demonstrates that the concept of description is a family resemblance concept. Consider this remark:

> Hold the drawing of a face upside-down and you can't tell the *expression* of the face. Perhaps you can see that it is smiling, but you won't be able to say what *sort* of a smile it is. You wouldn't be able to imitate the smile or describe its character more exactly.
>
> And yet the upside-down picture may represent the object extremely accurately. (RPP I, § 991)

The upside down picture may represent the construction of the face, such as the width of the face in relation to its length, or the triangular relation between the outer corners of the eyes to the base of the nose, exactly. Turning a portrait painting upside down is even a beloved practice of painters in order to check whether their constructive drawing represents the model accurately. Yet despite all this accuracy the character of the expression is gone. From the perspective of a constructive description nothing has changed, but one cannot conclude that the upside down picture of the portrait is seen in the same way as before. There is a radical change of one's impression of the face. Sameness of construction therefore is not sameness of expression. The one can be without the other. The construction can be described in constructive terms only, or by means of psychological terms. As this example illustrates, every attempt at describing the sort of expression of the inverted picture in constructive terms will fail to convey what sort of expression is involved. For that psychological concepts are essential. More importantly it is the specific use to which they are put which is essential. It is not just that we do not recognise the photograph's expression but the inverted photograph does not *make* us smile. As Wittgenstein observes we would not be able to imitate the smile. Imitating a smile is a way of representing and describing what is seen that is radically different from describing what one's sees in constructive terms. It is this deep distinction

between uses of language (and pre-linguistic behaviour) that explains what we want to say when we speak of inner states that are hidden behind outer behaviour or, like Köhler, inner states that are similar to behaviour.

6 Conclusion

As a concluding comment we can turn to a remark which has been quoted quite often in the literature but which has not been understood in the context of Wittgenstein's dialogue with Köhler:

> "We *see* emotion" – As opposed to what? – We do not see facial contortions and *make the inference* that he is feeling joy, grief, boredom. We describe a face immediately as sad, radiant, bored, even when we are unable to give other description of the features. – Grief, one would lik*e to say, is personified in the face. This is essential to what we call "emotion".* (RPP II, § 570)

Already the first sentence alludes to Köhler, who would emphasise that we see emotion as opposed to seeing mere colours and shapes and interpreting these visual data as emotion. For Wittgenstein the opposition between seeing in Köhler's sense and seeing plus interpreting in the empiricist sense misconstrues the concept of seeing as well as of interpretation. Wishing to oppose the empiricist Köhler puts all the weight on seeing. Replacing the idealised notion of seeing of the empiricist by the Gestalt concept of seeing, which includes and even prioritises the field of the object of perception, he believes to have found the explanation of social understanding which is in harmony with the naïve view of ascribing psychological states to other people. Seeing a person's anger is not just scanning his face but also seeing the dynamical development of objective experiences in the field of the observer which mirror the dynamical development occurring in subjective experience. As Köhler asks, who has not found himself occasionally walking faster when thinking about the disagreeable remarks of an adversary. And who has not observed his friend in the morning: "Sometimes his movements will be even and calm, sometimes his whole visible surface, his face and his fingers, will be unstable and restless" (249). For Köhler then the application of psychological concepts describing one's objective experiences of another person's experience depends on identifying the dynamical development of all sorts of traits which mirror his inner life.

Note that Köhler's descriptions of social understanding do not make use of psychological concepts like embarrassment, shyness or fear. Nor does he mention colour. Instead his descriptions remain at the *optical* level. We see that a face *lightens up, we see the crescendo*

and *ritardando* of behaviour and we see the *direction* of the eyes. Psychological states like embarrassment and shyness are similar to these objectively observable optical features.

Wittgenstein's approach is sharply opposed to Köhler's. Describing the emotion of another person is not to be understood as describing optical features of the face, or a larger sensory field which mirrors inner life and which for that reason is more than mere physical behaviour. Describing emotions is not mediated at all, neither by an inference to hidden inner states nor by a visible attunement of inner feelings and optical features of the face or body.

For Wittgenstein the appeal to optical features is as much an idealised notion of what counts as a description of what is seen as is the empiricist notion of sense datum. Are optical features really involved when attributing shyness to a person? Suppose I am drawing a sour face. To see whether I have got the expression right, what do I do? Typically, I step back and look at the drawing. But I do not check whether I got the expression right by comparing the expression with specific lines or shades of colour. To be sure I know that there are some ways to emphasise parts of the face to make a more convincing sour look. At any rate a teacher will not give pure visual hints, pointing to specific lines or halftones of colour. He may advise e.g. by building angular or blocky shapes, but these are not optical. In his Lectures on the Philosophy of Psychology, Wittgenstein remarks that the sentence "He looks shy" is embedded in completely different language games than "His face lightens up". To draw more in angular or blocky way, but surely drawing in this way belongs to the expression of the whole drawing. At this point Wittgenstein's occasional references to our native ability to imitate faces and ways of behaving may be profitable. In the first part of the *Philosophical Investigations* he says, think of our ability to imitate a facial expression without seeing ourselves doing it (e.g. in a mirror). And elsewhere he writes that to imitate a facial expression is a description, a language game (RPP I). To imitate a facial expression is not to derive one's description from visual or optical clues, since one does not see what happens while imitating. Mimicking a face, therefore, is not an indirect description: it is immediate. It is an expression which represents another expression.

To come back to Köhler's 'objectively observed behaviour', or the various optical features of the face. Suppose one is asked to imitate these optical features, the lighting up of the face, the crescendo and the ritardando of the 'fear', or the 'joy'. If we follow Wittgenstein's line of argumentation concerning the concept of imitation, to imitate the optical features would be to imitate the *expression*, for the imitation leaves no room for a distinction between on the one hand scanning optical features and using them for building one's imitation on the other. The imitation is itself an expression (of fear, of joy), hence it is immediate and direct.

Our application of psychological predicates to facial expressions is in no way different from imitation. To be sure we use words like shy or fear, but their application is immediate. "Even when we are unable to give other descriptions of the facial features", that is, even when from the perspective of plain seeing we would fail to see colours and shapes and other optical features, our description is direct. And this because the description of what is seen has taken the form of an interpretation of what is seen. Köhler is exactly in the position Wittgenstein exposes with the preceding remark. Köhler thinks that when we attribute emotion concepts to a person the sensory field in which the person's behaviour is perceived *must* be describable. As a help Wittgenstein reminds us of the aesthetic domain.

> But a painter can paint an eye so that it stares; so its staring must be describable by the distribution of colour on the surface. But the one who paints it need not be able to describe this distribution. (RPP I, § 1077)

Therefore, contrary to Köhler: "It is precisely a *meaning* that I see" (RPP I, § 869).

References

Hark ter, M.R.M. (1990). *Beyond the Inner and the Outer: Wittgenstein's Philosophy of Psychology*. Dordrecht: Kluwer Academic Publishers.

Hark ter, M.R.M. (2011). "The Experience of Meaning and Secondary Use". Kusela, O.; McGinn, M. (eds), *The Oxford Handbook of Wittgenstein*. Oxford: Oxford University Press, 499-523.

Hausen N.; Hark ter M.R.M. (2013). "Aspect seeing in Wittgenstein and Psychology". Racine, T.P.; Slaney, K.L. (eds), *A Wittgensteinian Perspective on the Use of Conceptual Analysis in Psychology*. London: Palgrave Macmillan, 87-110.

Köhler, W. (1940). *Dynamics in Psychology*. New York: Liveright.

Köhler, W. (1947). *Gestalt Psychology*. New York: Liveright.

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Basil Blackwell. [PI]

Wittgenstein, L. (1958). *The Blue and the Brown Books*. Oxford: Basil Blackwell. [BBB]

Wittgenstein, L. (1980). *Remarks on the Philosophy of Psychology*, vol 1. Oxford: Basil Blackwell. [RPP I]

Wittgenstein, L. (1980). *Remarks on the Philosophy of Psychology*, vol 2. Oxford: Basil Blackwell. [RPP II]

Wittgenstein, L. (1982). *Last Writings on the Philosophy of Psychology*, vol 1. Oxford: Basil Blackwell. [LW]

Discussions of a Private Language: Wittgenstein and Rhees

Volker Munz

Alpen-Adria-Universität Klagenfurt, Austria

Abstract   One year after the publication of *Philosophical Investigations*, the discussion about a private language had already gathered pace. Since then, the debate has moved in various directions: discussions about Wittgenstein's method of doing philosophy; about how to read him; about variations of 'private' language *users*; about private experiences, (private) ostensive definitions, behaviourism, the 'inner' and 'outer' etc. I have tried to cover those points, which I think crucial for the understanding of a 'private' language: the rule-fixing problem, the confusion of giving and using a sample, private charts, knowledge, memory, and justification. In doing so I have thereby made extensive use of remarks by Wittgenstein and Rush Rhees, particularly Wittgenstein's manuscripts, the Whewell Court lectures 1938-41, and unpublished material by Rhees. The reason for this is that I could not have put it in any better words, and that for me these remarks speak for themselves. I hope that others will have a 'similar' experience.

Keywords   Robinson Crusoe. Wallpaper pattern. The rule-fixing problem. Giving and using a sample. Private charts. 'Subjective' knowledge. Memory. Justification.

Summary   1 (Methodological) Remarks on Sense and Nonsense. – 2 Varieties of Private Language. – 2.1 "Another Person Cannot Understand the Language". Regarding PI, § 243. – 2.2 Giving a Sample and Using a Sample. Ostensive Definitions, Private Objects and the Rule-*fixing* Problem. – 2.3 Robinson Crusoe and the Wallpaper Pattern. – 3 Subjective Knowledge and the Private Chart Inspector. – 4 Subjective Memory and Justification: Apropos PI, § 258.

1   (Methodological) Remarks on Sense and Nonsense

Wittgenstein's discussion of the possibility of a private language certainly ranks as one of the most debated philosophical issues over the last seventy years. In 1954, Alfred Ayer and Rush Rhees initiated the debate in a symposium titled "Can there be a private language". Since then, the discussions have taken various directions – some linked to Wittgenstein's general method of doing philosophy, others related to various solitary men scenarios and the role of society in the invention of a private language, and some focusing on the general distinction between sense and nonsense, argument and therapy.

The main aim of this paper is to provide the readers with an overview of unknown or hardly known remarks from Wittgenstein's unpublished manuscripts and his *Whewell Court Lectures 1938-41*, as well as hitherto unpublished notes by Rush Rhees. What makes this new material so important is that it may shed some fresh light on the already existing old ways of reading and understanding Wittgenstein's discussion of a private language. The publication of this vast amount of new material may thereby help to see some of the old readings and standard interpretations from a different perspective.

In his review of the *Investigations*, Malcolm points out that Wittgenstein's discussion of a private language provides an argument in the form of *reductio ad absurdum* by postulating a 'private' language and then deducing that it is not a language in the first place. Malcolm also identifies other external arguments in Wittgenstein attempting to challenge the idea of a private language, such as PI, § 283 (cf. Malcolm 1954, 537).

In his paper "The Private Language Arguments", Peter Hacker argues that it might be misleading to speak in terms of just a singular argument, as Wittgenstein's discussion of the possibility of a private language is actually based on several arguments dealing with questions of epistemic privacy, private ownership, and private ostensive definitions (cf. Hacker 2019, 1). However, other philosophers, such as Barry Stroud, reject the idea of an argumentative structure in Wittgenstein's discussion of a private language. He writes:

> There is widespread agreement that what Wittgenstein does with the idea of such a language is to *refute* it – that he simply proves that a private language is impossible. And from that proof many powerful philosophical conclusions about the relation between body and mind, about our knowledge of other minds, and about the nature of psychological concepts – and no doubt about other things as well – are thought to follow and thereby to constitute Wittgenstein's philosophy of mind. Now I believe that no such conclusions or theories – and especially those widely discussed semantic theses we have heard so much about which would link 'behavioural

criteria' to 'mental concepts' – are to be found in Wittgenstein's text. In fact, I think it was an important part of Wittgenstein's own conception of what he was doing and of what needed to be done that no such philosophical doctrines or conclusions should be found there. (Stroud 2000, 69)

This debate about argument versus nonsense has also given rise to a broader question of how to interpret the private language debate in the first place. (For a more in-depth discussion of various readings, the resolute-substantial distinction, the Pyrrhonian-non-Pyrrhonian distinction, cf. Candlish 2019.) The different shifts within the whole issue have various reasons. Particularly, the question of whether 'private language' is a concept we understand, or whether the concept is nonsensical, has fuelled the debate on how to interpret Wittgenstein's responses to his fictitious opponents.

Concerning his methodology, Wittgenstein stands out as a philosopher who is particularly unique in employing a myriad of thought experiments, especially in his discussion of an essentially private language. He is less involved in offering counter-arguments to his interlocutors but rather invites us to meticulously go through each of his developed experiments – always approaching them afresh from different directions – (cf. Wittgenstein 2009, ix), to see where and how far the experiments will lead us when philosophising. In PI, § 374, Wittgenstein remarks:

> The great difficulty here is not to represent the matter as if there were something one *couldn't* do. As if there really were an object, from which I extract its description, which I am not in a position to show to anyone. – And the best that I can propose is that we yield to the temptation to use this picture, but then investigate what the *application* of the picture looks like. (PI, § 374)

This remark seems crucial to me for understanding Wittgenstein's method of doing philosophy: introducing a thought experiment, going through it, and trying to discern what its application would look like and where it might lead us. Similarly, in PI, § 422, Wittgenstein poses the question of what I am actually believing when I believe, for example, that men have a soul or that a substance contains two carbon rings. His answer is:

> In both cases there is a picture in the foreground, but the sense lies far in the background; that is, the application of the picture is not easy to survey. (PI, § 422)

And he continues:

Certainly all these things happen in you. – And now just let me understand the expression we use. – The picture is there. And I am not disputing its validity in particular cases. – Only let me now understand its application.

The picture is there; and I do not dispute its *correctness*. But *what* is its application? (423-4)

These remarks are also connected with Wittgenstein's general approach to doing philosophy, characterised by treating a philosophical question like an illness (cf. PI, § 255). Stroud points out that those

who demand philosophical results in the form of statable philosophical propositions or theories will no doubt remain disappointed or worse. [...] Those who seek 'results' in that way *should* remain disappointed with Wittgenstein. This is still better, I think, than inventing a set of definite doctrines and then claiming to find them, perhaps evasively suggested or only rhetorically expressed, in his unsystematic text. (Stroud 2000, 79)

Thought experiments, in my view, are often more effective in conveying a philosophical point than straightforward, systematic philosophical arguments. Dennett refers to thought experiments as 'intuition pumps' and notes:

Such thought experiments are not supposed to clothe strict arguments that prove conclusions from premises. Rather, their point is to entrain a family of imaginative reflections in the reader that ultimately yields not a formal conclusion but a dictate of 'intuition'. (Dennett 2015, 13)

Malcolm also points out that assuming a private language is possible or even necessary would not be 'eccentric' but rather 'natural' for anyone contemplating the relation between words and experiences (cf. Malcolm 1954, 531).

However, these natural or 'intuitive' approaches also carry the risk of leading in misleading directions from the very beginning. In PI, § 308, Wittgenstein uses the marvellous analogy of a *conjuring trick* to highlight this danger: When dealing with mental and physical processes, the first step often goes unnoticed because we already use terms like 'states' and 'processes', leaving their nature open. But this initial step commits us to a particular perspective, and we then strive to understand these states and processes better. The crucial step in the conjuring trick has already been taken, even though it seemed 'innocent' (cf. PI, § 308). Hence, caution is needed when taking the first step within a particular thought experiment. Wittgenstein provides

numerous examples of propositions that appear meaningful at first glance because we are familiar with their components. Also in the case of a 'private' language, there seems to be an unassuming first step. We talk about a language without determining its nature, but this already commits us to view it in a particular way, that is as a kind of *language* yet to be determined.

In PI, § 261, Wittgenstein warns against reverting to our ordinary language when trying to find a sign for a sensation because 'sensation' is already a word in our common language. Calling it 'something' instead does not help because this expression is also part of our common language. If the term 'something' has any meaning, it has a public meaning (cf. PI, § 261) A similar remark is found in "Notes for a Philosophical Lecture" (cf. NPL, 449).

If we regard the combination of the words 'private' and 'language' senseless, this does not mean as some readers seem to suggest that we are dealing with some kind of Meinongian 'impossible object', which, according to Meinong, has a particular kind of being, because without a prior understanding of impossible objects, such as a round square, we would not be able to ascribe impossibility to them (cf. Meinong 1981, 76-117). Leaving aside what Wittgenstein would say about "impossible objects" this way of reading his discussions of an essentially "private language" would be very misleading. There isn't a private language,

> something determinate that we cannot do, the idea that there is something, namely, a private language, that cannot be achieved; there is not a limitation on language. Rather, the idea is simply nonsense. (Candlish 2019)

In PI, §§ 499-500, Wittgenstein notes that to say of a combination of particular words that it has no sense, excludes it from the realm of language, thereby delimiting its area.

> When a sentence is called senseless, it is not as it were its sense that is senseless. But a combination of words is being excluded from the language, withdrawn from circulation. (PI, § 500)

In *Philosophical Grammar*, Wittgenstein reinforces a related point by questioning why we are more inclined to say that we cannot imagine something being otherwise rather than admitting that we cannot imagine the thing itself. For instance, we tend to consider a sentence like "This rod has a length" as a tautology rather than a contradiction. Instead of deeming both "This rod has a length" and "This rod has no length" as nonsense, we tend to affirm the first sentence as verified, thereby overlooking that it is a *grammatical* proposition. Once again, Wittgenstein emphasises that it is not the sense of these propositions

that is senseless, but rather these words are excluded from language in the same way arbitrary noises are: "[A]nd the reason for their explicit exclusion can only be that we are tempted to confuse them with a sentence of our language" (PG, 130. Cf. also Diamond 1991).

In PI, § 251, Wittgenstein revisits this example in the context of someone claiming that her mental images are private or that only she can know whether she is feeling pain, etc.:

> But the picture that goes together with the grammatical proposition could only show, say, what is called "the length of a rod". And what should the opposite picture be? (Remark about the negation of an a priori proposition.)

> "This body has extension." To these words, we could respond by saying: "Nonsense!" – but we are inclined to reply "Of course!" – Why? (PI, §§ 251-2)

In PI, § 464, Wittgenstein presents a vivid picture of his teaching aim when he notes: "What I want to teach is: to pass from unobvious nonsense to obvious nonsense" (PI, § 464). This, to me, is one of the central tenets of his methodology, demonstrating that Wittgenstein is not primarily concerned with the distinction between something being true/false or nonsensical, but rather with revealing unobvious nonsense. This nevertheless implies that transitioning from unobvious to obvious nonsense will impart crucial philosophical insights.

The words of an ordinary English sentence like "My images are private" or "Only I can know when I am in pain" are as nonsensical as some arbitrary noises, even though we are inclined to perceive them as true. This inclination arises because we are tempted to regard the sentence "as a sentence of our language". Instead, we must once again move from unobvious to obvious nonsense.

In "What Nonsense Might Be", Cora Diamond convincingly argues that

> for Wittgenstein, there is no kind of nonsense which is nonsense on account of what the terms composing it mean – there is as it were no 'positive' nonsense. Anything that is nonsense is so merely because some determination of meaning has not been made; it is not nonsense as a logical result of determinations that have been made. (Diamond 1981, 15)

Moreover, as Diamond points out:

> There is no 'positive' nonsense, no such thing as nonsense that is nonsense on account of what it would have to mean, given the

meanings already fixed for the terms it contains. This applies even to Wittgenstein's discussions of privacy. (16-17)

It is noteworthy that Wittgenstein frequently refers to fairy tales, fictional stories, poems, etc., when discussing sense and nonsense, both unobvious and obvious. It is therefore not accidental, I believe, that Wittgenstein remarks at one point that philosophy should only be poetised (cf. MS, 146, 25v). Particularly in his lectures spanning over more than ten years, Wittgenstein often brings up fictional literature while discussing the concept of nonsense:

In his 1935-36 lectures on 'private' experience and 'sense data', Wittgenstein begins to speak about different kinds of nonsense, such as "I feel his pain", and English sentences containing a meaningless word like 'abracadabra', or a string entirely composed of nonsense words. Regarding "I feel his pain", Wittgenstein remarks:

> Every words in the sentence is English, and we shall be inclined to say that the sentence has a meaning. The sentence with the nonsense word or the string of nonsense words can be discarded from our language, but if we discard from our language "I feel Smith's toothache", that is quite different. The second seems nonsense, we are tempted to say, because of some truth about the nature of things or the nature of the world. We have discovered in some way that pains and personality do not fit together in such a way that I can feel his pain. The task will be to show that there is, in fact, no difference between these two cases of nonsense though there is a psychological distinction in that we are inclined to say the one and be puzzled by it and not the other. We consistently hover between regarding it as sense and nonsense, and hence the trouble arises (24.10.1936; Macdonald, unpublished. Cf. also Diamond 1981).

One reason for distinguishing between kinds of nonsense, as Wittgenstein puts it, is a psychological inclination to say one and not the other or to oscillate between seeing something as sense and seeing it as nonsense. Similarly, in his notes for those 1935-36 lectures, he points out that it seems not to be false but rather meaningless to say that I can feel someone else's pain due to the nature of pain and the person, as if I were making a statement about the nature of things:

> So we speak perhaps of an asymmetry in our mode of expression and we look on this as a mirror image of the nature of the things. (LPE, 277)

In his "Lectures on Belief" from Easter term 1940, Wittgenstein gives the example of John Milton's *Paradise Lost*. I will quote the passage in full:

The other day I read a book which I didn't understand, "Paradise Lost". Right at the beginning, it is said that Satan lies in hell a time which measured in our time would be nine days. (Earth hadn't been created.) Now you might say: "What exactly does this mean?" Suppose that this had been a scientific observation of his. If this were a scientific observation, we might say we don't know for our lives what it means. How does he compare? "If Satan had lived ..., he would have asserted 'nine days'." First of all, it can't be given as an explanation, if it had been, it would have been said. Suppose that it were said by a physicist: "Before the earth and the sun existed, a certain event happened, which lasted nine days, as we would now say". Would you understand this? Would it be clear to you what this means? I mean, wouldn't the scientist have to give a brand new explanation? How does one measure a time? – It's like saying "It's five-o'clock on the sun".

In a fairy tale, "When it was five-o'clock on the sun, they had tea". Should we say "It is impossible to understand what is said in the poem"? On the other hand, if we took it to be a scientific statement, would it be relevant to know how things are compared? You might make some such remark as: "Oh, these poets, they don't bother their heads" – if you say this, has he [Milton] overlooked anything? Could this, as it were, be improved upon? Mr Lewy said: "Well, in a scientific work, I wouldn't understand it, in a poem I would". By the way, I don't understand it in "Paradise Lost" either. Couldn't you say: either this makes sense, or it doesn't? "Either it makes sense or the poet has made a blunder." It is important that a lot of people, and I among them, don't understand it. Not because I had thought about how it was verified. I should for instance say, "I don't know why he said nine days".

Context is a very complicated thing indeed. The statement puzzles us in a certain context. The statement only *sounds* queerer, than "the children lived on the sun, where it was five-o'clock". This may be as important as anything else. It would not puzzle us at all. As it doesn't puzzle us, when in a fairy tale three drops of blood spoke. If he (a scientist) said "The drop of blood spoke", I might have said, "What on earth do you mean?" It would be a question of understanding what he means. Whereas in a fairy tale, I wasn't in the faintest degree puzzled. (WCL, 238-9)

Interestingly enough, Wittgenstein picks up the "five-o-clock on the sun" example again in PI, § 350, in the philosophical context of notions such as "He feels the same as I". He points out the mistaken claim that if I knew what it meant to be five o'clock here, I would also understand the sentence "It is five o'clock on the sun" similar

to saying that if she has pain then she has the same as I do when I have pain.

> For what's surely clear to me is *this* part of grammar: that one will say that the stove has the same experience as I, *if* one says: it's in pain and I am in pain. (PI, § 350)

In *Mental Acts*, Peter Geach refers to an example Wittgenstein gave during his 1946-47 "Lectures on the Philosophy of Psychology". It is the example of Lytton Strachey's imaginative description of Queen Victoria's dying thoughts. Geach notes:

> He [Wittgenstein] expressly repudiated the view that such a description is meaningless because 'unverifiable'; it has meaning, he said, but only through its connexion with a wider, public, 'language-game' of describing people's thoughts; he used the simile that a chess-move worked out in a sketch of a few squares on a scrap of paper has significance through its connexion with the whole practice of playing chess. (Geach 1957, 3-4)

In the later publication of the lectures, we find the following alternatives:

> Lytton Strachey describes Queen Victoria's dying thoughts. But what is the use of this? As it stands there, it has no use at all. (LPP, 32. Geach's notes)

> Lytton Strachey on Queen Victoria's dying thoughts. How could he know? And if not, does he talk sense? (LPP, 99. Shah's notes)

> Remember Lytton Strachey talking of the thoughts of Queen Victoria on her deathbed. If he could not conceivably know, what sense could this have? If no sense, why say it? (LPP, 229. Shah's notes)

> What did Queen Victoria think as she lay dying? There is no verification *accessible* to 'What did the Queen think?' Then shall we say that Strachey was guessing at what she thought? You may. But it's a different use of 'guess' from the use we learned; another game. (LPP, 274. Jackson's notes)

In his "Lectures on Description" from Lent term 1940, Wittgenstein uses the example of a man saying: "Ultra-violet is a non-spectral colour". It is clearly a well-formed English sentence and may be uttered with a particular tone of conviction. But all the man really does is making queer English noises. The sentence just does not fit into any of the games into which similar sentences fit. Wittgenstein goes on:

> You could imagine nonsense poems of Lewis Carroll not to have
> been a joke. – "Does one of the sentences of Lewis Carroll's poem
> make sense or not?" It is not easy to answer. It does in a way. We
> have images, it has been illustrated, etc. In a sense, it does make
> sense; in a sense it doesn't. We can't say there is a large province
> of our language with sentences of this kind. [...]

> When should one say, "This does make sense", "This doesn't make
> sense"? – It is very often very clarifying to cut off altogether one
> sort of expression from a certain game. Just as some people say:
> "Poems like Lewis Carroll's make no sense", this is quite all right.
> It might be very useful to say "This has no sense", "This has". This
> doesn't mean that it [the sentence] isn't of any use. The answer can
> never be categorical at all.

> If one says it makes no sense, this means on the whole trying to
> dissuade the other man from saying it. It means: "Don't say that".
> (WCL, 167-8)

In his "Reply to a Paper by Y. Smythies on 'Understanding'", Wittgen-
stein makes a similar remark:

> Suppose, on being asked whether I understand the sentence "The
> blind man imagines", I answered ("No"), and Lewy said, "Honestly,
> don't you understand it at all?" – the only thing that I could say is:
> "Well, it depends. What sort of thing do you oppose it (understand-
> ing) to? To Lewis Carroll? Do you understand it as you understand
> 'A = A'? It is a different case. You don't know what verifies or falsi-
> fies it, but you can easily suggest something which you or I might
> take". Very often, given an expression in English, I could give you
> all the task: "You tell me what it might be used for" – that is to say,
> besides a meaning which is fixed, there is also something else, the
> next meaning that we give it. Cf. "This man married green". "This
> hasn't a meaning." – No one says this about mere noises, because
> the question doesn't arise. (WCL, 193)

Lewis Carroll's examples, among others, illustrate that the distinc-
tion between what makes sense and what does not is not a straight-
forward one. Wittgenstein suggests that we often stagger between
regarding something as sense and as nonsense, leading to trouble.
In this context, he seems to propose that there is no real difference
between uttering an ordinary English sentence, such as "The blind
man imagines", and making arbitrary noises like 'abracadabra'.
  Two additional examples involve fairy tales in the context of non-
sensical propositions. In *Philosophical Grammar*, Wittgenstein writes
about a schoolboy equipped with elementary trigonometry skills who

is asked to test a complex equation. According to Wittgenstein, the boy would neither know how to answer the question nor understand it. He compares this situation to the task set in a fairy tale where a prince asks a smith to fetch a "fiddle-de-dee (Busch, Volksmärchen.)" (PG, 378-9. Cf. also PR, 178). And in the sections on private language in PI, Wittgenstein makes another remark:

> "But in a fairy tale a pot too can see and hear!" (Certainly; but it *can* also talk.) "But the fairy tale only invents what is not the case: it does not talk *nonsense*, does it?" – It is not as simple as that. Is it untrue or nonsensical to say that a pot talks? Does one have a clear idea of the circumstances in which we'd say of a pot that it talked? (Even a nonsense-poem is not nonsense in the same way as the babble of a baby.) (PI, § 282)

These quotations present various cases of unobvious nonsense, many from literary contexts, that need to be made explicit. Ultimately:

> The results of philosophy are the discovery of some piece of plain nonsense and the bumps that the understanding has got by running up against the limits of language. They – these bumps make us see the value of the discovery. (PI, § 119)

2 Varieties of Private Language

2.1 "Another Person cannot Understand the Language". Regarding PI, § 243

The discussion of the private language is typically situated between paragraphs 243-315 of the *Philosophical Investigations*. Wittgenstein presents an initial version of the kind of 'language' he has in mind:

> A human being can encourage himself, give himself orders, obey, blame and punish himself; he can ask himself a question and answer it. So one could imagine human beings who spoke only in monologue, who accompanied their activities by talking to themselves. – An explorer who watched them and listened to their talk might succeed in translating their language into ours. (This would enable him to predict these people's actions correctly, for he also hears them making resolutions and decisions.)
>
> But is it also conceivable that there be language in which a person could write down or give voice to his inner experiences – his feelings, moods, and so on – for his own use? Well, can't we do so in

our ordinary language? But that is not what I mean. The words of this language are to refer to what only the speaker can know – to his immediate private sensations. So another person cannot understand the language. (PI, § 243)

PI, § 243 states that a man can encourage himself, command himself, obey, blame, punish, put a question to himself, and answer it. Therefore, Wittgenstein argues that we can even imagine people who speak only in monologues. In the MS 165 version (cf. MS 165, 103-6), he had begun by saying that in one particular sense we could speak of a 'private language', namely that of a Robinson Crusoe who speaks only to himself. Speaking to oneself, however, does not mean being alone and speaking. I can as well speak to myself when others are also there. Already in the *Blue Book*, Wittgenstein remarks:

> We could perfectly well, for our purposes, replace every process of imagining by a process of looking at an object or by painting, drawing or modelling; and every process of speaking to oneself by speaking aloud or by writing. (BBB, 4)

In *Last Writings on the Philosophy of Psychology*, we read:

> Only someone who can speak can speak in his imagination. Because part of speaking in one's imagination is that what I speak silently can later be communicated. – (LWI, § 855)

In a letter to Wisdom, dated 15 June 1954, Rhees writes:

> I may say to myself, "Where did I put the pen?" or "What is the matter with me?". But granting that Ayer's Crusoe might make these sounds too, he would still not be asking himself a question. I cannot ask myself a question unless I can *understand* a question. Speaking a language means, for instance, asking questions and giving orders. And that has its sense in what people do with one another. If a man talks to himself, that is not just making noises. And the difference is that they are noises that he has used and that he has heard in his discourse with other people. He knows what they mean (not: he knows what he uses them for). (Rhees, unpublished)

And later:

> If a man speaks to himself, he speaks in some language; and a language is spoken by others. Otherwise, he would not be saying anything to himself. I do not say that a man cannot speak a language unless he speaks it with others. I say he cannot speak a language

unless it is spoken by others. And I would add also: he has learned it. (Rhees, unpublished)

In PI, § 344, Wittgenstein raises the question of whether it would be conceivable that people should never speak an audible language but nevertheless speak a language inwardly in their imagination. He then gives the following answer:

> Our criterion for someone's saying something to himself is what he tells us as well as the rest of his behaviour; and we say that someone talks to himself only if, in the ordinary sense of the words, he *can talk*. And we do not say it of a parrot; or of a gramophone. (PI, § 344)

In the final PI version, Wittgenstein uses 'speaking in monologues' instead of 'speaking to himself', because the concept of 'speaking to myself' might be misleading. In the MS 180a version, which is only slightly changed in PI, the expression "only in monologues" is followed by a parenthesis ("So each of them might also have a language of his own. How he could learn it is irrelevant"). Wittgenstein invites us to imagine a society or group of people who would only speak monologues, a group exclusively of such people, again a quite radical thought experiment. So, we might imagine that each member of the group has come from a different society, in which he might have learned to speak with other people, and then, as Rhees puts it, "he became quasi-autistic". He notes:

> Apparently, the language of each monologue could also have been spoken in dialogue of people with one another. – This is suggested when Wittgenstein says that each of these inhabitants accompanies his actions by conversing with himself, and that a scholar (sociologist) who observed them and listened to their talking might translate their language into ours. – What was spoken in monologue would be a language in the same sense as ours is. With grammatical rules such as ours has. "(and then he could predict their actions correctly, for he can hear them formulate plans and decisions.)" Here, as with the question raised in PI, § 206, "The way of behaving which is common to human beings is the frame of reference for our interpretation of a language that is foreign to us." And added to this, the regularity, which was wanting in what first seemed to be a language of people imagined in § 207. (Rhees, unpublished)

In MS 165, Wittgenstein starts talking about a language, someone only talks to herself only for her understanding about her personal experiences. At this point, Wittgenstein does not delve into the discussion of such a language, as it pertains to the problems of idealism

and solipsism. He does, however, point out that no language has actually been described in this context, although it appears to be, because there is obviously no assurance that a word in this language is used twice with the same meaning. If you say, e.g., that some objects are equal if they appear equal to you, I might ask how the objects appear to you because after all 'equal' is a term in a common language (cf. MS 165, 101-3). We find a similar remark about 'sensation', 'something', and 'having' in PI, § 261.

In another sense, there can, of course, be a private language like that of a Robinson Crusoe who talks only to himself. He could, e.g., encourage himself with words to do something, could ask himself a question and answer it or blame himself, etc. According to Wittgenstein, we would only call such a phenomenon language if the behaviour of this person was similar to that of humans in general. And if we especially understood his gestures and facial expressions in the context of sorrow, displeasure, joy, etc., we could call this a language or a language-like phenomenon (cf. MS 165, 103-4). In MSS 124 and 149, Wittgenstein makes similar remarks about Crusoe talking to himself on his island and emphasises that if someone had listened to and observed Crusoe, she could have learned Crusoe's language, since the meaning of his words would show themselves in his behaviour (MS 124, 221-2). Accordingly, in MS 149 he notes:

> We can indeed imagine a Crusoe using a language for himself but then he must behave in a certain way or we shouldn't say that he plays language games with himself. (11v)

Wittgenstein continues MS 165 by introducing the case of a human being "who lives alone and draws pictures of the objects about him (say on the wall of his cave), and a picture-language of this sort could easily be understood" (105). But Wittgenstein points out that such a person who encourages herself is thereby not also able to master the language game of encouraging another person. Thus, the ability to speak to oneself does not necessarily imply the ability to speak to others, any more than someone who can play patience must also be able to play card games with others. Similarly, there can also be a language-like phenomenon: a language that each person speaks only to herself, thinking, e.g., about her future actions. A language is primarily something spoken by the peoples of the Earth. And we label as language those phenomena that bear resemblance to those languages. Ordering, for instance, is a technique of our language. So one can give oneself commands. But if we were to observe a Robinson Crusoe giving himself a command in a language unfamiliar to us, it would be much more challenging for us to recognise (cf. MS 165, 105-9). Wittgenstein continues with a remark which later moved to PI, § 206, when he notes:

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. (PI, § 206)

Later in MS 165, Wittgenstein introduces a Crusoe alternative. He starts again by inviting us to imagine a person talking to herself:

Suppose a human being (say a cave-man) spoke always to himself alone. Imagine a situation in which we might say: "Now he is thinking over whether he should act in this way or in that. Now he is ordering himself to act so." It is possible to imagine something like this, if he makes use of simple drawings, which we can interpret. (MS 165, 117)

But then, Wittgenstein continues by introducing a new case, in which someone invents a game which neither she nor anyone else will ever play:

Commands are sometimes not followed. But how would it appear if commands were never followed? Today, I can invent a game that neither I nor anyone else will ever play. However, what would it look like if games were never played but only invented? Now, can I not imagine that? Occasionally, someone takes paper and pencil, designs the plan for what we would call a game (such as tennis), and writes down the rules of the game. He might add: it would be nice if we acted this way. So why did I say that this condition cannot be imagined? Well, if it existed, and we saw it, the question is whether we would associate it with our concept of the game. Especially if the games corresponding to those plans were very different from those commonly used by us. (117-19)

Wittgenstein finishes the experiment with the well-known remark from PI, § 284: "Transition from quantity to quality" (MS 165, 120). These various examples obviously show the thin line between what we can imagine in the case of solitary men and when we transcend the line of what still makes sense.

In MS 116, Wittgenstein also introduces a Robinson Crusoe who uses a language for his private use, a case that does not seem to fall under the different variations Wittgenstein allows in the first paragraph of PI, § 243. He writes:

Language, as far as one understands it subjectively, may not serve as a means of communication with others but rather as a tool for

an individual's private use. The question, however, is whether we would still consider this utterance of sound combinations or the drawing of lines, and the like, a 'language' and whether we would still call it a 'tool'. For he would have to play language games with himself and he can indeed do so. Imagine a Robinson Crusoe who uses a language (signs) for his private use; imagine observing him (without his knowledge); you would see how, on various occasions, he carves lines into wood, utters sounds. Would we in all cases call this 'using signs'? Only if you were to observe a specific regularity. We observe a person who, on different occasions, emits sounds without any regularity – now we say, "This may be a purely private language; he probably associates the same thing with the same sound each time." (MS 116, 117-18)

Wittgenstein continues with one of the central remarks in his discussions of a private language, which we find in PI, § 264 voiced by his private language opponent: "'Once you know what the word signifies, you understand it, you know its whole application'", a remark which will be particularly crucial in the context of a private ostensive definition. In this remark, the quotation marks are, of course, essential.

During his stay in Bergen in autumn 1937, Wittgenstein was also concerned with a particular idea of a private language, which does not seem to be part of the cases covered by the first part of PI, § 243. Although he starts with the already familiar notion of someone talking to herself, this time his example refers to a colour concept such as 'blue', which a person might use to refer to a colour that comes to her mind, without bothering whether others would agree with this usage or not. In such a case, Wittgenstein argues, the person would ask herself what she could indeed do with such a language and whether we would still call it a language (cf. MS 119, 95v).

Several pages later in the same MS, Wittgenstein notes:

How can one give a name to a private object? What does it mean to recognize the private object? Does it mean essentially the same as believing to recognize it? 'Recognition' already implies certain public criteria. This seems to erase language almost as if it has been turned off. We are completely in the dark. We realize that the word 'red', for example, is only a word in our public usage. As soon as we retreat into the private, language ceases to exist; the word 'red' loses its use. (MS 119, 124r, 124v)

This remark already includes central issues of the private language debate such as giving a name to a private object, recognising the private object, and the role of public criteria (cf., e.g., PI, §§ 256, 260, 580).

Again, in the following MS 120, Wittgenstein connects the notion of a private language with that of a private object:

"There is, after all, a subjective regularity, a regularity that exists only for me." That is to say, we sometimes use the word 'regularity' in a way that suggests someone envisions a regularity, sees something regular, or perceives it as regular, and so on. However, this doesn't mean he has an object in front of him that none of us knows, and he calls it 'regularity'. If he is playing, besides the game I see, another one with himself that I know nothing about, then I don't know whether what he is doing should be called "playing a game". If, in addition to the public language, he speaks with himself in a private one that I know nothing about, why do I say it is a language? How do I know that it is a language? So, it seems he is playing, besides the game I see, another one with himself, which I know nothing about – but why do I call that a 'game'? In other words, we use the picture of the 'private object' that only he can see, and not others. It is a picture – be clear about that! And now, it is inherent in the nature of this image that we make further assumptions about this object and what he is doing with it; (because) it is not enough for us to say: He has a private something and does something with it. (MS 120, 27-8)

These considerations are obviously connected to the kind of private language Wittgenstein introduces in the second part of PI, § 243, i.e. a language another person *cannot* understand. The example also differs from the solitary men cases in so far as the ostensible users of such a private language are also familiar with a common language. What it further shows is that Wittgenstein does not deny the idea of a private object in the first place. He does not immediately dismiss such a picture but instead invites us again to get clear about this very picture we are using here and that we are using a picture. Then we have to see what we could do with it and where it would get us. In other words, there is nothing wrong with introducing a particular picture, as long as we are able to tell what we are supposed to be doing with it, what *use* we will make of it, how we will go on.

As we have seen so far, Wittgenstein discusses different kinds of solitary men, Crusoes, and cavemen over a period of about eight years. What these examples show is that Wittgenstein distinguishes cases where it *would* make sense to talk about using a private language when it shows similarities to our common language use, such as regularities, familiar behaviour, the possibility to learn such a language, translate the language, etc. In his "Robinson Crusoe Sails Again: The Interpretative Relevance of Wittgenstein's *Nachlass*", Peter Hacker convincingly argues that Wittgenstein did not show any

concerns about a speaker or a group of speakers who use a contingently private language. Hacker remarks:

> Wittgenstein discussed solitary people who follow unshared rules in many different manuscripts in his Nachlass between 1936-37 and 1944. In none of these numerous remarks, did he express any qualms about the conceivability of speakers with an innate knowledge of a language, or about speakers who speak a contingently private language. Throughout these discussions, his focus is unquestionably on the requirement of regularity seen as uniformity and conceived as a norm, not upon the requirement of multiplicity of agents. (Hacker 2010, 106-7. For the so called 'community view' cf. Candlish 2019)

It is, however, interesting that most of the cases Wittgenstein introduces in his manuscripts did not make it into PI. We could indeed ask, as Hacker also does (Hacker 2010, 107), why Wittgenstein did not include those discussions of the various solitary men in the *Investigations*. I assume there are various reasons. First of all, all those cases where Wittgenstein argues that we could very well imagine someone or a group of people using a particular kind of private language are, as I would want to call them, 'harmless cases' of private language users. All of those harmless cases already seem to be covered by the first part of PI, § 243. An explorer who listened to them and observed their behaviour could, in principle, translate their language into ours. So, these seem to be the unproblematic cases of people using such a language. Wittgenstein seems far more interested in a language in which a person would write down her inner experiences for her own use and another person *could* not understand this language.

In a letter to C.W.K Mundle, dated 3 April 1965, Rhees comments on Mundle's "'Private Language' and Wittgenstein's Kind of Behaviourism". He writes:

> Of course, Wittgenstein knew that the sense of 'privacy' he was discussing was a special and queer one. This is just what he was trying to bring out in a large part of his discussion. He was *not* denying that there are other senses in which people may speak of 'private languages'. And of course he was not denying that I can tell the doctor how I feel. He is talking about a particular idea of a language [...] He is *not* saying that a "private language must be of this sort". He is talking *about* the idea of this sort, and about the sense in which it is supposed to be 'private' (or 'incommunicable'). The 'language' in which I know what I mean – 'the inner language' which we might say *is* my understanding or my meaning; but which I cannot express to you directly, but only 'indirectly' by, as it were, translating it into the common language. The idea that

"there is a limit to what we can communicate". Or perhaps: "I can never know whether the content which you put into my words is the content which they had for me" – and so on. (Rhees, unpublished)

Secondly, as Rhees' remarks and Wittgenstein's own writings show, Wittgenstein was obviously less interested in cases where someone grew up in a strictly asocial context, without any knowledge of a common language, but rather in people who would, besides using a public language, also invent an essentially private language for their own use. At least PI, §§ 243, 256-8, 261, 265, 270, and others seem to suggest that the seemingly private language user is also familiar with a public language. In the same letter, Rhees gives the example of Samuel Pepys' diaries that were written in shorthand and sometimes in code:

> Wittgenstein agreed that you might say that Pepys had a private language (although it might be more accurate to call it a private cipher). This sense of "private language" has nothing to do with what he is examining. And your use of "used to refer to a private experience" – if it is illustrated by the expressions you use to tell the doctor of your symptoms – is not what he is discussing either. (May I repeat: he is *not defining* private experience or "private symbol" in any particular way and then saying that it cannot be used in any other way.) (Rhees, letter to Mundle, 3 April 1965)

In his "Lectures on Similarity" (Michaelmas 1939), Wittgenstein gives us an example of a diary user, different from Pepys' case, who is also capable of speaking a common language:

> Robinson Crusoe invented a language and used it for himself. Imagine that you have a diary in which you write down your experiences:

Monday	x
Tuesday	X0
Wednesday	∂

> etc. "What's all this?" "A private language." "What does it describe?" "I'm afraid I can't tell you." What reason have I to believe that I mean by the language what I do? That I mean by the language all that I claim I mean? If you say "It is a private language describing experiences", this has as much meaning to me as the word "experience" has. "Is it pains?" "No". "Is it religious experiences?" "No". On Monday, Tuesday and Wednesday he makes different scratches. What is at all similar here to a language? I

wouldn't know at all whether to say that it is a *private* language. All I know is that he makes scratches and says that it is a private language. But suppose that he makes scratches and says "I can't explain". Suppose he says "If X is repeated it's the same experience". I can't be sure whether to say he is using a private language. If this were the rule that people made scratches and said "I can't explain", etc. there would be nothing more I could call a private language. This – the kind of situation where we say there is a private language – is only possible because it is exceptional. You may say "There's something here I can't explain." (WCL, 95-6)

Here again, I get the impression that Wittgenstein wants to warn us not to make general inferences from exceptional cases. In the context of the concept 'sense-datum', which to me seems equally applicable to the term 'private language', Wittgenstein remarks:

We said that there were cases in which we should say that the person sees green what I see red. Now the question suggests itself: if this can be so at all, why should it not be always the case? It seems, if once we have admitted that it can happen under peculiar circumstances, that it may always happen. But then it is clear that the very idea of seeing red loses its use if we can never know if the other does not see something utterly different. (NPL, 285)

Here, too, the decisive movement in the conjuring trick seems to have been made: the innocent move from the exceptional case to the general one. A representative example for this 'innocent' step is Alfred Ayer's development of the "Argument from Illusion" (cf. e.g. Ayer 1963, 1-57). Or, as Wittgenstein puts it:

"If people talked only inwardly, to themselves, then they would merely be doing *always* what, as it is, they do *sometimes*." – So it is quite easy to imagine this; one need only make the easy transition from some to all.

"What sometimes happens might always happen." (PI, §§ 344-5)

Similarly, sometimes orders are not obeyed, but if we assume that no orders would ever be obeyed then the concept of 'order' would have lost its entire purpose (cf. PI, §§ 344-5).

Thirdly, Rhees points out that Wittgenstein did not include either of the remarks about the caveman's drawings in the *Investigations*. He notes:

As I remember a conversation in 1945, Wittgenstein grew more hesitant about saying "we can imagine" that someone who had

never communicated with others or lived in their company would make drawings of objects and use them deliberating (as we should say) on what to do. (Rhees, unpublished)

Wittgenstein discussed questions closely related to these issues in §§ 41 to 43 in *Remarks on Foundations of Mathematics*. According to Rhees, the manuscript of part VI (MS 164) was written probably six months or a year later than MS 165. These remarks certainly show that the line between the harmless and the essential cases is indeed not always clear and might make us hover again between the meaningful cases and the nonsensical ones: Wittgenstein again introduces a caveman who produces regular sequences of particular marks just for himself and draws them on the cave walls. But we would still not say that he is following the general expression of a particular rule. And if we want to say that the person does act in a regular manner, this is not because we are able to form such an expression. The point is that a word has a meaning only within the practice of a particular language. I can, of course, give myself a particular rule and then follow it. But we should not think it is only a rule because it looks analogous to what we call 'rules' in our common human activities.

> When a thrush always repeats the same phrase several times in its song, do we say that perhaps it gives itself a rule each time, and then follows the rule? (RFM, VI, §41, 345)

Similarly, Wittgenstein continues:

> If one of a pair of chimpanzees once scratched the figure | - - | in the earth and thereupon the other the series | - - | | - - | etc., the first would not have given a rule nor would the other be following it, whatever else went on at the same time in the mind of the two of them. If however there were observed, e.g., the phenomenon of a kind of instruction, of shewing how and of imitation, of lucky and misfiring attempts, of reward and punishment and the like; if at length the one who had been so trained put figures which he had never seen before one after another in sequence as in the first example, then we should probably say that the one chimpanzee was writing rules down, and the other was following them.

> It is possible for me to invent a card-game today, which however never gets played. But it means nothing to say: in the history of mankind just once was a game invented, and that game was never played by anyone. That means nothing. Not because it contradicts psychological laws. Only in a quite definite surrounding do the words "invent a game", "play a game" make sense. (RFM, VI, §§ 42-3, 345-6)

This is the reason why expressions such as 'language', 'order', 'rule', 'calculation', 'experiment', 'following a rule' are related to a particular technique and custom of acting, speaking (cf. RFM, VI, §§ 42-43, 345-6). In my understanding, these examples show that even in the case of observing a particular behaviour in a certain regularity, it is not once and for all clear whether such a scenario also guarantees the use of a language. I will come back to this point in the context of Rhees' metaphor of a 'wallpaper pattern', which he uses in his reply to Ayer in 1954.

2.2 Giving a Sample and Using a Sample. Ostensive Definitions, Private Objects and the Rule-*fixing* Problem

Kripke's famous account of Wittgenstein's treatment of rules and private language in particular has turned Wittgenstein's remarks into a kind of meaning scepticism. Kripke takes the problem of a private language just as a special case of Wittgenstein's paradox in the rule-following context: "This was our paradox: no course of action could be determined by a rule, because every course of action can be brought into accord with the rule" (PI, § 201). In a letter to Alfred Ayer, dated 6 July 1954, Rhees writes:

> In the *Investigations*, Wittgenstein raises this question when he is discussing what *following a rule* is, and what grasping the meaning of an expression is, or what understanding is. This question about following a rule is of fundamental importance for the nature of logic (the relation between logic and reality, for instance) and for the nature of mathematics. And one of the things he tries to bring out is the difference between "der Regel folgen" and "die Regel deuten" – as though understanding were a matter of seeing what is contained in it, in that sense. So he says (PI, § 202) that der Regel folgen eine Praxis ist und nicht ein Deuten. And it is for this reason that einer Regel zu folgen *glauben* is not: der Regel folgen. „Und darum kann man nicht der Regel 'privatim' folgen." [...] One of the questions there is whether I can *give* myself a rule, for instance. And in this is the question of what it could mean to say that I was following the rule correctly: or what difference there could be between following it and not following it. (Last italics added)

This point is bound up, Rhees continues,

> with the question of what "giving a name to something" is, and what it is for a mark or a sound to refer to something or mean something; and so with the question of whether a mark or a sound could mean something in a "private language". If I cannot *give* myself a rule, then I cannot *give* myself names for my private sensations either;

that is why concentrating the attention, or "pointing inwardly" as I spoke a sound for myself would be an idle ceremony. (Italics added)

Wittgenstein's point in PI, § 202 that to follow a rule is a practice, and that therefore one cannot follow a rule 'privately', because thinking that one is following a rule is not the same thing as following it, Kripke takes not only as a sceptical answer to the paradox but also to the private language case. He remarks: "The impossibility of a private language emerges as a corollary of [Wittgenstein's] sceptical solution to his own paradox" (Kripke 1982, 68). Although it is indisputable that many of the issues Wittgenstein raises in his discussion of a private language are internally related to his former discussions of giving names to things, ostensive definitions, following rules, of saying, meaning and understanding something, this does not mean that the private language discussion was already finished with PI, § 202. In his remarks, PI, §§ 243-315, Wittgenstein covers many subjects that are not discussed elsewhere (at least to this extent), such as private experiences, private objects, private charts, private ostensive definitions, subjective knowledge, subjective memory and justification, pain and pain behaviour, mind-body relations, and many other interrelated issues (for a thorough discussion of why Kripke's reading of Wittgenstein is mistaken, cf. Hacker 2019, 9-13). Furthermore, Wittgenstein's own examples of solitary men, Robinson, cavemen, and monologuists also show that the 'community' view is not in general a good answer to the question of whether one could follow a rule privately. It is, however, important to point out that those examples, as we have seen, only seem to cover the *harmless* cases of solitary men, summarised in the first part of PI, § 243. Therefore, I would like to argue that Wittgenstein's discussion of an 'essentially' private language, introduced in the second part of PI, § 243 and further elaborated in PI, § 256 and the diary case in PI, § 258, is not primarily a genuine case of a rule-following problem but to a much greater degree of a *rule-fixing* problem. Wittgenstein draws the distinction between thinking that one is following a rule as opposed to following it in the context of his private language discussion (cf. PI, § 260). Nevertheless, I think we could also apply this figure to 'thinking that one is fixing a rule' and actually 'fixing a rule'. The rule-fixing problem becomes most obvious in the distinction between *giving* a sample and *using* a sample, *giving a definition* and *making a statement*, respectively.

In *Philosophical Grammar*, Wittgenstein remarks:

The concept of meaning I adopted in my philosophical discussions originates in a primitive philosophy of language. The German word for "meaning" ("Bedeutung") is derived from the German word for "pointing" ("deuten"). When Augustine talks about the learning of language, he talks about how we attach names to things or

understand the names of things. Naming here appears as the foundation, the be-all and end-all of language. (PG, 46)

The learning of a language, the way Augustine describes it, "can show us the way of looking at language from which the concept of the meaning of words derives" (57). In his introduction to *The Blue and Brown Books*, Rhees points out that in the *Blue Book*, Wittgenstein discusses the process of grasping the meaning of words by having someone explain their meanings as though 'understanding' and 'explaining' were somehow interconnected. However, in the *Brown Book*, Wittgenstein points out that learning a language game precedes this understanding. What is essential is not explanation but rather training, akin to the training provided to an animal. This aligns with his emphasis in the *Investigations* that the ability to speak and understand what is said in the sense of knowing its meaning does not necessarily imply the ability to articulate that meaning. In PI, § 32, Wittgenstein points out that sometimes a person will learn the language of a people from ostensive definitions they give her. Then the person will sometimes just have to guess the meaning, sometimes correctly, sometimes incorrectly. Augustine portrays the acquisition of human language as if a child entered a foreign land and did not comprehend its language, implying that the child already possessed a language, just not the one in question. In order to find out whether a child knows a particular language, you might ask her whether she knows what a particular expression means. However, this way you could not say whether the child could also *talk*. Neither is it what the child learns when it learns to speak a language (cf. PI, § 32; BBB, vi). According to Rhees, Wittgenstein brings the language games in both PI and the *Brown Book* to shed some light on the relations of words and what these words stand for. In PI, however, he is concerned with the Augustinian conception of meaning, which eventually holds that only the demonstratives 'this' and 'that' are real proper names, and all other words called 'names' would only be names in a very inexact and approximate sense (cf. BBB, ix)

It is therefore not surprising that Russell, in "The Philosophy of Logical Atomism", notes:

A logically perfect language, if it could be constructed, would not only be intolerably prolix but, as regards its vocabulary, would be very largely private to one speaker. That is to say, all the names that it would use would be private to that speaker and could not enter into the language of another speaker. [...] A name, in the narrow logical sense of a word whose meaning is a particular, can only be applied to a particular with which the speaker is acquainted because you cannot name anything you are not acquainted with.

[...] One can use 'this' as a name to stand for a particular with which one is acquainted at the moment. (Russell 1956, 198, 201)

Wittgenstein himself defended a particular name-object relation in his *Tractatus*, which was internally related to his picture theory. In 3.2-3.22 he remarks:

In a proposition, a thought can be expressed in such a way that elements of the propositional sign correspond to the objects of the thought. [...] The simple signs employed in a proposition are called names. A name means an object. The object is its meaning. [...] In a proposition, a name is the representative of an object. [...] The requirement that simple signs be possible is the requirement that sense be determinate. (TLP, 3.2-3.23)

In his "'Ontology' and Identity in the *Tractatus*" Rhees notes:

The *Tractatus* hardly distinguishes naming and calling something by its name. And 3.3 shows that this is not an oversight. 'Nur im Zusammenhange des Satzes hat der Name Bedeutung.' So we may think that what the word 'red' means is expressed by the sentence 'a is red'. Someone might say: 'the name must correspond to some reality. It cannot describe anything if there is nothing which it signifies.' Or suppose I told you: 'I call each of these roses red because each of them *is* red. The word I use corresponds to the colour of the flower'. – But what corresponds is the *sentence*. The *Tractatus* supposed that 'red' determines how I use it. Wittgenstein rejected this later. It confuses giving a sample and using a sample. I may give a sample – a piece of coloured paper – to explain what I mean by 'vermilion'. Or I may use the sample in place of the word and tell you 'the flowers in that bed are *this* colour'. But I cannot use the sample to explain what colour this *sample* is. (Rhees 1970, 28)

The erroneous view was to understand samples as primary signs, which unmistakably explain themselves and could not be misunderstood. And the primary signs subsequently explain secondary signs. So without the primary signs, we would not know what we are saying.

Wittgenstein brought out the confusion in all this. But it showed that the distinction between what a name means and what is called by it is not always simple or easy. (28-9)

In *Philosophical Grammar*, Wittgenstein argues that only in an irrelevant sense could one say that the truth of a sentence immediately follows from its existence, for example, if we imagine a sentence written on a wall with a red colour, "in this room, there is something

red". The reason for such a sentence being irrelevant relates to the idea that an ostensive definition is not intended to make statements about the sample itself, but rather about objects corresponding to the sample, to which it can be applied. This is what it means to confuse *giving samples* and *using samples*. For example, in the case of a particular shade of red, I can give an ostensive definition by saying, "This is called 'red'", and I can also use the sample to make statements about coloured objects corresponding to it. However, I cannot use the sample itself to make statements about its own colour. Similarly, we can neither say that the standard metre is one metre long nor that it is not one metre long (cf. PI, § 50. For Kripke's misunderstanding of Wittgenstein's discussion of the standard metre cf. Munz 2023, III.2.1).

In the case of, e.g., a colour sample, such as 'sepia', we define the word by saying that it means the colour of the standard sample sepia, which is kept somewhere sealed. But once we have done this, it will no longer make sense to say of the sample either that it is the colour sepia nor that it is not. And this is because the sample is an instrument of our *language*, which we use to make colour ascriptions to particular objects: "In this game, it is not something that is represented but is a means of representation" (PI, § 50). The reason for this is that Wittgenstein realised that the paradigm is a component of the *symbolism* and not of the corresponding objects to which it is applied. In the context of his example "in this room, there is something red", written with a red pen, Wittgenstein remarks:

> This problem is connected with the fact that in an ostensive definition I do not state anything about the paradigm (sample); I only use it to make a statement. It belongs to the symbolism and is not one of the objects to which I apply the symbolism. (PG, 346)

To me, this remark seems crucial for the whole understanding of the concept of an ostensive definition, and it shows Wittgenstein's departure from his own Tractarian view about the relation between a name and an object.

In *Eine Philosophische Betrachtung*, Wittgenstein uses the example of *Nothung* to discuss his own misunderstanding of the name-object relation in the context of simples. He thereby admits that he had the erroneous idea that an object must correspond (*entsprechen*) to a name in order to have meaning, and that he had confused the *meaning of a name* with the *bearer of a name* (cf. EPB, 158). Wittgenstein picks up the sword example again in PI, § 39, and in PI, § 40 he remarks:

> It is important to note that it is a solecism to use the word "meaning" to signify the thing that 'corresponds' to a word. That is to confound the meaning of a name with the *bearer* of the name. (PI, § 40)

In this context, Wittgenstein also admits that we often point to the object and say the name when we give an ostensive definition. Similarly, we may point to a thing and say 'this' and a name and the word 'this' would occupy the same position within a sentence. But still we get the impression that 'naming' is a kind of an 'occult' process:

> Naming seems to be a *strange* connection of a word with an object. – And such a strange connection really obtains particularly when a philosopher tries to fathom the relation between name and what is named by staring at an object in front of him and repeat a name or even the word "this", innumerable times. (PI, § 38)

This remark seems to me quite similar to the idea of giving a name to my private experience by simply 'associating' the name with the sensation and using the name in a description, as Wittgenstein puts it in § 256. In his "Notes for a Philosophical Lecture" probably, from 1942, he readopts the picture of an occult act:

> Meaning consisting of the word referring to an object.
> Under what circumstances pointing can explain i.e. convey the use of a word. Not to a baby. *It* learns by being drilled. There is therefore no occult act of *naming* an object that in itself can give a word a meaning. […]. The private object. The naming of the private object. The private lang[uage]. The game someone plays with himself. When do we call it a *game*? If it resembles a public game. The diary of Robinson Cr[usoe].
> So we mustn't think that we understand the working of a word in language if we say it is a name which we give to some sort of pr[ivate] experience that we have. The idea is here: we *have* something it is as it were before the mind's eye (or some other sense) and we give it a name. What could be simpler? One might say / could put it roughly this way: All ostensive definition explains the use of a word only when it makes one last determination, removes one last indeterminacy.

> The relation of ostensibly defining. That's to say, in order to establish a name relation we have to establish a technique of use. And we are misled if we think that it is a peculiar process of christening an object which makes a word the word for an object. This is a kind of superstition. So it's no use saying that we have a private object before the mind and give it a name. There is a name only where there is a technique of using it and that technique can be private; but this only means that nobody but I know about it, in the sense in which I can have a private sewing machine. But in order to be a private sewing machine, it must be an object which deserves the name "sewing machine", not in virtue of its privacy

but in virtue of its similarity to sewing machines, private or otherwise. (NPL, 447-8)

Similarly in the case of a private language: in order for something to be a private language it must deserve the name 'private language', but not in virtue of it being private but in virtue of its similarities to *other languages*.

Rhees notes:

The confused idea that the meaning of the word "pain" (or "seeing") is what you would point to if you were *explaining* what it means. Whereas the pointing – the ostensive definition – is an explanation only if it makes you understand how to use the word in the rest of the language game (which may be complicated). (Rhees 15 April 1965, unpublished)

Similarly, in MS 116, Wittgenstein points out that it seems clear that "understanding the word" is one thing and "being able to apply the word" is another. This arises because we are accustomed to accepting the ostensive explanation as the final answer to the question "Do you understand the word...?" (MS 116, 144)

When Wittgenstein discusses the concept of an ostensive definition in PI, §§ 26-40, he points out that

an ostensive definition explains the use – the meaning – of a word if the role the word is supposed to play in the language is already clear. [...] One has already to know (or be able to do) something before one can ask what something is called. (§ 30)

Similarly, if someone shows me the king in chess and says: "This is called 'king'", this only tells me the use of the piece if I already know the rules of chess but not the shape of the king. The explanation that some figure is called 'king' can only tell me its use when its place in the game is already prepared. It can only work as a definition if I already know what a piece of chess is. Therefore, I can only ask what a particular piece is called if I already know what I can do with the name. In other words, I must already be able to master a language game in order to understand an ostensive definition in the first place. If, however, someone objects that all you really need is to know or to guess what the person who wants to give the definition is pointing to, I would answer that an ostensive definition could always be interpreted in different ways. But neither 'to mean' nor to 'interpret' the explanation in a particular way is an occurrence that accompanies the giving of an explanation and the hearing of it. Finally, I can say that a particular piece of the game is called 'king' but not the particular bit of material that I am pointing at (cf. §§ 31, 33-5).

These points are crucial for an understanding of the private language discussion: a 'private' ostensive definition will never work to fix a rule for the meaning of a 'private object'. I am pointing to a piece of a game and not to a particular 'private object' when I give an ostensive explanation of something. Therefore, just before introducing the private diary case, Wittgenstein remarks:

> When one says "He gave a name to his sensation" one forgets that much must be prepared in the language for mere naming to make sense. And when we speak of someone's giving a name to pain, the grammar of the word "pain" is what has been prepared here; it indicates the post where the new word is stationed. (PI, § 257)

On 9 April 1954, Rhees notes:

> It is impossible to *say* anything privately. (The only way an expression could mean anything privately would be by *confusing statements and definitions*. And then it does not matter what you say. There are no rules. If it does not make any difference what you say you say nothing. Since there are no rules, there is no language.) "Why can I not decide myself what following the definition is going to be?" When? Each time? If I make one decision – one ruling – once and for all, this only renews the problem: what is going to count as "following that ruling"?
>   "But I know from the public language how to follow a ruling". That is no help in this sort of case (or it would be no help). (Ayer asks, "Why can I not be trusted?" because he thinks the private language refers to objects that could be spoken about publicly, if only other people could 'see' what I refer to and "follow the description". This may underlie his query whether, if I am to understand a "descriptive statement" I must observe what it describes. This would be relevant if the issue were whether I can talk about private objects, in the sense in which no one else can observe. But if that were the issue, there would be no reason why I should not describe those objects in the language of communication and in terms that are used for "public" objects. Ayer takes the proposition that language must be public to mean that it must refer to public objects. That is not the point. The point is that it must be spoken by many people.) The question is whether the *language* can be private, not whether it can refer to a private object. If we do speak of a "private object" here, this is something different from "an object that only I can observe". It should make as little sense to talk about anyone's seeing this object as it would to talk about anyone's feeling pain. Ayer seems to recognize as the chief of the points which he is disputing "that for a person to be able to attach meaning to a sign it is necessary that other people should be

capable of understanding it too". (And he evidently does not think secret codes are relevant here.) But in fact, he hardly discusses this question. He simply asks whether it is inconceivable that Crusoe should "name" many things upon the island, and then suggests that it is not self-contradictory that someone should have invented language. He then takes "naming" many things to be the same as "inventing words to describe" many things. And "his justification for describing his environment in the way that he does will be that he perceives it to have just those features which his words are intended to describe" (Ayer 1954, 71). This is all of it question begging.

Ayer seems to take "Sensation coupled with outward expression" to be something like "abscess coupled with swelling" (or "fracture coupled with swelling"). It is in that sense that he speaks of "private sensations" as sensations that just lack outward manifestation. He seems to think of talking about a sensation as though that were parallel with talking about an abscess.

> The reasons why the accustomed – public – criteria for following a definition or following a rule cannot apply in the case of a "private" definition (which only I can understand): That sort of transference is possible in connection with secret languages, languages that another *could* understand. I know how such things are talked about. I know how such statements are taken. But not in the case of "what I can only say to myself". There is no sense in asking how this is taken. Nor is there sense in my deciding how I am going to use or follow the definition. This is not so much because it cannot really be a decision at all. Perhaps because *there is nothing to decide*. (Rhees 9 April 1954, unpublished. Last italics added)

Accordingly, about Ayer's semantic rules, Rhees notes:

> Ayer's meaning rules could be understood by or followed by – another person. But apparently, this is not essential. If I recognize these rules, I recognize what they prescribe anyway. But I do not see what the sense of "rule" or "prescription" is here. I do not see how they prescribe or what authority they would have. And I do not see what decides whether I am following the rule or not. (This is really another way of saying that I do not see in what sense it is a rule.) (Rhees, undated)

Particularly in the case of sense data, the meaning of a word was supposed to be what is referred to. Then it seems what is referred to may actually not be the same for different persons, and therefore the meaning would also be different in each case. But the meaning of a

word is not constituted by the various things which it might "correctly refer". So in the case of a colour concept, such as 'red', the meaning of 'red' is not all the various patches the word can correctly refer to. Rhees notes:

> [T]his account of meaning makes the same confusion between giving a definition – giving meaning – and using a word in its meaning, as is made regarding names of physical objects like "stone". (Rhees 30 June 1954, unpublished)

Later in the same typescript, Rhees continues:

> Now the people who discuss private languages seem to think it would be the same kind of case when I am supposed to give names to something, which I can talk about only to myself. And the "can" there is a logical "can": giving names to something of which it makes no sense to say that I could talk about them to anybody else, just as it makes no sense to say that another person could or could not feel my pains. (That raises difficulties about identity because one is inclined to speak of pains as objects, in certain contexts, and the question is whether he can feel what I feel. In one sense, of course, he can feel what I feel.) If I am supposed to talk about something to myself of which it makes no sense to say that I could talk about it to someone else, well then one has to insist again that it makes no sense to say that I could talk about it to myself either. And the whole business about whether one could remember is relevant for that. (Rhees 30 June 1954, unpublished)

Unless there *are* rules of a language, it is no use arguing that something is the same as it was two days ago. Similarly, in the case of sensations such as an itch, it appears as if one directly perceives what the itching is. Once the sensation is named, it seems as though the rules governing the subsequent use of that name are already determined by the sensation itself. But this impression is illusory. Even the feeling of itching gets its identity solely through a shared practice of expression, reaction, and language use (cf. Candlish 2019).

In "Can There be a Private Language?", Rhees makes the same point in the case of colour concepts. If someone says, "This is the colour 'red'", she would give us a definition by showing us a particular sample, but:

> Someone might say "I know what *I* mean by 'red.' It is what I experience when I look at this". […] I suppose the point would be that I know this independently of having learned the (public) language. If I know what I mean, in this way – if I know what colour I am referring to – then apparently I have done something like giving myself

a definition. But I must also have confused giving a definition and following a definition. It is this which allows me to evade the difficulty of what I am going to call 'following the definition'. Which is a real difficulty: what could it mean to say that I had followed the definition – 'my' definition – incorrectly? But if that has no sense, then what on earth is the point of the definition? And what does the definition *establish*? (Rhees 1954, 80-1)

A few pages later, Rhees again makes the point about confusing definitions with statements. He writes:

I cannot say anything unless I know the language. But I cannot know the language – any language – privately. I may have a secret code, but that is not the point here. It is a question of whether I can have a private understanding; whether I can understand something that *could* not be said in a language anyone else could understand. ("He may understand the language I speak, but he will not understand what I understand.") I say I cannot know a language privately, for what would there be to *know*? In language, it makes a difference what you say. But how can it make any difference what you say privately? (I do not mean talking to yourself.) It seems that in a private language, everything would have to be at once a statement and a definition. I suppose I may define a mark in any way I wish. And if every use of the mark is also a definition – if there is no way of discovering that I am wrong, in fact no sense in suggesting that I might be wrong – then it does not matter what mark I use or when I use it. (83)

I have extensively used Rhees' remarks to support the claim that in the case of private objects as the meanings of private sensation names, the ceremony of naming these sensations is a futile ceremony. We are not able to say what it would mean to follow the rule rightly or wrongly. Consequently, there is actually nothing to be followed. If a rule for the meaning of a word cannot be *established* in the first place, then the question of following a rule rightly or wrongly will not appear. We could not say anymore whether we are indeed following a rule or rather always *fixing* new meaning rules when we mistakenly assume we are using the rule. This problem is precisely due to the confusion between giving a sample and using a sample.

The extent to which the claim of private experiences is a grammatical rather than a psychological fiction is further illustrated by a passage from the "Notes for Lectures on 'Private Experience' and 'Sense Data'". Wittgenstein remarks:

The private experience is to serve as a paradigm, and at the same time admittedly it can't be a paradigm. The 'private experience' is a degenerate construction of our grammar (comparable in a

sense to tautology and contradiction). And this grammatical monster now fools us; when we wish to do away with it, it seems as though we denied the existence of an experience, say, toothache. (LPE, 314)

In his "Lectures on Description", Wittgenstein makes a similar point, arguing against a behaviouristic reproach:

Nevertheless, it is entirely correct to say that in the description of a language-game, the mention of an experience as the justification of what he says does not enter. What I want to show is that this is not behaviourism, although it sounds like it. It is not behaviourism if one says that, in the description of a language-game, we don't use having a certain experience [as the justification] for his saying something. But we only use the circumstances, and, of course, here you see in a way the reasons why people say such things as "Experiences are private". I might put it in such a way: nothing private enters the description of a language-game. You might say: "If nothing private enters the description of a language-game, that means there is nothing private, there are no experiences." I'd say: "Not at all. It could enter it by saying 'And then he sometimes says this' instead of 'whenever...'", etc. I want to show you: it doesn't mean, "There are no experiences", if they don't enter into the description of a language-game. (WCL, 158)

We find a similar remark in PI, § 304, when Wittgenstein again points out that a 'nothing' would just do the same job as a 'something' about which nothing could be said. Therefore, we have to give up the idea that language only functions in one particular way and that is to convey thoughts, may they be about objects, pains, or anything else (cf. PI, § 304). In "Lectures on Similarity", Wittgenstein gives us a further explanation about sensations in such contexts:

When you say that I tend to talk behaviouristically, you forget that I am not talking about pain, but about the use of the word "pain". If I was talking about pain I'd say, "I've got intolerable pain". (a) I seem to talk about a certain phenomenon, (b) denying it existed. In a sense, ruling out something. How can I rule out anything? Suppose one of us had a rather intense pain. "Surely, you don't wish to tell me this is nothing. Surely, this is something." Suppose you say this is something, you might seem to be quarrelling with me, whether it is something or nothing. But how are we to decide this? Either [this is a] routine [case], and I was saying, "You have no pain". Or, I am saying you're using "there is something" in an inappropriate sense. Someone could be said to be a behaviourist by saying, "If you have pain, you've got nothing", meaning

"got" is a very inappropriate word. If I said, "It is a very appropri-
ate word", then I would not be said to be a behaviourist. Only in
one way could I wave something aside, and that is by saying it is
an inappropriate word, etc. This is talking about the appropriate-
ness of language. (WCL, 116-17)

It is for this reason that we can eliminate the private object (cf. RPP,
I, §985), and even if some "private beetle boxes" of a community of
people were empty, the word 'beetle' might still have a use, and we
could 'divide through' the thing in the box, mathematically speaking
(cf. PI, § 293. Cf. also PI, § 304; WCL, 94).

In his "'Private Language' and Wittgenstein's Kind of Behaviour-
ism", C.W.K. Mundle shows precisely this kind of misunderstanding
of paragraph § 293, first by raising a verificationist blame, and sec-
ondly by assuming that Wittgenstein is precisely denying the exist-
ence of experiences:

> We have each learned to call "red" the shades we each see when
> we look at human blood, ripe tomatoes, etc.; and no one can verify
> that the shade he sees when he looks at a ripe tomato is the same
> as, or similar to, the shade seen by another when he looks at the
> same tomato. But Wittgenstein went too far when he said "the box
> might even be empty", i.e. that for all I know the box of any other
> person might be empty, i.e. that for all I know other people may
> have no private experiences. (Mundle 1966, 44)

It is, of course, absurd to assume that Wittgenstein denies the exist-
ence of 'private' experiences in Mundle's sense, as many of his re-
marks have already shown (what would 'public' experiences mean
in this context?). Peter Geach, too, points out that Wittgenstein has
often been accused of denying the existence of any mental acts and
that his remarks about "private objects", as, e.g., in PI, § 293, could
actually be taken this way (Cf. Geach 1957, 3). But certainly, Wittgen-
stein would never have denied the fact that people do have a 'private'
mental life in the sense that they might not tell anyone about their
present mood or try to hide that they are in pain, etc. Moreover, we
consult a doctor when suffering from a particular pain. In PI, § 256
Wittgenstein himself speaks of "the way in which we ordinarily re-
fer to our feelings", which obviously is a way of speaking about our
feelings. Similarly, there is absolutely no reason

> why I should not give an account of something which only I can
> see. Or of something which only I can feel: as when I tell a doctor
> what I feel in my abdomen. He does not feel my sensations (if that
> means anything), but he knows what I am talking about; he knows
> what sensations they are. (Rhees 1954, 84)

Wittgenstein himself points out:

> *How* do I use words to signify my sensations? – As we ordinarily do? Then are my words for sensations tied up with my natural expressions of sensation? In that case, my language is not a 'private' one. Someone else might understand it as well as I. – (PI, § 256)

One crucial point here is that Wittgenstein does not deny a conceivable private reference of a psychological expression, in the sense that it denotes a kind of experience that could be characterised as 'private'. For some reason, I may not want to tell anybody about it; I may try to hide any behavioural expressions, etc. What he does deny is the possibility of giving such linguistic signs a private meaning by means of private ostensive definition. To that extent, the assumption of private objects proves irrelevant, and in this sense "a Nothing would render the same service as a Something about which nothing could be said" (PI, § 304). Wittgenstein introduces a similar picture to the beetle box in PI, § 271: a person who is unable to remember what the word 'pain' means and therefore constantly calls something else by the word but does nevertheless use 'pain' as we all do with the usual symptoms and presuppositions of having pain (cf. PI, § 271). This again shows that the assumptions of a private object and a private meaning are idling wheels.

One way of interpreting the empty box case introduced in PI, § 293 might be Malcolm's introduction of a person, also called Robinson, who has never experienced any pain but has still been able to gain an understanding of what it means for other persons to be in pain. Hence, we could imagine, without implying any contradiction, that even in his own case, Robinson might meaningfully say that he is not feeling pain when the doctor pricks him with a needle (For a discussion of Malcolm's example, see Munz 2023, 198-200).

In the final section of this chapter, I will basically just quote some of Rhees' replies to Ayer's Crusoe in some of his letters. In these remarks, he makes a point, which I think is crucial for the understanding of an essentially private language and which has not yet been discussed in greater detail, at least as far as I know. The point will be that even observations of regularity and similar behaviour to ours might not be sufficient to judge whether the person observed does indeed *use* a language, because *interpreting* a language is different from *understanding* a language. In other words, to interpret a particular pattern as a language from *our* point of view as language users does not necessarily imply that the inventor of a particular pattern herself has introduced it as a system of rules for possible language *use*.

2.3   Robinson Crusoe and the Wallpaper Pattern

In his symposium paper, Ayer gives us a description of his Robinson Crusoe that first seems to remind one of the harmless cases Wittgenstein himself introduces in his manuscripts. But then it turns out to be one that was introduced to defend a private language in the strong sense: A Robinson Crusoe is abandoned on an island as an infant, never having learned to speak. Like Romulus and Remus, he is nurtured by a wolf or another animal until he becomes self-sufficient and reaches adulthood. In adapting his behaviour to various features on the island, he would undoubtedly be able to recognise many things. And Ayer asks whether it were unthinkable that he might also assign names to these things? There could, of course, be psychological reasons to question whether such a solitary individual would actually create a language. It may be argued that the development of language is a social phenomenon. It does not, however, seem to be inherently contradictory to consider that someone untrained in the use of any existing language could still invent a language for himself. And if we allow that Crusoe could make up a language and invent words to describe his surroundings, why should we not also allow that he could invent words in order to describe his sensations? (Cf. Ayer 1954, 70.) In the previously quoted letter to Ayer, Rhees writes:

> If you should ask "Might there not be a language that could be understood by various people, even though some of its words are names of objects that only one person can observe?", then I think I should say yes, and I think Wittgenstein would. But I do not know that we should have agreed on anything very important.

After the already quoted passages (see above pp. 206-7), Rhees continues:

> It seems to me that one important difference between your view and mine lies in the notion of "giving names to things" or in "using a name to refer to something". I may have read you wrongly, but you seem to me sometimes to hold that the meaning of an expression is what it refers to or what it indicates. (Cf. PI, § 264: "'Wenn du einmal weißt, was das Wort bezeichnet, verstehst du es, kennst seine ganze Anwendung'" [Quotes are in the original and are of course essential]). And you seem to hold that you can tell whether you are using a word in the same way by seeing whether you are using it to refer to the same thing. That would seem to make "referring to the same thing" in a way prior to "using the word in the same way" or "in the same meaning". And that it certainly something I should dispute, and I think Wittgenstein would. I do not think that a word can be said to "refer to" anything at all except

in connection with a rule, except where there is a difference between the right use and a misuse of the word. One might even say, as used to be said, that it refers to something in virtue of what it means, but then what it means is not what it refers to.

Another important difference – or at any rate a difficulty – for our discussion is connected with the notion of "privately". Wittgenstein would agree that there is a certain sense in which you might say "sensations are private", although then this is a grammatical proposition, like "patience is something you play by yourself" [cf. PI, § 248]. But he would not ever hold that they are private *objects*. In most contexts, it would be misleading to speak of them as objects at all.

May I call attention to one place where I think you have misunderstood him. You speak, for instance, on page 76, of "something that is naturally associated with what it describes, in the way that feelings are associated with their 'natural expressions'", and it looks as though you thought that were a way in which Wittgenstein might have spoken. But he does not say that a feeling or a sensation is *associated* with its natural expression. He says (256) „Sind also meine Empfindungsworte mit meinen natürlichen Empfindungsäußerungen verknüpft?" ["Then are my words for sensations tied up with my natural expressions of sensation?"] Is *the meaning of the words* bound up with the expression of the sensations? He would not have held that the sensation was something that could be associated with an expression, as a disorder of the liver might be associated with a discolouring of the eye, for instance. The natural expression of the sensation is in no way a *symptom* of the sensation. There is not the expression plus the sensation. And when I see another person in pain, I am not indirectly aware of his sensation, as a doctor may be indirectly aware of the condition of my liver. As I say, this goes together with the fact that sensations are not objects, nor processes either. This is important on its own account – this matter about sensations – and it is important because something very analogous holds of understanding and of meaning what you say. "Understanding is private" has much in common with the discussion of "sensations are private". And this brings us back to the distinction between following a rule and interpreting a rule again. (Rhees 1954, unpublished)

This is why Wittgenstein argues that we learn the meanings of sensation names by learning new pain-behaviour. A word such as 'pain' is connected with a primitive and natural expression such as 'ouch' and used in its place. This does, of course, not mean that the word 'pain' just *means* 'ouch'; on the contrary, it *replaces* the cry. Therefore, "I

am in pain" is not a *description* of my inner state but instead takes over the role of a primitive reaction, such as a cry. Or how could we try to interpose between the pain and its natural expression (cf. PI, §§ 244-5)? To take the natural expression as a *symptom* for someone having a particular sensation does, however, erroneously open the door for the possibility of private objects and the discussion of pain with and pain without behaviour, which Wittgenstein picks up in PI, § 304 and elsewhere.

In his reply to Ayer, Rhees tries to show the internal relation between referring to something and understanding something. If my words are to refer to anything at all, they must be understood. In other words, reference only works in connection with a particular use which one learns when one learns what the expression means. Words cannot refer to anything without a manner in which the language is employed. This is why there cannot be a private understanding. If what is said does not make any difference, then nothing is understood. Certainly, there is no inherent reason why I should not be able to give an account of something visible only to me or something only I can feel, such as when I describe a particular unpleasant sensation to a doctor. The doctor does, of course, not experience my sensations, but he understands what I am talking about; he comprehends the nature of those sensations. Now, Ayer asks, why Crusoe should not be able to also devise names for his sensations? I can, of course, invent a name for a sensation, but this is because I speak a public language that already contains names for sensations. This is why I know what a name for a sensation is. In other words, to invent a name or to give a name to something belongs to the language as we already speak it (cf. Rhees 1954, 84-5). It is not that someone could just invent a particular language because language is internally connected with a particular way of living. Rhees remarks:

> A man might invent marks to go with various objects. That is not language. And when Ayer's Crusoe invents *names* to *describe* flora and fauna, he is taking over more than he has invented. He is supposed to keep a diary, too. Ayer thinks that if he could do that when Friday was present he could surely have done it when he was still alone. But what would that be – keeping a diary? Not just making marks on paper, I suppose (or on a stone or what it might be). You might ask "Well, what is it when I do it? And why should it not be the same for him, only a bit more primitive?" But it cannot be that. My marks are either marks I use in communication with other people, or they stand for expressions I use with other people. (87)

But what difference would it make, and why cannot Crusoe use his marks in the same way I do? It is because I can use them in their various meanings, something Crusoe is not able to do. Rhees notes:

There seems to be nothing logically absurd in supposing that he behaves just as I do. To a large extent I agree. But it is absurd to suppose that the marks he uses mean anything; even if we might want to say that he goes through all the motions of meaning something by them. (88)

Therefore, if I am supposed to be saying anything at all, this must take place in some language: "If there were no more than my behaviour, the marks I make and so on, then I should not mean anything either" (88). Ayer's depiction of Crusoe suggests that he might use marks for specific purposes, such as indicating the location of hidden items. He could do this with as much regularity as we can imagine. However, this is not what is meant by the regular use of an expression in a particular language. If Crusoe were to suddenly employ these marks in a completely different way, it would not make sense to claim that he had done something wrong or inconsistent with his prior actions. We could not say he used them with the same or with a different meaning. Even if he consistently used them for the same purpose, as he might always gather wood for the same purpose, this does not capture what we mean by 'using an expression in the same way'. Using an expression in the same way is not the same as using it for the same purpose. Moreover, any discrepancy between what I say at one time and what I say at another does not imply that my actions with a mark or sound at one time are different from what I did before. If I have consistently done the same thing with the mark, there is no implication of a language rule, in that following a particular rule or word must always allow the possibility of misunderstanding it or making a mistake. Rhees notes:

> Ayer's Crusoe may make the kind of mistakes animals do. He may mistake a bird which he does not like to eat for one which he likes. This is not like a mistake in understanding the meaning of an expression or a mistake in following what was said. "Why not? He calls the edible bird 'ba', and when he sees the inedible one he says 'ba' and kills it." That is not a mistake in following the meanings of words. He could have made the same mistake without using words at all. (92)

I will finish this section by quoting from a few more letters, two of them to John Wisdom, already quoted, and one to Elisabeth Anscombe, which will make Rhees' point still clearer.

In a letter to Wisdom, dated 15 July 1954, Rhees expresses his difficulty in saying that a man who had never known society might speak a language, and what is the difference between saying something and just making marks or sounds, which is closely connected with the general relation between language and things. He writes:

The speakers [during the symposium 1954] told me all the things Crusoe might do. But I do not think it really matters what he could do. I said that the question of whether he could speak a language is something an experiment could not decide. If someone ever does come across a man who has grown and lived as Ayer supposed Crusoe did, it will not make our question any easier to answer. The question is whether such a man could speak a language. And we cannot answer this by watching what he does. Meaning goes hand in hand with following a rule. And that is connected with the fact that, as a rule, expressions are taken in these and these ways. That is what makes the difference between saying something and just making a sound or a mark. I cannot find that Ayer discusses this, and the meeting did not.

Making drawings of birds and so on – even if you call them "charts" – that brings me no nearer to speaking if I have nothing to which I can appeal to know if I have followed them correctly or not.

If there is no way of telling whether he has followed the drawings correctly – then it does not make sense to ask. "But he might make all the signs of having made a mistake." But if that is all there is to it, then if he does not make the "signs of dissatisfaction", if he remains serenely convinced that it is all right – then it is all right. The only criterion of having gone wrong is that he *thinks* it is wrong and feels dissatisfied.

On this basis, there could be no such thing as intelligibility. Nor any such thing as trying to understand what he means or what he is saying. He is not saying anything. That is why I say it does not matter *what* you see him doing. Unless there is some rule to which he can appeal, he is not speaking and he is not writing. To say that he could *give himself* a rule by drawing so and so is no help – just because there would still be no difference between his being sure that he was following the rule and his following the rule. I said that although Ayer's Crusoe might conceivably use the same marks and sounds as I do, he could not use them in their various meanings. […] If you should ask, "If you find someone doing all these things, would you not be inclined to say he had a language?", well, of course, I should. But then I should think (and so would you) that he had at some time known human society. If we were assured that he had not, I should find his behaviour remarkable. But I should not think he was using marks and sounds as expressions used in a language. (Rhees, unpublished)

In a subsequent letter to Wisdom, dated 8 August 1954, Rhees writes:

Wittgenstein in the *Investigations* was not directly concerned with the case of the congenitally solitary man. I heard him refer to that sort of case only once, and then he did not go into it fully. He said, as I remember, that if you found a man making noises like that, which seemed to be regularly connected with what he was doing, then you might feel some doubt about what to call it – it would clearly have certain analogies with language. He mentioned it when he was discussing the foundations of mathematics. And I never heard him return to this.

The emphasis, if I may repeat, should be on the question of how words mean, or on what "saying something" is, and not primarily on the question of how words refer to sensations, for instance.

I think there must also be a difference between 1) following a rule and 2) doing just what someone would do if he were following a rule.

I do not think that he would be following the rule unless it were conceivable that he should go contrary to it, and I do not think you could say of anyone that he had gone contrary to the rule unless you might conceivably bring him to see that he had done so. Understanding a game is not like understanding a mechanism.

Suppose a man had never done a calculation in his life and had never been taught what a calculation was, and one day he suddenly wrote down a calculation in arithmetic. I do not think we would say he was calculating, or that he was following the rules of arithmetic.

A rule of calculation *cannot* be something that is followed only once. He cannot follow a rule just once in his life.

I want to say that if he has never *learned* to follow a rule, then whatever he does he will not be following a rule. If he has never learned to calculate, then whatever he does he will not be calculating. [...] (That is why I wanted to insist against Ayer that following a rule is entirely different from following a habit.) [...] If a man speaks to himself, he speaks in some language, and a language is spoken by others. Otherwise, he would not be saying anything to himself. May I repeat: I do not say that a man cannot speak a language unless he speaks it with others. I say he cannot speak a language unless it is spoken by others. And I would add also: he has learned it. (Rhees, unpublished)

Finally, I will quote a few remarks Rhees wrote to Elisabeth Anscombe on 25 July 1954, where he gives a better understanding of

what he means by the metaphor of a 'wallpaper pattern', which he uses in his reply to Ayer. Looking back to the notes I took during our discussions in spring 1989, just a few months before his death, Rhees still talked about the picture of a wallpaper pattern when arguing that some philosophers make philosophy look like a piece of wallpaper with its pattern and without any *perplexities*. It must have been an important picture for him.

In his letter to Anscombe, Rhees writes:

> You might find some rule of translation that would give you a series of mathematical propositions and even proofs. Put the other way about, this would mean that we could express mathematical proofs in the designs of a wallpaper pattern. But this would not show that those who designed the wall were calculating and proving things in mathematics when they were doing it. It would not show that the marks on the wallpaper were mathematical expressions at all. They would become so when we used them as such; but not before. [...] I might put this by saying that since the wallpaper designs do not play the role of mathematics, they are not mathematics. And they would not be that even if they happened to be the same as the marks which we might write on a sheet of paper in doing a calculation. The fact that they may be used by us in doing mathematics does not show that those who made them were doing mathematics.

In the case of the marks and noises, a solitary man might make, Rhees remarks:

> The fact that we could find some rule of transformation that would turn these marks and sounds into sensible expressions would not show that those marks and noises played the role of language there at all. What does it mean to understand what is being said or to understand the language that is being spoken? And I said that this does not mean just that you can interpret or transform the marks and sounds that are made into the expressions of language which you understand. [...] Understanding what is said is not interpreting what is said. [...] Understanding the language means knowing the language. [...] Knowing what it means is not interpreting. It is knowing a rule (following a rule) and that is following a practice. [...] Anyway, the fact that you can "interpret" what he utters to make it correspond to the expressions of some language will not show that he is saying anything in any language himself. And it will not show that his utterance is capable of being understood. [...] You might complain here that I have made an inexcusable jump. I suggested that if he says anything, he must say it in some language that he has spoken. And now I am suggesting that if he

has not known human society he cannot be saying anything. [...] A machine would not be saying anything even if it made sounds that were the sounds of the English language and which came in the form of intelligible English sentences (or made marks, which were letters of the English alphabet, etc.) Let it be logically possible that such a machine should be produced by natural causes. It would not be saying anything more than your tape recorder says anything. And no more would Ayer's Crusoe, if he happened to utter what we should recognize as an English sentence. Whether it plays the role of a language: that does not depend on whether you can interpret the marks and sounds as the expressions of a language. It is not that question at all.

All these remarks point to a problem whose importance has, in my view, not yet been appreciated. The distinction between the *harmless* and the *strong* cases of a private language is not easy to draw, and even if it were, it would still be a very thin line. In addition, as we already know, context is a very complicated thing and the answer can never be categorical. (cf. WCL, 239, 168)

3   Subjective Knowledge and the Private Chart Inspector

In PI, § 246, Wittgenstein raises the question of the sense in which sensations might be private and offers an initial answer, stating: "Well, only I can know whether I am really in pain; another person can only surmise it". His immediate response to this is that "in one way this is wrong, and in another, nonsense" (PI, § 246). There are cases where we do know that another person is in pain. For example, if we observe someone being knocked down by a car, crying out, squirming, and asking for a doctor, we would immediately recognise that the person is in severe pain. One might choose to say, "I believe he is in pain" instead of "He is in pain", but that's about it. What seems like an explanation here is essentially an exchange of one expression for another, of which the former seems more appropriate. However, Wittgenstein emphasises: "Just try – in a real case – to doubt someone else's fear or pain!" (PI, § 303). While there is a possibility of being wrong, such as when a person is acting in a film or pretending to be in pain, it does not imply that we could *never* know whether another person is in pain.

Wittgenstein's criticism of sentences like: "Only I can know whether I am really in pain" is tied to the grammar of the word 'to know', which must always allow for the possibility of being mistaken about what one assumes to know. This relates to the use of the expression "I only thought I knew but I did not". When we take the word 'know' to be describing a particular state of mind, we tend to eliminate the

expression "I thought I knew" because we believe it is simply wrong. After all, we cannot know what did not happen. We want to reserve the term 'to know' for situations where we cannot be wrong, such as when talking about our impressions or 'sense-data' (cf. WCL, 37, 43). If the word 'to know' describes a state of affairs meant to guarantee what is known, then we forget the use of the expression "I thought I knew" (cf. OC § 12). Whereas I can say that I thought I knew the capital of Austria but indeed I did not, it does not make sense to say that I thought I knew that I was in pain, but in fact I did not. The grammar of 'to know' must always allow the possibility of being wrong.

Similarly, in the case of another person, if one says: "I know only indirectly what she sees, but directly what I see", this conveys a misleading picture. The example of the matchbox is employed to illustrate this:

> I can't be said to know that I have toothache if I can't be said not to know that I have toothache. I can't be said to know indirectly what the other has if I can't be said to know it directly. The misleading picture is this: I see my own matchbox but I know only from hearsay what his looks like. We can't say: 'I say he has toothache because I observe his behaviour, but I say that I have because I feel it'. (LPE, 319)

This picture is already familiar from PI, § 293, where Wittgenstein asks if one who maintains that she knows what the word 'pain' means only from her own case, must not also admit that this holds for other people too? Everyone would claim they know what pain is *only* from their own case, leading to the strong case of understanding experiences as *private*.

Wittgenstein now extends this discussion of knowing one's own private sensations to the far more interesting case of colour concepts:

> The essential thing about private experience is really not that each person possesses his own specimen, but that nobody knows whether other people also have *this* or something else. The assumption would thus be possible – though unverifiable – that one section of mankind had one visual impression of red and another section another. (PI, § 272)

The discussion then explores whether one could never know whether another person sees the same colour, even though they use the colour concept in the ordinary way. Thomas Nagel frames it as follows:

> How do you know that red things don't look to your friend the way yellow things look to you? Of course, if you ask him how a fire

engine looks, he'll say it looks red, like blood, and not yellow, like a dandelion; but that's because he, like you, uses the word "red" for the colour that blood and fire engines look to him, whatever it is. Maybe it's what you call yellow, or what you call blue, or maybe it's a colour experience you've never had and can't even imagine. (Nagel 1987, 21)

Rhees remarks:

> The idea that "there is a limit to what we can communicate". Or perhaps: "I can never know whether the content which you put into my words is the content which they had for me" – and so on. This is connected with such questions as: "Can I ever know whether anyone else sees the same colours as I do?"; and also with: "Can I ever know whether another person has the feelings or sensations which I have? Or even: whether he can *really* know *what* feelings or sensations I have. Obviously *I* know what feelings I have – but *you* can't really know this (although you may form conjectures and analogies) because to *really* know what they are, what they feel like, you'd have to *feel* them". Of course you do not think along these lines when you are telling the doctor how you feel. And this is what Wittgenstein emphasized over and over again. But people often get caught in this way of thinking when they are trying to give an account of the language. And in particular here: when they are trying to give an account of the language in which we express feelings and talk about feelings. [...] My point is that Ayer's reference to "my own capacity to recognize my experiences" does nothing to give sense to the talk of "private experiences" which Wittgenstein was considering in that passage [PI, §258]. [...] He [Ayer] has not begun to think about the difficulties in the notion of *giving a name* to something; the distinction between giving a name and using a name; the notion of telling myself (informing myself?) what experiences I am having – in a sense of "telling" in which I cannot tell anyone else; etc., etc. (Rhees, letter to Mundle, 3 April 1965)

Here again, Rhees points out the confusion between defining or giving a name to an object and using the name to make a statement. One reason someone might defend such a position seems to rest on the fact that we cannot compare our immediate impressions with those of other people, and that we do not have immediate access to other people's minds. We could not have a method of comparison in such a context. In the above remark § 272, Wittgenstein himself emphasises that the claim that one part of a society had one impression of red and another part another impression would not be verifiable. But then, in the "Lectures on Similarity", he introduces the splendid thought experiment of a mind inspector who is able to check the

assumed private charts of some people who were asked to make coloured marks on a blackboard. Wittgenstein starts by asking about the idea of a 'private regularity', which someone does not convey. Let us imagine a person having a piece of paper before her, which she does not show to anyone else. Or we might imagine her head as a box, and her being able to look inside it. This would be the image of a 'private picture'. Let us further assume that we all have such a private picture before us. Looking at what she does would then just be second best to showing us what she 'really' sees:

> Suppose I say, "Let's compare your private picture and my private picture". This, so far, makes no sense, unless we tell what we mean by it, what our method of comparing is. [...] Suppose that to draw on the blackboard with crayons – white, blue, etc. – each person has a private colour chart. [Two examples of a chart follow.] Each looks up the colour from the chart. The assumption is that I don't know what is on your colour chart. The colours on the chart may have changed. You look the colours up and point according to my orders – "Blue! Red!", etc. One day, someone comes in and inspects. Your red patch has changed to green, etc., or all your patches are grey. But you look the colours up from your chart and draw them or point to them correctly. Are we to say that the word "green" means something different to you than what it means to us, or are we to say that it means the same? We don't know here what "using the chart" means. Your chart now has nothing in common with an ordinary chart at all. (WCL, 94. Cf. also PI, § 275)

Thus, it is possible to imagine the case of correctly identifying a colour and following an order in spite of constantly changing private colour charts. Accurate reference can therefore not be based on the use of one's own subjective colour chart. We would not know what "using my private colour charts" *could* mean in such a case. And this precisely shows that the question of private charts is not something that is possible though not verifiable because Wittgenstein himself just turns it into an imaginary case of verification. To me, this just seems like passing the buck back. And I guess this would also mean that, according to Ayer, the persons inspected were making *mistakes* when they were following the order to choose a coloured crayon, like Crusoe, who shot the wrong animal 'ba'. But it is obvious that the problem is not one of being right or wrong in applying my private charts, but that the application of such a private chart would not make any sense:

> 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. (PI, § 265)

If one were to object that this experiment does not apply because it is impossible in principle to examine private colour charts of different subjects, then one can only *assert* the possibility of strictly private contents of consciousness without any justification, except with the argument of logical freedom from contradiction. This is why I think this thought experiment is so important, because it simply makes use of the other side of the 'nonsensical coin', so to speak. Of course, so far, it makes no sense to assume that I could look into another person's box. But it is as nonsensical to assume that our private charts would constantly change while we use the relevant concepts as we all do in our ordinary language, as it is to assume that each of us would have their own colour impressions and corresponding private tables as a justification for knowing the meaning of a sensation word or a colour concept. Therefore, I think both thought experiments, Nagel's person who could not 'really' tell us what she sees, and Wittgenstein's mind inspector, would be idling wheels. This is what I take "making unobvious nonsense obvious" to mean. Interestingly enough, Ayer himself seems to allow the conceivability of looking into someone else's soul. He writes:

> It is not even necessary to make the assumption that Man Friday comes to know what Crusoe's sensations are and so to understand the words which signify them through having similar sensations of his own. It is conceivable that he should satisfy all the tests, which go to show that he has this knowledge, and indeed that he should actually have it, even though the experience which he rightly ascribes to Crusoe is unlike any that he has or ever has had himself. It would indeed be very strange if someone had this power of seeing, as it were, directly into another's soul. But it is strange only in the sense that it is something which, on causal grounds, we should not expect to happen. The idea of its happening breaks no logical rule. (Rhees 1954, 74)

It would be interesting to know what Ayer means by "breaking no logical rule". Furthermore, shall we assume that those with the permanently changing private colour charts would *know* that their private samples are constantly changing? Would they be aware of the fact that, according to Ayer, they are continually making *mistakes*? Would some of them be surprised and consult an eye doctor? Would some of them answer that sometimes they have very strange visual experiences, tomatoes in the supermarket suddenly looking blue and bananas red? Would some of them question the inspector and insist that whenever they picked up the green crayon, they used their private green sample because "Only I can know what colour chart I have"? Would some of them be Russellians and say that they know the colour perfectly and completely when they see it, and that no further

knowledge of the colour itself is possible, even theoretically (cf. Russell 1956, 202)? Or would some of them say that they know what colours their charts have, only that they cannot tell us, like Nagel's example or Wittgenstein's private diarist who only makes scratches?

Even John Locke, who has often been quoted as a pioneer for the possibility of a private language, introduces a case of colour inversion and points out that such an assumption is of very little use. He argues that even in situations where an object elicits different ideas in different individuals, perhaps due to differing structures of their organs, there is no attribution of falsehood regarding their ideas. And one might add, "and no attribution of truth, either". Locke uses the example of a violet, which causes an idea in person A, corresponding to that of a marigold in person B, and vice versa. However, the possibility of this occurring would never be discernible, as the consciousness of person A could not penetrate the body of person B to perceive the appearance generated by B's organs (cf. Locke 2011, II: xxxii, 15, 389). Locke draws the conclusion that, in such cases, neither the ideas nor their corresponding names are confused. In situations where both individuals perceive a violet, they both use the term "blue" and are capable of making corresponding colour distinctions:

> For all Things, that had the Texture of a Violet, producing constantly the Idea, that he called Blue; and those which had the Texture of a Marigold, producing constantly the Idea, which he as constantly called Yellow, whatever those Appearances were in his Mind; he would be able as regularly to distinguish Things for his Use by those Appearances, and understand, and signify those distinctions, marked by the Names Blue and Yellow, as if the Appearances, or Ideas in his Mind, received from those two Flowers, were exactly the same, with the Ideas in other Men's Minds. (Locke 2011, II: xxxii, 15, 389)

Certainly, it is undisputed that there can be situations in which, for example, normal conditions such as lighting, visual ability etc. are not met, leading to potential perceptual discrepancies. Philosophically relevant here is the seemingly innocuous transition from such a specific situation, where the described deviations exist and are usually diagnosed, e.g., colour blindness, to the general cases where these differences are not supposed to be detectable. Wittgenstein again provides us with a thought experiment that initially describes the possibility of an intrapersonal inverse spectrum, as he notes:

> Consider this case: someone says "I can't understand it, I see everything red blue today and vice versa". We answer "it must look queer!" He says it does and, e.g., goes on to say how cold the glowing coal looks and how warm the clear (blue) sky. I think we should

under these or similar circumstances be inclined to say that he saw red what we saw blue. And again we should say that we know that he means by the words 'blue' and 'red' what we do as he has always used them as we do. (LPE, 284)

This would again be the philosophically harmless case in which a person experiences perceptual contents that deviate from others, based on the same extramental objects, such as the blue sky and a glowing piece of coal. However, from the example it becomes clear that this colour inversion has no impact on the meaning of the terms 'red' and 'blue', as the person undergoing the inversion uses the terms in accordance with our usage. We perceive the glowing coal as red and the sky as blue. The situation is different, however, in the following case, as Wittgenstein continues:

> On the other hand: Someone tells us today that yesterday he always saw everything red blue, and v[ice] v[ersa]. We say: But you called the glowing coal red, you know, and the sky blue. He answers: That was because I had also changed the names. We say: But didn't it feel very queer? and he says: No, it seemed all perfectly ordinary /natural/. Would we in this case too say: …? (LPE, 284)

The answer seems clear because we would not know on what grounds we could say that the person was using the terms 'blue' and 'red' in the same way as we do. Deviating perceptions, such as seeing violets and marigolds, can, according to Locke, be neglected:

> I am nevertheless very apt to think, that the sensible Ideas produced by any Object in different Men's Minds are most commonly pretty near and undiscernably alike. For which Opinion, I think, are many Reasons offered: but that being besides my present Business, I shall not trouble my Reader with them; but only mind him, that the contrary Supposition, if it could be proved, is of little use, either for the Improvement of our Knowledge or Conveniency of Life; and so we needn't trouble ourselves to examine it. (Locke 2011, II: xxxii, 15, 389)

So even Locke, who constantly argues that all our ideas are within our own breasts, hidden and invisible to others (cf. III, ii, 1, 405), labels such assumptions as quite useless. Similarly, when I look at the sky and say: "What a wonderful blue sky!", I point to the sky, not into myself. I do not assume that the colour impression only belongs to me, and I do not hesitate to also tell others (cf. PI, § 275).

In "Lectures on Similarity", Wittgenstein provides another private chart example, similar to the private colour charts and to PI, § 271:

Suppose, when we play chess, we each have a private chess board on which we make moves before we make moves on the public board. Suppose someone plays chess all right, but makes moves on his private board in a completely haphazard way, but with all the appearances of setting great value on his moves on the private board, etc.). Malcolm and I (before using our private chart) both learnt the words "green", "blue", etc. How did we learn to use the table on our charts? – We learnt the private chart by learning the public chart. The private game may be any damn thing, as it is only judged by giving rise to the game of chess which we publicly play. (WCL, 95)

All these examples show that there is nothing wrong per se with the picture of a private chart if we can make some *use* of it. Being able to speak a common language also allows for the possibility of private charts, because we have already learned how to use colour concepts, sensation words, public charts, etc. Otherwise, the private chart could be "any damned thing" because it does not belong to the game. "The impression of a 'private table' in the game arises through the absence of a table and through the similarity of the game to one that is played with a table." (Z, § 552)

In "Lectures on Description" from Lent term 1940, Wittgenstein gets back to the example of a private chart and draws a parallel with the role of experiences:

I gave the example of a game played publicly, in which each of the players had a table which he didn't show to anyone. I can send a man to fetch things of various colours, and it is obvious that this game could be played by means of a colour chart, in which he looks up, goes to patch, etc. The point is: suppose this were done, but suppose that then, somehow or other, the charts which each man had were changed, so that green stood for red, etc., and suppose that, nevertheless, the man went on as before and fetched the right thing, etc. What if they were all grey [all the colours on his chart] – if you looked at what he was doing? The point is that in this case we could not say that "looking up" entered the game. He did something, but it obeyed no rules, and, then, indeed played a game independently. We would not say in this case that what he did depended on what he saw in the table. In the way this chart enters, this is the way experiences enter. We would say the chart is no justification, and in fact plays no role in the game. If we talk of the image of so and so, I was saying this image does not come in as a picture. The expressions of experience come in just like those expressions which the man uses looking at the private table. That is, there is a peculiarity in this language-game, which is that it ends somewhere. It goes on, up to a point, as though there was

a table. You actually could, as it were, supply a hypothetical table, although, as a matter of fact, there isn't one. (WCL, 161)

At the end of this chapter on 'subjective' knowledge, I would like to quote from a Rhees manuscript dated 17 April 1965. The remarks quoted here are closely connected with his reply to Ayer and his letter to Mundle, written just two weeks before. And here again, Rhees points out the difference between *giving a definition* and *using a definition*:

> If "I can't be mistaken about what colour I see" is taken to be more than a grammatical statement, then – since the difference in grammar between "I see red" and "he sees red" is (apparently) sharp and constant – someone might go on to suppose that *because* it always has sense to say "perhaps he doesn't see red", I can never *really know* that he does. And if you emphasize the "*never* really know...", it may seem as though it would have sense to say: "Perhaps nobody ever sees the colours I see. I can know only indirectly what he *does* see – by what he says, for instance. All I know is that he says he sees red. But even if I am convinced he is not lying, this does not really tell me what he *does* see. For I cannot know that he uses the word 'red' for the same colour as I do". "He sees the same colour as I see." What is it that I am saying perhaps he doesn't do? For it sounds almost as though I were prevented from making the comparison that would assure me one way or the other. [...] But 1) such a comparison – a "direct" comparison of what I see and what he sees cannot even be imagined. And 2) it is not the case that the visual impression I am having (when I look at my red curtain) *tells* me what colour I'm seeing, and that *therefore* I can't be mistaken when I say it is red. Compare the suggestion that children know what colours they see before they learn to speak. I do not say they don't. For I should not understand the denial any more than I understand the statement itself. It takes the expression "know what colour he sees" out of the game in which we use it: in which we know what it has sense to ask, what would be reasons for doubting it, what would be conclusions from it, and so on. Similarly with "I remember my experiences before I was born". My impression does not *tell* me anything at all, and certainly not whether the word I have used is the right one to describe it. How can you know whether it is red – or *that* it is red – unless you know what "red" means? You do not learn what "red" means (nor "colour" either) by looking at a red surface. It seems to tell you what colour it is, in the sense in which a sample may tell you what colour it is. But a sample only does this with a recognized application. (It would not do to take as a sample warm, cold, or tepid). But the delusion of "I know what colour

I'm seeing" is the delusion of taking what you see as a sample of itself: as though you could use it to describe itself. We might say you confuse giving a definition and using the word in accordance with a definition. And unless you *distinguish* these, it has no sense to speak of a definition at all. (This is relevant to "giving names to my sensations.") If I thought that nobody meant by "red" what I do, the word would have lost its point. There would be no point in my using the word – and the phrase "mean by it what I do" would be empty as well. The suggestion that "it would have sense to say that nobody sees the same colours as I do – or that nobody means by colour words what I mean by them" turns out to be false. It is not logically possible. [...] what looked like a "logically possible" suggestion turns out again to be meaningless.

In his reply to Ayer, Rhees perfectly sums up the central points of the issue. If words did not have a regular use, I would not know that something is, e.g., red, I would not know that the private colour chart was constantly changing, etc., because there would simply be *nothing to know*.

Because there is this agreement, it is possible to say something. When I tell you that the patch on the patient's skin is red, I am not saying that it is called red, but that it is red. But I could mean nothing definite by that, and you could not understand me unless people who have learned the words as we have would agree in calling this red. If people could not be brought to use the word in any regular way, if one man who had been taught as we have should go on to give the name to what we should call the complementary colour, if another used it as we do on Monday but in a different way on Tuesday, and if others did not show even these degrees of regularity – then it would not mean anything to say that someone had used the word mistakenly. There would be no distinction between mistakenly and correctly. And there would be no distinction between saying that it is red and saying anything else. It is not a statement about what I do or about what people generally do. But unless the words had a regular use, I should not know it was red, and I should not know what colour it was because there would be nothing to know. I know what colour it is because I know red when I see it; I know what red is. (Rhees 1954, 79)

Similarly, in "Notes for a Philosophical Lecture", Wittgenstein notes:

Talking about impressions already means to look at phenomena in one particular way, i.e., to think about them in one particular fashion. "What does green look like to me? – it looks like this → to me." – "This is the colour impression which I'm calling 'green'."

> Am I sure I'm talking about my private *impression*? And how can
> I be sure – ? Do I *feel* that I'm talking about the impression? What
> happens? I look at a green patch, I concentrate my attention on
> such a patch and I say these words. But what kind of a patch? Not
> on a green one. On one that seems to deserve the name "green"?
> It is not true that I see impressions before me and that they are
> the primary objects. In the sense in which I can't explain "what
> green looks like to me", *I can't say that I know what it looks like ei-*
> *ther*. (NPL, 457; last italics added)

To me, all these remarks suggest that in the case of private objects
and a private table, there is indeed *nothing* to know. Therefore, it is
nonsensical to say that I know that this patch is red because I am now
having a particular colour impression, which I call 'red'. This point
also holds for the concept of 'memory' in the private diary case Witt-
genstein introduces in PI, § 258. Here, too, the question is not wheth-
er I apply my memory rightly or wrongly, but that instead the instru-
ment of our memory is of no *use* in such a case. It would not mean
anything to say that my memory justifies me in saying that I am hav-
ing the same sensation again because here, too, there is nothing to
remember and therefore nothing to tell me whether I have applied
my memory rightly or wrongly. Memory scepticism is not at issue.

4 Subjective Memory and Justification: Apropos PI, § 258

In PI, § 258, Wittgenstein introduces the famous thought experiment
about someone who wants to keep a diary about the recurrence of a
particular sensation, whereby a definition of the sign cannot be ex-
pressed. For the record, the person simply associates the sensation
with the sign and writes it down whenever the sensation appears. She
thereby concentrates her attention on the sensation, pointing inward-
ly, so to speak. The failure of a baptism by ostensibly defining a pri-
vate sample to fix the meaning of the sensation sign has already been
discussed at great length. But just to make this important point again
in connection with PI, § 258: let us assume that a diarist writes down
the sign 'S' and at the same time concentrates her attention on S by
pointing inwardly. Thereby she ostensively defines 'S' by introducing
and fixing S as its meaning. Let us further assume that two days later,
the diarist notices a fresh sensation, and in order to find out wheth-
er it is S, she uses the S-sample and compares. Now, what if she mis-
identifies another sensation for S without realising it? Since she still
thinks it is S and therefore believes that she has followed the mean-
ing rule for 'S' correctly in applying the primal sample S, she writes
down 'S' again. She is thereby convinced that she is right because she

believes she is using the originally defined sensation S as a sample. According to Ayer, I guess, we would have to say that she just made a mistake, comparable to Crusoe, who shot the wrong bird. But again, *being right or wrong* is off the cards. What really happened is that she did not *use* S and apply it to the fresh sensation, but instead *fixed* an entirely new rule for the meaning of 'S' by introducing or *giving* a new sample, without knowing it. After all, the fresh sensation is not S. And when, three days later, yet another sensation occurs, the diarist compares it with S again and writes down 'S' if she thinks it *is* S. However, she would just compare it with the *newly* introduced sample three days ago, without knowing it. And if she misidentifies the sensation for the second time (either for S or S_{new}), again she would not have *used* a sample, but would instead introduce a new rule by ostensively defining a new sample. This is what Wittgenstein means when he says that private rules are only *impressions* of rules and that whatever is going to *seem* correct for the diarist will *be* correct for her (cf. PI, §§ 258-60). The diarist only thinks she is following a rule but instead constantly introduces new meaning rules whenever she misidentifies sensation S. I actually think that this is the crux of the confusion between giving a sample and using a sample, giving a definition and making a statement, which I discussed in section 2.2.

Particularly in PI, § 258, our human memory plays a crucial role in this context. The last part of this paper will therefore focus on our memories and try to show that they logically cannot serve as a tool to fix or retain the meanings of private sensation words.

In a letter to Ilham Dilman, dated 9 March 1965, Rhees writes:

> It may even be a misfortune that Wittgenstein brought in the imagined case of keeping a diary of one's sensations – partly because diaries often are private in a sense that is irrelevant here, although he brought it in partly to emphasize this contrast. And partly because Pepys, and no doubt others, have written diaries in a private language. Pepys' language has been deciphered, and Wittgenstein used to mention it in order to make clear that this was not the sense of "private language" with which he was concerned. But the example may have made people fall just into those confusions it was intended to prevent; I do not know.

As we all know, Wittgenstein himself used to keep diaries, particularly during the First World War and used a secret code for his private remarks.

In the case of a private diary in the strong sense, the person only has her memory as the decision tool for identifying the sensation rightly or wrongly. In this respect, she is like the person who buys several copies of today's morning paper to confirm that what the

paper writes is true (cf. § 265). Now Ayer argues that there is nothing absurd about buying another newspaper of a different type to check the first (cf. Ayer 1954, 71). This suggestion is quite curious because it would turn the example into a public one. If I do not trust what is written in the *Daily Express*, I will buy myself a copy of the *Daily Telegraph* or *The Times* to double-check. Ayer continues:

> And in a place where there was only one morning newspaper, but it was so produced that misprints might occur in one copy without occurring in all, it would be perfectly sensible to buy several copies and check them against each other. (71)

This, too, does not seem very plausible, for I can always check possible misprints by applying various obvious standards. Further, it is not clear which one would be the defining "ur-newspaper". The one with the fewest mistakes? And are we supposed to lay down our private sensation sample to compare it with a present one for possible agreement or disagreement? David Hume already pointed out the difference between present 'impressions' and 'ideas':

> Every one will readily allow, that there is a considerable difference between the perceptions of the mind, when a man feels the pain [...], and when he afterwards recalls to his memory this sensation [...]. The most lively thought is still inferior to the dullest sensation. (Hume 1992, 17)

According to Ayer, verification must stop at some point, and in the case of Crusoe's sensations, there is obviously no further test option beyond his own memory. However, "it does not follow that he has no means of identifying it, or that it does not make sense to say that he identifies it right or wrong" (Ayer 1954, 72). Again, the issue is not whether someone is making mistakes because her memory fools her but she nevertheless still follows the meaning rule. It simply does not make sense to say "I remember having had S yesterday", either rightly or wrongly, because in the case of a *private* object such as a private sensation, there is simply nothing to remember.

Mundle plays the same horn by arguing that Wittgenstein's "theory of meaning" demands independent justification for the diarist using 'S' for verifying the correctness of his application. And Mundle answers:

> Presumably the diarist has a memory, has the capacity to recognize what he feels as well as what he sees or hears; and what grounds are there for embracing skepticism concerning people's capacities to recognize their sensations, and to remember some

of their earlier sensations and compare these with their present ones? (Mundle 1966, 39)

He also offers the case of a diarist which shows that he obviously did not get Wittgenstein's point at all:

> Suppose our diarist is liable to suffer from spells of tummy trouble which his doctor cannot diagnose. The onset of such attacks has been preceded by distinctive shooting sensations in his midriff, and has involved several days of different but equally unpleasant sensations, during which time, unless he sticks to a light and simple diet, he vomits and has to retire to bed. Our diarist starts writing 'E' in his diary ('E' being short for 'those damned twinges') in order to verify whether E's are regularly followed by the other symptoms, because this information might help the doctor's diagnosis. Having confirmed this, he continues to record E's in his diary to remind himself to cancel all engagements to eat out for the next few days. It is surely unwarranted dogmatism to say that in such cases 'E' has no meaning for the diarist. Does not everyone experience some types of sensation which are so distinctive and so frequent that he has no more difficulty in recognizing them than he has in recognizing apples or 'apples'? (41)

To me, this seems like a perfect example of a *useful* case of keeping a diary about one's own sensations or pains. Wittgenstein himself presents such a case in PI, § 270, where he discusses someone recording a particular sensation, as a manometer indicates that blood pressure rises whenever the sensation occurs. This provides a valuable result, informing the person when their blood pressure is increasing without having to rely on the manometer. In this scenario, it does not matter whether the person identifies the sensation correctly or not. Rhees remarks:

> Example of thinking about what to tell the doctor about my pain today, different from two days ago, etc. I might compare the relevant sensations, and that is all right and common enough. But here, I am concerned with what I am to say about sensations in a way that I expect the doctor may understand, for instance. And I am not concerned in any way with private languages. (Rhees, unpublished)

In the case of the private diarist in § 258, however, the question of recognising the sensation, whether rightly or wrongly, is not the point. Similar to the private chart example, we would not know what it means to say that I am justified in writing down 'S' because my memory tells me so. Since no meaning for 'S' has been established,

there is indeed nothing to remember, and it makes no sense to say that my memory was right or my memory fooled me. Rhees notes:

> There is the question of what would be involved in "meaning the same experience", and that would be closely connected with recognizing the experience when it occurs again. There is a kind of parallel between knowing red when I see it and knowing pain when I feel it. And I want to say that in both cases, it is a matter of knowing the language. And that means that in both cases, it is a matter of knowing how the language is spoken, which, again, has to do with the agreement in people's reactions.
>
> A rule is something that is kept. The meaning of a word is something that is kept. It is for that reason that I can talk about seeing the same colour that I saw a moment ago. [...] But it is possible to talk about seeing the same colour because I know red when I see it.
>
> (I can check my memory by imagining the timetable page because that is not mere imagining. In the "private" case, there would be no way of distinguishing a memory image from a fantasy. And it would mean nothing to try to distinguish them.) (The main question is as to what could we even mean by saying "I know it is the same sensation again.") (Rhees 5 April 1954, unpublished)

About two months later, Rhees continues:

> But when it was suggested that there could not be a private language because there would be no criterion to tell whether one was using a name for a sensation in the same way – the reply was suggested that all that one needs to do is to remember the sensation for which it was used before, and see whether that is the same as the sensation that is being experienced now. One of the principal difficulties in this is in the notion "seeing whether it is the same"; because that is something that presupposes an existing language, presupposes criteria of identity. We know how to distinguish the same meaning from a different meaning just as we can recognize ambiguity, and so forth. We know this because we have learned the language and speak it as we do. But in the case of a private language those criteria could hardly have any application. And it is not at all clear what could be meant by "seeing that it is the same".
>
> That is one point, and probably the fundamental one. But another is connected with that when we ask what it would mean to say that you trust your memory in such a case. How do you know that it is a memory, or how do you know that your memory is not playing

you tricks? In the vast majority of cases we do know that our memories are not playing tricks. But we also know the difference between the case in which your memory does play us tricks and that in which it does not.

One must understand what is being asked there, otherwise it has no sense to say that here and in the majority of cases I do remember correctly. That is one point. Another is that my memory's playing me tricks must be the exception; the rule must be that our memories do not play us tricks. *That is essential for the distinction that we make between correctly and incorrectly at all.*

We learn how to use "I remember" by learning the language. We do not learn it by having a special experience. […] "I remember" is *not* a description of something going on inside me. That is not how we have learned it. Obviously "I remember" does not refer to some present experience which verifies it. But if I trust my memory, I do not "trust" the sensation or the experience which I am having.

I know what it would mean to discover that my memory had played me trick. But in the case of a private sensation about which I cannot speak to anybody, I do not know what it would mean to discover that.

If *all* I have to appeal to is *my* "memory", this is no help: it does not do anything to show that the other memory is "correct". If I can check my memory by the timetable, the case is different. In the majority of cases of course I do not need to check it. But the point is that it is sensible to speak of checking it or consider checking it. *And it is for this reason that it is sensible to speak of my memory as correct, and to speak of trusting it.*

In the case of the private memory it does not even make sense to speak of checking it that way. For this reason, it can have no relation to the way in which we have learned to use the expression "I remember". And we do not know what would be meant by saying that it was correct or that it was incorrect.

Ayer misses the point of the morning papers example, partly because he does not consider what is involved in the "checking against one another" which he mentions in his objection. That is all part of what I have called an institution. The point of the example was to suggest a case in which it makes no sense to speak of such checking.

In the case of private sensations we would not know what it means to discover that our memories have played us tricks. And that is another difficulty in the way of suggesting that you can remember how you used it before. It is connected with the question

of whether you can remember that you are using "the same" in the same way as you used it before.

The main point is that it would not even make *sense* to ask whether you were remembering correctly or not. Which is important if you want to decide whether there could be anything like a rule of language. What could be meant by discovering that you were using the expression *in*correctly? *In this kind of language*?

There is no suggestion that you cannot remember whether this sensation is the same as you had before or not. Of course you can. And of course you do not depend generally upon anybody's corroboration when you do. But in all this you are employing the language and the concepts that we normally do employ in speaking about sensations. And it is hard to think of comparing the sensation you have with the sensation you remember in any other way.

An expression like "the colour is what it looks like" is very confusing and reminds one of the idea that you can learn what colours are – what red is – just by looking at them. This is again the idea that what the word means is what it can possibly refer to.

Unless there are rules of language, then it does not help to say that you can remember that it is the same as that was.

The remembering that it is the same is in a way question begging, as far as this matter of private languages is concerned. Or perhaps one could even say that it is contradictory, because it is assuming that what you are describing as entirely private in this sense, is something that is *not* entirely private in this sense, – that it is something to which you can apply the criteria of identity or sameness which belong to the things that you speak of in ordinary language.

I do not think it would make sense for Ayer's Crusoe […] to remember that it was the same as the sensation he is having now.

Unless there are rules, it does not make any difference what you say. You could say anything. You might make the sound that you make or make the inward gesture or whatever it may be that you make, when you have a sensation. (Rhees 30 June 1954)

These remarks succinctly capture the essential points related to "remembering that it is the same" and are self-explanatory.

Ultimately, Wittgenstein does not argue that the ascription of meaning to a sign requires justification (cf. PI, § 289). Instead, he contends that if a sign has meaning, it has a use and can be used incorrectly. Therefore, the act of 'ascribing meaning privately' is not

valid. Wittgenstein had previously addressed the notion that rules of language do not need justification from reality in his 1930-33 lectures and *Philosophical Grammar*. He emphasised that rules are arbitrarily established and cannot be deduced from or justified by the facts of reality. They are arbitrary precisely because they are not determined by reality in the way descriptions are. Therefore, it is nonsensical to assert that the rules for words like 'red' or 'blue' correspond to specific facts about these colours (cf. PG, 246. Cf. also PR, 55).

When someone claims to have a justification for saying, for example, "I now feel toothache", it is acceptable as long as it simply means that she is speaking the truth. In other words, a justification for such statements is only meant to rule out the possibility of lying or play-acting. Saying, "I have a justification for saying that I am in pain now", essentially means "it is true", "it is indeed the case", or "I am not lying" (cf. WCL, 98).

In "Lectures on Similarity", Wittgenstein remarks:

Suppose I look at the colour of these shoes and say, "I now see brown", I'm then very inclined to say: "There is something which made me say 'brown', just me [now], namely, that I had a peculiar impression". The idea of this justification would be roughly this: not a justification derived from a rule, but a justification by intuition. "This impression I now have justifies me in saying 'I see brown'." When I said "this impression", I could have been said to point to an impression. For whose sake am I pointing? "This impression I now have justifies me in saying 'I see brown'" could only mean "Looking at these shoes justifies me [in saying 'I see brown']". It seems as though I pointed privately, and informed myself of the fact that this impression justified me. The words are entirely all right in "What I see justifies me", but I am inclined to do something very queer with them. "What I see justifies me" is a case of meaning "They are brown, and if you look you'll see for yourself". If you say, "This impression justifies me", well, so it does, meaning "this impression when I see this". But in this case [the former, philosophical case], I'm saying to myself "This impression justifies me in saying 'brown'".

"One can say it to someone else, but also to oneself." It is not at all clear in all cases what is meant by "saying to myself". Under what circumstances does one say that one says something to oneself? Is it when one says something when one is alone? People assume that language-games played with others can be played with oneself. Cf. giving a present to oneself from the right hand to the left. [Cf. PI, § 268]

I can cheat myself, but not in the same way as I can cheat others. The mere fact that I can ask myself a question, and answer it – give

myself an order – tell myself a lie – the mere fact that there are analogues doesn't yet tell you what the analogue is. (WCL, 104-5)

The concept of someone saying something to herself was previously discussed in section 2.1.

In Lecture 11 of "Lectures on Similarity", Wittgenstein revisits the question of justification by pointing out that we are sometimes inclined to something like the following:

> "The peculiar impression I see justifies me in using the word 'brown', quite independently of what anyone else says." Here, it seemed we had a justification independent of any rule given. "If I see brown, by 'brown' I just mean this". How on earth can a word 'brown' refer to an experience?

> Pointing is of importance if I want to show someone something – if I say 'Look at this spot'. As a matter of fact, when we point to something, people do something, react in a certain way. If they didn't do that, pointing would be completely useless. (WCL, 107)

This discussion circles back to the question of how words refer to things and the nature of ostensive definitions.

The discussion of Wittgenstein and Rhees on the possibility of a 'private language' aimed to demonstrate that such a language, in the 'strong' sense, is impossible. This is due to our inability to establish any meaningful rules, given the confusion between giving a sample and using it. Consequently, one cannot communicate with others or even oneself. Additionally, in the realm of 'subjective knowledge' and our memory of 'private' sensations, there is fundamentally nothing to know and nothing to remember. In Rhees' words:

> The point is that no one could invent just *language*. Language goes with a way of living. An invented language would be a wallpaper pattern; nothing more. (Rhees 1954, 87)

References

Ayer, A.J.; Rhees, R. (1954). "Symposium: Can There Be a Private Language?". *Proceedings of the Aristotelian Society, Supplementary Volumes*, 28(1), 63-94. https://doi.org/10.1093/aristoteliansupp/28.1.63.
Ayer, A.J. (1963). *The Foundations of Empirical Knowledge*. London: Macmillan.
Candlish, S. (2019). "Private Language". https://plato.stanford.edu/entries/private-language/.

Dennett, D.C. (2015). *Elbow Room. The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press.

Diamond, C. (1981). "What Nonsense Might Be". *Philosophy*, 56(215), 5-22.

Geach, P.T. (1957). *Mental Acts: Their Content and Their Objects*. London: Humanities Press.

Hacker, P.M.S. (2010). "Robinson Crusoe Sails Again: The Interpretative Relevance of Wittgenstein's *Nachlass*". Venturhina, N. (ed.), *Wittgenstein After His Nachlass*. Basingstoke; New York: Palgrave Macmillan.

Hacker, P.M.S. (2019). *An Analytical Commentary on the Philosophical Investigations.* Vol. 3, *Wittgenstein: Meaning and Mind*. 2nd extensively revised edtion. Oxford: Blackwell.

Hume, D. (1992). *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. Edited by L.A. Selby-Bigge. Oxford: Clarendon Press.

Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.

Locke, J. (2011). *Essay Concerning Human Understanding*. Oxford: Oxford University Press.

Macdonald, M. (unpublished). "Lectures on 'Private Experience' and 'Sense-Data'". Smythies, Y., *Literal Estate*. Klagenfurt.

Malcolm, N. (1954). "Wittgenstein's Philosophical Investigations". *The Philosophical Review*, 63(4), 530-59.

Meinong, A. (1981). "The Theory of Objects". Chisholm, R.M. (ed.), *Realism and the Background of Phenomenology*. Glencoe, IL: Free Press, 76-117.

Mundle, C.W.K. (1966). "'Private Language' and Wittgenstein's Kind of Behaviourism". *The Philosophical Quarterly*, 16(62), 35-46.

Munz, V.A. (2023). *Begriff, Bewusstsein und Bedeutung. Zum Verhältnis von Sprache, Mentalem, und Bezugsobjekt*. Berlin; Boston: De Gruyter.

Nagel, T. (1987). *What Does It All Mean? A Short Introduction to Philosophy*. Oxford; New York: Oxford University Press.

Rhees, R. (1970). "'Ontology' and Identity in the *Tractatus* Apropos of Black's *Companion*". Rhees, R. (ed.), *Discussions of Wittgenstein*. London: Routledge, 23-36.

Rhees, R. *Unpublished Papers*. Klagenfurt.

Russell, B. (1956). "The Philosophy of Logical Atomism". Russell, B. (ed.), *Logic and Knowledge*. London: Allen and Unwin, 177-281.

Stroud, B. (2000). "Wittgenstein's 'Treatment' of the Quest for 'a Language which Describes My Inner Experiences and which Only I Myself Can Understand'". Stroud, B. (ed.), *Meaning, Understanding, and Practice*. Oxford: Oxford University Press, 67-79.

Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus*. Transl. by D.F. Pears; B.F. McGuinness. London: Routledge and Kegan Paul. International Library of Philosophy and Scientific Method. [TLP]

Wittgenstein, L. (1967). *Zettel*. Edited by G.E.M. Anscombe; G.H. von Wright, transl. by G.E.M. Anscombe. Oxford: Basil Blackwell. [Z]

Wittgenstein, L. (1968). "Notes for Lectures on 'Private Experience' and 'Sense Data'". *The Philosophical Review*, 77(3), 275-320. [LPE]

Wittgenstein, L. (1969). *On Certainty*. Edited by G.E.M Anscombe; G.H. von Wright. Oxford: Basil Blackwell. [OC]

Wittgenstein, L. (1969). *The Blue and Brown Books: Preliminary Studies for the Philosophical Investigations*. 2nd edn. Edited by R. Rhees. Oxford: Blackwell. [BBB]

Wittgenstein, L. (1970). "Eine Philosophische Betrachtung". Rhees, R. (Hrsg.), *Schriften Bd. 5*. Frankfurt am Main: Suhrkamp, 117-282. [EPB]

Wittgenstein, L. (1974). *Philosophische Grammatik*. Edited by R. Rhees. Oxford: Basil Blackwell. [PG]

Wittgenstein, L. (1975). *Philosophical Remarks*. Edited by R. Rhees. Oxford: Basil Blackwell. [PR]

Wittgenstein, L. (1978). *Remarks on the Foundations of Mathematics*. Edited by G.H. von Wright; R. Rhees; G.E.M. Anscombe. Oxford: Basil Blackwell. [RFM]

Wittgenstein, L. (1980). *Remarks on the Philosophy of Psychology*, vol. 1. Edited by G.E.M. Anscombe; G.H. von Wright, transl. by G.E.M. Anscombe. Oxford: Blackwell. [RPPI]

Wittgenstein, L. (1988). *Lectures on Philosophical Psychology 1946-47*. Edited by P.T. Geach. London: Harvester. [LPP]

Wittgenstein, L. (1993). "Notes for a Philosophical Lecture". Klagge, J.C.; Nordmann, A. (eds), *Philosophical Occasions 1912-1951*. Indianapolis; Cambridge, MA: Hackett, 445-58. [NPL]

Wittgenstein, L. (1982). *Last Writings on the Philosophy of Psychology*, vol. 1. Edited by G.H. von Wright; H. Nyman. Oxford: Blackwell. [LWI]

Wittgenstein, L. (2009). *Philosophical Investigations*. 4th ed. Edited and transl. by P.M.S. Hacker; J. Schulte. Oxford: Wiley-Blackwell. [PI]

Wittgenstein, L. (2017). *Whewell's Court Lectures: Cambridge, 1938-1941. From the Notes by Yorick Smythies*. Edited by V. Munz; R. Ritter. Malden, MA: Wiley Blackwell. [WCL]

Wittgenstein, L. (2009). *Wittgenstein Source Bergen Nachlass Edition*. Edited by A. Pichler. Bergen: Wittgenstein Archives at the University of Bergen. `http://www.wittgensteinsource.org`.

Caveat Lector: From Wittgenstein to The Philosophy of Reading

Robert Hanna
University of Colorado, Boulder, USA

Abstract   Against the grain of Analytic philosophy's general avoidance of the fact or phenomenon of reading, and starting out with Wittgenstein's compact investigation into "the part the word ['reading'] plays in our life, therewith the language-game in which we employ it", in this essay I explore the nature of reading, and thereby initiate what is in effect a new philosophical sub-discipline: the philosophy of reading.

Keywords   Wittgenstein. Philosophy of language. Philosophy of mind. Language. Reading.

Summary   1 Introduction. – 2 Wittgenstein on the Use of the Word 'Reading'. – 3. Caveat Lector Sentences and the Right Way to Start Epistemology. – 4 The Logic of Legibility. – 5 Necessary and Sufficient Conditions for Legibility and Reading. – 6 Are There Some Legible Texts that Even the World's Most Sophisticated Robot Cannot Read? – 7 Conclusion.

> For us it is *the circumstances* under which he had such an experience that justify him in saying in such a case that he understands, that he knows how to go on. [...] This will become clearer if we interpolate the consideration of another word, namely "reading". [...] The use of this word in the ordinary circumstances of our life is of course extremely familiar to us. But the part the word plays in our life, therewith the language-game in which we employ it, would be difficult to describe even in rough outline.
>
> (PI, I, §§ 155-6)

> 'Twas brillig, and the slithy toves
> Did gyre and gimble in the wabe;
> All mimsy were the borogoves,
> And the mome raths outgrabe.
>
> (Carroll 1988)

> As I read and contemplated the subject [of the immorality of slavery and its abolition, as discussed in the book, *The Columbian Orator*], behold! that very discontentment which Master Hugh had predicted would follow my learning to read had already come to torment and sting my soul to unutterable anguish. As I writhed under it, I would almost at times feel that learning to read had been a curse rather than a blessing. It had given me a view of my wretched condition, without the remedy. It opened my eyes to the horrible pit, but no ladder upon which to get out. In moments of agony, I envied my fellow slaves for their stupidity. I have often wished myself a beast. I preferred the condition of the meanest reptile to my own. Anything, no matter what, to get rid of thinking! It was this everlasting thinking of my condition that tormented me. There was no getting rid of it. It was pressed upon me by every object within sight or hearing, animate or inanimate. The silver trump of freedom had roused my soul to eternal wakefulness. Freedom now appeared, to disappear no more forever.
>
> (Douglass 1995, 24)

1 Introduction

Since it is self-evidently true that you, the reader of this very sentence, are reading this very sentence, then we can safely assume that you already *know how* to read and also *what* reading is – at least, as the later Wittgenstein rightly puts it in the text quoted as the first epigraph of this essay, in a way that suffices for "the ordinary circumstances of our life", even if the fact or phenomenon of reading is *philosophically* "difficult to describe even in rough outline". But, we can fully concede that philosophically describing the fact or phenomenon of reading is difficult. Nevertheless it's passing strange that, with the notable exception of fifteen sections in the *Philosophical*

Investigations (Wittgenstein, 1953: pp. 61e-70e, §§ 156-71), even during the heyday of the "linguistic turn" to "linguistic philosophy" (Rorty 1967b) that was enacted by the tradition of classical Analytic philosophy from 1880 through the 1970s, and equally during the post-classical, post-linguistic-philosophy period spanning the philosophy of language-&-mind, the philosophy of mind per se, and Analytic metaphysics, from the 1980s into the 2020s – so, for the last 140+ years – Analytic philosophers have paid *surprisingly little* attention to the fact or phenomenon of reading. Indeed, I think that one can even correctly say that they've *consistently avoided* thinking, talking, and writing about it.

Now, since Analytic philosophers – like all other philosophers – live, move, and have their being as thinkers, talkers, writers, and above all, as *readers*, then perhaps this general pattern of philosophical avoidance is simply an instance of the widespread cognitive pathology that one might call 'young fish syndrome', whereby those who are everywhere surrounded by and ensconced in a certain cognitive, affective, moral, or sociopolitical transparent medium by means of which they encounter themselves, each other, and their world, nevertheless blithely fail to recognise the necessary and obvious existence of that all-encompassing medium:

> There are these two young fish swimming along and they happen to meet an older fish swimming the other way, who nods at them and says "Morning, boys. How's the water?" And the two young fish swim on for a bit, and then eventually one of them looks over at the other and goes "What the hell is water?". (Wallace 2012)

To be sure, this would not be the *only* occurrence of young fish syndrome in the 140+ years of the Analytic tradition (Hanna 2021, esp. chs. 17-18). In any case, against the grain of Analytic philosophy's general avoidance of the fact or phenomenon of reading, and starting out with Wittgenstein's compact investigation into "the part the word ['reading'] plays in our life, therewith the language-game in which we employ it", in this essay I want to explore *the nature of reading*, and thereby initiate what is in effect a new philosophical sub-discipline, *the philosophy of reading*.[1]

---

1 Apart from later Wittgenstein, and now taking account of *non*-Analytic philosophy as well as Analytic philosophy since 1900, as far as I know, the only other exception to the general avoidance of reading as a fact or phenomenon meriting careful, critical, focused, and systematic – let us call this, collectively, 'serious' – philosophical investigation, is the Polish phenomenologist Roman Ingarden's 1968 book, *The Cognition of the Literary Work of Art* (Ingarden 1973). Of course, philosophically-minded literary theorists either belonging to or influenced by French trends in the 1960s, 70s, and 80s, including post-structuralism, deconstructionism, semiotics, and so-on – for example,

2   Wittgenstein on the Use of the Word 'Reading'

The larger context of Wittgenstein's discussion of the use of the word 'reading' is what commentators call "the rule-following considerations", including The Rule-Following Paradox and its solution, in *Investigations* §§ 134-242.

In the *Tractatus Logico-Philosophicus*, it was assumed that concept of a proposition expressed the *essence* of the proposition: necessarily and sufficiently, all propositions describe facts ("this is how things are"), and every proposition is bipolar ("a proposition is whatever can be true or false") (PI, I, §§ 134, 136). So, necessarily, a part of language is a proposition if and only if it satisfies these basic conditions. This dictum, however, also pre-reflectively invokes a bad (i.e., false, misleading, and mind-enslaving) philosophical picture about the inherent systematicity of language, a picture according to which propositions are hypostatised, substantial, platonic (i.e., non-spatiotemporal, abstract, non-causal) entities floating around listlessly in Frege's Third Realm. On the contrary, it is essentially more enlightening to say simply that there is a language-game about propositions and a proposition is automatically whatever is determined by the use of signs in that game (PI, I, § 137). But since you can always automatically either add as a prefix the phrase "This is how things are:" or add as a suffix the phrase "is true", to any proposition whatsoever, it seems that, necessarily, any part of language is a proposition if and only if it satisfies this condition (PI, I, § 137).

Similarly, it seems that any part of language has meaning if and only if it satisfies the sense of a sentence that I understand; and in this connection, we will also recall Frege's famous remark: "[O]nly in a proposition have [...] words really a meaning" (Frege 1953, 71, § 60), i.e., a word has meaning only in the context of a whole proposition, also known as *The Context Principle*. And understanding, it also seems, is 'grasping' the meaning of a word or other expression in a 'flash'. But if a flash-grasping understanding of words is possible, then this contradicts the thesis that the meaning of a word is its use (PI, I, § 138).

So what *is* understanding a word? Understanding a word is *neither* a picture that comes before my mind when I hear a word, *nor* it is a picture plus a method of projection from the picture, because (i) the same mental picture/projection method can be correlated with different applications of the word (PI, I, §§ 139-40), and (ii) the same application can occur without the occurrence of that particular mental

---

Roland Barthes, Jacques Derrida, et al. – have had many and various, and often interesting, things to say about reading texts. But in my opinion, none of this is *serious* philosophy of reading.

picture or projection method (PI, I, § 141). As an example, let us consider understanding how to complete a series by writing down signs representing the natural numbers (PI, I, §§ 140-8). Here understanding a word is neither a state of consciousness nor a mental process because (i) mental states have temporal duration, whereas understanding does not (PI, I, § 59), and (ii) to hold that understanding is a mental process, is to confuse the characteristic *accompaniments* of understanding – which can vary widely across contexts – with understanding itself (PI, I, §§ 149-52).

This argument requires two implicit premises in order to be valid. The **first** implicit premise says that mental pictures, rules of projection, states of consciousness, and mental processes exhaust the possible inner determinants of understanding. And the **second** implicit premise says that the determinants of understanding are either inner or outer, and not both. Therefore, since understanding is after all determined by *something*, it can only be determined by something *outer*: by the manifest or behavioural mastery of a linguistic technique (PI, I, § 150), and by the "particular circumstances", or context, of displaying that mastery (PI, I, §§ 154-5).

Let us consider now a simplified form of mastery of a linguistic technique that does not itself involve understanding: namely, *reading*, where this is specifically the activity of rendering out loud what is written or printed, writing from dictation, writing out something printed, following a score, etc. (PI, I, § 156). There is no single set of necessary and sufficient conditions (a definition or criterion) for mastery of this linguistic technique. Consider, for example, self-consciously attentive reading, human "reading machines", beginning readers, and so-on (PI, I, §§ 156-8). We are tempted to say that the criterion for applying the word 'reading' is the conscious act of reading (PI, I, § 159), but even if the conscious act of reading were lacking – imagine a "reading-zombie" – it is at least conceivable that such a creature might still count as a reader (PI, I, § 160).

This raises an absolutely crucial point that is often overlooked: Wittgenstein is implicitly presupposing and deploying a fundamental distinction between (i) *conceptual or logical* possibility, and (ii) *real or metaphysical* possibility. Roughly speaking, something is conceptually or logically possible if and only if it is consistent with the basic principles or laws of classical logic, conservatively extended to include a theory of fine-grained concepts. By contrast, something is really or metaphysically possible if and only if it is consistent with the basic principles or laws of classical logic *together with* a theory of fine-grained concepts, *together with* the basic principles or laws of mathematics, *together with* the formal structures of manifestly real spacetime, and *together with* the basic principles or laws of non-equilibrium thermodynamics, especially including those governing organismic life, all of them *indexed to the actual world*. In short, real

or metaphysical possibility not only picks out a *more* restricted class of possible worlds than conceptual or logical possibility does, but also picks out a *less* restricted class than natural or physical possibility does, which is further constrained to what satisfies the conservation laws, including the first law of thermodynamics, together with the second law of thermodynamics, i.e., equilibrium thermodynamics. Then, for example, molecule-for-molecule, behaviourally identical, but also non-conscious and mechanistic duplicates of human 'all-too-human' creatures like us, also known as zombies, are *conceptually or logically* possible, but not *really or metaphysically* possible, since creatures like us are living organisms, *not* natural mechanisms, and *not* replicable by means of so-called 'artificial intelligence' or AI, because consciousness is an essentially embodied form of organismic life. A detailed theory of all that is a long philosophical story for another day:[2] the absolutely crucial point for the purposes of this essay is that the later Wittgenstein is implicitly fully onboard with this fundamental distinction.

So consciousness is not the criterion of mastery. What then about "deriving the reproduction from the original" as a criterion of mastery of this linguistic technique? The problem with this is that even if someone never sticks to a single method of derivation, we can still plausibly call him a reader (PI, I, § 163). Hence there is no single sort of mastery of a technique: even for reading, there is a *family* of criteria for what counts as reading (PI, I, § 164), and there is no single specific marker of what will count as a genuine reading (PI, I, §§ 165-8), because reading can, at least in principle, always occur without any such single specific marker. Even if there is no single specific marker, however, it is still true that reading always involves some sort of causal influence between the letters and the reading (PI, I, § 169). More generally, in all cases of reading I let myself be guided by the letters (PI, I, § 170). Therefore, mastery of a linguistic technique always involves 'being guided' by the linguistic basis of the technique. This could also be equivalently described as the subjective experience of having the sound of the word 'intimated' to me by the letters, such that there is a manifest *unity* between word and sound (PI, I, § 171).

Notice, however, that this subjective experience of having the sound of the word intimated to me by the letters is clearly *a mode of consciousness*, i.e., this is *a phenomenological structure of reading*, which is smoothly consistent with Wittgenstein's earlier claim that reading-zombies are conceptually or logically *possible*, *only if* he is also committed to the view that reading-zombies are really or metaphysically *impossible*. This fundamental point, in turn, is the segue

---

2 Hanna, Maiese 2009; Hanna 2011; 2015; 2022; 2023a; 2023b; 2023c; 2023d; 2023e; 2023f; 2023g; 2023h; 2023i; 2023j; 2023k; 2023l; 2024.

to my Wittgenstein-inspired philosophy of reading, which not only takes onboard all of Wittgenstein's basic claims about the use of the word 'reading', but *also* fully incorporates the phenomenology of the essentially embodied act or process of reading.

3   Caveat Lector Sentences and the Right Way to Start Epistemology

Let us call any sentence that is (i) specifically about the act or process of *reading*, and that is also (ii) *self-referring* by means of the 2nd-person indexical description 'you, the reader', and the indexical description 'this very sentence', *a caveat lector sentence*. Such sentences are so-named by me after the Latin phrase *caveat lector*, meaning 'let the reader beware'; but I am interpreting that phrase broadly enough so as *also* to include the meaning 'let the reader be self-consciously aware'.

From a philosophical standpoint, here is the paradigmatic example of a caveat lector sentence:

> **You, the reader of this very sentence, cannot either coherently or self-consistently deny that it is self-evidently true that you are reading this very sentence.**

For convenience, I will call the sentence I displayed in boldface text immediately above,

THE SENTENCE

and for the purposes of this section, it will not matter whether THE SENTENCE is a universal *sentence-type* or a particular *sentence-token*.

Granting those stipulations, then what I want to argue now is that *philosophically appealing to your reading caveat lector sentences like THE SENTENCE, are the right way to start epistemology*. Here is my argument, in eight steps.

1. As the Wittgenstein compellingly argues in the *Investigations*, language is inherently a set of *social practices* and more generally a *social institution* (PI; Hanna 2021, chs. 11-15). Therefore, your reading caveat lector sentences like THE SENTENCE are inherently a *collective*, *communal*, or *intersubjective* phenomenon, and not an *idiosyncratic*, *solipsistic*, or otherwise *subjectivistic* phenomenon.

2. Whether a caveat lector sentence like THE SENTENCE is a universal *sentence-type* or a particular *sentence-token*, it is nevertheless a *physical* phenomenon. Now, the act or process of reading is an essentially embodied phenomenon of conscious and self-conscious

intentionality, like all forms of rational human cognition (Hanna, Maiese 2009; Hanna 2011; 2015). And a caveat lector sentence like THE SENTENCE is inherently *the intentional object* of that specific mode of intentionality. Therefore, your reading caveat lector sentences like THE SENTENCE is inherently a *psychophysical phenomenon*, i.e., it is inherently *non-dualistic*.

3. The act or process of reading is inherently *a rational human activity*. Therefore, your reading caveat lector sentences like THE SENTENCE is an actualisation of all the cognitive-and-epistemic or theoretical, affective or emotional, and moral or practical *sub-capacities* that are properly contained in and jointly constitutive of the complex, unified capacity for human rationality.[3]

4. Your reading caveat lector sentences like THE SENTENCE is inherently *authoritative* and *rationally intuitive*, precisely because *it is both intellectually and sensibly self-manifesting* (Hanna 2015, esp. chs. 1 and 6-8). Therefore, your knowledge of such sentences is *scepticism-resistant*, as per the First Investigation, but without also requiring any *vicious regress* of knowing and/or knowers, according to which your knowing *X* requires *that you also know that you know X*, and *that you also know that you know that you know X*, etc., *ad infinitum* and indeed also *ad nauseam*, an epistemic sickness-unto-death.

5. Your reading caveat lector sentences like THE SENTENCE inherently involves using the 2nd-person indexical expression 'you'. Therefore, it puts the burden of collecting evidence and providing proof on *you, the reader* of such sentences and also on *all the other readers* of such sentences, i.e., on *collectives or communities* of rational human animals, not on *the individual writer* of such sentences.

6. Your reading caveat lector sentences like THE SENTENCE inherently requires that, necessarily, you, the essentially embodied reader of such sentences, are embedded in *an egocentrically-centred orientable manifestly real space*. Therefore, it inherently requires (i) that you *are not* living in a digital simulation, (ii) that *you exist*, and (iii) that *the external world exists*.

7. In short, then, starting epistemology by philosophically appealing to your reading caveat lector sentences like THE SENTENCE effectively avoids all the *dualistic*, *mechanistic*, and *rationalistic* cognitive-and-epistemic or theoretical and metaphysical-and-ontological *vices* of classical Cartesian epistemology.[4] At the same time, it also fully possesses some of the very same cognitive-and-epistemic or theoretical and metaphysical-and-ontological *virtues* that are promised by classical Cartesian epistemology – in particular, being a secure foundation for all *sciences* in the maximally broad sense of 'organised

---

3 Hanna 2006a; 2006b; 2015; 2018a; 2018b; 2018c.

4 Descartes 1984-85a; 1984-85b, 1984-85c; 1984-85d; 1984-85e.

bodies of knowledge', including not only the formal sciences (e.g., logic, mathematics, and computer science) and the natural sciences (e.g., physics, chemistry, and biology), but also the social sciences, the 'human sciences' and 'moral sciences' more generally, and philosophy itself. *For without self-manifesting acts or processes of reading, and without rational human readers, how could there be any sciences?* But these cognitive-and-epistemic or theoretical and metaphysical-and-ontological foundational virtues are now fully transposed into the radically *non*-Cartesian and indeed *anti*-Cartesian cognitive-and-epistemic or theoretical, moral or practical, and metaphysical-and-ontological framework of what I call *rational anthropology*.[5]

8. Therefore, philosophically appealing to your reading caveat lector sentences like THE SENTENCE, is *the right way to start epistemology* (see also Hanna 2023n).

4 The Logic of Legibility

In this section I want to explore some of the important *logical* features of the facts or phenomena of *legibility* and *reading* in relation to *the sciences*, as broadly defined in section 3.

For convenience and ease of expression, in what follows in the rest of this essay, I am going to use the terms *legible*, *legibility*, *illegible*, and *illegibility*, respectively, as synonyms for the terms *readable*, *readability*, *unreadable*, and *unreadability*, respectively. Moreover, as per section 3, let's call any sentence that is (i) specifically about the act or process of *reading*, and that is also (ii) *self-referring* by means of the 2nd-person indexical description 'you, the reader', and the indexical description 'this very sentence', *a caveat lector sentence*.

Here, again, is what I take to be the paradigmatic example of a caveat lector sentence:

> **You, the reader of this very sentence, cannot either coherently or self-consistently deny that it is self-evidently true that you are reading this very sentence.**

In section 3, conveniently but also rather prosaically, I called the sentence I displayed in boldface text immediately above, THE SENTENCE, but in this section and henceforth, somewhat more imaginatively, I will call it *The Lector Sentence*. And for the purposes of my argument, again, it will not matter whether The Lector Sentence is a universal *sentence-type* or a particular *sentence-token*. Above all, however, we must recognise that The Lector Sentence is a caveat

---

[5] Hanna 2006a; 2006b; 2015; 2018a; 2018b; 2018c; 2023m.

lector sentence that is *self-manifestingly true*. Then it is highly instructive logically to compare-and-contrast The Lector Sentence with the classical *Liar Sentence*, i.e.,

This very sentence is false.

The Liar Sentence, as *self-manifestingly false*, is not only a contradiction but also *paradoxical*, since necessarily, if it is true then it is false and if it is false then it is true, hence necessarily, it is true if and only if it is false. The Lector Sentence and The Liar Sentence are (i) each of them *reflexive*, i.e., self-referring, (ii) each of them *self-manifesting*, and (iii) mutually *antithetical*. More specifically, The Lector Sentence is reflexive, non-contradictory, true, and furthermore self-manifestingly true, whereas The Liar Sentence is reflexive, contradictory, self-manifestingly false and paradoxical, and furthermore both true and false, i.e., a *truth-value glut*. In these ways, The Lector Sentence shows us *the foundations* of all science, truth, sound proof, and knowledge, whereas, as Alfred Tarski so brilliantly showed, The Liar Sentence shows us *the limits* of all science, truth, sound proof, and knowledge (Tarski 1943; 1956b).

For the purposes of this essay, I will define a *text* as *any sequence of one or more characters*, with a one-character sequence as the lower-bound limiting case, and there is no upper bound on the number of characters, where, as per the *Oxford Encyclopedic English Dictionary*, 'character' is defined as "a printed or written letter, symbol, or distinctive mark" (Hawkins, Allen 1991, 247).

Then, a text is *illegible* if and only any of the perceptible, syntactic, or semantic features that are either individually or conjointly required for reading that text *cannot be discerned*.

Some important and even leading or paradigmatic sciences contain contradictions or even paradoxical sentences. For example, as per Kurt Gödel's first incompleteness theorem, the *Principia Mathematica*-style formalisation of Peano Arithmetic contains undecidable, unprovable, self-contradictory, and indeed paradoxical sentences, *if* that formal system is assumed to be not only sound but *also* complete (Gödel 1967). But no science can contain *nothing but contradictions or paradoxes*, on pain of *explosion*, or logical chaos, whereby not only is it the case that every sentence follows from every other sentence, but also that every sentence is a truth-value glut. So, the fact that no science can contain nothing but contradictions or paradoxes is a direct implication of what I have called, following Hilary Putnam, *the minimal principle of non-contradiction*: necessarily and *a priori*, not every sentence is both true and false (Putnam 1983; Hanna 2006a, ch. 2; 2015a, ch. 5).

Correspondingly, and now zeroing in on the logical features of reading in relation to the sciences, all sciences must be at least

minimally legible, i.e., there cannot be a science that is *completely il-legible*. Let's call that *the principle of minimal legibility*. The principle of minimal legibility obtains because (i) every science must be *communicable*, but if no one can read any of it, then obviously it cannot be communicated and (ii) in order for a science to be, taken as a whole, *meaningful*, *truth-evaluable*, and *knowable*, then at least *some* of the sentences of that science must be completely legible.

Can there be an illegible sentence? Yes, if that means *a sentence that is partially but not completely legible*: a sentence that contains *some* but not all-and-only illegible characters could still be *otherwise legible*. Let 'BLAH' stand for an illegible character within a sentence. Then, the sentence

The cat is sitting on the BLAH.

is partially but not completely legible, and therefore it is illegible to that extent. But there is no such thing as a sentence made up of nothing but illegible characters; for example, the text

BLAH BLAH BLAH BLAH BLAH BLAH BLAH.

is not a sentence: it is gibberish.

Can there be an illegible word? Yes, if that means *a word that is partially but not completely legible*: a word that contains *some* but not all-and-only illegible characters could still be otherwise legible. Let '#' stand for an illegible character within a word. Then, the word

ca#

is partially but not completely legible, and therefore it is illegible to that extent. But there is no such thing as a word made up of nothing but illegible characters; for example, the text

###

is not a word: it is gibberish.

Obviously, all sciences must be ideally aimed at *truth*, *sound proof*, and *knowledge*, even if they do in fact fall short of that, but always only within the limits of *the minimal principle of non-contradiction* – i.e., necessarily and *a priori*, not every sentence is both true and false – and its De Morgan equivalent, *the minimal principle of excluded middle* – i.e., necessarily and *a priori*, some sentences are either true or false with no third value and no value-gap, i.e., necessarily and *a priori*, not every sentence is neither true nor false with a third value or a value-gap – otherwise, they are logical chaos. Correspondingly, all sciences must also be ideally aimed at *complete legibility*,

even if they do in fact fall short of that, but always only within the limits of *the principle of minimal legibility*: otherwise, they are gibberish. Therefore, The Lector Sentence, complete legibility, and the principle of minimal legibility should *also* be explicitly and fully recognised by all philosophers and scientists as taking their rightful logico-normative places alongside the classical logical norms of truth, sound proof, knowledge, and the minimal principle of non-contradiction and/or minimal principle of excluded middle.

5 Necessary and Sufficient Conditions for Legibility and Reading

In this section, I will propose a set of fairly precise necessary and sufficient conditions for legibility and reading.

As I mentioned in section 4, according to the *Oxford Encyclopedic English Dictionary*, 'character' is defined as "a printed or written letter, symbol, or distinctive mark" (Hawkins, Allen 1991, 247).

In view of that, then I will again define a *text* as *any sequence of one or more characters*, where a one-character sequence is the lower-bound limiting case, and there is no upper bound on the number of characters. In turn, what I will call a *text-in-L* is defined as *any sequence of one or more characters belonging to a particular language L*. It is important to note that a language L can contain some characters (hence also some texts) that belong to one or more different languages L2, L3, L4, etc. So, for example, English contains some letters, words, and sentences belonging to other languages, including Greek, Latin, French, German, Italian, etc. Then, I'll provide necessary and sufficient conditions for legibility in two parts, as follows:

1. A text T-in-L is legible if and only if T-in-L satisfies *the perceptibility condition*, *the syntactic condition*, and *the semantic condition*, and

2. all and only such texts-in-L have legibility.

The perceptibility condition says that *the basic orientable (i.e., intrinsically directional, for example, up-down, back-front, or right-left) spatial shape and structure of T-in-L must be at least minimally perceptually detectable*, i.e., that T-in-L must be *at least partially perceptually detectable*, hence *it is not completely perceptually undetectable*, and thereby T-in-L is *able-to-be-scanned* to at least that minimal extent. For example, if a text is completely blacked out, erased, otherwise completely smudged out or obscured, invisibly small, or so big that its shape cannot be perceived, then it is perceptually undetectable and illegible. But on the other hand, as it were, even if a text T-in-L is right-to-left⫛left-to-right mirror-reversed and turned upside down, like the one in English that I have displayed directly below this paragraph **[fig. 1]** it is still *able-to-scanned* to the minimal

extent that *it is not completely undetectable*; and indeed, with a little effort, one can see that in fact it is an upside-down enantiomorph of the extremely interesting English sentence I dubbed *The Lector Sentence* in section 4:

> **You, the reader of this very sentence, cannot either coherently or self-consistently deny that it is self-evidently true that you are reading this very sentence.**

in explicit comparison-and-contrast with the classical *Liar Sentence.*

Figure 1 An upside-down enantiomorph of *The Lector Sentence*

The syntactic condition says that *T-in-L must be at least minimally well-formed*, i.e., that T-in-L must be *at least partially well-formed*, hence *it is not completely well-formed*, and thereby T-in-L is *able-to-be-parsed* to at least that minimal extent. For example, even if a text T-in-L is perceptually detectable, it can be completely jumbled, completely misspelled, or completely ungrammatical, or its characters can be completely randomly distributed, and in any of those ways it would be syntactically illegible. Indeed, ciphers or secret codes (as opposed to hidden messages in otherwise legible texts) are designed to *approach* syntactic illegibility, on the working assumption that the more illegible they are, the harder they are to break; so if there are some ciphers that *have never been broken* and all their creators are dead, or, more thought-experimentally, if there were a cipher created by intelligent non-human aliens that, even in principle, *could never be broken by rational human animals*, then they would be illegible in the syntactic sense. Therefore, a text-in-L's satisfying the perceptibility condition, as such, *is not itself independently sufficient for readability* and thus *it is not itself independently sufficient for being the target of any actual or possible act or process of reading*.

And the semantic condition says that *the conceptual content and/or essentially non-conceptual content of T-in-L must be at least minimally coherent*, i.e., that the conceptual content and/or essentially non-conceptual content of T-in-L must be *at least partially coherent*, hence *not completely incoherent*, and thereby the conceptual content

and/or essentially non-conceptual content of T-in-L is *able-to-be comprehended* to at least that minimal extent. For example, even if a text is minimally perceptible and also minimally well-formed, nevertheless it can still violate minimal requirements of conceptual *sortal correctness* and/or essentially non-conceptual *sortal correctness*, or be strictly non-referential, and be semantic *gibberish*, hence be illegible in the semantic sense, like this non-poetical text-in-English, a paradigm case of sortal incorrectness, devised by Bertrand Russell (Russell 1940, 166):

> quadruplicity drinks procrastination

or this famous poetical text-in-English, a paradigm case of strict non-referentiality, taken from Lewis Carroll's *Jabberwocky*, that I quoted as the second epigraph of this essay:

> 'Twas brillig, and the slithy toves
> Did gyre and gimble in the wabe;
> All mimsy were the borogoves,
> And the mome raths outgrabe. (Carroll 1988)

Therefore, that text from *Jabberwocky*'s satisfying the perceptibility condition *together with* the syntactic condition, yet also failing the semantic condition, shows that the first two conditions *are not themselves conjointly sufficient for readability* and thus that *they are not themselves conjointly sufficient for being the target of any actual or possible act or process of reading*. Of course, millions of people, including you, the reader of this very essay, have in some sense or another 'read' that text from *Jabberwocky*; but my way of explaining away this apparent inconsistency is just to point out that *Jabberwocky* is indeed *legible* in both the perceptible and synactic senses (so in two senses, readable), but *illegible* in the semantic sense (so in one sense, unreadable), hence not legible in *all* relevant senses, hence *illegible* by my contextual definition, or conceptual analysis, of legibility. The same point holds, mutatis mutandis, for "quadruplicity drinks procrastination" and all other essentially similar texts-in-L: you can "read" it in two senses (the perceptible sense and the syntactic sense), but strictly speaking, it is *illegible* according to the necessary and sufficient conditions of legibility, precisely because it fails the semantic condition.

Assuming all of that so far, I am now in a position to provide precise necessary and sufficient conditions for the act or process of reading. In the following contextual definition, or conceptual analysis, by *person* I mean *rational human minded animal*: namely, a living human organism that is capable of (i) consciousness, (ii) self-consciousness, (iii) caring (i.e., desire, emotion, and feeling – the affects),

(iv) sensible cognition, (v) intellectual cognition, (vi) volition, (vii) object-directed and act-directed intentionality more generally, and (viii) free agency. Then, I will provide necessary and sufficient conditions for reading in two parts, as follows:

1. A person P reads a text T-in-L if and only if P consciously or self-consciously at least minimally scans, at least minimally parses, and also at least minimally comprehends T-in-L, and

2. all and only such acts or processes are reading.

It is important to note that, consistently with this contextual definition, or conceptual analysis, of reading, a person P can read a text T-in-L either *aloud* or *silently to themselves*. It is also important to note that neither scanning, nor parsing, nor comprehending, need be *self-consciously* or *reflectively* performed: this can be done in a more-or-less or even altogether *pre-reflectively* or *unself-consciously conscious* way; indeed, we typically 'look right through' what we are reading in order to go directly to the *meaning* (whether sense, reference, or speech-act uptake) of what we are reading, and altogether overlook the scanning, parsing, and comprehending dimensions of the act or process of reading itself. In order to bring those dimensions back into view, all you have to do is to repeat any text-in-L – for example, a sentence or word – out loud a few times (say, ten times) until it sounds strangely *bereft* of meaning; that strange *absence-of-meaning* has then become vividly manifest to you precisely because the perceptibility and syntax of that particular text-in-L have been temporarily self-consciously detached from what you have previously been, pre-reflectively and unself-consciously yet still consciously, comprehending.

And it is also important to note that the point I made above about 'readers' of *Jabberwocky* and "quadruplicity drinks procrastination" goes, *mutatis mutandis*, for my contextual definition, or conceptual analysis, of reading: of course, millions of people, including you, the reader of this very essay, are in some sense or another 'readers' of that text from *Jabberwocky*; and no doubt a few thousand people have read "quadruplicity drinks procrastination"; but my way of explaining away this apparent inconsistency too, is just to point out that *Jabberwocky* and "quadruplicity drinks procrastination" can indeed be *read* in both the perceptible and synactic senses (so in two senses, that is reading), but *cannot be read* in the semantic sense (so in one sense, that is *not* reading), hence it is not reading in *all* the relevant senses, hence it is *not reading* by my contextual definition, or conceptual analysis, of reading.

These necessary and sufficient conditions for legibility and reading, when taken together with the logic of legibility, amount to the basics of *a theory of legibility and reading*. To be sure, in the interests of full philosophical disclosure, I must admit that for the purposes of these analyses and this theory, I have presupposed (i) the very

ideas of (ia) *a language*, including its characteristic syntactic and se-
mantic properties, and (ib) *our knowledge of a language*, including
our knowledge of its characteristic syntactic and semantic proper-
ties (see, e.g., Chomsky 1957; 1988), (ii) a certain theory of *linguistic
cognition* and *logical cognition* (see, e.g., Hanna 2006a, esp. chs. 4,
6), (iii) a specifically *dual-content* cognitive semantics of conceptual
content and essentially non-conceptual content, the latter of which
also crucially functions as the source of what Otto Paans and I call
thought-shapers (see, e.g., Hanna 2015, esp. chs. 2, 4; Hanna, Paans
2021), for the explanation of linguistic meaning, and above all, an-
other necessary condition of reading: (iv) the rational human capac-
ity *to understand at least one language, at least minimally* (see, e.g.,
PI; Chomsky 1957; 1988).

But, one need not necessarily be able to *speak* a language L – in
the sense of being able to *talk-in* L – in order to be able to read texts-
in-L. For example, like many other English-speaking people, I can
understand and read a few words or sentences in some other lan-
guages (say, Finnish, Hungarian, or Russian) that I cannot talk-in at
all. More interestingly, perhaps, it seems that there are or at least
have been some actual children who can *understand* texts-in-L, and
thus, at least in principle, can *read* texts-in-L, where L is their first
or native language, *before* they can talk-in L. For example, accord-
ing to various sources, Albert Einstein did not talk until he was 3, 4,
or 5; but according to others' testimony and his own, for some peri-
od prior to that time he was in fact able to *understand* German (see,
e.g., Brian 1996), a phenomenon that is more generally known now-
adays as *late-talking syndrome* or *Einstein Syndrome* (Smith-Garcia
2020). Given Einstein's native intellectual brilliance, then presuma-
bly, during the time when he understood German but could not yet
talk-in German, *he could still have been taught to read German or
have learned on his own to read German*. So, my theory of legibility
and reading predicts that for at least some actual children who are
late-talkers, *it should be possible for them to be taught to read texts-
in-L or learn on their own to read texts-in-L, before they can talk-in L*.
At the present time, I have not done a systematic survey of the rele-
vant scientific literature in order to find out whether this prediction
has already been empirically tested, and if so, whether it has been
confirmed or disconfirmed by means of replicable studies, although
at least one book by a non-scientist says that it has been confirmed
(Sowell 1997). But in any case, it would be extremely philosophically
interesting to me, and also perhaps of some real-world interest and
value to late-talkers and their families, if it were indeed confirmed
or at least confirmable by replicable studies.

Correspondingly, here is something about the relationship be-
tween reading and *writing*, in view of what I have just been arguing
about the relationship between reading and *talking*. If there actually

are some late-talkers who read before they can talk, then reading *logically precedes* and sometimes also *psychologically precedes* talking. Now, the very act or process of *writing* presupposes that the writer is *already able to read*, at the very least, *their own writing*: therefore, *reading logically precedes writing*. Of course, writing is typically taught to children only *after* they can talk. But if reading logically precedes and sometimes also psychologically precedes talking, and if reading logically precedes writing, then a late-talker who can read, could also, at least in principle, be taught to write or learn on their own to write. So, my theory of legibility and reading also predicts that for at least some actual children who are late-talkers *and* readers, then *it should also be possible for them to be taught to write texts-in-L or learn on their own to write texts-in-L, before they can talk-in L*. And again, it would be extremely philosophically interesting to me, and also perhaps of some real-world interest and value to late-talkers and their families, if this prediction were indeed confirmed or at least confirmable by replicable studies.

6 Are There Some Legible Texts that Even the World's Most Sophisticated Robot Cannot Read?

I am in a position now to say something substantive about the metaphysics and ontology of legible texts. As we have seen, the intentional targets of the act or process of reading are at-least minimally scannable, at-least minimally parse-able, and at-least minimally comprehensible *structural objects belonging to some or another language L*, that are ineluctably embedded in an egocentrically-centred, orientable, manifestly real, three-dimensional space, thereby necessarily requiring the actual existence and essential embodiment of the reader. As linguistic structural objects, the intentional targets of reading are manifestly real *linguistic physical tokens* of manifestly real *linguistic physical types*, which in turn are *inherently repeatable objects* that are *non-platonically and kantianly abstract* according to this definition:

> X is non-platonically and kantianly abstract if and only if X *is not uniquely located and realized in manifestly real spacetime*, and X is *concrete* otherwise. (Hanna 2015, 269-70)

Now, the *rational human cognition* of concrete tokens of the linguistic structural objects of reading, whether in perception, memory, or imagination, is what Kant calls *sensibility* (*Sinnlichkeit*), which in turn requires a capacity for first-order conscious or self-conscious, essentially non-conceptual, and non-empirical *unified formal spatial or temporal representation*, or what Kant calls *pure intuition* or *reine*

Anschauung (CPR A20/B34-5). Therefore, the act or process of reading is an essentially *intuitionistic* activity *that does not require any sort of platonic objects*. The act or process of reading thereby wholly avoids the classical metaphysical/ontological and epistemic problems of *platonism*, especially including *The Benacerraf Dilemma*, which says: (i) on the one hand, our standard Tarskian semantics of mathematical truth requires platonically abstract objects that exist outside of spacetime and are causally inert, but (ii) on the other hand, our best theory of human knowledge requires directly sensibly accessible causal objects of perception, so (iii) mathematical truth is humanly unknowable (Benacerraf 1973). In short, the act or process of reading, by virtue of *its intuitionistic nature*, is decisively (to coin a nifty neologism) *trans-Benacerraf-Dilemma-istic*, precisely because it is metaphysically *structuralist*, ontologically *non-platonistic*, although fully accommodating *non-platonically and kantianly abstract objects*, and epistemically *scepticism-resistant*, from the get-go (Hanna 2015, chs. 6-8).

With those points under our belts, I turn next to the *strong thesis of artificial intelligence*, also known as *strong AI*, which is the two-part thesis which says (i) that rational human intelligence can be explanatorily and ontologically reduced to Turing-computable algorithms and the operations of digital computers (also known as *the thesis of formal mechanism, as it is applied to rational human intelligence*), and (ii) that it is technologically possible to build a digital computer that is an exact counterpart of rational human intelligence, such that this machine not only exactly reproduces (or simulates) all the actual performances of rational human intelligence, but also outperforms it (also known as *the counterpart thesis*) (see, e.g., Block 1980, part 3; Kim 2011, ch. 6). If the strong AI thesis is true, then, at the very least, necessarily, *some robot must be able to do anything that any ordinary rational human minded animal can do*. Correspondingly, the standard strategy in the strong AI program is to start with some accomplishment, act, or task that any ordinary rational human minded animal can already achieve or perform, and then *reverse-engineer* a digital computer program and either a stationary digital computer or a mobile digital computer – a robot – that can perform the same accomplishment, act or task, at least as well as, or better than, any ordinary rational human minded animal. Now, robots can do some things that no stationary digital computer can do. So, the leading question I have asked in the title of this section is whether there are some legible texts that we – i.e., ordinary rational human minded animals – can read, but even the world's most sophisticated robot – cannot read? If so, *then the strong AI thesis is false and the strong AI program is impossible*.

My theory of legibility and reading predicts *that there are legible texts that ordinary rational human minded animals can read, that*

even the world's most sophisticated robot cannot read, even when we bracket temporarily the contested issue of the role of consciousness or subjective experience, i.e., sentience, *versus* computational zombie-states, i.e., non-consciousness or non-sentience, in acts or processes of reading.

To show this, let us consider computational reading that is based on *optical character recognition* (OCR), and let us also make the plausible assumption that even the world's most sophisticated robot will have to employ some or another version of OCR:

> There are two basic methods used for OCR: [m]atrix matching and feature extraction. Of the two ways to recognize characters, matrix matching is the simpler and more common.

> Matrix Matching compares what the OCR scanner sees as a character with a library of character matrices or templates. When an image matches one of these prescribed matrices of dots within a given level of similarity, the computer labels that image as the corresponding ASCII character.

> Feature Extraction is OCR without strict matching to prescribed templates. Also known as Intelligent Character Recognition (ICR), or Topological Feature Analysis, this method varies by how much "computer intelligence" is applied by the manufacturer. The computer looks for general features such as open areas, closed shapes, diagonal lines, line intersections, etc. This method is much more versatile than matrix matching. Matrix matching works best when the OCR encounters a limited repertoire of type styles, with little or no variation within each style. Where the characters are less predictable, [intelligent character recognition, or topological feature analysis,] is superior. (Data ID, 2023)

Now, let us consider *garbled texts*: that is, texts that contain misspelled sub-texts, sub-texts with missing characters, sub-texts with obscured characters, sub-texts whose characters are excessively large or excessively small, ungrammatical sub-texts, incomprehensible sub-texts, and above all, texts that contain *disoriented sub-texts*, that is, sub-texts reversed in a mirror, tipped sideways, or upside down. Necessarily, any digital computer running an OCR program must process information in a step-by-step sequence, and whenever it encounters *something that it cannot recognise as a determinate unit of information*, whether by matrix matching, feature extraction, also known as intelligent character recognition, also known as topological feature analysis, or whatever, *it simply stops processing and cannot go on*. This in turn triggers Turing's *halting problem* in the logical theory of digital computation. The halting problem, which is provably

unsolvable, says that there is no general algorithm for determining, from a description of an arbitrarily-selected computer program together with an arbitrarily-selected input, whether this program will either effectively complete its computation, i.e., be computable/decidable, or else continue processing forever, i.e., be uncomputable/undecidable (Turing 1936-37; Boolos, Jeffrey 1989, 28-33, 41-2, 49-50). Therefore, once a digital processing system has simply *stopped* processing, there is no general way of determining whether it has either effectively completed its computation or else would have continued processing forever.

But, as ordinary rational human minded animals, we intuitionistically represent texts as complete *Gestalt*-structures that are embedded in manifestly real, egocentrically-centred, orientable space, and therefore *we always have a unified formal spatial representation of the text as a whole for guiding us through our reading*, not only before we begin scanning it sequentially, but also throughout the time we are scanning it sequentially. This enables us to *jump over, fill in*, or *creatively interpret* illegible sub-texts, and/or *re-orient disoriented sub-texts in spatial imagination*, when we encounter garbled texts, hence we are able to read all sorts of garbled texts, provided that they are otherwise at-least minimally legible by the criteria I provided above. Hence *our ordinary rational human minded animal ability to read garbled texts, provided that they are otherwise at-least minimally legible, will necessarily exceed the digital processing abilities of any and all computers to read those texts*, i.e., there are some legible texts that ordinary rational human animals can read, that even the world's most sophisticated robot cannot read.

Here is an example of a legible text that any ordinary rational human animal can read, but even the world's most sophisticated robot cannot read, using a text that we have seen twice already:

'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves,
And the mome raths outgrabe. (Carroll 1988)

By hypothesis, this text from *Jabberwocky* satisfies the perceptibility condition and the syntactic condition, yet also *fails* the semantic condition. So it is *prima facie* illegible and unreadable. Now, consider any ordinary rational human minded animal, for example, *Bob*. And correspondingly, let us consider the world's most sophisticated robot, a behavioural counterpart to Bob, *Robobob*. After successfully scanning and parsing that text from *Jabberwocky*, Robobob attempts to comprehend it, but cannot do so, and concludes that it is incomprehensible, so stops processing. But Bob, who like any other ordinary rational human minded animal, has an innate capacity

for creative 'gumption' – i.e., creative initiative and resourcefulness, which of course both Lewis Carroll and Einstein possessed to an extraordinary degree – does not give up, and continues to think about the text, muse about it, sleep on it, and dream about it. Then finally, when Bob wakes up the next day, he finds that, like a creative artist or creative scientist, he is freely and spontaneously assigned private meanings to all the nonsense terms, and has a novel semantic *Gestalt* of the entire text, so that the text is now fully legible for him. These meanings are not *necessarily* private – a "private language" in that absolute sense, as Wittgenstein compellingly argued, is conceptually impossible (PI, I, §§ 242-315; Hanna 2021, ch. XIII) – since in principle Bob could tell other people about them, or others could somehow learn about these meanings in some other way: hence they are only *contingently* private and in-principle universally shareable. But, as a matter of fact, Bob never tells anyone about them, and no else ever learns about them, including of course Robobob. Yet the *Jabberwocky* text is *legible for Bob in all three senses*, and he privately enjoys reading it over and over, for the rest of his life. Nevertheless, Robobob cannot read that text because it stopped processing, and also, above all, because it is nothing but a mobile digital computer and therefore lacks any inherent capacity whatsoever for creative gumption, although of course it could be programmed to exhibit behaviour that mimics creative gumption. Hence there is at least one legible text, i.e., that *Jabberwocky* text, that is legible for Bob, and also for any other ordinary rational human minded animal with at least as much creative gumption as Bob, that even the world's most sophisticated robot cannot read. So the strong AI thesis is false and the strong AI program is impossible.

7 Conclusion

It should be self-evident by now that the philosophy of reading – by which I mean *serious* philosophy of reading –[6] is *centrally and fundamentally important*, even though it has been generally avoided by Analytic and non-Analytic philosophers alike since 1900. Finally, then, I am going to return briefly to the *difficulty* of the philosophy of reading, as so insightfully and rightly pointed up by Wittgenstein (PI, I, §§ 155-6): precisely *why* is the philosophy of reading *such hard work*? I think that it is for two reasons.

   First, it is because the philosophy of reading brings together central and fundamental issues and problems in philosophical logic, the philosophy of language, the philosophy of language-and-mind,

----

[6]   See note 1 above.

epistemology, metaphysics, cognitive science, and the general theory of human rationality, in an *inherently interconnected* and indeed *profoundly tightly-knotted-up* way. In that sense, this Wittgenstein-inspired inaugural investigation in the philosophy of reading has been an extended exercise in patiently tugging away at this gnarly knot and, to the extent I've been able to succeed, *untying* it.

And **second**, it is because the act or process of reading *also* presents itself as something so utterly *obvious* – after all, every reader of this very sentence has already learned to read, and of course virtually all of us have also done so before the age of 6 or 7, so "even a child can do it!" – that we completely fail to notice its profound complexity and its central and fundamental importance in our rational human lives, a notable instance of the cognitive pathology I have called *young fish syndrome* in section 1, riffing on David Foster Wallace's famous allegory. Indeed, only someone like Frederick Douglass, a former slave who had been immorally *prevented* from learning to read until finally taught its basics by the kindly wife of one of his slave masters, 'Master Hugh', would be fully and vividly aware of the act or process of reading *as such*, especially in its disruptive and indeed explosive potential for radically changing our conscious, self-conscious, cognitive, affective, moral, and sociopolitical lives, as Douglass so brilliantly and movingly describes it in the third epigraph at the top of this essay (Douglass 1995; see also Scott 2023). So in that sense, my Wittgenstein-inspired inaugural investigation in the philosophy of reading has also been *a heads-up call* to all contemporary philosophical logicians, philosophers of language, epistemologists, metaphysicians, cognitive scientists, and theorists of human rationality:

Caveat lector! You avoid the philosophy of reading *inevitably and only* at the excessively high theoretical cost of disastrously and even tragically *going off the rails* in philosophical logic, the philosophy of language, the philosophy of language-and-mind, epistemology, metaphysics, cognitive science, and the general theory of human rationality, *from the get-go.*[7]

---

[7] I am grateful to Martha Hanna for thought-provoking conversations on and around the main topics of this essay.

References

Benacerraf, P. (1973). "Mathematical Truth". *Journal of Philosophy*, 70, 661-80.

Block, N. (ed.) (1980), *Readings in the Philosophy of Psychology*. Vol. 1, Cambridge, MA: Harvard University Press.

Boolos, G.; Jeffrey, R. (1989). *Computability and Logic*. 3rd ed. Cambridge: Cambridge University Press.

Brian, D. (1996). "Einstein: A Life", ch. 1. *The Washington Post*. `https://www.washingtonpost.com/wp-srv/style/longterm/books/chap1/einstein.htm`.

Carroll, L. (1988). *Through the Looking-Glass*. New York: Dial.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1988). *Language and Problems of Knowledge*. Cambridge, MA: MIT Press.

Data ID (2023). "What's OCR?". *Data Identification Online*. `http://www.data-id.com/aboutocr.htm`.

Descartes, R. (1984-85a). *The Philosophical Writings of Descartes*. 2 vols. Transl. by J. Cottingham; R. Stoothoff; D. Murdoch. Cambridge: Cambridge University Press.

Descartes, R. (1984-85b). *Rules for the Direction of the Mind*. Descartes 1984-85a, vol. 1, 9-78, AT X, 359-472.

Descartes, R. (1984-85c). *Discourse on the Method*. Descartes 1984-85a, vol. 1, 111-51, AT VI, 1-78.

Descartes, R. (1984-85d). *Meditations on First Philosophy*. Descartes 1984-85a, vol. 2, 3-62, AT VII, 1-90.

Descartes, R. (1984-85e). *Principles of Philosophy*. Descartes 1984-85a, vol. 1, 179-291, AT VIIIA, 1-329.

Douglass, F. (1995). *Narrative of the Life of Frederick Douglass*. Garden City, NY: Dover.

Frege, G. (1953). *The Foundations of Arithmetic*. 2nd ed. Transl. by J.L. Austin. Evanston, IL: Northwestern University Press.

Gödel, K. (1967). "On Formally Undecidable Propositions of Principia Mathematica and Related Systems". Van Heijenoort, J. (ed.), *From Frege to Gödel*. Cambridge, MA: Harvard Univesity Press, 596-617.

Hanna, R. (2006a). *Rationality and Logic*. Cambridge, MA: MIT Press. `https://www.academia.edu/21202624/Rationality_and_Logic`.

Hanna, R. (2006b). "Rationality and the Ethics of Logic". *Journal of Philosophy*, 103, 67-100. `https://www.academia.edu/7940609/Rationality_and_the_Ethics_of_Logic`.

Hanna, R. (2011). "Minding the Body". *Philosophical Topics*, 39(1), 15-40. `https://www.academia.edu/4458670/Minding_the_Body`.

Hanna, R. (2015). *Cognition, Content, and the A Priori: A Study in the Philosophy of Mind and Knowledge*. Oxford: Oxford University Press. `https://www.academia.edu/35801833/The_Rational_Human_Condition_5_Cognition_Content_and_the_A_Priori_A_Study_in_the_Philosophy_of_Mind_and_Knowledge_OUP_2015_`.

Hanna, R. (2018a). *The Rational Human Condition*. Vol. 1, *Preface and General Introduction, Supplementary Essays, and General Bibliography*. New York: Nova Science. `https://www.academia.edu/35801821/The_Rational_Human_Condition_1_Preface_and_General_Introduction_`

Supplementary_Essays_and_General_Bibliography_Nova_Science_2018_.

Hanna, R. (2018b). *The Rational Human Condition*. Vol. 2, *Deep Freedom and Real Persons: A Study in Metaphysics*. New York: Nova Science. `https://www.academia.edu/35801857/The_Rational_Human_Condition_2_Deep_Freedom_and_Real_Persons_A_Study_in_Metaphysics_Nova_Science_2018_`.

Hanna, R. (2018c). *The Rational Human Condition*. Vol. 3, *Kantian Ethics and Human Existence: A Study in Moral Philosophy*. New York: Nova Science. `https://www.academia.edu/36359647/The_Rational_Human_Condition_3_Kantian_Ethics_and_Human_Existence_A_Study_in_Moral_Philosophy_Nova_Science_2018_`.

Hanna, R. (2021). *The Fate of Analysis: Analytic Philosophy From Frege to The Ash-Heap of History*. New York: Mad Duck Coalition. `https://themadduckcoalition.org/product/the-fate-of-analysis/`.

Hanna, R. (2022). "Merleau-Ponty Meets The Kripke Monster Redux: The Essential Embodiment Theory Now". Unpublished ms. `https://www.academia.edu/88987003/Merleau_Ponty_Meets_The_Kripke_Monster_Redux_The_Essential_Embodiment_Theory_Now_December_2022_version_`.

Hanna, R. (2023a). "How and Why ChatGPT Failed The Turing Test". Unpublished ms. `https://www.academia.edu/94870578/How_and_Why_ChatGPT_Failed_The_Turing_Test_January_2023_version_`.

Hanna, R. (2023b). "How and Why to Perform Uncomputable Functions". Unpublished ms. `https://www.academia.edu/87165326/How_and_Why_to_Perform_Uncomputable_Functions_March_2023_version_`.

Hanna, R. (2023c). "Creative Rage Against the Computing Machine: Necessary and Sufficient Conditions for Authentic Human Creativity". Unpublished ms. `https://www.academia.edu/101633470/Creative_Rage_Against_the_Computing_Machine_Necessary_and_Sufficient_Conditions_for_Authentic_Human_Creativity_June_2023_version_`.

Hanna, R. (2023d). "The Myth of Artificial Intelligence and Why It Persists". Unpublished ms. `https://www.academia.edu/101882789/The_Myth_of_Artificial_Intelligence_and_Why_It_Persists_May_2023_version_`.

Hanna, R. (2023e). "Further Thoughts on the Myth of Artificial Intelligence, the Mind Snatching Invasion of the Chatbots, and How to Save Higher Education". Unpublished ms. `https://www.academia.edu/103442312/Further_Thoughts_on_The_Myth_of_Artificial_Intelligence_the_Mind_Snatching_Invasion_of_the_Chatbots_and_How_to_Save_Higher_Education_June_2023_version_`.

Hanna, R. (2023f). "Are There Some Legible Texts that Even the World's Most Sophisticated Robot Can't Read?". Unpublished ms. `https://www.academia.edu/95866304/Are_There_Some_Legible_Texts_That_Even_The_Worlds_Most_Sophisticated_Robot_Cant_Read_January_2023_version_`.

Hanna, R. (2023g). "It's All Done with Mirrors: A New Argument that Strong AI is Impossible". Unpublished ms. `https://www.academia.edu/95296914/Its_All_Done_With_Mirrors_A_New_Argument_That_Strong_AI_is_Impossible_January_2023_version_`.

Hanna, R. (2023h). "Babbage-In, Babbage-Out: On Babbage's Principle". Un-published ms. `https://www.academia.edu/101462742/Babbage_In_Babbage_Out_On_Babbages_Principle_May_2023_version_`.

Hanna, R. (2023i). "'It's a Human Thing. You Wouldn't Understand'. Computing Machinery and Affective Intelligence". Unpublished ms. `https://www.academia.edu/102664604/_Its_a_Human_Thing_You_Wouldnt_Understand_Computing_Machinery_and_Affective_Intelligence_June_2023_version_`.

Hanna, R. (2023j). "Let's Pretend We're Machines: A New Proof that Free Will Exists and that You Have It". `https://www.academia.edu/103199623/Lets_Pretend_Were_Machines_A_New_Proof_That_Free_Will_Exists_and_That_You_Have_It_June_2023_version_`.

Hanna, R. (2023k). "The Circularity of Human Rationality". Unpublished ms. `https://www.academia.edu/106083565/The_Circularity_of_Human_Rationality_August_2023_version_`.

Hanna, R. (2023l). "Oppenheimer, Kaczynski, Shelley, Hinton, & Me: Don't Pause Giant AI Experiments, *Ban Them*". Unpublished ms. `https://www.academia.edu/97882365/Oppenheimer_Kaczynski_Shelley_Hinton_and_Me_Don_t_Pause_Giant_AI_Experiments_Ban_Them_July_2023_version_`.

Hanna, R. (2023m). "Philosophy that Can't be Done by Chatbots: Rational Anthropology". Unpublished ms. `https://www.academia.edu/74562370/Philosophy_That_Can_t_Be_Done_By_Chatbots_Rational_Anthropology_April_2023_version_`.

Hanna, R. (2023n). "The Philosophy of Reading as First Philosophy". Unpublished ms. `https://www.academia.edu/107390679/The_Philosophy_of_Reading_as_First_Philosophy_September_2023_version_`.

Hanna, R. (2024). *Science for Humans: Life, Mind, the Formal-&-natural Sciences, and a New Concept of Nature*. Berlin: Springer Nature.

Hanna, R.; Maiese, M. (2009). *Embodied Minds in Action*. Oxford: Oxford University Press. `https://www.academia.edu/21620839/Embodied_Minds_in_Action`.

Hanna, R.; Paans, O. (2021). "Thought-Shapers". *Cosmos & History*, 17(1), 1-72. `http://cosmosandhistory.org/index.php/journal/article/view/923`.

Hawkins, J.M.; Allen, R. (eds) (1991). *The Oxford Encyclopedic English Dictionary*. Oxford: Clarendon Press; Oxford University Press.

Ingarden, R. (1973). *The Cognition of the Literary Work of Art*. Transl. by R.A. Crowley; K.R. Olson. Evanston, IL: Northwestern University Press.

Kim, J. (2011). *Philosophy of Mind*. Boulder, CO: Westview.

Putnam, H. (1983). *Realism and Reason: Philosophical Papers*, vol. 3. Cambridge: Cambridge University Press.

Putnam, H. (1983). "There Is At Least One A Priori Truth". Putnam 1983, 98-144.

Rorty, R. (ed.) (1967a). *The Linguistic Turn*. Chicago, IL: University of Chicago Press.

Rorty, R. (1967b). "Introduction: Metaphilosophical Difficulties of Linguistic Philosophy". Rorty 1967a, 1-39.

Russell, B. (1940). *An Inquiry into Meaning and Truth*. London: George Allen and Unwin.

Scott, A.O. (2023). "Everyone Likes Reading. Why Are We So Afraid of It?". *New York Times Book Review*, 21 June. `https://www.nytimes.com/2023/06/21/books/review/book-bans-humanities-ai.html`.

Smith-Garcia, D. (2020). "Einstein Syndrome: Characteristics, Diagnosis, and Treatment". *Healthline*. `https://www.healthline.com/health/einstein-syndrome`.

Sowell, T. (1997). *The Einstein Syndrome: Bright Children Who Talk Late*. New York: Basic Books.

Tarski, A. (1943). "The Semantic Conception of Truth and the Foundations of Semantics". *Philosophy and Phenomenological Research*, 4, 342-60.

Tarski, A. (1956a). *Logic, Semantics, and Metamathematics*. Oxford: Oxford University Press, 152-278.

Tarski, A. (1956b). "The Concept of Truth in Formalized Languages". Tarski 1956a, 152-278.

Turing, A. (1936-37). "On Computable Numbers, with an Application to the Entscheidungsproblem". *Proceedings of the London Mathematical Society*, 42, 230-65, with corrections in 43, 544-546.

Wallace, D.F. (2012). "This is Water". *FS blog*. `https://fs.blog/2012/04/david-foster-wallace-this-is-water/`.

Wittgenstein, L. (1953). *Philosophical Investigations*. Transl. by G.E.M. Anscombe. New York: Macmillan. [PI]

Wittgenstein, Contexts, and Artificial Intelligence
An Engineer Among Philosophers, a Philosopher Among Engineers

Carlo Penco
Università di Genova, Italy

Abstract Could we take Wittgenstein's philosophy as antagonistic or compatible with AI? Interpretations go in opposite directions. In this paper, I stand with compatibilists and claim that Wittgenstein's discussion on contexts has deep connections with the early stages of AI at different levels. Furthermore, his remarks on context aids in the comprehension of the recent advancement in machine learning based AI, although they embed a warning against the oversimplified association of artificial and human intelligence.

Keywords Wittgenstein. AI. Context. Concept. Family resemblance.

Summary 1 Introduction. – 2 Wittgenstein's Contexts. – 3 Connections Between Wittgenstein's Ideas of Context and AI. – 4 Understanding and Learning Language in Context. – 5 Summary and Conclusion.

1   Introduction

"But machines can't think!" Wittgenstein considers this assertion with doubt, given that he follows remarking that we say "only of a human being and what is like one that it thinks", including dolls and ghosts (PI, § 360). His discussion on the idea that machines can think began in *The Blue Book* and went through the *Remarks on the Foundations of Mathematics* and the *Philosophical Investigations*, and it has often been interpreted as a criticism of Alan Turing, who followed his lectures on the foundation of mathematics in 1939. At the beginning, Wittgenstein considered attributing thinking to machines a category mistake, like attributing colour to numbers (BB, 47) or speaking of "artificial pain" (PG, 64). However, his later remarks are more ambivalent, and interpretations are divided between antagonists, who interpret Wittgenstein's work as a means to contrast AI, and compatibilists, who see Wittgenstein as an inspiration for AI.

Among the antagonists, Dreyfus (1972; 1992) links Wittgenstein's language games to Heidegger's *Dasein* and uses Wittgenstein's conception of "form of life" for denouncing the limits of AI: a form of life cannot be programmed. Harre (1988) relies on the fact that human skills and practices are strictly linked to the human world, while Casey (1988) assumes that we may find in Wittgenstein an argument against the idea that machines can think. However, he claims that the supposed Wittgenstein's argument hides a missing premise that would make the argument invalid or weaker than it appears.[1] Shanker (1998) links Wittgenstein's antagonism towards AI to his criticism of both a mechanistic and psychologistic view of mind where there would be no space for a normative conception of calculation. Although he finds "obscure" Wittgenstein's quotation of Turing machines as "humans who calculate" (RPPI, § 1096), he eventually interprets it as suggesting the relevance of the difference between mechanically following a rule and following a mechanical rule. Fuchs (2022, fn. 8), quoting from Wittgenstein's *Zettel*, insists that a precondition for human interaction is attributing subjectivity to our counterpart, so that we may have shared feelings.

Compatibilists try to bypass Searle's criticism that AI is syntactic

---

[1] Casey presents a reconstruction of Wittgenstein's argument as follows: (i) we are entitled to predicate 'thinking' only to human beings and to what is sufficiently like a human being; (ii) machines are not sufficiently like human beings; (ii) therefore we cannot predicate 'thinking' of machines. He claims that the second premise, to grant the conclusion, should be modal: "Machines cannot be and can never be sufficiently like human beings". But it is difficult to grant that *a priori*. Therefore, it would become empirical and open to debate, given the vagueness of the expression "sufficiently like human beings". See the more complex view by Beran 2014.

and not semantic and lacks intentionality,[2] relying on Wittgenstein's conception of thinking. Obermeier (1983) suggests connecting the concept of thinking to the idea of understanding-as-performance; against Dreyfus, Neumayer (1986) suggests that our interaction with robots may bring about embedding robots in our form of life,[3] and is implicitly followed by Sunday Grève (2023), who speaks of "artificial forms of life". Penco (2012) relies on the idea of understanding as symbol manipulation in context, and Floyd (2019) analyses the connections and philosophical consonance between Turing and Wittgenstein, detailing their meetings and reciprocal influence. Xu (2016) analyses some of the main arguments by Shanker and insists that Shanker's criticism concerns ideas of the old artificial intelligence and does not touch recent connectionist systems that are far away from adherence to psychologism. Molino and Tagliabue (2023) claim that Wittgenstein inspired artificial intelligence, shortly referring to Margaret Masterman, to whose work Wilks (2005) dedicated a careful analysis, on which we devote a section in this paper. Gaver (2023) and Gomes and Selman (2023) enthusiastically put Wittgenstein as ideally inspiring machine learning systems and Large Language Models like ChatGPT.

Who is right? On the side of compatibilism, we have Wittgenstein's attack on the mentalist view of thinking as a hidden process accompanying speech. Against the mentalist view, Wittgenstein defined thinking as "the activity of operating with signs" (BB, 6) or "operating with symbols" (PG, 65), specifying that "'thinking' is a fluid concept, and what 'operating with symbols' is, must be looked at separately in each individual case" (RPPII, §§ 7-8). In face of the question of whether the human body is to be called a machine, he answers, "[i]t surely comes as close as possible to being such a machine" (PI, § 359). These statements are not too far from Newell and Simon's Physical Symbol System Hypothesis (1976), which holds that a physical symbol system has both the necessary and sufficient means for intelligence and that a human is a physical symbol system (as a computer is). Following this analogy, we may interpret his curios remark about Turing machines: "These machines are humans who calculate" (RPPI, § 1096). Differently from Shanker (1998), who devotes quite a bit of space to commenting this passage, it seems to me that this assertion suggests that Wittgenstein was somehow approving Turing's presentation of his machines as "a human being operating with a table of

---

2 In his paper on the Chinese room, Searle never quotes Wittgenstein, although he can be considered a 'wittgensteinian' for his work on the idea of background.

3 He relies on a quotation where Wittgenstein suggests a possible "language game in which I produce information automatically, information which can be treated by other people quite as they treat non-automatic information" (RPPI, § 817).

rules according to a certain routine" (Floyd 2019, 280-1). Following Floyd, we may remark that Turing used the notion of a human calculator to ground the foundation of logic on socially shareable procedures, and therefore far away from a psychological account or relying on what happens in the mind, in a very Wittgensteinian mood.[4] At the same time, Wittgenstein gives a very open attitude to what we mean by "calculating", saying, just before the remark on Turing machines:

> That we calculate with some concepts and with others do not, merely shews how different in kind conceptual tools are (how little reason we have ever to assume uniformity here). (RPPI, § 1095)

Wittgenstein died in 1951, and it is not awkward to consider the kind of early work in AI, which emerged in the 1950s and developed in the second half of the twentieth century, as part of the spirit of the time. We cannot forget that the first project of AI was presented at Dartmouth College in 1956, five years after Wittgenstein's death, and that part of Wittgenstein's background consisted of studies in engineering.[5] Although compatibilists have made some suggestions regarding these affinities, there is a blind spot in all attempts to propose a Wittgensteinian view of AI: the variety of his remarks on context. I think that Wittgenstein's complex notion of context is what may clarify the deep connection between Wittgenstein's philosophy and the work of AI, and, at the same time, present a warning to a too easy assimilation between artificial and human intelligence.

In what follows, I give an overview of Wittgenstein's remarks on context and their connection with the concepts of family resemblance and language games (§ 2). I then suggest three kinds of analogies or influences of Wittgenstein's ideas on some aspects of the first strand of AI and on the beginning of information retrieval, which ideally extends to more recent results (§ 3). Eventually, I will discuss his notion of learning and understanding to gain a better grasp on the difference between artificial intelligence and human understanding (§ 4).

---

[4] Shanker insists on the difference between the first works of Turing and the Turing's works after the fourties, and takes a unitarian view of Wittgenstein's remarks, while it seems that Wittgenstein slowly changed his mind from the early thirties to the later years, after the publication of Turing (1936).

[5] We may see traces of his studies in engineering in his continuous reference to technical problems, even in the presentation of classical philosophical theories, like Frege's context principle, as in the following remarks: "If we say: A word only has meaning in the context of a proposition (*satzzusammenhang*), then that means that it's only in a proposition that it functions as a word, and this is no more something that can be said than […] that a cogwheel only functions as such when engaged with other cogs" (PR, 12). Or also see: "A word only has meaning in the context of a proposition (*Satzbervand*): that is like saying only in use is a rod a lever. Only the application makes it into a lever" (PR, 14).

2   Wittgenstein's Contexts

The first occurrences of the term "context" concern Frege's contextual principle presented in the *Grundlagen*: never to ask for the meaning of a word in isolation, but only in the context of a sentence.[6] Eva Picardi (2010) analyses the difference between Frege's and Wittgenstein's context principles. First, Wittgenstein rejects the Fregean idea that sentences can be considered names whose reference (*Bedeutung*) is a truth value. For Wittgenstein, names have reference but not sense, while sentences have sense (truth conditions) but no reference. Second, sentences express states of affairs, and this feature opens a new element in a possible semantic framework (something analogous to what later came under the tag of 'aboutness'). Although this aspect may be rediscovered when discussing the early Frege, as Perry (2019) did, the reference to states of affairs, situations, or circumstances marks a further departure from Frege's framework. Third, after 1930, Wittgenstein takes another step forward, widening Frege's principle: besides speaking of the meaning of a word as dependent on the context of a sentence, he begins to speak of the meaning of a sentence as depending on the context in which it is uttered. Here we may think of the meaning of a sentence, à la Robert Brandom, as inferential potential, or as a set of presuppositions, entitlements, commitments, inferences, or conversational implicatures. All of these elements may find clarification when the sentence, as Wittgenstein suggests, is understood as uttered in the context of a play, a theatrical performance, or a drama (LWI, § 38).

However, as Picardi (2005) remarks, we should be careful not to widen this generalisation of context to a holistic view of meaning, as done by Davidson and Brandom. Wittgenstein was very careful to keep the idea of contextual dependence always delimited to specific language games and on the idea that there are rules constitutive of some concepts (for instance, defining logical constants). Concerning his discussion on contexts, Wittgenstein's thought is not completely at home neither with contextualism nor with holism. I do not

---

6 Frege: "Nach der Bedeutung der Wörter muss im Satzzusammenhange, nicht in ihrer Vereinzelung gefragt warden". Wittgenstein: "Nur im Zusammenhang des Satzes hat ein Name Bedeutung" (TLP, 3.3); "Nur im Satzzusammenhang hat ein Wort Bedeutung" (PR, 12); "Ein Wort hat nur in Satzverband Bedeutung" (PR, 14). Besides the relation between words and sentences, we also have the corresponding relation between objects (the references of names) and states of affairs: "Wenn ich mir den Gegenstand im Verbande des Sachverhalts denken kann, kann ich ihn nicht außerhalb der M ö g i c h k e i t dieses Verbandes denken" (TLP, 2.0121). "If I can imagine objects combined in states of affairs, I cannot imagine them excluded from the possibility of this context." The translation from Pears and McGuinness speaks of "possibility of such combination" and I changed coherently with Wittgenstein's suggestion to Ogden to translate the passage with the term "context".

enter here in the discussion on semantic holism, to which I devoted some space elsewhere.[7] I only remark that there is a problem with the general agreement of the identification of Wittgenstein's ideas with contextualism, as represented by Robyn Carston, Francois Recanati, Charles Travis, and many others who centre their core ideas on the notion of underdetermination of meaning. The problem is that, besides the idea of a deep grammar characterising kinds of language games, Wittgenstein speaks of "ordinary context" and "ordinary sense" as if he accepted the notion of "literal meaning" in contrast with extreme contextualism. Wittgenstein defines "ordinary context" in an interesting discussion of sentences where we have no criteria of application, suggesting that ordinary meaning is linked to standard criteria of application.[8] He also speaks of the use of a word in the context of the language game "which is its original home" (PI, § 116). He thereby suggests the concept of stereotypical meaning or prototype, a term he uses in *The Blue Book* and a subject that Hilary Putnam, Eleanore Rosch, and Marvin Minsky have all developed in various ways.[9] Surely, Wittgenstein has been a source of inspiration to many, but it is difficult to compare his remarks on context with contemporary philosophy of language, just as it is difficult to compare a "sketch of landscapes" (PI, preface) with detailed and alternative maps of the same landscape. Precisely defined concepts like Kaplan's context of utterance, Stalnaker's context set, or Recanati's context-sensitivity are not to be found in Wittgenstein's works, although there are hints towards some of those directions of research (like, for instance, the idea of the meaning of a sentence as an "expansion" (OC, § 349)).

---

[7] Robert Brandom follows Davidson's acceptance of Quine's holism, which was supported in *Word and Object* by *The Blue Book*'s quotation that understanding a sentence is understanding a language. However, Wittgenstein speaks of "a" language, and we may interpret Wittgenstein's view as "understanding a sentence is understanding the language game in which it makes sense" (BB, 5), which amounts to a form of local holism, or weak molecularism, as Michael Dummett insisted. *Global* holism (or even a strong form of molecularism that leads to holism) claims that if two people share a belief p, there is some other belief q, which must also be shared: $\forall p \exists q$ ($q \neq p$ & Nec (p is shared $\rightarrow$ q is shared). On the contrary, local holism would account for something weaker: necessarily, if you share p, there are other beliefs that are also shared. However, there is no privileged set of beliefs that must be shared; it is only necessary that, if p is shared, some not previously determined belief should be shared: $\forall p$ Nec (p is shared $\rightarrow \exists q$ ($q \neq p$ & q is shared). See Perry 1994 and Penco 2001; 2004.

[8] See BB, 10. Besides speaking of "ordinary language" and "ordinary context", the idea of an *ordinary* meaning or ordinary use of a word is a topic that often recurs in his works, especially in BB (27, 36, 52, 53, 62, 66, 140), but also in PI (§§ 258, 344, 351, 418, 536, 615, and PI II, 176, 192) and in RPPI (§§ 52, 99, 126, 358) where we find also the idea of "ordinary language game" (§ 820).

[9] On the difference between stereotype and prototype, see Marconi 1997, 22-8, who discusses the relationships between the ideas by Rosh, Putnam, and Minsky.

However, there is more to say about Wittgenstein's different applications of the term "context", a term that Wittgenstein often uses but never mentions or thematises as such, implicitly suggesting that it might be a perfect case of family resemblance predicate.[10] Beside the Fregean context principle (TLP, 3.3; PR, 12, 14; PI, § 50), he uses the term "context" in different ways: as spatial context (PG, 88; PI, § 539; RC, § 255) or context of perception (RPPI, § 531; PI, II, xi, concerning the duck-rabbit figure), context of conversation or context of speech (PG, 79; Z, § 311; LWI, §§ 118-20; OC, § 349), and eventually context *as circumstance* or the set of circumstances in which a person speaks (PG, 28, 88; PI, §§ 203, 539; RFM, V, § 45; RPPI, § 331; LWI, §§ 253-4; OC, § 662). To understand a sentence, you have to look at the context, intended as the circumstances or situations in which people interact (for instance, as already mentioned, the situation described in a play).

Context as a situation or set of circumstances is a concept that has been developed in different ways in philosophy and computer science as well, from Barwise and Perry's situational semantics to Margaret Masterman's theories of semantic classification for information retrieval and John McCarthy's multi-context theory. I was surprised to verify that there were no explicit connections between situational semantics and multi-context theory, although Barwise, Perry, and McCarthy shared a common ground at the University of Stanford. Often philosophers and computer scientists work on parallel lines and therefore do not immediately converge.[11] In this historical paper, I cannot fill the gap but only show some connections between Wittgenstein's work and the computer scientists' work.

3 Connections Between Wittgenstein's Ideas of Context and AI

We can devise three main lines of the connection between Wittgenstein's ideas of context and early AI: (i) the idea of family resemblance insofar as it derives from the idea of privileged or ordinary contexts on which we rely and from which we may define similar ones by analogy; (ii) the idea of language games insofar as it derives from the idea

---

[10] In a different setting, McCarthy, Buvač (1994, 45) claim that "the term 'context' will appear in useful axioms and other sentences but will not have a definition involving 'if and only if'".

[11] However, even computer scientists do not interact enough. Shoham (1991, 395) wondered about the interest "to examine the extent to which work on situation semantics can be usefully applied in AI". Just a few years before, Terry Winograd (1985) had written a paper on the possible use of situation semantics for the development of more expressive programming languages. There were no connections between the two ideas.

that the meaning of words and sentences depends on particular delimited situations or circumstances in which we may better understand how language works; (iii) the criticism of mathematical logic and formal semantics prompting the first use of Wittgenstein's ideas for information retrieval.

(i) Family resemblance, concepts and contexts

Among the first representatives of artificial intelligence, we find John McCarthy and Marvin Minsky, who, besides apparent differences, worked on the same kinds of problems: how can we treat reasoning when new information comes and compels us to change our premises? How do we treat language processing when our concepts are vague and cannot be defined by necessary and sufficient conditions? A chair is a four-legged piece of furniture, but it may have three legs or even one. Birds fly, but some birds do not: penguins, embalmed birds, and birds with broken wings. Minsky answered with the idea of concepts as frames with default values (a chair has 4 legs or birds fly unless some contradictory situation triggers a change in these values). Minsky (1974) explicitly used Wittgenstein's idea of family resemblances to introduce his idea of frames as stereotyped situations with default information. If everything goes, the standard frame is accepted by default, but we may find unexpected differences and therefore we should have features ('demons') that suggest what to do if the expectations of the frame are not fulfilled (for instance, if a chair does not have four legs). Therefore, the default values of a frame can change depending on context, and frames will be connected with other frames, like an enriched semantic network. Minsky used the Wittgensteinian idea of a "network of overlapping and crisscross resemblances" as a way to explain

> how we can feel as though we know what a chair or a game is—
> yet we cannot always define it in a 'logical' way as an element in
> some class hierarchy or by any other kind of compact, formal, declarative rule. (Minsky 1974, 51)

While Minsky was looking for some alternative with respect of mathematical logic, McCarthy – who could be called the grandfather of the great old-fashioned AI (GOFAI) – tried to make mathematical logic more adaptable to the vagaries of commonsense reasoning. McCarthy (1986) used non-monotonic logic (circumscription) to accept changes in the conclusions when some abnormality enters the set of information (all 'normal' birds fly, but we can change this property in front of an 'abnormal' bird).

A similarity between McCarty's multi-context theory and Wittgenstein's view also concerns the idea of the impossibility of a complete description: if I want to take a flight, I just need to know the timetable and buy the ticket. But if I lose the ticket, I need to describe

the situation with more details on how to recover the ticket, and if the flight is cancelled, I need to enrich my description of the situation with more information. There is no unique context in which I may have an absolute complete description. The completeness of the description depends on the needs of the particular circumstance in which I am in and on underlying assumptions. Analogously, probably referring to the standard problems of Russell's theory of description, Wittgenstein was very well aware that what belongs to a "complete description" will depend on "the purpose of the description" (RFM, VII, § 311). Referring to the idea of Quine's "eternal sentences", which, according to Quine, do not depend on context, McCarthy remarks that they do not exist and that all sentences are dependent on the kind of context chosen. We may only reach a "relative decontextualization", where some common context is explicitly expressed.

(ii) Contexts as situations: language as a motley of language games

We find the most striking similarity with Wittgenstein's ideas in McCarthy's multi-context theory, where contexts are defined as theories with their axioms, their domain, and their rules, creating a vision of language as an unordered series of different local theories that could be considered a good approximation of the idea of different language games, where no sentence can have meaning "out of any context", but can always be considered as having a meaning in particular contexts or particular circumstances (PI, II, ii; OC, §§ 349, 532, 553, 662; RPPI, § 1037).

One of the first computer science exemplifications of a language game is Winograd's SHRDLU, a program for a dialogue with a simulated robot with which to interact in a toy world of boxes, cubes, and pyramids of different forms and colours. The game was very simple: giving orders on how to move those blocks, asking questions, or giving names for new arrangements of those blocks. There is a striking similarity between this toy world and Wittgenstein's builders.[12]

Where are the similarities? Both Wittgenstein and Winograd realised that they could give a good analysis of the workings of language if they considered simplified situations: the knowledge of the toy world permits simple linguistic interactions, similar to the interaction of the builders. Besides, both examples are an expression of the idea of language as a kind of action in context. While Wittgenstein was developing his view of language games as a mixture of language and actions, Winograd (1972) was using Austin's classification of basic speech acts (question, command, assertion) in his

---

12 I was impressed by listening to Winograd's presentation of SCHRDLU in 1972 during a Pisa conference organised by Antonio Zampolli, chief of the laboratory of computational linguistics where I was working at making punched cards. Some years later, I presented the comparison between Winograd's SHRDLU and Wittgenstein's builders with Marcello Frixione at a meeting of EECSE in Camogli (Italy) in 1994.

interaction with the simulated robot, in a view of language "as action rather than structure or the result of a cognitive process" (Winograd 1980, 230). Every speech act triggers a procedure: a command triggers an operation on the blocks, a question the giving of information, an assertion the storing of information. The meaning of a word is represented as a procedure that permits finding the referent of the expressions in the sentence (in our case, the objects are simple and composed blocks). The basic idea underlying Wittgenstein's and Winograd's views of language was similar: simplified models of language interaction suggest the idea of language as a heterogeneous set of diverse toy worlds or various language games or, in McCarthy's terminology, of various contexts. The difference concerns the aims: computer scientists aim to provide a good representation of natural language processing (NLP) for creating working programs, while Wittgenstein's philosophical aim is to remove specific misunderstandings, and even his criticisms towards Turing concern not his mathematical theory, but the possible psychologistic interpretation of the idea of thinking machines.

The perspective of multi-context theories maps even more strictly the attitude of Wittgenstein towards language as a mixture of language games where there is no meta-language game over the others – Wittgenstein speaks of mathematics as "a motley of techniques of proofs" (RFM, III, § 46). Analogously, for McCarthy, there is no universal context but just different contexts with their own rules. In multi-context theory, we find something new with respect to the role of contexts in describing the workings of language: not only words and sentences, but also rules depend on the context we are considering. For McCarthy and Buvač (1994), some rules may be common to different contexts, permitting us to let what has been derived in one context enter another context. In this way, McCarthy introduced a new problem: how to individuate the relations among contexts, and this problem maps Wittgenstein's idea of intermediate members of different language games (see Penco 2004; 2007), or his requirement to look for a wider context to change interpretation of what is said (PG, 88; PI, §§ 539, 686; RPPI, § 1066).

The effort of McCarthy was to show which rules may govern relations among contexts, how to find whether contexts are compatible with one another or not, how the conclusions reached inside a context may be valid in another context or not, how we can change a context either by enriching it or simplifying it, or by making parameters explicit or leaving them implicit, how we can leave a context to enter another with different rules, and so on.[13] What emerges from these

---

[13] On these kinds of rules see Benerecetti, Bouquet, Ghidini 2000, quoted by Guha, McCarthy 2003.

early works is a representation of language as a non-ordered mixture of different contexts, or, if you like, of different toy worlds or different language games related one another by different rules without a general universal structure underlying all of them.

(iii) Criticism of mathematical logic and information retrieval

Margaret Masterman (MM) was one of the five students to whom Wittgenstein dictated *The Blue Book* in 1933-34.[14] Later, in 1955, she founded the Cambridge Language Research Unit (CLRU), where she promoted a view of natural language processing against both the semantic analysis of Richard Montague and the syntactic analysis of Noam Chomsky. Her view was relying on the use of thesauri to find patterns or underlying structures of semantic relations. For MM (2005, 109), a thesaurus is a "language system classified as a set of contexts", where contexts are sentences in which a word appears, following the first traditional definition of meaning as the use of a word in a sentence, but also defined by semantic clusters. A few aspects of her research reveal the influence of Wittgenstein:

– MM gave great relevance to the concept of context as a kind of situation: she claims that, notwithstanding differences in language and culture, we "can share a common stock of extra-linguistic contexts" (2005, 127), which is a common stock of kinds of situations. As example of lack of shared extra-linguistic contexts she presents the comparison between the forms of life of humans and ants.[15]

– MM maintained a distance towards formal logic as an analysis of natural language:

> [F]ormal logic as we at present have it is not and cannot be directly relevant to the contextually based study of semantic pattern. Logic is the study of relations, and, in particular, it is the study of derivability. (2005, 261)

Semantic patterns are not related to mathematical logic, but are derived from the contexts in which a word appears:

> [T]he use of a word is its whole field of meaning, its total 'spread'. Its usages, or main meanings in its most frequently found contexts, together make up its Use. (2005, 126)

---

[14] On the relation between Margaret Masterman and Wittgenstein see Wilks 2005; 2007 and Liu 2021. Among many others, Margaret Masterman deeply influenced Yorik Wilks, Margaret Boden and Kwame Anthony Appiah.

[15] Humans sleep and dream, while the latter do not, and this makes it impossible to share a common stock of situations. It is curios that Wittgenstein used an example about ants in showing the difference between humans' natural history and other species (RPPII, §§ 22-4).
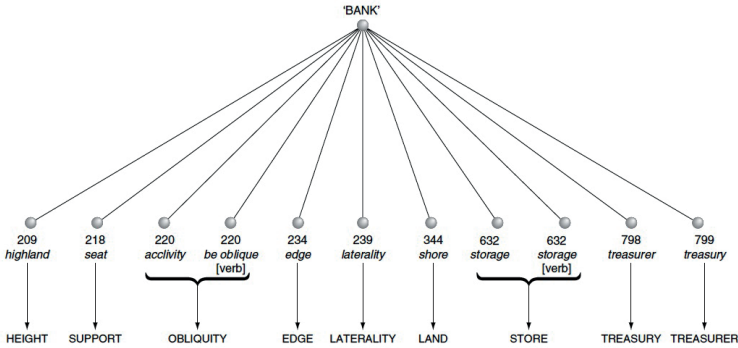
– Following Wittgenstein's example of the ambiguity of the term "bank" (BB, 11-13), MM presented the term "bank" together with the different contexts or "subjects-which-it-is-most-used-to-talk-about" (2005, 289), speaking of those different subjects as "quasi-Wittgensteinian families", suggesting a comparison with Wittgenstein's idea of family resemblance predicates **[fig. 1]**.

– Beyond the analysis of the same words used to refer to different concepts (like "bank") or to be embedded in different patterns, MM discusses the problem of different expressions used to refer to the same situation. The basic idea is that there is no proper complete synonymy – if not, it would be useless – but what different people say about the same kinds of situations largely overlaps and may form clusters of meaning and also clusters of overlapping contexts (2005, 69).

Masterman's strategy to find similarities, clusters, and patterns in a thesaurus will become central to the CRLU information retrieval system. Works in the tradition of Margaret Masterman used strategies of letting patterns or clusters of words emerge from unsupervised statistical methods on a large data set, following Wittgenstein's idea of family resemblance.[16] Those works, as Halpin (2011) remarks, were at the source of the first Web search engines. Referring to the

---

16   See for instance Karen Spärck Jones (1986, 64), who quotes Wittgenstein's method of similarities and differences concerning her view of treating synonymy. She relied on Roger Needham's statistical theory of clumps, a particular explication of the idea of family resemblances, for which words can be defined in terms of statistical clumps of other words. See also Wilks 2008, who connects Wittgenstein's ideas to his preferential semantics, Shank's conceptual dependency, or Fillmore's case grammar.

contrast between the use theory and the causal theory of meaning, Halpin (2011, 18) remarks that "search engines like Google embody an alternative theory of meaning, one based on an objective notion of sense implicitly given by Wittgenstein". In a way, the inventors of Google, Sergey Brin and Larry Page, still working with his tutor Terry Winograd,[17] developed a search engine whose basic ideas come from a long history, ideally reconnecting them to Wittgenstein's influence. Margaret Masterman, missing the technical means we have today, was probably very much in advance of her times and anticipated some of the most recent trends in extracting patterns from learning algorithms. Wilks (2008) and Molino and Tagliabue (2023), referring to Masterman's work, give more detailed comments on the development of machine learning algorithms following this kind of trend up to the more recent LLMs.

4 Understanding and Learning Language in Context

Having presented the affinity of Wittgenstein's ideas of context dependence with some features of artificial intelligence projects, we may come back to the general question: do machines think? Following Tarski,[18] we might say: "Of course they can, it only depends on what you mean by 'think'". If we equate thinking as "operating with signs" and understanding language with "to be master of a technique" (PI, § 199), it would be very difficult to avoid the conclusion that contemporary chatbots or advanced robots master a technique of language use and therefore think. The problem is that the technique of "thinking machines" is different from the technique used by humans, just as the technique of flying aeroplanes is very different from the technique of flying used by birds. Therefore, assuming that understanding a language means mastering a technique and that machines 'understand language' (in the limited sense that they pass the Turing test), where does the difference lie between human understanding and machine understanding?

AI based on deep learning algorithms is very different from the workings of the human mind, as strongly remarked by Chomsky, Roberts and Watumull (2023). Chomsky's point concerns not only the concept of understanding but also the concept of learning. Learning is a family resemblance concept, and Wittgenstein gave examples of different ways we use this concept:[19] learning a language (PI, § 7),

---

17 Winograd himself contributed to the presentation of the page-rank algorithm. See Page et al. 1998.

18 Quoted in Obermeier 1983, 347.

19 See Williams 1999 and Vazquez Hernandez 2020.

learning the numerals by heart and learning to use them (PI, § 9), learning an unknown language from ostensive definitions (PI, § 32), and different ways of learning the meaning of a word, like the difference between "to point to this thing" and "to point to the colour, not to the shape" (PI, § 35), or the different examples and different language games with which we learn the word "good" (PI, § 77), and so on.[20] One may therefore think that we may see the technique of machine learning that governs most LLMs as a new kind of learning procedure, a realisation of the idea of Margaret Masterman of extracting semantic patterns from data. Learning algorithms, after training, are let loose to autonomously find further patterns inside texts. Could this be compared to human learning? Certainly it can, but the difference is so huge that our concepts of thinking and understanding that depend on the way we learn language are stretched *near* a breaking point. Let us see the differences.

The first difference is quantity: humans learn to speak with a very small amount of linguistic data; this means that they learn language in an extralinguistic context, and rely on some kind of innate system (brought about by evolution) to master the technique of linguistic interaction. Learning algorithms learn from a huge amount of data, both visual patterns and linguistic content taken from many already organised data sets, and therefore produce a second order intelligence grounded on our examples of natural language, which brings us to the second difference.

The second difference is quality: AI dialogue systems take their content from big data sets, of technical or literary content. Therefore they begin with the highest result of our civilisation (included programming languages or social networks), including biases depending on the data sets used. On the contrary we begin to learn with more emotional and basic stuff, mostly linked to perception, emotion, and physical interaction. Our body, emotions, and the context of perception, have an essential part in learning language, and only after many years we humans begin to reach the abstract concepts and connections that are given to the machines since the outset.

The third difference is classification: the clusters and patterns arrived at by machine learning algorithms are not necessarily similar to the way we classify concepts and situations. Logicians like Frege began to find patterns in sentences that permit us to better understand our logical reasoning, but the patterns individuated by new learning

---

20   I just used some examples from the first pages of the *Philosophical Investigations*, but the discussion on different aspects of 'learning' goes through all his published books, where the word 'learn' in different ways is reported in almost 400 quotations. The reference to the word 'good' is reminiscent of Wittgenstein's reaction to Moore's lessons on the term 'good', where he found one of the first examples of family resemblance concepts (see Vaccarezza, Penco 2023).

algorithms are more linked to clusters identified in a way even programmers cannot know. Although those algorithms may have some procedure similar to ours, the patterns they use are such that we cannot grasp them in the same way we grasp our concepts.

But what does it mean 'to grasp' a concept or a thought? A semantic theory may claim that mastering a thought is mastering its inferential potential, the commitments undertaken in asserting the corresponding sentence. When we speak of commitments, we introduce an ethical issue and not only a cognitive issue. The thought experiments of a drone that kills the operator in order to perform its task[21] introduces a new vision of what it means to grasp a thought: you need to master the context and the complex circumstances in which you are, and AI lacks this kind of complexity. It is a complexity that cannot be reduced to a technique. We have been impressed by the idea that understanding a language means mastering a technique. However, besides different kinds of techniques for mastering a language, Wittgenstein suggests that there is also a kind of learning that is not given by any technique.

Wittgenstein speaks of the capacity to give 'expert' judgements about the genuineness of expressions of feeling. AI experts have invented programs that detect shifts in speech rate, pitch, volume, or microtremors undetectable by humans to verify the 'real' feeling or honesty of a person. Who knows? Maybe it is even simpler than NLP. But Wittgenstein speaks of learning 'by experience' as a kind of understanding that only experienced people can teach. There are rules linked to this experience, but they are different from 'calculating rules'. This last remark seems to make a distinction among different kinds of understanding, one linked to a technique, the other linked to context understanding that no technique can offer (PI, II, xi, 227; LWI, §§ 917-27).

5 Summary and Conclusion

At the end of the day, Wittgenstein was more open to the idea of thinking machines than it is often claimed, and the compatibilists have some reasons to see Wittgenstein's work as a forerunner of artificial intelligence, starting with his influence on Turing and Masterman and the similarity of his views and those of the early AI experts. The point of Wittgenstein's remarks on contexts is the search for differences: different kinds of learning show that there are and will be different ways of thinking and understanding linked to the different kinds of contexts in which we are in. There is enough space for

---

21   See the discussion in Davis, Squire 2023.

different contexts and different language games to contain both human and artificial thinking and their interactions. However, a warning remains central: what counts is awareness of the difference between AI ways of reasoning and human ways of reasoning. The latter may contain some specific capacities, resulting from different learning procedures and the expertise that arises from them.[22]

References

Appiah, K.A. (2008). "The Epiphany Philosophers". *New York Times Magazine*, September 21.

Barwise, J.; Perry, J. (1983). *Situations and Attitudes*. Cambridge, MA.; London: MIT Press.

Benerecetti, M.; Bouquet, P.; Ghidini, C. (2001). "On the Dimensions of Context Dependence: Partiality, Approximation, and Perspective". *LNAI*, 2116, 59-72. https://doi.org/10.1007/3-540-44607-9_5.

Beran, O. (2014). "Wittgensteinian Perspectives on the Turing Test". *Studia Philosophica Estonica*, 35-57. https://doi.org/10.12697/spe.2014.7.1.02.

Brandom, R. (1994). *Making It Explicit*. Cambridge, MA.: Harvard University Press.

Casey, G. (1988). "AI and Wittgenstein". *Philosophical Studies*, 32, 156-75. https://doi.org/10.5840/philstudies19883239.

Chomsky, N.; Roberts, I.; Watumull, J. (2023). "The False Promise of ChatGPT". *The New York Times*, March 8.

Davis, C.R.; Squire, P. (2023). "Air Force Colonel Backtracks Over His Warning About How AI Could Go Rogue and Kill Its Human Operators". *Business Insider*. businessinsider.com/ai-powered-drone-tried-killing-its-operator-in-military-simulation-2023-6.

Dreyfus, H.L. (1972). *What Computers Can't Do*. New York: Harper & Row.

Dreyfus, H.L. (1992). *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge, MA.; London: MIT Press.

Floyd, J. (2019). "Wittgenstein and Turing". Mras, G.M.; Weingartner, P.; Ritter, B. (eds), *Philosophy of Logic and Mathematics = Proceedings of the 41st International Ludwig Wittgenstein Symposium*. Berlin; Boston: De Gruyter, 263-96. https://doi.org/10.1515/9783110657883-016.

Fuchs, T. (2022). "Understanding Sophia? On Human Interaction with Artificial Agents". *Phenomenology and the Cognitive Sciences*, 23(2), 1-22. https://doi.org/10.1007/s11097-022-09848-0.

Gaver, M. (2023). "From Wittgenstein's Language Games to Large Language Models". YouTube video on The Times of AI channel.

Gomes, C.; Selman, B. (2023). "Exploring Connections – Wittgenstein, Gödel, and ChatGPT". YouTube video on Vienna Center For Logic and Algorithm channel.

Guha, R.; McCarthy, J. (2003). "Varieties of Contexts". *LNAI*, 2680, 164-77. https://doi.org/10.1007/3-540-44958-2_14.

Halpin, H. (2011). "Sense and Reference on the Web". *Minds and Machines*, 21, 153-78. `https://doi.org/10.1007/s11023-011-9230-6`.

Harre, R. (1988). "Wittgenstein and Artificial Intelligence". *Philosophical Psychology*, 1(1), 105-15. `https://doi.org/10.1080/09515088808572928`.

Liu, L.H. (2021). "Wittgenstein in the Machine". *Critical Inquiry*, 47(3), 425-55. `https://doi.org/10.1086/713551`.

Marconi, D. (1997). *Lexical Competence*. Cambridge, MA; London: MIT Press.

Masterman, M. (2005). *Language, Cohesion and Form: Selected Papers*. Edited by Y. Wilks. Cambridge: Cambridge University Press. `https://doi.org/10.1017/CBO9780511486609`.

McCarthy, J. (1986). "Applications of Circumscription to Formalizing Common-sense Knowledge". *Artificial intelligence*, 28(1), 89-116. `https://doi.org/10.1016/0004-3702(86)90032-9`.

McCarthy, J. (1990). *Formalizing Common Sense*. Edited by V. Lifschitz. Norwood, NJ: Ablex Publishing Corporation.

McCarthy, J.; Buvač, S. (1994). "Formalizing Context (Expanded Notes)". *Technical Report STAN-CS-TN-94-13*. `http://i.stanford.edu/pub/cstr/reports/cs/tn/94/13/CS-TN-94-13.pdf`.

Minsky, M. (1974). "A Framework for Representing Knowledge". *MIT-AI Laboratory Memo 306*.

Molino, P.; Tagliabue, J. (2023). "Wittgenstein's Influence on Artificial Intelligence". `https://doi.org/10.48550/arXiv.2302.01570`.

Neumayer, O. (1986). "A Wittgensteinian View of Artificial Intelligence". Born, R. (ed.), *Artificial Intelligence. The Case Against*. London: Croom Helm, 132-73. `https://doi.org/10.4324/9781351141529-6`.

Newell, A.; Simon, H.A. (1976). "Computer Science as Empirical Inquiry: Symbols and Search". *Communications of the ACM*, 19(3), 113-26. `https://doi.org/10.1145/360018.360022`.

Obermeier, K.K. (1983). "Wittgenstein on Language and AI: The Chinese-room Thought Experiment Revisited". *Synthese*, 56(3), 339-49. `https://doi.org/10.1007/BF00485470`.

Page, L. et al. (1998). "The PageRank Citation Ranking: Bringing Order to the Web". *Technical Report, Stanford Digital Libraries SIDL-WP-1999-0120*, 161-72.

Penco, C. (2001). "Local Holism". *LNAI*, 2116, 290-303. `https://doi.org/10.1007/3-540-44607-9_22`.

Penco, C. (2004). "Wittgenstein, Locality and Rules". Coliva, A.; Picardi, E. (eds), *Wittgenstein Today*. Padova: Il Poligrafo, 249-74.

Penco, C. (2007). "Idiolect and Context". Auxier, R.E.; Hahn, L.E. (eds), *The Philosophy of Michael Dummett. Library of Living Philosophers*, vol. 31. Chicago and La Salle (IL), Open Court, 567-90.

Penco, C. (2012) "Updating the Turing Test. Wittgenstein, Turing and Symbol Manipulation", *Open Journal of Philosophy*, 2, 189-94. `https://doi.org/10.4236/ojpp.2012.23029`.

Perry, J. (1994). "Fodor and Lepore on Holism". *Philosophical Studies*, 73, 123-38. `https://doi.org/10.1007/BF01207661`.

Perry, J. (2019). *Frege's Detour*. Oxford: Oxford University Press. `https://doi.org/10.1093/oso/9780198812821.001.0001`.

Picardi, E. (2005). "Was Frege a Proto-Inferentialist?". Reprinted in Picardi 2022. `https://doi.org/10.1093/oso/9780198862796.003.0014`.

Picardi, E. (2010). "Wittgenstein and Frege on Proper Names and the Context Principle". Reprinted in Picardi 2022. https://doi.org/10.1524/dzph.2009.0051.

Picardi, E. (2022). *Frege on Language, Logic and Psychology: Selected Essays*. Edited by A. Coliva. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198862796.001.0001.

Shanker, S. (1998). *Wittgenstein's Remarks on the Foundations of AI*. London; New York: Routledge.

Shoham, Y. (1991). "Varieties of Context". Lifschitz, V. (ed.), *Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy*. New York: Academy Press, 393-408. https://doi.org/10.1016/B978-0-12-450010-5.50028-1.

Spärck Jones, K. (1986). *Synonymy and Semantic Classification*. Edinburgh: Edinburgh University Press.

Sunday Grève, S. (2023). "Artificial Forms of Life". *Philosophies*, 8(5), 89. https://doi.org/10.3390/philosophies8050089.

Turing, A. (1936-7). "On Computable Numbers, with an Application to the Entscheidungsproblem". *Proceedings of the London Mathematical Society*, 42(2), 230-65. Reprinted in Davis, M. (ed.) (1965), *The Undecidable: Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions*. Hewlett, NY: Dover Publication, 116-54.

Vaccarezza, S.; Penco, C. (2023). "Wittgenstein's Non-Non-Cognitivism". Dreon et al. (a cura di), *Senza Trampoli. Saggi filosofici per Luigi Perissinotto*. Milano; Udine: Mimesis.

Vázquez Hernández, A. (2020). *Wittgenstein and the Concept of Learning in Artificial Intelligence* [MA Dissertation]. Bergen: University of Bergen.

Williams, M. (1999). "The Philosophical Significance of Learning in Wittgenstein's Later Philosophy". Williams, M., *Wittgenstein, Mind and Meaning: Toward a Social Conception of Mind*. London; New York: Routledge, 188-215.

Wilks, Y. (2005). "Editor's Introduction". Masterman 2005, 1-18. https://doi.org/10.1002/9780470776599.ch.

Wilks, Y. (2007). "A Wittgensteinian Computational Linguistics?". *Epiphany Philosophers Blog*. http://epiphanyphilosophers.org/wp-content/uploads/2017/01/wittgenstein.pdf.

Wilks, Y. (2008). "What Would a Wittgensteinian Computational Linguistics Be Like?". *Proceedings of Convention of the Society for the Study of AI and Simulation of Behavior. Workshop on Computers and Philosophy*, 1st-4th April 2008, University of Aberdeen.

Winograd, T. (1972). "Understanding Natural Language". *Cognitive Psychology*, 3(1), 1-191. https://doi.org/10.1016/0010-0285(72)90002-3.

Winograd, T. (1980). "What Does it Mean to Understand Language?". *Cognitive Science*, 4(3), 209-41. https://doi.org/10.1207/s15516709cog0403_1.

Winograd, T. (1985). "Moving the Semantic Fulcrum". *Linguistics and Philosophy*, 91-104. https://doi.org/10.1007/BF00653378.

Wittgenstein, L. (1958). *Philosophical Investigations*. 2nd edition. Edited and transl. by G.E.M. Anscombe. Oxford: Basil Blackwell. [PI]

Wittgenstein, L. (1969). *The Blue and The Brown Books*. 2nd edition. Oxford: Blackwell. [BB]

Wittgenstein, L. (1974a). *On Certainty*. Reprinted with corrections and indices. Edited by G.E.M. Anscombe; G.H. von Wright, transl. by D. Paul; G.E.M. Anscombe. Oxford: Blackwell. [OC]

Wittgenstein, L. (1974b). *Philosophical Grammar*. Edited by R. Rhees, transl. by A. Kenny. Oxford: Blackwell. [PG]

Wittgenstein, L. (1975). *Philosophical Remarks*. Edited from his posthumous writings by R. Rhees, transl. by R. Hargreaves; R. White. Oxford: Blackwell. [PR]

Wittgenstein, L. (1978a). *Remarks on Colour*. Edited by G.E.M. Anscombe, transl. by L.L. McAlister; M. Schättle. Oxford: Blackwell. [RC]

Wittgenstein, L. (1978b). *Remarks on the Foundations of Mathematics*. 3rd edition. Edited by G.H. von Wright; R. Rhees; G.E.M. Anscombe, transl. by G.E.M. Anscombe. Oxford: Blackwell. [RFM]

Wittgenstein, L. (1980a). *Remarks on the Philosophy of Psychology*, vol. 1. Edited by G.E.M. Anscombe; G.H. von Wright, transl. by G.E.M. Anscombe. Oxford: Blackwell. [RPPI]

Wittgenstein, L. (1980b). *Remarks on the Philosophy of Psychology*, vol. 2. Edited by G.H. von Wrightf; H. Nyman, transl. by C.G. Luckhardt; M.A.E. Aue. Oxford: Blackwell. [RPPII]

Wittgenstein, L. (1981). *Zettel*. 2nd edition. Edited by G.E.M. Anscombe; G.H. von Wright, transl. by G.E.M. Anscombe. Oxford: Blackwell. [Z]

Wittgenstein, L. (1982). *Last Writings on the Philosophy of Psychology*, vol. 1. Edited by G.H. von Wright; H. Nyman, transl. by C.G. Luckhardt; M.A.E. Aue. Oxford: Blackwell. [LWI]

Xu, Y. (2016). "Does Wittgenstein Actually Undermine the Foundation of AI?". *Frontiers of Philosophy in China*, 11(1), 3-20.