For BWS 2017 16. 7. 2017

WITTGENSTEIN’S LEGACY:

THE PRINCIPLES OF THE PRIVATE LANGUAGE ARGUMENTS

1. *An overview*

In the absence of a crystal ball, one cannot say what Wittgenstein’s role in 21st century philosophy will be. But one can attempt to articulate his philosophical legacy to the 21st century. As he remarked early in the 1930s, ‘a new method has been found’. Of course, that a new method has been found does not mean that people will use it, or even understand how to apply it. That a cure for itches has been found does not mean that people will cease to have itches or will stop scratching themselves. But it does mean that it is incumbent upon those who understand the new method or methods and who know how to use them to explain them to the next generation, to apply them to the problems of philosophy, both old and new, and to sciences which are enmeshed in conceptual difficulties. For this much is clear. Wittgenstein showed us the ubiquity of conceptual confusion, both among philosophers, among empirical scientists, and within the culture at large. Wittgenstein was the first philosopher to show us clearly why philosophy has a right, if not indeed a duty, to be a tribunal of sense before which scientists may be arraigned when they transgress the bounds of sense in their reasoning, in so called cognitive science, cognitive neuroscience, empirical psychology, economics, biology and physics. This provides a huge field for critical reflection among philosophers who have understood Wittgenstein’s achievements and methods.

To elaborate Wittgenstein’s manifold methodological insights, his conceptual cartography in the domains that he investigated, and his critical achievements in blocking what appear to be irresistible or deeply appealing lines of thought is far too large a subject for a single lecture. What I wish to present to you today is a survey of methods and principles evident in the private language arguments that are elaborated in *Investigations* §243-§315. They have widespread application in contemporary philosophy and in contemporary sciences of mind and brain.

The target of the private language arguments is far larger than the *prima facie* puzzling question of whether there could be a language that no one other than its speaker can understand. It is rather a certain conception of human nature that has dominated European philosophical thought (especially since the seventeenth century) and that informs a multitude of distinct, indeed often conflicting, philosophical theories concerning the nature of the mind, the nature of the body and of the relation between mind and body, the logical analysis of psychological concepts, and their Janus-faced character. Wittgenstein occasionally refers to it as the 'inner/outer' picture. He observed that a common flaw in philosophical reflection is the failure to put the question marks deep enough down – failure to dig down to the roots of the problems that beset us in philosophy (CV 62). He himself dug deeper than his predecessors or contemporaries. No one, in the history of our subject prior to Wittgenstein, saw that philosophical theories as far apart as representative idealism (Descartes, Locke, Frege, Russell), problematic (reductive) idealism (Berkeley, Hume, phenomenalism), materialism (both the quasi-materialism of Diderot and d'Holbach as well as the neuroscientific materialism of such Nobel prize winning neuroscientists as Crick and Kandel) all share the common presupposition that psychological predicates are names of psychological attributes. No one of them thought to investigate, let alone to challenge, that presupposition. In general, it was thought, we are immediately acquainted with these attributes by introspection. Such immediate acquaintance was conceived to be sufficient, or at any rate necessary, for possession of the ideas or concepts of such psychological attributes. It is also evident that *these* conceptions have in turn yet deeper roots in the Augustinian conception of language, namely that the fundamental function of words is to name things, and that the fundamental function of sentences is to describe how things stand; names have meanings, which are the things they stand for, and words are connected with *nominata* by acts of meaning. So the 'inner/outer' conception of the mind and of the relation of the mind to the body and to behaviour, itself tacitly presupposes that ubiquitous notion of linguistic meaning.

Wittgenstein's private language arguments not only exemplify his radicalism – his quest for the roots of philosophical confusion, they also instantiate an equally profound principle of investigation in philosophy. This is the principle, which I have discussed elsewhere, that when debates on a particular problem continue for centuries, and polarize into classical dichotomies, such as mentalism as opposed to behaviourism, idealism by contrast with materialism, nominalism versus realism, the way to proceed is *not* carefully to examine the arguments on both sides, weigh them carefully and opt for the side that is supported by better arguments. Rather, it is to dig down to the deepest presuppositions of the debate, presuppositions that *are shared by both sides*, and to challenge these. Hence the respective challenges to the referentialism of the Augustinian conception of meaning and the attendant idea of naming, and the focus upon the idea of a private language. In the philosophy of psychology, one might say, the idea of a private language is the heart of representative idealism, idealism, and solipsism in philosophy, as well as so called theory-theory and simulation theory in psychology, and the semantic marker hypothesis in cognitive neuroscience. It is this beating heart that one must stop if one is to extirpate such intellectual monsters.

To follow Wittgenstein's footsteps through the jungle of our psychological concepts, we should constantly keep in mind two further guidelines. First, we must seek out the grammatical and pictorial analogies that hold our imagination in a vice and lead us astray. Secondly, we must take pains not to conflate empirical exclusion of knowledge consequent upon ignorance or doubt, with logical exclusion of knowledge consequent upon *grammar*.

That a wide range of philosophical confusions – conceptual entanglements – are due to our being mesmerized by misleading analogies in the use of language is a principle that occurred to Wittgenstein in the early 1930s and that remains constant in all his subsequent thought. In *The Big Typescript* he wrote:

> If I rectify a philosophical mistake and say that this is the way it has always been conceived, but this is not the way it is, I must always point out an analogy, according to which one had been thinking, but which one did not recognize as an analogy.
>
> The effect of a false analogy accepted into language: it means a constant battle and uneasiness (a constant irritation, as it were). It is as if something seems to be a human being from afar, because at that distance we don't perceive certain things, but from close up we see that it is a tree stump. The moment we move away a little and lose sight of the explanations, one figure

> appears to us; if we then move more closely, we see a different figure; now we move away again, etc., etc. (BT 408f.)

What the philosopher must do is to identify the false analogy that holds us in sway, and to find the 'liberating word' that will enable us to grasp what, until then, had constantly weighed on our consciousness (ibid.). Only when we acknowledge the analogy presented to us as the source of our thought can we be liberated from intellectual bondage.

The importance of this should be evident. We are, when philosophizing about experience and its objects, mesmerized by the analogy between having a chattel and having an experience. We think correctly that when we have a house, we own the house, that it is our property, that a relationship of ownership obtains between us and our house. So too, we are prone to think (as Frege and Strawson so clearly did) that when we have an experience (such as pain, or a visual experience, or an emotion), we own the experience, that a relationship of ownership holds between us and the experience. The only difference is that in these cases, we think, the relationship is one of 'logically non-transferable ownership', whereas in the case of the house it is a matter of legal and transferable ownership. So obsessed are we with this analogy that we subsume what we think of as the relation between ourselves and our bodies under it too. For, after all, don't we speak of *having* a beautiful, healthy, ailing or feeble, body? Don't we talk of *my* body, *his* or *her* body? And so powerful is this false analogy that women now speak of owning their body, of their body as belonging to them, taking this as a ground for their right not to be sexually assaulted and of their right to an abortion. But arguments in favour of the right to abortion and sexual choice should not rely upon false analogies with ownership. We do not *own* our bodies. Our bodies do not *belong* to us. We *are* bodies (living spatio-temporal continuants) and our talk of *having* such-and-such a body is no more that talk of our somatic characteristics. The analytic task is to exhibit the connective analysis of 'having a pain' and 'having a penny' in order to show how superficial the grammatical analogy is. (And, by parity of reasoning, to show how misleading it is to construe our talk of having a beautiful or ageing body on the model of having a beautiful or ageing house). Light dawns when we realize that we have confused the matter of ownership with the form of ownership.

We are similarly taken in by the apparent analogy between a public ostensive definition and a putatively private ostensive definition. To be sure, non-philosophers have never heard of this term of art. But we are all aware of the fact that we often explain what we mean by a word by pointing at an object and saying ‘That is an O’, and we are naturally inclined to think that in exactly the same way we can, for ourselves alone, point ‘mentally’ at whatever experience we are undergoing and say to ourselves ‘This is E’. ‘Of course I know what the pains of childbirth are’, a mother may naturally say, ‘I have undergone them’; ‘I don’t really know what jealousy is – I’ve never been jealous’, or ‘The blind can’t really know what “red” means, since they have never had the experience of seeing red’. To be sure, one cannot literally point at one’s pains, emotions, or visual experiences of colour. But one can concentrate one’s attention on the experience, and is that not a kind of pointing? So here too, Wittgenstein’s task is to disabuse us of these false analogies. One must examine the disanalogies between pointing and concentrating one’s attention, and between concentrating one’s attention on a genuine sample, such as a metre ruler or a colour sample, and concentrating one’s attention on one’s headache or on one’s visual image.

We are very readily inclined to agree that we can’t really know what experiences others have. A variety of false analogies force themselves upon us here. There is a perfectly correct use for such utterances as ‘I can’t see what you see’, for example, when you look through your binoculars and say ‘I can see an osprey’, and I, looking through mine very carefully and not spotting the osprey, say, ‘I can’t see it’. Or, when you say ‘I can see Jack over there’, and I, looking with screwed up eyes, say ‘I can’t see him, your eyes are better than mine’. Or ‘I can’t see what you see, let me get up onto the ladder’, when one can see the object only from the top of the ladder. These, and many other similar uses are in order. But they provide false analogies for the idea that I can never experience what another person experiences, and so can *never know*, or *really* know, what experience he has (LPE 277f.). We transform the idea of a viewing point into that of a point of view, and then draw an analogy between looking at something from a certain point of view, and looking at something with one’s eyes (or, with conceptual abandon, through one’s eyes), concluding that you *cannot* (metaphysically) perceive things from my point of view.

When something goes on behind a curtain, we often want to know what it is. When someone thinks something, we often want to know what it is that he is thinking. So we are inclined to wonder what is going on behind that face or behind that furrowed brow. In the former case, we rightly suppose, the event is hidden behind a curtain, in the latter case it seems to be hidden behind the face or brow. We fail altogether to see the disanalogy in the analogy. For while we can draw the curtain to see what is behind it, we cannot draw his brow, but only ask him. And, to be sure, he can tell us. So now we think that he can see something (by introspection) which we cannot see, and report it to us. Now that we have scanners, we are inclined to think that we *can* at last see thinking – the subject sees his own thinking from the inside and we see it on the scanner-screen from the outside. But we are victims of a bad analogy. We cannot *see* what we think, and to say what we are thinking is not to report something only we can see. When someone reports what he thinks, he is neither describing what he thinks, nor is he describing anything he sees. Moreover, an fMRI scan can only show minute increases of oxygen consumption in various parts of the brain, and to see an image of *that,* constructed by a series of algorithms on the scanner-screen, is not to see thinking at all.

A second guideline Wittgenstein gives us is constantly to bear in mind the distinction between empirical exclusion of possibilities and logical exclusion of possibilities. I have tried to capture this with the redeeming word that *a logical or grammatical impossibility is not a possibility that is impossible*. We think correctly that when all grounds for doubt have been excluded (blocked) by reliable evidence, then we can be certain how things are, were, or will be. When confronted with a case in which doubt is *impossible*, where *nothing* could make us doubt, we are inclined to think that we can be *absolutely* certain how things are. But now doubt is precluded (ruled out) not by evidence but by grammar: no sense has been assigned to attaching the operator 'I doubt whether' to such a sentence in such circumstances. So the resultant sentence, for example 'I know I am in pain', is nonsense. Furthermore, if doubt is logically precluded, then so too is certainty. If it is impossible to doubt, *then it is likewise impossible to be certain*. For the negation of a nonsense is nonsense. If there is no such thing as doubting in such cases, then there is no such thing as being certain either. For certainty excludes doubt, but if doubt is already precluded by grammar, there is nothing for certainty to exclude, no conceivable work for it to do, and so no function for the operator ‘I am certain that …’ to fulfil. So, if there is no such thing as doubting whether one is in pain, then there is no such thing as being certain that one is in pain. Similarly, if there is no such thing as being in pain and not knowing it, then equally there is no such thing as being in pain and knowing it. Here ignorance is precluded by grammar, and there is no role for the operator ‘I know that’ to perform. But we are prone to think that if one cannot be in pain and not know one is, then surely whenever one is in pain, one *must* know it. And such knowledge seems singularly firm, clear and distinct, as Descartes averred. When someone denies that this is a case of knowing at all, we are almost irresistibly inclined to think that he is asserting that when people are in pain they don’t know that they are, that they are ignorant of their pain, do not know that they are in pain. What is so difficult here is to realize that when we are in pain, we neither know we are, nor are we ignorant of being in pain: *both* knowledge *and* ignorance are logically precluded. Similar reasoning applies to such verbs as ‘recognize’ or ‘be conscious of’. Since to be conscious of something in one’s perceptual field is to have one’s attention caught and held by it, one cannot be conscious of what one is *intentionally* attending to. But of course, one cannot not be conscious of it either – *both* are precluded by grammar. This insight is of pivotal importance in rejecting foundationalist accounts of empirical knowledge, in which the foundations of knowledge are held to be sense-datum statements of how things subjectively seem to one to be. The error is similarly responsible for standard accounts of self-knowledge as knowledge of how things currently are with one.

We are confused here not merely by failure to distinguish empirical, evidential, exclusion of doubt from logical or grammatical preclusion of doubt, but also by failure to apprehend the grammatical character of certain propositions that we conceive to be descriptions of *de re* necessities. So ‘Only I can have my sensations (mental images, experiences)’ appears to be a ‘synthetic a priori’ truth, or a metaphysical proposition. But all it is is a grammatical proposition in the guise of a description of an objective necessity. What it amounts to is no more than the *rule* that ‘My sensation’ = df. ‘the sensation I have’. A consequence of that is that ‘I have my sensation’, *a fortiori* ‘Only I have my sensation’ has no sense.

A third paramount guideline is to place little trust in categorial classification, but to examine each concept and its use in its own right. There are two reasons for this.

(a) Categorial terms, such as 'state', 'object', 'act', 'activity', 'event', 'process' are soft, not hard, let alone, as he himself had thought when writing the *Tractatus*, the hardest of the hard. They are not akin to indexed variables in a calculus in which the names, monadic predicates, and n-adic predicates that are their values are sharply defined. These categorials are vague and ill-defined. Moreover, it would be mistaken to assume that physical and psychological states are two species of a genus 'state', or that material objects and mental objects are coeval species of objects, and so forth. A mental state is *logically* unlike a physical state, being characterized, as Wittgenstein explained, by genuine duration as opposed to mere duration,. Furthermore a *mental state* is not to be assimilated to a *state of mind*, which is akin to a frame of mind rather than to a state. The danger, to which numerous philosophers succumb, is to employ such categorials *without investigating their application* either in general or in the particular case:

> How does the philosophical problem about mental processes and states and about behaviourism arise? — The first step is the one that altogether escapes notice. We talk of processes and states, and leave their nature undecided. Sometime perhaps we'll know more about them –we think. But that's just what commits us to a particular ways of looking at the matter. For we have a certain conception of what it means to learn to know a process better. (The decisive movement in the conjuring trick has been made, and it was the very one that seemed to us quite innocent.) – And now the analogy which was to make us understand our thoughts falls to pieces. So we have to deny the as yet uncomprehended process in the as yet unexplored medium. (PI §308)

One of the most influential categorial mischaracterizations of a psychological attribute in the last century was the claim that understanding is an activity or process of the mind – without examining the grammar of the verb 'to understand'. So philosophers (including the author of the *Tractatus*) and theoretical linguists committed themselves to a particular way of looking at the matter, and raised the question of how one can understand a sentence one has never heard before. For if understanding the utterances of others is a process or activity of the mind, then it seems legitimate to ask how it is done or engaged in. This question (raised in the *Tractatus*) was declared to be *the* pivotal question in philosophy of language and linguistics (Chomsky, Dummett). Strikingly, its sibling: 'How can one understand sentences one utters which one has never uttered or heard before?' was declared to be beyond the powers of the human intellect to solve.[1] But if understanding has a kinship to abilities rather than to processes, then the questions make no more sense than 'How does one do being able to play chess?'

(b) Many psychological predicates may occur in one context in a form that approximates one category, for example, a mental state, and in another context in a form that approximates another, for example, a disposition (e. g. 'She is very anxious about Tom', and 'She has a very anxious personality'). Some psychological attributes cannot be correctly subsumed under *any* categorial (e. g. belief, which is standardly and mistakenly said to be a mental state).[2] Others can be characterized only as *approximating* a given psychological category (as understanding approximates an ability rather than a process or activity).

The moral of all this is that much greater care needs to be taken over categorial classification, bearing in mind that the categories are themselves vague, that the concepts that one is trying to categorize are often categorially indeterminate, that some psychological attributes may not belong to any superior genus, and that some may have no more than a kinship with a given category (as understanding has a kinship with abilities, and that thinking has only a kinship with activities.

2. *Fundamental insights*

Delaying for a moment Wittgenstein's criticisms of classical conceptions of the mind and the mental, of the body and the behavioural, I shall turn first to Wittgenstein's most fundamental insights into the logico-grammatical structures he deals with in *Investigations* §§243-315. At the level of grand-strategy, as it were, one might reasonably select four unprecedented insights into the structure of our thought or conceptual scheme, general ways of describing the logico-grammatical relations between concepts of the mental and the somatic, and concepts of the psychological and the behavioural.[3]

---

[1] N. Chomsky, *Language and Problems of Knowledge* (MIT Press, Cambridge, Mass., 1988), chapter 5.

[2] Demonstrated in P. M. S. Hacker, *The Intellectual Powers: a Study of Human Nature* (Wiley-Blackwell, Oxford, 2013), ch. 5.

First, Wittgenstein realised (a) that the psychological is a logical category, not a metaphysical one; (b) that the mental (what we think of as 'the inner') is not only empirically connected with behaviour (what we think of as 'the outer') but also logically; (c) that the root of the problems surrounding psychological concepts lies with the inner rather than with the outer.

Philosophers, in particular early modern philosophers and their heirs, conceived of the mental as attributes of an immaterial substance whose identifying and defining property is thinking. Thinking subsumed everything of which one was conceived to be necessarily conscious *in foro interno* (Descartes, Locke) or of which it is necessarily possible for one to be conscious (Kant). But these distinctions and attributions do not rest on empirical investigation or discovery. Rather they are *metaphysical*, and so too was the category of the mental. Wittgenstein, however, held the mental to be a distinctive *logical* category of attributes. They are neither attributes of the mind nor attributes of the body, let alone attributes of the brain, but attributes of the living human being which have a distinctive array of *logico-grammatical features*, very different from those of the somatic attributes of human beings.

The asymmetry between first- and third-person psychological propositions had been noted by all philosophers who reflected on the psychological or mental. In general, it was the first-person case that seemed to be philosophically unproblematic. It seemed evident that one has an idea of what sensations, perceptions, emotions, and desires are merely by having them or experiencing them, and that one masters their names by associating a name with a given kind of experience. One possesses the concept of a sensation, perception, emotion or desire in virtue of experiencing the mental phenomenon and subsuming it under appropriate concepts. It is sufficient for having an idea or concept of thinking, believing, knowing, or understanding by experiencing these *mental states* and attending to them by means of introspection. What appeared to be problematic was applying the same ideas or concepts to others. There had to be some kind of empirical link between a psychological idea or concept of an experience or mental state of one's own, and the ascription of mental states and experiences to others. The question, it seemed, was whether

---

[3] The boundaries of the psychological are blurred. The relation between the psychological and the mental are unclear. The psychological and the mental are anything but logically uniform (e. g. immunity to subjective doubt and mistake applies to pains, but not to knowing or understanding). I shall disregard this below for the sake of simplicity of exposition, leaving it to the reader to make appropriate adjustments where necessary. Wittgenstein's discussions of the 'inner' by and large present it as roughly coinciding with Cartesian 'thoughts' or *cogitationes*.

(i) one generalizes to others from the link, which one observes in one's own case, between any given psychological phenomenon or state and one's behaviour in appropriate circumstances. Or whether

(ii) one's ascription of psychological predicates to others depends upon seeing an analogy between one's own inner response to a stimulus and the kind of behaviour this causes and the observable behaviour of others exposed to similar stimuli. Or whether

(iii) one empathetically puts oneself in another's shoes, imagining what one would feel or experience in their circumstances, and projects that onto the other person.

Wittgenstein, however, held that although there are doubtless innumerable empirical links between the mental and human behaviour, there are also *logical* links. Moreover, the latter are a precondition of the former. Behaviour (which is not merely colourless physical movements) in given circumstances, is a *logical criterion* of the mental.

Far from the ascription of mental attributes to others posing the salient problem to understanding the nature of the mental, it is the *conditions of the possibility* of first-person thoughts or utterances that are what is deeply problematic. Wittgenstein showed that resolving *this* problem provides the key to unlocking the nature of the mental in general. Neither first-person nor third-person sentence has conceptual priority over the other, and the mastery of one cannot be conceptually detached from mastery of the other. Both the concepts of introspection and of consciousness were misconceived in the philosophical tradition. The employment of psychological predicates in the first-person does *not* depend upon an inner sense, or on introspection, or on being conscious of what passes in one's mind. There is no facultative source of knowledge of immediate experience (indeed, there is no *knowledge* of immediate experience, although there is no ignorance either). Rather our utterances, avowals, and reports of immediate experience *have no grounds at all*. It is not merely that the answers to the question ‘How do you know that things are thus-and-so with you?’ are awry. We have no inner sense analogous to outer senses, introspection is misconstrued as looking into our mind to see what is going on there, and consciousness of how things are with us plays a quite different role from that of a source of knowledge). So it is *the question* that is awry. The new questions that now have to be confronted are: what are the conceptual preconditions for a human being to avow or assert *groundlessly* how things are with him? How is it possible for one’s utterance to have a special status, quite unlike the status of descriptions others may give of how things are with one? In the case of a first-person utterance, truthfulness guarantees truth, but sincere ascription of psychological attributes to others is no guarantee of truth. How can one say how things are with one without knowing how they are and indeed without being ignorant either? These questions now transform the scene completely, and the proposal is radical. But, as Wittgenstein observed towards the end of his life:

> If one doesn’t want to SOLVE philosophical problems – why doesn’t one give up dealing with them. For solving them means changing one’s point of view, the old way of thinking. And if you don’t want that, then you should consider the problems unsolvable. (LW II, p. 83)

Secondly, Wittgenstein solved the problem that he posed with two powerful moves. (a) He showed how the first- and third-person uses of psychological expressions are welded together into a single coin of conceptual currency; and (b) he located the use of these expression-types in the characteristic language-games to which they belong.

The criterion for knowing what these words mean, and hence the criterion for possession of the concepts they express involves mastery of both the groundless first-person use and of the well-grounded third-person use. The thought that one might know what the word ‘pain’, or ‘see’, or ‘sad’ means when used in application to oneself, and only *then* go on to master its application to others is incoherent. Groundless self-attribution is possible *only on condition of recognition of the criteria warranting other-ascription.* As Wittgenstein remarked in §288: ‘if I assume the abrogation of the normal language-game with the expression of a sensation, then I need a criterion of identity for the sensation, and then the possibility of error also exists’, i. e. if we assume the abrogation of the normal language-game by arguing that it is only from my own case that I know what the word ‘pain’ means, then we sever the conceptual connection between the concept of pain and pain-behaviour. But if we sever the connection between the concept of pain and the behavioural criteria for pain, then we deprive ourselves of our ordinary criteria of identity for pain. But we need a criterion of identity for pain if we are to possess a concept of pain at all. So we are forced into the supposition that a criterion of identity is furnished by reference to a private ostensive definition and a mnemonic image of a metal sample. However, if that were intelligible, then the possibility of correct and incorrect application of the definiendum would exist. And if that were so, it would be possible for me to make a mistake about whether I am pain or not, and it would be possible for me to doubt whether I am in pain, but not to be sure. But none of these are genuine possibilities. So this route is closed.

The groundless first-person use and the criterially grounded third-person use are inseparably welded together. But more has to be said to shed light on the first-person use, without which it may well appear to be mysterious or even magical in those cases in which it has immunity to doubt or mistake. And Wittgenstein did indeed have much more to say. If we locate first-person utterances in the language-games to which they belong, any appearance of mystery will evaporate. In the case of pain, we are asked to view a pain-utterance as grafted onto natural pain-behaviour of screaming in pain, crying out in pain, crying in pain. Its groundlessness is no more mysterious, no more open to question or doubt, than is an infant's scream of pain or crying with pain, or an adult's crying out in pain. Pain-utterances are acculturated extensions of natural pain-behaviour, and are themselves criteria for others to ascribe pain to the speaker. But, of course, pain-utterances occur in certain behavioural and circumstantial contexts – of painful stimuli, of injury or illness, of non-linguistic pain-behaviour. If we remove these, then a pain-utterance on an occasion ceases to be intelligible. Pain-stimulus, pain, pain-utterance, pain-behaviour, and pain-circumstances (injury, illness) belong together. There are two important qualifications: First, room must be made for reports no less than for utterances – and Wittgenstein does indeed make room for them. Secondly, each psychological attribute is a case in its own right. The elucidation of the concept of pain does not provide a template for all other psychological predicates. The concept of wanting, in the primitive language-game from which it grows, is not rooted in primitive utterances but in conative behaviour of *trying to get* in a context of perceived desiderata. Perceptual concepts are learnt as operators on observation-sentences in the context of questions, such as ‘How do you know?’, ‘Can you see …?’, as well as instructions, such as ‘Look! – what do you see?’, or ‘Listen? – what can you hear?’. The roots of expressions of intention are different again. They are learnt not as expression or reports on how things are with one, but as heralding actions – if one says ‘I’m going to …’, then one has to go on to … . No one learns how to use the verb ‘to dream’ and its cognates by identifying dreams and naming them. Nor does one say what one dreamt by reading a description off the facts. Rather, the child wakes up crying in fear and relates what he thinks he just underwent. We teach him to prefix to his utterance the operator “I dreamt that …’, on the understanding that what follows is not a true description of antecedent events. Rather, it is only an expression of residual impressions on waking.[4] And so on.

Thirdly, Wittgenstein had the Aristotelian insight that the intelligibility of ascribing a psychological attribute to a subject is mereologically constrained. It is, he saw, a widespread mistake among philosophers, psychologists and neuroscientists to ascribe psychological attributes to parts of a human being, rather than to the living human being as a whole. It is not one’s mind or soul that is conscious or unconscious; it is not one’s brain that sees or is blind; and it is not one’s frontal cortices that think or are thoughtless – it is the human being. Nor is it one’s body (the body one has) that walks or stands still, that talks or is dumbstruck – it is the human being (for walking or talking are not somatic features of a human beings). This is not an arbitrary stipulation. Nor is it merely an appeal to how we in fact talk (unless corrupted by bad philosophy and confused cognitive neuroscience). It is a reminder of the grammatical web within which psychological predicates are positioned. The meanings of psychological expressions are partly determined by the character of the behavioural criteria for their application in the third-person case. It is pain-behaviour in appropriate circumstances of pain-stimuli and injury or illness that warrants ascribing pain to others. It is the use of one’s eyes, their orientation and the manner in which one follows a moving object of vision, one’s movements among the objects in one’s environment and one’s descriptions of visibilia that warrant ascribing sight to other human beings. It is tears and lamentations, facial expression and mien in circumstances of grievous loss that warrant saying of someone that they are grief stricken. But neither the mind and the soul nor the brain exhibit pain-behaviour, see or fail to notice something, exhibit grief and sorrow, and the body one has is not a subject of these predicates precisely because they are not somatic characteristics.[5] Pointing out the mereological fallacy, and its cousin the homunculus fallacy, has far reaching consequences for philosophy of psychology, metaphysics and theology.

---

[4] One must bear in mind that these observations are not armchair learning theory, but logical investigations into what has to be mastered before a person can be said to possess the relevant concepts.

The mental is intrinsically connected with the behaviour of the subject, but it is possible to think without expressing one's thoughts, to perceive without manifesting one's perceiving, and to want something without showing that one does. The mind is empirically (contingently), but not logically (conceptually or grammatically) connected with the brain, even though it is true that we would have no mental powers without an appropriately functioning brain. It is precisely because of this that neuroscientific investigation into the correlations between brain events and processes, on the one hand, and psychological powers and their exercise, on the other, has to presuppose the grammar of our ordinary psychological expressions. For there can be no inductive correlation without independent identification of the relata, and there can be no identification of the mental without acknowledging the constitutive behavioural criteria for psychological attributes.

The fourth and final very general grammatical principle that Wittgenstein discloses is that the limits of the intelligible application of psychological predicates to a being are specified by reference to possibilities rather than actualities. It is intelligible to say *truly or falsely* of a creature that it is in pain or feels no pain, that it sees or is blind, that it desires or does not desire something, that it is pleased or indifferent, that it is angry or cheerful, not merely if it *actually* displays certain kinds of behaviour, nor only if it has a *disposition* to display a certain kind of behaviour in certain circumstances, but only if certain kinds of behaviour *lie within its normal behavioural repertoire*. That is to say, the condition of *intelligible* ascription of a psychological attribute to a creature is that it is *possible* for the creature (in the normal exercise of its unconstrained powers) to manifest such behaviour as warrants such ascription. It makes sense to ascribe pain to a being if and only if it is possible (logically, not circumstantially, possible) for that being to exhibit pain-behaviour. It makes sense to say of a creature that it perceives something in its environment if and only if perceptual-discriminatory behaviour is within its behavioural repertoire. It may see something, but not show that it sees, but it must be logically possible for it to manifest its seeing that thing. It does not make sense to say of trees and plants that they see or hear (even if they grow better to the sounds of music) *or* that they are blind and deaf, because nothing that they *can* do would show auditory discrimination and response. It does not make sense to say of a cat or dog that they now expect something to happen next Sunday, since although they can now exhibit expectation for something here and now, there is nothing that they *can* do that would show that they now expect to be fed next week. This principle has striking ramifications, which Wittgenstein reveals after the private language arguments, in his investigations into thought and thinking (PI §§316-62). For this principle shows its strength in the investigation of the limits of thought, and of the relation between thought and language. Applied to thinking and reasoning, it is clear that the limits of thought are the limits of the behavioural expression of thought. It makes sense to say of a creature that it thinks this or that only if some behaviour that lies in its behavioural repertoire would count as manifesting its thinking or having thought what it is said to have been thinking or having thought. But the only behaviour that manifests thinking (reasoning), drawing conclusions from premises, weighing reasons and making decisions on the basis of reflection, is linguistic behaviour.

---

[5] In distinguishing here between the body I am and the body I have (see P. M. S. Hacker, *Human Nature: the Conceptual Framework*, chap. 9, I am going beyond Wittgenstein. But that distinction is perfectly compatible with everything he says and, I believe, complements it.

This has dramatic implications. For it immediately illuminates venerable puzzles concerning what non-human animals can think or experience, and whether they can reason or act for reasons. It sheds no less light on the question of whether machines can think, or whether we shall be able, sometime in the future, to make machines that can think. It is not surprising that Wittgenstein exclaimed ‘How much one must be able to do before one can be said to think’. It has equally dramatic consequences for theology, both for the belief that we survive, bodilessly, after death and for the idea that there can be a supreme thinking being that has no body.

It is an obvious corollary of these four fundamental grammatical insights that the whole inner/ outer picture of human nature to which we are all so attracted is misconceived. Wittgenstein does not explicitly draw this conclusion in the *Investigations*. But in his very last writings on philosophy of psychology towards the end of his life, he was perfectly explicit. ‘The “inner’ is an illusion. That is: the whole complex of ideas alluded to by this word, is like a painted curtain drawn in front of the scene of the actual use of the word’ (LW II, p. 84). Of course, that does not imply that we do not all have an ‘inner life’, but only that Wittgenstein denies that the picture of the inner and the outer with all its ramifications gives us a correct idea of the use of our psychological vocabulary (cp. PI §305).

3. *Consequences and confusions*

Wittgenstein has shown that there is no such thing as a private language, that the phrase ‘a private language’ or more specifically ‘a language the words of which refer to the speaker’s own private experiences that only he can know’ is excluded from the language. This is not unlike a mathematician’s impossibility proof. The proof that in Euclidean geometry one cannot trisect an angle with a compass and rule in effect proves that the phrase ‘to trisect (or ‘a trisection of’) an angle with compass and rule’ is excluded from Euclidean plane geometry. There is no such construction, and the sequence of words does not describe a possibility. But it does not describe an impossibility either – for a logical or mathematical impossibility is not a possibility that is impossible. The sequence of words has been shown to have no sense – to be a nonsense. So too with the form of words ‘a private language’.

This, then, is the drama within which a variety of subplots are interwoven. In the course of his prolonged and convoluted arguments, Wittgenstein establishes exceedingly important conceptual truths expressed by grammatical propositions. Experience is not owned by its subject, a fortiori not privately owned. Present subjective experience is not known to its subject, a fortiori not privately known. There can be no such thing as a private ostensive definition. Rules for the use of an expression are essentially public, i.e. shareable. Mental images are not possible objects for comparison. Mental images cannot fulfil the role of defining samples. Psychological attributes are ascribable only to a living animal as a whole, not to its parts, nor to anything that is not a living animal. So they are not ascribable to minds or brains. Nor are they attributes of the body (with the exception of verbs of sensation). Machines are not subjects of cognitive attributes. In particular, they cannot be said to think.

These are important conclusions. To be sure they are grammatical, in Wittgenstein's idiosyncratic sense of the word. That does not derogate from their importance, for they determine the limits of what is intelligible within the framework of our conceptual scheme (just as the rules of chess determine the possible moves that can be made in any game of chess). Determination of the limits of what makes sense is hardly trivial, because what lies beyond those limits are not hitherto unimagined multiverses, but nonsense – that is to say: nothing. Of course, we can lay down new rules and change old ones. There is no commitment to linguistic conservatism in Wittgensteinian philosophy. But (i) recommending such changes is not the task of philosophy, but of the empirical and mathematical sciences. The task of philosophy is to clarify the conceptual scheme we have, not some other imaginable conceptual scheme. It falls to philosophy to disentangle the knots we inadvertently tie in the web of connected concepts, both in ordinary discourse and in the sciences. (ii) If one changes the rules, then one is no longer employing the same concepts but different ones, and no longer clarifying the concepts that gave rise to the problem that one is confronting, but replacing them without having eliminated the problem. (One can change the rules of chess and no longer have a queen, but then the whole game changes dramatically). While doing this may be profitable in the empirical sciences and mathematics, it can serve no purpose in philosophy.

The consequences of Wittgenstein's investigations into the issues associated with a private language are far reaching, both for philosophy, and for the natural sciences. Within philosophy, his arguments and conclusions have a direct bearing on the venerable controversy concerning primary and secondary qualities. For if there is no such thing as defining an expression by reference to a mental sample stored in the reproductive imagination, then names of secondary qualities cannot be defined or explained by reference to the contents of experiences. They are explained by public ostensive definitions or explanations that employ public samples of colours, sounds, tastes and smells, and of such tactile qualities as hot and cold, rough and smooth, wet and dry, and so forth. If that is correct, then the whole bizarre picture of the world we inhabit as colourless, soundless and odourless, eagerly advocated by physicists since Galileo, by psychologists and by cognitive neuroscientists, needs radical rethinking. Similarly, the causal theories of perception advanced by philosophers, psychologists, and cognitive neuroscientists rest on unstable and infirm conceptual foundations. The philosophical theories need to be abandoned, and the scientific theories need radical recasting with clean conceptual tools.

Similarly, philosophical theories concerning our knowledge of the mental attributes of others, theories that have infected the sciences of psychology and cognitive neuroscience, are to be rejected. The classical theory of the argument from analogy (Berkeley, Hume), empathetic projectionist theories of the nineteenth century (Theodor Lipps), and reductionist behaviourist theories of the first half of the twentieth century (Watson, Carnap, Feigl) can all be dismantled and assigned to the history of conceptual confusions. Contemporary psychological theories such as the so called 'theory of mind theory' and 'simulation theory' patently need to be uprooted. For both presuppose the intelligibility of grasping psychological concepts in their first-person application independently of grasping their third person use. This, by the way, is anything but a criticism of the product of language idling, since these misconceptions are hard at work in practical reasoning about psychological treatment of forms of mental impairment, such as autism.

Philosophical, theoretical and cognitive scientific investigations into natural languages and human speech are likewise affected by Wittgenstein's private language arguments (as indeed they are by his philosophy of language). No cogent reflections on language and speech, on linguistic understanding, and on the neural conditions that make our linguistic competence possible can afford to disregard Wittgenstein's conclusions. In particular, they have to take into account the incoherence of the following ideas (i) that the brain follows rules; (ii) that, as Chomsky avers, there are rules deeply buried beyond the reach of consciousness in the mind/brain (including, one must presume, the samples by reference to which sensible qualities are ostensively defined); (iii) that a human being can intelligibly be said to follow rules he has never heard of and could not understand were they introduced to him; (iv) that, as was originally suggested by Anne Treisman, there is a mental dictionary or lexicon in the brain (currently a prominent idea among eminent psycholinguists such as M. Coltheart) that correlate words with the concepts they name; (v) that concepts are storables.

I shall say little about the bearing of the private language arguments on cognitive neuroscience. I have discussed this issue at length in two books, co-authored with Professor Maxwell Bennett: *Philosophical Foundations of Neuroscience* (2003) and *History of Cognitive Neuroscience* (2008). Precisely because cognitive neuroscience investigates the neural structures and activities that make it possible for us to have and exercise our cognitive, cogitative and volitional powers, it is a science that straddles the behavioural, the psychological and the neural. Hence it cannot avoid the use of the concepts that define these domains. Consequently it cannot sidestep the conceptual problems that beset the logical relations between behavioural, psychological and neuroscientific concepts. These problems are manifold, and the mistakes made throughout the history of cognitive neuroscience and still made today are numerous and anything but superficial.

Finally, it is evident that Wittgenstein's remarks on the limits of the psychological has a direct bearing on empirical research into animal powers and their exercise. For the ascribability of a psychological attribute to a creature is constrained by the creature's behavioural repertoire. It makes sense to ascribe to an animal a psychological attribute only if that animal can behave in ways that constitute criteria for the ascription of that attribute. Consequently, to ascribe attributes of intellect or will to non language-using creatures is misconceived. This has far reaching consequences for the study of animals, for it implies that ascription of self-consciousness to animals (e.g. on the grounds that they can recognize their faces in a mirror) makes no sense, that ascribing reasoning to animals (e.g. on the grounds that they can solve problems) involves failure to understand both what reasoning is, and what doing something for a reason presupposes, and that ascribing to apes a theory of mind on the grounds that they stealthily conceal food from other apes who they see have seen their food is nonsense.

These methodological guidelines, these conceptual insights and cartography, and their accompanying dialectic – the associated logic of conceptual illusions, are a part of Wittgenstein's legacy to the 21st century. It is our task, and the tasks of our successor, to ensure that they are understood and to cultivate their application.[6]

---

[6] This paper is a fragment of a longer essay for the extensively revised new edition of volume 3 of the commentary, *Wittgenstein: Meaning and Mind* (forthcoming, Wiley-Blackwell, 2018).