# Wittgenstein's Remarks on Technology and Mental Mechanisms

Thomas Raleigh, Ruhr University Bochum

**ABSTRACT**: This article provides a survey of Wittgenstein's remarks in which he discusses various kinds of technology. I argue that throughout his career, his use of technological examples displays a thematic unity: technologies are invoked in order to illustrate a certain *mechanical* conception of the mind. I trace how his use of such examples evolved as his views on the mind and on meaning changed. I also discuss an important and somewhat radical anti-mechanistic strain in his later thought and suggest that Wittgenstein's attitude to mechanistic explanations in psychology was ultimately quite ambivalent.

## 1. INTRODUCTION

It is well known that Wittgenstein had a more than merely amateur interest in technology. He was originally an aeronautical engineering student in Manchester, where he worked on designing a jet engine and a propeller, before becoming interested in the foundations of mathematics and thence philosophy. He is reported to have built a working model of a sewing machine, constructed from wood and wire, at the age of just ten[1]. And during the second World War, whilst working first at Guy's Hospital in London and then at the Royal Victoria Infirmary in Newcastle, Wittgenstein used his technical skills to produce zinc oxide ointment ('Lassar's Paste') and paraffin histology sections – he even invented a new device that improved the way that pulse pressure was recorded in experiments on rats[2]. But his attitude to technology, like his attitude to the influence of science on 20th century culture in general, was by no means uniformly positive. Consider, for example, the wonderfully lofty pronouncement he once made in conversation with David Pinsent: "Music came to a full stop with Brahms. And even in Brahms one can begin to hear the sound of machinery."[3]

Scattered throughout his philosophical writings, both early and late, there are various references to technologies and machines of various kinds. The aim for the present paper will be to chart what I take to be the central, recurring theme or purpose of these technological examples – which is to illustrate a certain conception or model of the mind as a kind of mechanism for encoding and processing information by manipulating symbols. The idea that the brain stores and processes information by performing operations on representational elements of some kind, is these days utterly familiar and orthodox – though it is certainly not universally accepted[4]. Of course Wittgenstein was writing in an era before electronic computers, but the model of the mind which he illustrates by appeal to such technologies as the Jacquard Loom, a pianola, or a vinyl gramophone record, is, I shall argue, fundamentally the same model that we would nowadays label 'The Language of Thought' hypothesis, and which would standardly be explained by comparison with a computer running a programme written in some operating language or code.

---

[1] See McGuiness (1988, p45), Monk (1991, p13).

[2] See Henderson (1973), Teive, Silva & Munhoz (2011).

[3] The remark is reported in (ed.) Rhees (1984).

[4] For some recent, dissenting views see Hutto & Myin (2017), Chemero (2009), Ramsey (2007).

As we will see, Wittgenstein used technological examples and analogies to explore this conception of the mind both in his earlier work (**section 2**, below) and in his later work. But very soon after returning to philosophy in the early 1930's, he came to realise that this sort of model or conception of the mind cannot, by itself, answer certain fundamental philosophical questions about the nature of meaning, or the grounds of intentionality (**section 3**). However, that a mechanistic or computational conception of the mind in itself does not answer the sorts of problems about meaning that Wittgenstein famously raised in the 'rule-following remarks', does not show that such a model is totally worthless or false. Indeed I will argue in **section 4** that even in his later period Wittgenstein continued to feel *some* measure of attraction to this conception of the mind – though in **section 5** I will turn to discussing a number of his remarks which seem to be more hostile to this form of explanation. So then as well as illustrating how Wittgenstein's remarks about technology are centrally concerned with examining a mechanistic conception of the mind, and sketching how his views on this topic evolved, a central aim of this paper will be to show that Wittgenstein's later writings evince an important degree of *ambivalence* towards the mechanical model of the mind and towards mechanistic explanations of psychological phenomena. I trust that exploring the reasons for Wittgenstein's ambivalent attitude here may be of wider philosophical interest and not just an exegetical exercise. After all, this mechanistic/computational model remains a central guiding theoretical assumption for much contemporary cognitive science and much current philosophy of mind.

## 2. THE PICTURE THEORY & REPRESENTATIONAL MEDIA

The only place in the *Tractatus* where Wittgenstein uses a technological example is around section 4.014, where he discusses the relation between a gramophone record, a written musical score, sound waves and the *musical thoughts* of the composer:

> 4.014 A gramophone record, the musical thought, the written notes, and the sound-waves, all stand to one another in the same internal relation of depicting that holds between language and the world.
> They are all constructed according to a common logical pattern.
> (Like the two youths in the fairy-tale, their two horses, and their lilies. They are all in a certain sense one.)

> 4.0141 There is a general rule by means of which the musician can obtain the symphony from the score, and which makes it possible to derive the symphony from the groove on the gramophone record, and, using the first rule, to derive the score again. That is what constitutes the inner similarity between these things which seem to be constructed in such entirely different ways. And that rule is the law of projection which projects the symphony into the language of musical notation. It is the rule for translating this language into the language of gramophone records. (Wittgenstein, 1922, 23-24)

These remarks flow from his general 'picture theory' of meaning, according to which a thought or representation gets to be about some state of affairs – i.e. have a propositional content – in virtue of sharing a common combinational structure. The elements in the representation are combined in a certain manner that is isomorphic to the manner in which target elements in the represented state of affairs are combined. (And moreover, the permitted range of combinatorial possibilities of the representing medium correspond to all the range of possible relations that the elements of the represented scene can stand in to one another.)

Consider, for example, a vinyl gramophone record of Brahms's 3<sup>rd</sup> String Quartet[5], a musical score of the quartet, and the sound waves produced in a performance of the quartet; clearly these must all share some kind of structural isomorphism (or homomorphism), and so there are 'rules of projection' that would allow us (in theory) to derive (at least to close approximation) any one of these from any of the others. Just as there is a rule of projection that allows musicians to derive the correct notes to play from a written score, so we could also (in theory) derive the correct notes to play from the grooves on a vinyl record. Or, conversely, we could (in theory) derive what pattern of grooves in the vinyl would be needed to make a record of the piece from the written score. Likewise, for Wittgenstein at this time, the musical thoughts of the composer – there must be some rule of projection from the thoughts to the written score and in turn to the record and the sound waves (and vice-versa)[6]. Wittgenstein was thus effectively thinking of the rules of combination for the elements within a representational medium as a kind of *syntax*, and of the rules of projection as part of its *semantics*, which take us from the representational medium to some other structurally isomorphic arrangement of elements.

At this stage, Wittgenstein seems to have thought that it was a matter for psychology to determine what exactly the elements are, which supposedly get combined in thought. His job as a philosopher was just to point out that there must *be* some such combinatorial elements. Likewise his job as a philosopher was to point out that reality, all facts, must ultimately be composed of ontologically independent simple elements. It was then up to someone else – maybe the physicists, maybe the phenomenologists – to figure out what these elements are exactly. And likewise too for the rules of projection, i.e. the specific relations, that the psychical elements that compose a thought bear to the states of affairs that they represent. In this Tractarian period, Wittgenstein presumably thought that the specific nature of these rules of projection – i.e. mind-world representational relations – was also an empirical matter for psychology to investigate. The philosophical issues are settled once he has established merely that there must be some such relations. For example in the Notebooks from 1914-16, he records the following remark from a conversation with Russell:

> Does a *Gedanke* [thought] consist of words?' No! But of psychical constituents that have the same sort of relation to reality as words. What those constituents are I don't know. (Wittgenstein, 1961, 156–7)

We have here an explicit statement of how close the Tractarian model of thought was to what philosophers would these days call the "Language of Thought" hypothesis[7]. Notice also that already in this early period, Wittgenstein was making a comparison between a piece of information storing technology (a vinyl gramophone record), some symbolic notation governed by syntactic rules (the musical score), and a thought (the composer's musical idea) in order to illustrate and explore this model of the mind.

## 3. MECHANISMS AND MEANING

Wittgenstein famously came to abandon the 'picture theory' of meaning. And one of the crucial changes that results in Wittgenstein's philosophy as we move from the 'early' to the 'later' periods is his realisation that the relations (or 'rules of projection') between sentences or thoughts and the worldly states of affairs that they are about is not simply a matter of empirical investigation for psychology to discover. Rather, Wittgenstein came to see that

---

[5] George Steiner claimed that Wittgenstein once told Norman Malcolm that the slow movement of this quartet had twice led him back from the brink of suicide (Steiner, 1996, 125).

[6] Section 4.002 of the *Tractatus* makes clear that these rules of projection need not be *known* by the subject, the subject simply *embodies* these combinatorial elements of thought.

[7] See e.g. Harman 1973, Fodor 1975, Davies 1991, Rey 1995, Schneider 2011.

understanding how it is determined *which* of indefinitely many possible relations, or rules of projections, is the *correct* one, the one which accords with the meaning or content of the sentence or thought, is perhaps the fundamental philosophical problem concerning meaning and representation – i.e. intentionality in the philosophical sense that we inherit from Brentano (1874).

Consider the following passages from the *Philosophical Grammar*, which not only illustrate this deep change in Wittgenstein's views about thought and content, but also show how he continued to have recourse to mechanical or technological examples as his preferred comparisons when thinking through these issues.

> "The sentences that we utter have a particular purpose, they are to produce certain effects. They are parts of a mechanism, perhaps a psychological mechanism, and the words of the sentences are also parts of the mechanism (levers, cogwheels, and so on.) The example that seems to illustrate what we're thinking of here is an automatic music player, a pianola. It contains a roll, rollers, etc., on which the piece of music is written in some kind of notation (the position of holes, pegs, and so on). It's as if these written signs gave orders which are carried out by the keys and hammer. And so shouldn't we say that the sense of the sign is in its effect? But suppose the pianola is in bad condition and the signs on the roll produce hisses and bangs instead of the notes. – Perhaps you will say that the sense of the sign is their effect on a mechanism in good condition, and correspondingly that the sense of an order is its effect on an obedient man. But what is regarded as a criterion of obedience here?" (Wittgenstein, 1974, 69-70)

At the start of this passage, Wittgenstein gives voice to the idea that the sentences of natural language operate as the input to a mental mechanism, which then processes this input according to some fixed/mechanical procedure (i.e. an algorithm). As in the Tractatus passage quoted in the previous section, Wittgenstein provides an example of musical technology to illustrate this idea – though here the specific technology in question is the pianola. The input is a roll of paper with punched holes, which the mechanism operates upon in order to produce the movements of the hammers and keys and so in turn to produce the intended piece of music. The rigid, mechanical nature of the player piano can make it seem obvious and unproblematic what the 'rules of projection' are that relate the punched holes in the piano roll to the internal movements of the machine or to the notes of the original musical score. But as Wittgenstein points out, this assumes that the pianola is in good enough condition that it is working *as intended* – i.e. this takes for granted that there is a *correct* way for the pianola to respond to the pattern of punched holes. Likewise then we cannot explain the intentionality of natural language just by charting the effects of hearing a sentence on a person's internal psychological mechanisms – for this would require us to be able to distinguish the correct operation of such mechanisms from their incorrect working. And that would just be to already take the meaning of the sentence in question for granted.

> 'How does thought work, how does it use its expression?' This question looks like 'How does a Jacquard loom work, how does it use the cards?'
> In the proposition 'I believe that p is the case' we feel that the essential thing, the real process of belief, isn't expressed but only hinted at; we feel it must be possible to replace this hint by a description of the mechanism of belief, a description in which the series of words 'p' would occur in the description as the cards occur in the description of the loom. This description, we feel, would at last be the full expression of the thought.
> (ibid., 102)

Wittgenstein here shows that he understood and wanted to scrutinise the strong allure of the idea that the natural language expression of a thought is a translation or projection from a sentence in the 'language of thought'. And that if we could adequately investigate the internal operations of the relevant psychological mechanism, we would find the real or fundamental

form/syntax of the thought in question. Just as the pattern of textile that a Jacquard loom[8] produces only indirectly shows us its internal operations, whereas knowing the series of punched cards that it operated on gives us a more direct and complete understanding of the technology.

> We ask: 'What is a thought? What kind of thing must something be to perform the function of thought? This question is like: 'What is a sewing machine, how does it work? – And the answer which would be like ours would be 'Look at the stitch it is meant to sew; you can see from that what is essential in the machine, everything else is optional. So what is the function, that makes thought what it is? –If it is its *effect*, then we are not interested in it. We are not in the realm of causal explanations, and every such explanation sounds trivial for our purposes." (ibid., 105)

Here again Wittgenstein recognises the strong temptation to think that the mysterious *intentionality* of thought may become better understood by investigating the causal-psychological mechanisms that are involved in, say, giving the correct verbal response to a question. This would be to treat a thought as an internal mechanism that produces behaviour as a sewing machine produces a stitch. But Wittgenstein warns us that such causal understanding will not shed light on the mysterious *aboutness* that thoughts have. To repeat: understanding the causal processes that are in fact operative in people when they respond to a question cannot distinguish the *correct* responses from the *incorrect* unless we are already taking for granted the *meaning* (intentionality) of these responses[9].

The example of the pianola would later be re-used in both the *Brown Book* (§66, p118) and also in the *Investigations* (§157, discussed in section 5, below). Both a pianola and a Jacquard loom take as their 'inputs' a pattern of punched holes in paper/card, which interacts with their internal mechanisms so as to produce a specific output: a pattern of struck notes on the pianola, a pattern of brocaded textile from the Jacquard loom. The similarity to the (imaginary) roll of tape, divided into cells, which forms the input to a Turing Machine is obvious. (We will return to the notion of a Turing Machine in section 4, below.)

We can see in these passages from the early 1930's that Wittgenstein had already identified the fundamental problem with appealing to an inner mental mechanism when explaining the meaning/content of thoughts/sentences – how are we to specify the *correct* functioning or operating of the mechanism without already taking for granted the meaning that the thought/sentence is supposed to have. The mechanism is only a 'picture' of the meaning/intention if we already have a grasp of the correct way of projecting the mechanism's behaviour into the future. This is, I take it, fundamentally the same issue that would later be pursued in the celebrated 'rule-following remarks' in the *Investigations*: how are we to distinguish between correct and incorrect interpretations of a rule for applying a concept/term without already taking for granted the meaning of that concept? How are we to distinguish between correct and incorrect dispositions on the part of the subject to apply a concept or term without already taking for granted its meaning?

---

[8] The Jacquard loom was invented in 1804 by Joseph Marie Jacquard, though the great inventor Jacques de Vaucanson had already, in 1745, designed an automated loom that could be 'programmed' using punch cards. In a Jacquard loom these punched cards could be chained together to form sequences of any length, thus allowing for something like a primitive kind of binary programming language.

[9] Moreover, merely causal accounts of the psychological mechanisms involved in thought do not seem able to explain the ability of thoughts be about items that have no causal connection to the subject – such as future items, non-existent items, abstract items.

Near the start of the *Blue Book*, Wittgenstein is quite clear that even a complete causal understanding of our neural/psychological 'mechanisms' would not answer these sorts of *philosophical* questions about intentionality and the basis of determinate mental content:

> "We are tempted to think that the action of language consists of two parts; an inorganic part, the handling of signs, and an organic part, which we may call understanding these signs, meaning them, interpreting them, thinking. These latter activities seem to take place in a queer kind of medium, the mind; and the mechanisms of the mind, the nature of which, it seems, we don't quite understand, can bring about effects which no material mechanism could. Thus e.g. a thought (which is such a mental process) can agree or disagree with reality; I am able to think of a man who isn't present; I am able to imagine him, 'mean him' in a remark which I make about him, even if he is thousands of miles away or dead.' (Blue Book, 3-4)

If the production of genuinely meaningful claims does essentially involve the 'handling of signs' – i.e. formally/syntactically defined operations – then this is something that an 'inorganic' machine or technology could also do. So then, as Wittgenstein recognizes, we are tempted to think that there must be something else added to the story, such as a special organic-mental medium[10], to account for the 'queer'-ness of intentionality – the ability of a genuinely meaningful claim to point beyond itself to distal parts of the world, or even non-existent parts of the world, etc. For it seems to us that a merely 'material mechanism' could not produce these peculiar 'effects' just by shuffling symbols. But then on the next page, Wittgenstein renounces the idea that it is 'curious effects' that need explaining:

> '…we are tempted to say 'the mechanism of the mind must be of a most peculiar kind to be able to do what the mind does'. But here we are making two mistakes. For what struck *us* as being queer about thought and thinking was not at all that it had curious effects which we were not yet able to explain (causally). Our problem, in other words, was not a scientific one; but a muddle felt as a problem.' (Blue Book, 5-6)

I take it that the two mistakes Wittgenstein refers to here are: (a) thinking that the mysterious nature of intentionality should be thought of as a property/phenomenon which produces strange effects that we so far lack a causal explanation for, (b) thinking that the nature of the internal mechanisms in our heads must therefore be of a commensurately strange kind if they are producing such strange effects. But what is mysterious or 'queer' about intentional mental states is not so much the 'effects' that they produce – something that could indeed be investigated in a causal-scientific manner. Rather, it is the normativity of their meaning/content, the fact that they can be rendered correct or incorrect, satisfied or unsatisfied, according to the state of distal parts of the world; the fact that the meaning of concepts 'reaches out' to a potential infinity of future cases.

In the Investigations, the most extended discussion connecting the technological notion of a machine with questions about meaning and rule-following, occurs in sections §193 and §194.

> "The machine as symbolizing its action: the action of a machine—I might say at first— seems to be there in it from the start. What does that mean?—If we know the machine, everything else, that is its movement, seems to be already completely determined. We talk as if these parts could only move in this way, as if they could not do anything else. How is this—do we forget the possibility of their bending, breaking off, melting, and so on? Yes; in many cases we don't think of that at all. We use a machine, or the

---

[10] This temptation to appeal to some kind of special feature of an 'organic medium' is, arguably, clearly illustrated by John Searle's somewhat infamous claim that original or intrinsic intentionality is a 'natural biological phenomenon' depending somehow on the specific biochemistry of the brain – see Searle, 1980, 1984.

drawing of a machine, to symbolize a particular action of the machine. For instance, we give someone such a drawing and assume that he will derive the movement of the parts from it. (Just as we can give someone a number by telling him that it is the twenty-fifth in the series 1, 4, 9, 16, . . . .)

… But when we reflect that the machine could also have moved differently it may look as if the way it moves must be contained in the machine-as-symbol far more determinately than in the actual machine." (PI §193)

The machine-as-symbol conception takes the correct functioning of the mechanism for granted, ignoring the possibility of malfunction. An internal, neural mechanism may seem to already contain all its future possibilities for producing utterances or behaviour and so it may seem as if we can, as it were, read-off the *meaning* of the internal mechanism from its manner of operating. But this seems so only because we are already taking for granted the distinction between correct and incorrect functioning, between how it *will* in fact empirically behave and how it *ought* to behave.

These same issues concerning the fundamental grounds for determinate content, which Wittgenstein first began wrestling with in the 1930's, came to centrally occupy many recent theorists working with a broadly functionalist or computational view of the mind and also many researchers in A.I. – e.g. Steve Harnad (1990) called this the 'symbol grounding problem'. Thus Fodor supplemented the Language of Thought hypothesis with an *independent* psycho-semantic theory of reference for the basic elements of mental syntax – his theory of asymmetric causal dependence[11]. And many others have offered competing psycho-semantic accounts designed to solve this same fundamental problem of somehow distinguishing correct from incorrect functioning of a putatively meaningful mental mechanism in a non-circular way – e.g. Ruth Millikan explicitly offered her influential teleosemantic theory as an answer to the problem raised by Wittgenstein's rule-following considerations[12].

## 4. INTERPRETABILITY AND THINKING MACHINES

Now, to repeat, the fact that a description of an internal, symbol-manipulating mechanism would not explain meaning's normativity or the foundations of intentionality, hardly shows that appealing to such an inner mechanism can have no explanatory role at all in the full story of how we produce meaningful thought and speech. Think again of Fodor: he clearly recognised that the Language of Thought hypothesis had to be supplemented with a further psycho-semantic theory, but even in the absence of such a theory, the recursive/compositional nature of the L.O.T. model is still (supposedly) able to explain how we can immediately understand endlessly many novel sentences and the limitless potential of human thought for producing novel sentences, despite our finite brains.

And Wittgenstein himself, at least at the time of the *Blue Book*, seemed happy to allow that as an empirical hypothesis about human thought, this sort of mechanistic L.O.T. hypothesis has at least a great deal of plausibility. For example in the *Blue Book* he writes:

"We may say that thinking is essentially the activity of operating with signs." (Blue Book, 6)

---

[11] See Fodor (1987)

[12] See Millikan (1990)

Wittgenstein at this stage was also apparently happy to allow that it might, in certain contexts and for certain purposes, be legitimate to identify the location of thoughts/thinking with the location of certain neurophysiological processes :

> "Now does this mean that it is nonsensical to talk of a locality where thought takes place? Certainly not. This phrase has sense if we give it sense. Now if we say "thought takes place in our heads", what is the sense of this phrase soberly understood? I suppose it is that certain physiological processes correspond to our thoughts in such a way that if we know the correspondence we can, by observing these processes, find the thoughts. But in what sense can the physiological processes be said to correspond to thoughts, and in what sense can we be said to get the thoughts from the observation of the brain? I suppose we imagine the correspondence to have been verified experimentally." (Blue book, p7)

> "If… we do use the expression "the thought takes place in the head", we have given the expression its meaning by describing the experience which would justify the *hypothesis* that the thought takes place in our heads, by describing the experience which we wish to call 'observing thought in our brain'." (Blue Book, 8)

Wittgenstein at this point in the development of his views, seems to be *allowing* that inner physiological mechanisms, whose operations correspond to syntactic rules for operating with signs, may well play an important role in enabling or realising thought in humans and that this may be something that we could eventually verify by experiment and observation.

So what exactly is Wittgenstein's own positive answer to the sort of foundations-of-meaning questions that the mechanistic, 'symbol-handling' model of the mind is not, in itself, equipped to solve? This has been a matter of much lively interpretative controversy, generating a huge literature[13], ever since Kripke's (1982) much-discussed 'sceptical solution' to the rule-following paradox, stated in section §201 of the Investigations[14]. I shall not try to definitively settle these deep and difficult issues here[15]. But I do think that one important component of any plausible interpretation of Wittgenstein's own answer to the rule-following 'paradox' of §201, is the crucial role, for Wittgenstein, of *inter-personal interpretability* as a necessary basis for determinate content. I would thus tend to read Wittgenstein as offering an account of the grounds of representational content that is in the same general ball-park as the accounts later given by Davidson (1973) and Dennett (1987) – both of whom constitutively link meaning to the interpretability of both linguistic and non-linguistic behaviour as expressive of broadly rational, comprehensible attitudes (beliefs and desires and so on)[16].

What is this notion of interpretability? In section §206, Wittgenstein imagines the situation of being confronted with an alien culture or species and asks: when would we ascribe detailed representational contents to the utterances (and thoughts) of the subjects we meet?

---

[13] See Miller & Wright (eds.) (2002) for a collection containing a number of the most important papers in this literature.

[14] For example: McDowell (1984, 2009) suggests that Wittgenstein espoused a kind of *quietism*: according to which *nothing* much can be said about why or in virtue of what, one way of applying a concept, or continuing a rule, or acting in response to an intention, is in accord with the meaning and other ways are not. Wright (1986, 1989) argues for a form of constructivist anti-realism, according to which the correct meaning of a term must somehow be determined or constituted by our best judgements about its meaning.

[15] See Kusch (2006) for a book-length treatment of these issues, to which the present paper is greatly indebted.

[16] See Hopkins (1999, 2004) for a detailed account that connects Wittgenstein's thought with Davidson's notion of radical interpretation. The present paper is also greatly indebted to Hopkin's account.

> "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)

As countless commentators have remarked, this passage is clearly making the same kind of move that Quine and Davidson[17] later made when they invoked the figure of a radical translator or radical interpreter – and it is also, I think, pretty clearly the same kind of move that Dennett makes with his notion of the 'intentional stance'[18]. Very briefly, the common idea here is that the words of a language can only be interpreted as meaningful given that they are employed together with comprehensible deeds (actions), as part of behaviour that is manifestly a broadly rational and purposive way of interacting with the surrounding environment and with other people. But equally, without something like the detailed, regular patterns of rule-governed language use, we could not ascribe anything like the same level of precise, determinate meanings as we do in fact ascribe to each other. Non-linguistic animals may display sufficiently rational and purposive behaviour that we can attribute them simple beliefs and desires, but without language to further articulate and differentiate these states, their content must remain extremely simple and non-specific. At a number of points in the *Investigations*, Wittgenstein considers the inevitable limitations on the mental ascriptions that we can sensibly make to non-linguistic animals, such as dogs:

> "Why can't a dog simulate pain? Is he too honest? Could one teach a dog to simulate pain? Perhaps it is possible to teach him to howl on particular occasions as if he were in pain, even when he is not. But the surroundings which are necessary for this behaviour to be real simulation are missing." (PI§250)

> One can imagine an animal angry, frightened, unhappy, happy, startled. But hopeful? And why not? A dog believes his master is at the door. But can he also believe his master will come the day after to-morrow?—And what can he not do here?—How do I do it?—How am I supposed to answer this? Can only those hope who can talk? Only those who have mastered the use of a language. That is to say, the phenomena of hope are modes of this complicated form of life. (PI, part IIi, p174)

To return to the fundamental question about intentionality: What makes a bit of behaviour into a correct or incorrect response to a rule? What gives an utterance a specific meaning? Whatever further complexities there may be, I think it is clear that a key part of Wittgenstein's answer is that the individual bit of behavior, or utterance, must form part of a wider pattern of behavior that exhibits the sort of regularity and rationality that allows for the interpretation and ascription of both meaningful talk and meaningful intentional mental states. For example: what makes writing '1002' the correct continuation of the series +2? Well, at the very least this response must be part of a wider pattern of behavior that has enough regularity to allow for this degree of determinate interpretation of the order '+2'. This is not *just* the mere fact that the subject, or the surrounding community, does in fact generally continue the series by writing: '1002, 1004, 1006 etc'. The response '1002' must also be part of a wider practice of counting, a practice that has a comprehensible use and purpose and in which people allow themselves to be *corrected* and *constrained* in their responses in various

---

[17] Compare the following passage from Davidson, where he deliberately employs a Wittgensteinian turn of phrase:
"The central point is that finding the common ground is not subsequent to understanding, but a condition of it. … If we understand their words, a common ground exists, a shared 'way of life'." (Davidson, 1982, 20)
[18] See Dennett, 1987.

rationally comprehensible ways. This in turn requires that both the subject and the surrounding community speak and behave in ways that allow the ascription of broadly rational/comprehensible beliefs and desires in the first place.

Now much more could and should be said about all this. The point I have emphasized here is just that, according to Wittgenstein, the ascription of determinate content to people's utterances and mental states requires, at a minimum, a certain kind of *pattern of regularity* in their verbal and non-verbal behavior. Wittgenstein is quite explicit about this in section §207, where he continues the 'radical interpreter'-style thought experiment:

> "Let us imagine that the people in that country carried on the usual human activities and in the course of them employed, apparently, an articulate language. If we watch their behaviour we find it intelligible, it seems 'logical'. But when we try to learn their language we find it impossible to do so. For there is no regular connexion between what they say, the sounds they make, and their actions; but still these sounds are not superfluous, for if we gag one of the people, it has the same consequences as with us; without the sounds their actions fall into confusion—as I feel like putting it.
> Are we to say that these people have a language: orders, reports, and the rest?
> There is not enough regularity for us to call it "language"." (PI §207)

The moral then for present purposes: Wittgenstein held that not only must a subject's behavior be interpretable as broadly rational and purposive, there must also be a certain kind of 'regular connection' between the subject's non-verbal behavior and her verbal responses in order for us to be able to ascribe complex, determinate contents.

And so to the extent that it is possible for machines also to participate in these kinds of humanly interpretable patterns of verbal and non-verbal behaviour, 'this complicated form of life', Wittgenstein actually seems to have been at least open to the possibility of something like 'Artificial Intelligence'. As early as 1933 Wittgenstein was already thinking about the technological idea of 'a prosthetic apparatus for thinking':

> 'If one thinks of thought as something specifically human and organic, one is inclined to ask 'could there be a prosthetic apparatus for thinking, an inorganic substitute for thought?' But if thinking consists only in writing or speaking, why shouldn't a machine do it?…
> 'But could a machine think?' – Could it be in pain? – Here the important thing is what one means by something *being in pain*.' (Philosophical Grammar, §64, p105)

Some of this same material was then later re-worked into the Investigations:

> Could a machine think?——Could it be in pain?—Well, is the human body to be called such a machine? It surely comes as close as possible to being such a machine.
> But a machine surely cannot think!—Is that an empirical statement? No. We only say of a human being and what is like one that it thinks. (PI §359-360)

Wittgenstein is insisting here that the difference between a genuine, thinking subject and a mere unthinking mechanism cannot be drawn just by looking at the kinds of internal processes and structures involved – for after all, a human body is surely just a very complex organic mechanism. In light of PI§207, quoted above, we should understand that a crucial part of what is required to be 'a human being and what is like one' is a regular, interpretable pattern of connections between verbal and non-verbal behaviour, which is fundamental for the ascription of meaningful language and thought. In principle, an inorganic piece of technology that displayed such a rich pattern of connections could be sufficiently like a human to count as genuinely thinking. Wittgenstein thus opposed the idea that mere mechanical/syntactic manipulation of symbols in accord with algorithmic rules – as in a

Turing Machine – could suffice by itself for genuine thought. This point is stated clearly, and with explicit reference to the technological notion of a 'calculating machine', in the *Remarks on the Foundations of Mathematics*:

> Does a calculating machine *calculate*?
> Imagine that a calculating machine had come into existence by accident; now someone accidentally presses its knobs (or an animal walks over it) and it produces the product 25 x 20.
> I want to say: it is essential to mathematics that its signs are also employed in *mufti*.
> It is the use outside mathematics, and so the *meaning* of the signs, that makes the sign-game into mathematics. (RFM, part V, §2, p257)

We know that Turing attended some of Wittgenstein's classes and that they conversed and disagreed about the foundations of mathematics. We also know that Wittgenstein possessed a copy of Turing's landmark paper 'On Computable Numbers' (Turing, 1936). And in fact there is one explicit reference to Turing Machines in the *Remarks on the Philosophy of Psychology*:

> Turing's 'Machines'. These machines are *humans* who calculate.' (RRP, vol 1, §1096)

Though the remark is somewhat cryptic[19], it is important to remember that Turing explicitly connected his hypothetical machine with a human performing a calculation[20]. Wittgenstein here seems to allow that a Turing Machine algorithm can be instantiated by a human. The converse formulation of this thought – that a human's calculation might, in some contexts, qualify as a merely mechanical action – is also stated in the *Remarks on the Foundations of Mathematics*:

> If calculating looks to us like the action of a machine, it is *the human being* doing the calculation that is the machine. (RFM, part IV, §20, p234)

> A human calculating machine might be trained so that when the rules of inference were shewn it and perhaps exemplified, it read through the proofs of a mathematical system (say that of Russell), and nodded its head after every correctly drawn conclusion, but shook its head at a mistake and stopped calculating. One could imagine this creature as otherwise perfectly imbecile. (RFM, part V, §3, p258)

In summary so far then: Wittgenstein's many discussions of technology in his later period are centrally concerned to combat the idea that identifying a particular kind of internal mechanism will allow us to discover what is essential or fundamental to *meaningful* thought and language. He understood the allure of thinking of the mind in terms of the operation of some internal mechanism or piece of technology, but he repeatedly directs us to consider the wider context and pattern that is required for mere physical behaviour to be normatively evaluable as rule-governed, to count as correct or incorrect, to be meaningful. However, we have also seen that Wittgenstein (in the Blue Book) explicitly allowed for the possibility that we might discover a regular 'correspondence' between internal, neural mechanisms and thoughts and so that it might, on occasion, be useful to talk of thoughts as taking place within the head. Nor did he think that there was something special about the kind of organic mechanisms found within the human skull – and in fact he seems to have been open to the

---

[19] For detailed discussion of this cryptic remark see Shanker (1998), Floyd (2012).

[20] In section 1 of the paper Turing writes: 'We may compare a man in the process of computing a real number to a machine which is only capable of a finite number of conditions…' (Turing, 1936, 231). And when in section 9 of the paper Turing writes 'The behavior of the computer at any moment is determined by the symbols which he is observing, and his 'state of mind' at that moment' (Turing, 1936, 250), he is referring to a *human* computer. Many thanks to Greg Lavers for helpful discussion.

possibility that an inorganic piece of technology might, *given* the right surrounding context and pattern and history, count as genuinely thinking or calculating.

## 5. AN ANTI-MECHANISTIC STRAND

Once we start thinking about this pattern of regular connections – between the subject's external circumstances, the subject making certain kinds of utterances and the subject behaving in certain rationally interpretable ways – it becomes very natural to assume that there must be *some kind* of physiological mechanisms in the brain and nervous system which *causally* explain such regularity. Fodor (1975) provocatively claimed that the L.O.T. hypothesis was 'the only game in town' when it came to providing such an explanation of the systematicity and productivity of our thought and language. Many have disagreed with Fodor's assessment and proposed alternative cognitive architectures – e.g. connectionist networks. One moral of the rule-following remarks is that we would have no basis to ascribe a determinate meaningful content to any such inner mechanisms *unless* they formed part of a bigger interpretable picture; a picture that plausibly requires inter-personal normative practices of various kinds. But the mere idea that there *are* some such inner mechanisms, which are the causal wellspring of our regularly patterned speech and behavior, is at least *consistent* with what I take to be Wittgenstein's account of the ultimate grounds of intentionality. And as the passages quoted above from the Blue Book illustrated, Wittgenstein was apparently, at least sometimes, comfortable with this immensely natural assumption.

However, there are other strands in his later writings, which suggest that Wittgenstein, sometimes at least, held a rather more hostile attitude to the idea that meaningful thought can be explained by appeal to the regular operation of inner mental mechanisms. In the *Investigations* this is perhaps clearest in his discussion of the ability to read:

> "…Now we should of course like to say: What goes on in that practised reader and in the beginner when they utter the word can't be the same. And if there is no difference in what they happen to be conscious of there must be one in the unconscious workings of their minds, or, again, in the brain.—So we should like to say: There are at all events two different mechanisms at work here.
> And what goes on in them must distinguish reading from not reading. —But these mechanisms are only hypotheses, models designed to explain, to sum up, what you observe." (§156)

> "But isn't that only because of our too slight acquaintance with what goes on in the brain and the nervous system? If we had a more accurate knowledge of these things we should see what connexions were established by the training, and then we should be able to say when we looked into his brain: "Now he has *read* this word, now the reading connexion has been set up".——And it presumably *must* be like that—for otherwise how could we be so sure that there was such a connexion? That it is so is presumably a priori—or is it only probable? And how probable is it? Now, ask yourself: what do you *know* about these things?——But if it is a priori, that means that it is a form of account which is very convincing to us." (§158)

Now if *all* Wittgenstein is saying here is that it is ultimately an empirical hypothesis that a regular ability to correctly read a word is correlated with the operation of some kind of internal mechanism or process in the brain – well, that is surely correct so far as it goes. Perhaps we are hasty to assume we can know *a priori* that it must be so. But still, given all that we do in fact now know about the brain, it is surely an *extremely* reasonable and plausible hypothesis. Consider the scientific study of dyslexia. What is it that these scientists are studying if not the inner causal mechanisms which crucially underpin the ability to read fluently?

Elsewhere in his later writings Wittgenstein was apparently prepared to go further than merely pointing out that *we cannot know a priori* that outward behavioural regularities are underpinned by the regular operation of inner causal mechanisms. For there are passages where he seems to assert that it is a reasonable, open hypothesis that there are *not* any such inner casual mechanisms whose regular pattern of operation is responsible for our intelligent purposive behavior. Consider his somewhat notorious example of two kinds of seeds. Wittgenstein used this example at a number of points in his notebooks – here is its occurrence in *Zettel*[21]:

> "No supposition seems to me more natural than that there is no process in the brain correlated with associating or with thinking; so that it would be impossible to read off thought-processes from brain-processes… The case would be like the following – certain kinds of plants multiply by seeds, so that a seed always produces a plant of the same kind as that from which it is produced – but *nothing* in the seed corresponds to the plant which comes from it; so that it is impossible to infer the properties or structure of the plant from those of the seed that comes out of it – this can only be done from the history of the seed. So an organism might come into being even out of something quite amorphous, as it were causelessly; and there is no reason why this should not really hold for our thoughts, and hence for our talking and writing."(Zettel §608)

> 'It is thus perfectly possible that certain psychological phenomena *cannot* be investigated physiologically, because physiologically nothing corresponds to them' (Zettel §609)

Though Wittgenstein couches these remarks in terms of 'supposition' and 'possibility', they surely strike us as expressing a remarkably stark and radical hostility to the idea that investigating neuro-physiological mechanisms will provide *any* kind of understanding or explanation of psychological phenomena.

> 'I saw a man years ago: now I have seen him again, I recognise him, I remember his name. And why does there have to be a cause of this remembering in my nervous system? Why must something or other, whatever it may be, be stored up there *in any form*? Why *must* a trace have been left behind? Why should there not be a psychological regularity to which *no* physiological regularity corresponds?' (Zettel §610)

> 'The prejudice in favour of psychophysical parallelism is a fruit of primitive interpretation of our concepts…' (§611)

Again, if all that Wittgenstein is trying to establish here is that one can consistently *imagine* that there is no kind of regular pattern amongst the processes going on inside the head which is even *correlated* with regular, intelligent behavioural ability – well, it should be conceded that there may be no internal, logical contradiction in imagining such a situation. But the idea that this might *actually* turn out to be the case, let alone that 'no supposition is more natural', will, in the light of modern neuroscience, surely strike most of us these days as a quite incredible claim – at least when it comes to the sorts of psychological abilities that Wittgenstein is discussing here: reading, remembering.

Now, there are various ways one might try reading these passages so as to make their message seem less radical. One might think that connectionism or parallel distributed processing could vindicate the idea that there is no one specific structure or process in the

---

[21] The example is also discussed in sections §903-906 in volume 1 of *Remarks on the Philosophy of Psychology* (Wittgenstein, 1980), and pp 401-11 of 'Cause and Effect: Intuitive Awareness' (Wittgenstein, 1976). See also the notes made by Rush Rhees, published as Rhees (2002, p17-18) and Wittgenstein, Rhees & Citron (ed.) (2015, p45). I am very grateful to Lynette Reid for help identifying these various instances of the 'seed' example.

brain, nor even a specific set of connection strengths, which is present whenever there is a specific kind of thought[22]. But the very existence of a flourishing connectionist program hardly fits with the claim that 'psychological phenomena cannot be investigated physiologically'.

One might think that externalism about mental content could vindicate the idea that two brain states/processes could be intrinsically identical and yet have different contents, thus explaining why it is impossible to 'read-off thought-processes from brain processes'[23]. But content externalism of the kind made famous by Putnam's Twin Earth example does not seem to fit the example of intrinsically identical seeds which are nevertheless disposed to develop into differently shaped and coloured adult plants, even under identical growing conditions. For even though the meanings (reference) of Oscar's and Twin Oscar's thoughts are supposed to differ, insofar as they are intrinsic duplicates of each other, the sounds coming out of their mouths are supposed to still be acoustically identical, and their bodies would both be disposed to react in physically identical ways in identical circumstances.

One might recall that it is increasingly popular these days for metaphysicians to hold that there can be *pure dispositions* (or 'powers')*, which are not grounded in any categorical basis. For example: the probabilistic disposition of certain radioactive particles to decay with a specific chance within a specific period of time, is often given as the standard example of a pure disposition. And so two intrinsically identical particles, in identical circumstances, may decay after totally different periods of time. But notice that this is *not* an example of two intrinsically/categorically identical items that have *different* pure dispositions or powers. It is rather an example of two intrinsically identical items with *the same non-deterministic* (probabilistic) dispositions. And in neither the example of the seeds and the flowers, nor in the case of the relation between brain processes and thought processes, does Wittgenstein seem to be thinking that the former *probabilistically* constrains the latter.

Although it is certainly wise to exercise the utmost charity and caution when interpreting these difficult passages from *Zettel*, ultimately I think it is a mistake to try to downplay the radical, revisionary intent of Wittgenstein's ideas here. For he was quite explicit that he meant to 'upset' our basic assumptions about causality:

> 'Why should there not be a psychological regularity to which *no* physiological regularity corresponds? If this upsets our concept of causality then it is high time it was upset." (Zettel §610)

> "Why should there not be a natural law connecting a starting and a finishing state of a system, but not covering the intermediary state?" (Zettel, §613)

> "Why should this order not proceed, so to speak, out of chaos?" (Zettel, 608)

Finally, there is another strand to Wittgenstein's hostility towards the notion of inner psychological mechanisms. Rather than suggesting that there literally may not be any kind of inner physical order, and that regular patterns of behaviour may emerge from inner 'chaos', Wittgenstein also seems to suggest at points that the presence or absence of certain physiological processes or mechanisms is just *irrelevant* to our conception/understanding of mental states and abilities. Consider this section §157 of the *Investigations*, which again invokes the technology of the pianola:

---

[22] See Mills (1993) for a suggestion along these lines.

[23] See McDonough (2004) for a suggestion along these lines.

"Consider the following case. Human beings or creatures of some other kind are used by us as reading-machines. They are trained for this purpose. The trainer says of some that they can already read, of others that they cannot yet do so.…
…in the different case of a reading machine which translated marks into sounds, perhaps as a pianola does, it would be possible to say: "The machine *read* only after such-and-such had happened to it—after such-and-such parts had been connected by wires; the first word that it read was ...." . But in the case of the living reading-machine "reading" meant reacting to written signs in such-and-such ways. This concept was therefore quite independent of that of a mental or other mechanism… The change when the pupil began to read was a change in his *behaviour*, and it makes no sense here to speak of 'a first word in his new state'." (PI §157)

Now, I should confess that I find it especially difficult to interpret Wittgenstein's remarks on reading that occur in the 140's and 150's of the *Investigations*. It is certainly true that we can ascribe the ability to read to ourselves and to others without knowing anything about what goes on inside their skulls. And it is also a very plausible functionalist thought that the ability to read could be multiply realized by all sorts of different physical mechanisms or 'architectures'. Indeed, those impressed by connectionist models might want to insist that there may be no cleanly identifiable or decomposable 'mechanisms' in the brain that are responsible for, say, the ability to read a text – the ability might reside in the pattern of synaptic connection strengths throughout the entire brain. But from the mere fact that we can competently ascribe a cognitive ability in complete ignorance of how the brain works, it clearly does not follow that the *ability* to read – as opposed to our *concept* of 'being able to read' – is *independent* of any mental mechanism.

One possible interpretation of §157 would be that Wittgenstein is implicitly relying here on something like the following principle:

- If we can know that the predicate F correctly applies to O, *without* knowing the truth of some proposition, p, then the satisfaction conditions for 'F' are totally independent of the truth or falsity of p.

But if this *is* the underlying principle, it seems straightforwardly false: for the criteria by which we come to know that O is F, might in fact have some nomic or necessary connection with p, though we remain ignorant of this connection.

In a similar vein, consider section §149:

"If one says that knowing the ABC is a state of the mind, one is thinking of a state of a mental apparatus (perhaps of the brain) by means of which we explain the manifestations of that knowledge. Such a state is called a disposition. But there are objections to speaking of a state of the mind here, inasmuch as there ought to be two different criteria for such a state: a knowledge of the construction of the apparatus, quite apart from what it does." (PI§149)

To repeat: I am uncertain that I fully understand what Wittgenstein is up to here. But, prima facie, it seems that here too he is pointing out that there are criteria which allow us to correctly attribute knowledge of the alphabet to a subject without knowing anything about the subject's inner mental 'apparatus'. And it seems that this is meant to be a reason *against* taking the subject's knowledge of the alphabet to be a dispositional state that is essentially grounded or realised by such an inner mental apparatus. If that *is* the point Wittgenstein is making here, I am afraid I cannot see the force of it. For we often attribute dispositions to things despite knowing nothing of the causal-physical structure that grounds the disposition. Thus, people correctly described glass as fragile and salt as soluble, long before anyone had any idea about how the molecular structure of these substances explains these dispositions.

As Snowdon (2018) points out, one standard *reason* for speaking in dispositional terms is precisely *because* we are ignorant of the intrinsic structure, or 'categorical grounds' that give rise to the object's dispositions. Indeed, this is frequently the case when we want to describe the behavior of some piece of technology whose inner workings we don't really understand – e.g. calling a computer 'glitchy' or a car's engine 'temperamental'[24].

## 6. SUMMARY

This article has gathered together many of Wittgenstein's explicit references to technology and/or machines in order to show how throughout his career, the main purpose for such technological comparisons was to illustrate both the appeal and the limitations of the now familiar model of the mind as a kind of symbol-manipulating mechanism. Though he soon came to realize that this model of the mind, in itself, can tell us nothing about the fundamental grounds or conditions required for meaning and intentionality, he remained – sometimes at least – open to the possibility that such a mechanistic model might yet have some explanatory worth. But I have also tried to demonstrate that there is another strand to his later thought that was much more hostile to the whole idea of making any kind of explanatory appeal to internal mechanisms. And so I think we should accept that Wittgenstein's attitude to a mechanistic conception of the mind in his later period was somewhat conflicted; alongside the deep and glorious insights of the rule-following discussions there are also occasional passages which may seem, at least to some 21st century eyes, to evince a rather radical and implausible wholesale rejection of the notion of inner mental mechanisms. But then Wittgenstein would surely not have been surprised to find that he was, in this respect as in many others, out of step with the mechanistic and technology-revering spirit of our times[25].

## *References*

- Armstrong, D. (1968) *A Materialist Theory of the Mind*. Routledge & Kegan Paul.
- Armstrong, D. (1981) *The Nature of Mind and Other Essays.* Cornell University Press.
- Brentano, F. (1874) *Psychologie vom empirischen Standpunkt*, Leipzig: Duncker & Humblot.
- Chemero, A. (2009) *Radical Embodied Cognitive Science*, MIT Press
- Davidson, D. (1973), 'Radical Interpretation', *Dialectica*, 27: 314–28
- Davidson, D. (1982) 'Expressing Evaluations', *The Lindley Lecture*, University of Kansas Press.
- Davies, M. (1991) "Concepts, Connectionism, and the Language of Thought," in *Philosophy and Connectionist Theory*, W. Ramsey, S.P. Stich and D.E. Rumelhart (eds.), Hillsdale, NJ: Lawrence Erlbaum.
- Dennett (1987) *The Intentional Stance*, MIT Press.
- Floyd, J. (2012) 'Wittgenstein's Diagonal Argument: A Variation on Cantor and Turing' in P. Dybjer et al. (eds.) *Epistemology versus Ontology*, Springer.
- Fodor, J. (1975) *The Language of Thought*, Harvard University Press, 1975
- Fodor, J. (1987) *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*, MIT Press.
- Harman, G. (1973) *Thought*, Princeton University Press.

---

[24] Compare a similar sort of criticism that Armstrong (1968, 1981) makes of Ryle's (1949) Behaviourism for ignoring the explanatory worth of identifying the categorical grounds for a disposition. Many thanks to Stefan Koller for drawing my attention to this similarity.

[25] A version of this paper was presented at the workshop 'Wittgenstein and Philosophy of Technology' at the University of Vienna. I am grateful to the audience on that occasion for their helpful questions. This paper was crucially shaped by a three-way correspondence with Lynette Reid and Stephan Koller, from whom I learned a huge amount and to whom I am greatly indebted. Finally, many thanks to two anonymous referees for this journal for their many valuable suggestions and comments.

- Harnad, S. (1990) "The Symbol Grounding Problem", *Philosophical Explorations* 42: 335-346
- Hélio A.G. Teive, Guilherme Ghizoni Silva, Renato P. Munhoz. "Wittgenstein, medicine and neuropsychiatry." Arquivos de Neuro-Psiquiatria 2011;69(4):714-716
- Henderson, J. R. (1973) 'Ludwig Wittgenstein and Guy's Hospital', Guy's Hospital Reports 122: 185-193.
- Hopkins, J. (1999) "Wittgenstein, Davidson, and Radical Interpretation", in F. Hahn (ed.), *The Library of Living Philosophers: Donald Davidson*. Open Court.
- Hopkins, J. (2004) "Wittgenstein and the Life of Signs", in Max Kölbel & Bernhard Weiss (eds.), *Wittgenstein's Lasting Significance.* London: Routledge.
- Hutto & Myin (2017) *Evolving Enactivism*, MIT Press.
- Kripke, S (1982) *Wittgenstein on Rules and Private Language*. Harvard University Press.
- Kusch, M (2006) *A Sceptical Guide to Meaning and Rules: Defending Kripke's Wittgenstein*, Routledge.
- McDonough, R. (2004) "Wittgenstein, German Organicism, Chaos, and the Center of Life", *Journal of the History of Philosophy*, Volume 42, Number 3, July 2004, pp. 297-326
- McDowell, J. (1984) "Wittgenstein on Following a Rule", *Synthese* 58: 325-364.
- McDowell, J. (2009) 'Wittgensteinian "Quietism"' *Common Knowledge* 15 (3): 365-372.
- McGuiness, B. (1988) *Wittgenstein: A Life.* Duckworth Press.
- Miller, A. & Wright, C. (eds.) (2002) *Rule-Following and Meaning*, Acumen Press.
- Millikan, R. (1990) "Truth, Rules, Hoverflies, and the Kripke-Wittgenstein Paradox", *Philosophical Review* 99 (3):323-53
- Mills, S. (1993) "Wittgenstein and Connectionism: A Significant Complimentarity," *Royal Institute of Philosophy Supplement* 34:137-157
- Monk, R. (1991) *Wittgenstein: The Duty of Genius*, Vintage Press
- Ramsey, W. (2007) *Representation Reconsidered*, Cambridge University Press.
- Rey, G. (1995) "A Not 'Merely Empirical' Argument for a Language of Thought," in *Philosophical Perspectives* 9, J. Tomberlin (ed.), pp. 201–222
- Rhees, R. (ed.) (1984) *Recollections of Wittgenstein*, Oxford: Oxford University Press.
- Rhees, R. (2002) 'Five Topics in Conversation with Wittgenstein', *Philosophical Investigations* 25:1, 1-19.
- Ryle, G. (1949) *The Concept of Mind*, Hutchinson & Co.
- Schneider, S. (2011) *The Language of Thought: A New Philosophical Direction*, MIT Press.
- Shanker, S. (1998) *Wittgenstein's remarks on the foundations of AI,* New York: Routledge.
- Searle, J. (1980) 'Minds, brains and programs', *Behavioural and Brain Sciences* 3: 417-457.
- Searle, J. (1984) 'Intentionality and Its Place in Nature', *Dialectica* 38(2-3): 87-99.
- Snowdon, P (2018) "Wittgenstein and Naturalism", in Cahill, K. & Raleigh, T. (eds.) *Wittgenstein and Naturalism*, London: Routledge.
- Steiner, G. (1996) 'A Reading against Shakespeare', in *No Passion Spent: Essays 1978-1995*, Yale University Press
- Turing, A. (1936) 'On Computable Numbers, With an Application to the Entscheidungsproblem', *Proceedings of the London Mathematical Society* 42(1): 230-265.
- Wittgenstein, L. (1922) *Tractatus Logico-Philosophicus,* C. K. Ogden (trans.), London: Routledge & Kegan Paul.
- Wittgenstein, L. (1953) *Philosophical Investigations*, G.E.M. Anscombe and R. Rhees (eds.), G.E.M. Anscombe (trans.), Oxford: Blackwell.
- Wittgenstein L. (1958) *The Blue and Brown Books*, Oxford: Blackwell
- Wittgenstein, L. (1961) *Notebooks 1914-1916*, G. H. von Wright and G. E. M. Anscombe (eds.), Oxford: Blackwell.
- Wittgenstein, L. (1967) *Zettel*, G. E. M. Anscombe and G. H. von Wright (eds.), G. E. M. Anscombe (trans.), Oxford: Blackwell.
- Wittgenstein, L. (1974) *Philosophical Grammar*, R. Rhees (ed.), A. Kenny (trans.), Oxford: Blackwell
- Wittgenstein, L. (1976) "Cause and Effect: Intuitive Awareness", *Philosophia*, Vol. 6 Nos. 3-4, pp 409-425
- Wittgenstein, L. (1980) *Remarks on the Philosophy of Psychology*, vol. 1, G. E. M. Anscombe and G. H. von Wright (eds.), G. E. M. Anscombe (trans.), vol. 2, G. H. von Wright and H. Nyman (eds.), C. G. Luckhardt and M. A. E. Aue (trans.), Oxford: Blackwell.

- Wittgenstein, L., Rhees, R., Citron, G. (ed.) (2015) "Wittgenstein's Philosophical Conversations with Rush Rhees (1939–50): From the Notes of Rush Rhees", *Mind*, 124 (493): 1–71
- Wright, C. (1986) "Rule-Following, Meaning and Constructivism", in Charles Travis (ed.), *Meaning and Interpretation*. Oxford: Blackwell.
- Wright, C. (1989) "Wittgenstein's Rule-Following Considerations and the Central Project of Theoretical Linguistics", in A. George (ed.), *Reflections on Chomsky*, Oxford: Blackwell.