

In Benedikt Paul Göcke and Ralph Stefan Weir eds. *From Existentialism to Metaphysics: The Philosophy of Stephen Priest*, (Berlin: Peter Lang, 2021) 105-122.

Private Language and the Mind as Absolute Interiority

Ralph Stefan Weir

For several decades, Stephen Priest has championed a picture of the mind or soul as a private, phenomenological space, knowable by introspection and logically independent of behaviour. Something resembling this picture once dominated Western philosophy, but it suffered a severe setback in the mid-twentieth century as a result of Wittgenstein's 'private language argument'. While Priest has written about the threat posed by Wittgenstein's argument to the picture of the mind that he favours, he has not explained how advocates of that picture should respond to Wittgenstein. The present essay takes up this challenge, defending the picture of the mind as a private phenomenological space against four lines of argument drawn from Wittgenstein's *Philosophical Investigations* § 243–315.

1. Priest and the Mind as an Absolute Interiority

The philosophy I met with as an undergraduate seemed, at the time, to come in three varieties. The first was dogmatically scientific. It aimed to make philosophy continuous with and, insofar as possible, indistinguishable from physical science. The second was more humanistic. It ignored physical science and drew on phenomenology, literature, history and wider culture. But it treated science and philosophy as nonoverlapping magisteria: philosophy need not serve as the handmaiden to science because philosophers do not try to describe the world in the way that scientists or traditional metaphysicians do. The third variety was not dogmatically scientific and it did not treat philosophy as radically discontinuous with science

either. It drew on both scientific and non-scientific sources and engaged in ambitious metaphysical theorising. But it had not been practiced since the eighteenth century.

As a graduate student, at Blackfriars Hall, Oxford, I discovered that I had been wrong. Stephen Priest's lectures were untouched by dogmatic scientism and drew heavily on non-scientific sources. Neither did they ignore science or avoid metaphysics. Rather, they engaged in philosophy in the ambitious, all-embracing way that was customary before Kant. They were regarded by many who attended as the best lectures on offer in Oxford at the time. I sought Priest as a tutor at the first opportunity and have incurred a great debt to him in the intervening period.

My undergraduate experience left me particularly perplexed about the philosophy of mind. I came away with the impression that the discipline was divided between two approaches, representing the two schools of philosophy still in operation: physicalism and Wittgensteinian behaviourism. Neither seemed to me even slightly plausible. Priest, I learned, was an unapologetic opponent of both.

An early version of Priest's response to the mind-body problem appears in his *Theories of the Mind* (1991). According to the account given there, thought and conscious experience are private, internal, nonphysical activities, undergone by subjects who are in an important sense, particular individuals (Priest 1991, pp. 210–222). Priest suggests that most of this is unmysterious, but he makes an exception for the individuality of subjects which is 'in a fairly precise sense, a "miracle"' (Priest 1991, p. 221). Priest has developed and amplified his position since that work. But some central features have remained the same. To understand the mind, we must posit a private phenomenological space—an 'absolute interiority' as Priest puts it—with which one is directly acquainted by introspection (see e.g. Priest 2012). Attending to this interior space is the starting point for Priest's phenomenology, metaphysics and theology.

Priest's picture of the mind as a private, interior space is an emphatic example of an approach that once dominated Western philosophy. In the mid-twentieth century, Wittgenstein's *Philosophical Investigations* subjected that approach to an immensely influential challenge, in the form of the 'private language argument'. This is not so much a single argument as a collection of arguments at *Investigations* § 243–315, purporting to show that the view of conscious experiences (or 'sensations', as they are usually called in the context of Wittgenstein's argument) as private, inner events, presupposes an incoherent theory of language. Priest describes the threat posed by the private language argument as follows:

The private language argument potentially has enormous eliminatory power in philosophy. If sound, it is a refutation not only of dualism, idealism (including solipsism), phenomenological philosophy and phenomenism but also renders senseless the posing of certain sceptical questions; for example, the suggestion that we cannot know that or what other people think and the suggestion that your experience may be utterly different from mine... Wittgenstein, if successful, must inflict severe damage on the positions of philosophers as diverse as Descartes, Locke, Berkeley, Hume, Schopenhauer, Husserl, Russell and Ayer. (Priest 1991, pp. 58–9)

Although Priest recognises the threat that the private language argument poses to the picture of the mind that he champions, he has not explained in detail how advocates of that picture should answer Wittgenstein. In what follows, I take up this challenge, responding to four lines of argument drawn from *Investigations* § 243–315. First I consider an argument at § 271 that might be taken to show that if sensations are logically independent of behaviour they cannot enter into the mechanism of sensation language. Secondly, I discuss an argument from

§ 293 that the possibility of interpersonal differences in sensations would leave no use for sensation terms in a publicly intelligible language. Thirdly I consider the claim, suggested by § 246 and § 302, that the privacy of sensations is incompatible with our ability to ascribe sensations to others. Finally I respond to the argument at § 256 ff that a private language is impossible.

2. Approaching Wittgenstein's Argument

An attempt to assess the significance of *Investigations* § 243–315 for present-day philosophical views such as Priest's faces two methodological challenges. First, interpreters debate whether *Philosophical Investigations* contains arguments of the kind found in traditional philosophy in the first place.¹ In the following I treat *Investigations* § 243–315 as presenting such arguments. And I assume that these arguments can be handled with the usual tools of analytical philosophy. In adopting this approach I make no argument for its interpretive superiority.² Rather, I take it as sufficient to justify the following discussion that *some* distinguished exegetes *do* testify to the presence of traditional philosophical arguments in *Investigations* § 243–315 and to the challenge that those arguments pose to approaches to the philosophy of mind such as Priest's.³

¹ Kahane, Kanterian, and Kuusela (2007, pp. 1–36) survey interpretive approaches to Wittgenstein. I have been guided mainly by Hacker (1986; 1990), Priest (1991) and Schroeder (2001; 2006).

² For such an argument see Glock (2007, p. 60).

³ For example, Hacker (1990, p. xvii), Priest (1991 pp. 57–9), Schroeder (2006, 202, fn. 26). Hanfling (2001, pp. 52–8) describes present-day philosophers of mind in general as suffering from a deleterious 'Wittgenstein amnesia'.

Secondly, even we assume that *Investigations* § 243–315 advances logical arguments concerning sensations and sensation language, it is not obvious what the target of these arguments is. Commentators frequently speak in broad terms of a ‘Cartesian picture’ or an ‘inner object view’ to which philosophers are drawn (cf. Hacker 1990, p. 15; Kenny 2006, p. 13; Schroeder 2006, p. 202, fn. 26). But characterisations of this picture tend to throw together a varied assortment of theses concerning sensations, sensation language and the philosophy of mind generally (e.g. Schroeder 2006, p. 203). This is unsatisfactory because it is unclear why the theses associated with the ‘Cartesian picture’ should stand and fall together. If they do not then a present-day philosopher might endorse many of the theses associated with that picture without being vulnerable to Wittgenstein’s arguments.

To avoid this problem I have selected for special attention three theses that are intuitive compelling, and that capture a significant proportion of the ‘Cartesian picture’ that is advocated by philosophers such as Priest, and which *Investigations* § 243–315 is supposed to refute. The three theses upon which I focus are:

- [1] Sensation terms are referring terms.
- [2] Sensations are logically independent from behaviour.
- [3] People are acquainted with their own sensations and nobody else’s.

By [1] I mean that a sensation term like ‘pain’ belongs to the same semantic category as names for physical objects and events. I suggest that the meaning of such terms is, in most contexts, captured by the conditions under which something falls in their extension. [2] could be formulated in several ways. I take it at least to entail that for any sensation *S* had by an agent *A* at a time *t* in the actual world *w*, there is some possible world *w*₁ that is the same as *w* with respect to behaviour, but it is not true that *A* has *S* at *t* in *w*₁. Two worlds, *w* and *w*₁, are

‘the same with respect to behaviour’ if and only if for any agent *A*, any behaviour, ϕ , and any time *t*, *A* exhibits ϕ at *t* in *w* if and only if *A* exhibits ϕ at *t* in *w*₁. [3] is intended to capture the idea that sensations are phenomenal states, known by introspection. It is the combination of [2] and [3] that makes sensations ‘private’. [1] is a linguistic precondition for [2] and [3] in that, if sensation terms are not referring terms, then there is no such thing as a sensation, private or otherwise.

[1]–[3] is a staunchly un-Wittgensteinian triad. Nonetheless I think it stands a good chance of surviving the arguments of *Investigations* § 243–315. Given the richness of the text it has been necessary to focus on what I take to be four central, lines of reasoning.⁴ I argue that [1]–[3] are at least safe from these.

3. Idle Wheels—PI § 271

It might be thought natural to begin by looking at Wittgenstein’s celebrated discussion of private language at § 256 ff. But that discussion is not obviously relevant to [1]–[3]. For these theses concern our *publicly intelligible* sensation terms. In this and the following two sections I therefore discuss arguments that *do* seem directly relevant to [1]–[3]. Then, in the final section, I return to § 256 ff and explain why the arguments of that section seem harmless to [1]–[3].

The first argument I examine appears at § 271 (cf. also § 270). Wittgenstein is discussing the meaning of publicly intelligible sensations-terms like ‘pain’. He asks us to imagine that the putative referent of the term ‘pain’ keeps altering due to a lapse in memory on the part of a speaker:

⁴ These approximate those selected for highlighted by Schroeder (2006, pp. 201–18).

“Imagine a person whose memory could not retain what the word ‘pain’ meant—so that he constantly called different things by that name—but nevertheless used the word in a way fitting in with the usual symptoms and presuppositions of pain”—in short he uses it as we all do. Here I should like to say: a wheel that can be turned though nothing else moves with it, it not part of the mechanism. (PI § 271)

Wittgenstein proposes that despite changes in the putative referent of the term ‘pain’ the agent’s linguistic behaviour might go on as usual. A similar scenario appears in Part II:

Always get rid of the idea of a private object this way: assume that it constantly changes, but that you do not notice the change because your memory constantly deceives you. (PI p. 207)

The scenarios described at § 271 and p. 207 suggest that if a sensation term refers to a ‘private object’, then that object could repeatedly change without this making any difference to our linguistic behaviour. On this basis, Wittgenstein suggests that the private object plays no role in the ‘mechanism’ of our public sensation language.

Do such scenarios pose a threat to [1]–[3]? In order to answer this question it is useful to speak in terms of possible worlds. Suppose a person *A* in the *actual* world *w* at a time *t* makes the first-person sensation ascription: ‘there is a sharp pain in my arm’. According to [1] sensation terms are referring terms. So a proponent of [1]–[3] will hold that if language is functioning successfully *A*’s utterance ‘sharp pain’ at *t* in *w* refers to some sensation *S* such that *A* has *S* at *t* in *w*. However according to [2] sensations are logically independent from behaviour. So it seems that a proponent of [1]–[3] must also hold that there is some possible world *w*₁ that is the same as *w* with respect to behaviour, and that it is nonetheless false that *A*

has S at t in w_1 . In w_1 , A 's linguistic behaviour proceeds just as it does in w , despite the fact that A 's utterance 'sharp pain' fails to refer to S in w_1 .

It seems then that a proponent of [1]–[3] is committed to the possibility of scenarios relevantly like those described at § 271 and p. 207. Whenever some person A successfully uses a term like 'pain' in a first-person sensation ascription, there is some possible world w_1 where A 's linguistic behaviour is just the same, despite the absence of the referent of 'pain' in w_1 . (We could explain A 's unchanged behaviour in w_1 by a lapse in memory as Wittgenstein proposes, but other explanations might also do.)

Nonetheless, the fact that a proponent of [1]–[3] is committed to the possibility of such scenarios does not show that they are committed to a view upon which the referent of the sensation term plays no role in the mechanism of sensation language. For Wittgenstein's mechanical metaphor is misleading. To see this consider the relationship between the external behaviour of a mechanical object, such as a watch, and its internal parts. For any part m in my watch, there is some possible world w_1 where my watch works just as it does in the actual world, despite m 's absence. For the absence of m could be made up for by a replacement; or by an alteration somewhere else in the mechanism, or (we are talking about logical possibility here) a miracle. But of course, this does not mean that m plays no role in the mechanism underlying my watch's behaviour. For although my watch's behaviour is *logically* independent of its parts, it is nonetheless *causally* dependant on them.

When we want to know whether some part plays a *causal* role in a mechanism, the correct question is not whether that mechanism *could* have gone on the same, in the absence of the relevant part, but whether it *would* have done so, were that item absent. And the normal way to answer *this* question is to consider, not whether there is *some* possible world where the part is absent and the mechanism behaves the same, but whether *at the closest* possible world where the part is absent, the mechanism behaves the same (cf. Lewis, 1973).

Proponents of [1]–[3] are committed to the existence of a possible world where the pain referred to by *A* is absent, but *A* still announces ‘there is a sharp pain in my arm’. But proponents of [1]–[3] are not committed to the claim that at *closest* world where the pain is absent, *A* nonetheless makes this announcement. For a world in which the absence of the pain is made up for by some further alteration, is further from the actual world than one where the pain is simply absent and *A* does not, therefore, make any mention of it.

If the scenarios described at § 271 and p. 207 pose a threat to the proponent of [1]–[3] it is not by showing that the putative referents of sensation terms play no role in the mechanism of language. Of course a proponent of the additional thesis that sensations cannot causally influence behaviour *would* be vulnerable to such an objection. And there is a genuine question about whether epiphenomenalism is consistent with our linguistic ability to refer to conscious experiences (cf. Chalmers 1996, pp. 201–3; Gomes 2005). But this is irrelevant to [1]–[3] since these theses do not entail epiphenomenalism.

Might the mere *possibility* of the scenarios described at § 271 and p. 207 threaten [1]–[3]? This seems doubtful. For suppose a person *A* in the actual world *w* refers to a physical object *O* at *t*. For example *A* says ‘there’s an oak tree in that field’. It is natural to suppose that there is a possible world *w*₁, that is the same as *w* with respect to behaviour, even though it is false that there is an oak tree in the relevant field at *t* in *w*₁. For example in *w*₁ *A* might suffer an hallucination. But this does not, I take it, compel us to conclude that *A* cannot actually refer to a physical tree that is logically independent of *A*’s behaviour. And if it did this would have nothing to do with the private, inner nature of trees.

4. Beetles—PI § 293

At § 293 Wittgenstein describes a famous thought experiment which is, in some respects, similar to the scenario described at of § 271.⁵

Suppose everyone had a box with something in it: we call it a “beetle”. No one can look into anyone else’s box, and everyone says he knows what a beetle is only by looking at his beetle.—Here it would be quite possible for everyone to have something different in his box. One might even imagine such a thing constantly changing.—But suppose the word “beetle” had a use in these people’s language?—If so it would not be used as the name of a thing. The thing in the box has no place in the language-game at all...the box might even be empty. (PI § 293)

Once again we are asked to imagine that the putative referent of a term changes, or is absented entirely, whilst linguistic behaviour continues unhindered.

However, the beetle scenario differs in at least two respects from those considered above. First, we are now asked to imagine not only *intrapersonal* but also *interpersonal* differences in the putative referent of ‘beetle’. Secondly, we are asked to suppose that ‘everyone says he knows what a beetle is only by looking at *his* beetle.’ This is meant to capture the idea that ‘it is only from my own case that I know what the word “pain” means’ (PI § 293a). (Stern (2007, p. 265) claims that the beetle argument targets *only* this thesis.)

So, leaving aside the intrapersonal changes in the contents of the box, we can consider whether the *interpersonal* differences in this scenario suggest a further argument against [1]—

⁵ Schroeder, 2006, presents § 271 and § 293 together under the title ‘The Idle Wheel Argument’. But it is possible to find two distinct arguments in these sections.

[3]. The passage quoted seems to suggest the following argument for the conclusion that 'beetle' can have no role on these people's language:

1. 'Beetle' refers to something in each person's box.
2. Nobody can look into anyone else's box.
3. (From 1 and 2) For any person *A*, *A* understands the word 'beetle' by looking into *A*'s box.
4. (From 3) For any person *A*, in order to share *A*'s understanding of 'beetle' one must look in *A*'s box.
5. (From 2 and 4) No two persons share their understanding of the word 'beetle'.
6. (From 5) The term 'beetle' has no use in these people's common language.

To assess the relevance of this argument for [1]–[3] we can replace 'beetle' with 'pain', and 'box' with 'mind' to produce the homologous argument:

- S1 'Pain' refers to something in each person's mind.
- S2 Nobody can look into anyone else's mind.
- S3 (From S1 and S2) For any person *A*, *A* understands the word 'pain' by looking into *A*'s mind.
- S4 (From S3) For any person *A*, in order to share *A*'s understanding of 'pain' one must look into *A*'s mind.
- S5 (From S2 and S4) No two persons share their understanding of the word 'pain'.
- S6 (From S5) The term 'pain' has no use in our common language.

The conclusion S6 is obviously false of the English word ‘pain’. So if the argument is valid we must reject either S1 or S2. However, this does not look possible for the proponent of [1]–[3]. For if we construe ‘looking into’ as introspection, and understand someone’s ‘mind’ to contain whatever items that person is acquainted with by introspection, S1 and S2 follow from [1] and [3].

If this is correct, and if the argument is valid, then at least one of the theses [1]–[3] must be false. Wittgenstein implies that we should drop [1], the claim that sensation terms are referring terms: ‘if we construe the grammar of the expression of sensation on the model of “object and name” the object drops out of consideration as irrelevant’ (PI §293c).⁶

I believe the argument given here constitutes a powerful objection to someone committed to the first three steps, S1–S3. However, while proponents of [1]–[3] must accept the premisses S1 and S2, it seems to be that they might resist the argument by questioning the move from S1 and S2 to S3. Why should the view that ‘pain’ *refers* to something with which people are acquainted by introspection only in their own case commit one to the thesis that people understand the meaning of the word ‘pain’ only through such introspection?

The obvious answer would appeal to the assumption in order to understand ‘pain’ one must be acquainted with the kind of sensation it refers to. However it is not clear why proponents of [1]–[3] should accept this. For [1] is supposed to commit them to the thesis that sensation terms are referring terms in the way that names for physical objects and events are. And we do not typically require anything like acquaintance, let alone by introspection, with the referents of terms for physical objects and events in order to understand them. For example, I understand the meaning of ‘pterodactyl’ ‘Leningrad’ and ‘plutonium’, but I am not acquainted with their referents.

⁶ I follow Hacker (1990, p. 2017) in rendering *Bezeichnung* ‘name’ here.

Perhaps it would be urged in reply that even if S3 does not follow strictly from [1] to [3], it comes with the intuitive territory. For there exist historical proponents of theses like [1]–[3] who appear to have endorsed S3. For example, Schroeder (2001, p. 177) attributes such a view to Locke (cf. Hacker 1990, pp. 94–5; Priest 1991, p. 59). But a proponent of [1]–[3] might counter that when people say they know the meaning of ‘pain’ only from their own case, the fundamental insight they are expressing is that they know *what pain is like* only from their own case. It is *this* that follows from [1] to [3]. Knowing what a sensation *S* is like is not a necessary condition for linguistic mastery of the term ‘*S*’.

We might say that understanding the term ‘pain’ stands to knowing what pain is like in something like the relation that understanding the term ‘water’ stands to knowing that water is H₂O. It is true that if you do not know that water is H₂O, you are missing a tremendously important item of information about the nature of water—you know only its accidents and are ignorant of its essence. But it does not follow that you do not understand the word ‘water’. Likewise, if you have mastered the term ‘pain’ but have never undergone a pain.

We might add that just as other people’s pains might be of a radically different nature to our own, so it might have turned out that the water in Japan is of a radically different nature to the water in England. Of course, proponents of [1]–[3] will also see an important disanalogy between water and pain here. For we can, with some effort and expense, check what the water in another region is like just as well as the denizens of that region, whereas we cannot check what other people’s pains are like just as well as the subjects of those pains. But this epistemological disanalogy does not yet appear to have the semantic consequences that Wittgenstein claims for it.

It seems, therefore, that whilst the beetle thought experiment presented at § 293 is suggestive of a powerful counterargument to a ‘Cartesian’ view comprising [1]–[3] *plus* S3, it does not present a challenge to [1]–[3] alone.

5. Ascribing Sensations to Others—PI § 246, § 303

Does the beetle scenario raise any further concern for [1]–[3]? In conjunction with § 246, it might. There, Wittgenstein points out that ‘if we are using the word “to know” as it is normally used... then other people very often know when I am in pain.’⁷ That is obviously correct. But it might be argued that the beetle scenario shows that proponents of [1]–[3] cannot accommodate this fact. For if nobody can look into anybody else’s box, it is not clear how they can know that it contain a beetle. It seems reasonable, therefore, to require that proponents of [1]–[3] explain how we can know about others’ sensations.

In order to explain this, I suggest that proponents of [1]–[3] can rely, as in section two, on the casual relationships between sensations and their behavioural effects, as well as their publicly observable causes. I will offer a (largely unoriginal) story about how, given [1]–[3] people can know of others’ sensations, and respond to two objections.⁸

A proponent of [1]–[3] who rejects epiphenomenalism will attribute (at least) two kinds of property to sensations: causal and phenomenal. So a proponent of [1]–[3] might

⁷ I do not discuss Wittgenstein’s subsequent claim that *I* cannot know that I am in pain. This does not seem directly relevant to [1]–[3] since [3] does not concern *knowledge that* one is in pain but acquaintance with what one’s pain is like. Hyman (2001) makes a compelling case against Wittgenstein’s claim that I cannot know that I am in pain (cf. however Hacker 2006).

⁸ Pagin (2000) offers a more detailed story of this sort.

class sensations solely in terms of the former, or in terms of the latter, as convenient. In order to keep these approaches distinct we might use '*F*-pain' for any sensation that stands in the appropriate nomic relations to pain causes and to pain behaviour. We might then say that for any person, *A*, an '*A*-pain' is any sensation that shares its phenomenal character with *A*'s *F*-pain. Because *F*-pain is defined by reference to its typical causes and effects, and *A*-pain by its relation to *F*-pain, proponents of [1]–[3] can tell the following story about how we know that others experience these.

First, it is part of the concept of *F*-pain that it is typically caused by events like illness and injury and typically causes behaviour like grimacing and howling. So publicly observable phenomena of this sort provide defeasible evidence that someone is in *F*-pain. This does not rule out sceptical scenarios. For typical causes of pain can fail to cause pain and typical effects of pain can have causes other than pain. But I take it the proponent of [1]–[3] is not obliged to rule out such sceptical scenarios, since these are not special to our knowledge of others' sensations, but are common to all empirical knowledge.

Secondly, since for any agent *A*, an *A*-pain is any sensation that shares its phenomenal character with *A*'s *F*-pain, evidence that *A* is in *F*-pain is equally good evidence that *A* is in *A*-pain. Hence, if [1]–[3] obtain, people other than *A* can still have empirical knowledge that *A* is undergoing *A*-pain.

Of course, if you do not happen to be *A* then this will not tell you what *A*-pain is *like*. And, there is a further question about the basis on which you might ascribe *A*-pain to somebody other than *A*. Applied to oneself this is the question, 'do *A*'s pains feel like my pains?' This is, of course, a widely discussed question amongst philosophers who endorse [1]–[3].⁹

⁹ See for example, Block (2007) for a relevant discussion.

This too, does not seem to be an insuperable problem however. For it is plausible that uniformity in the phenomenal character in one's own *F*-pains is evidence for at least some uniformity across the population. And it is plausible that uniformity in the neurological correlates of pain provide further evidence for phenomenal uniformity. So, there is likely to be a powerful, albeit defeasible, empirical case for ascribing *A* pains to human beings other than *A* who are undergoing *F*-pain, including oneself. Hence, if [1]–[3] obtain, we can still have empirical knowledge people other than *A* undergo *A*-pains.

I now consider two objections. The first is suggested by the following passage in the *Investigations*:

If one has to imagine someone else's pain on the model of one's own, this is none too easy a thing to do: for I have to imagine pain which I do not feel on the model of the pain which I do feel. (PI § 302)

Wittgenstein argues that one cannot imagine the pain of another person on the model of one's own. His reason for this has to do with the distinction between imagining someone else's pain, and imagining that 'I feel pain in some region of his body'.

If it is true that one cannot imagine another's pain on the model of one's own then it might be argued that it will be impossible to make sense of the idea of *A*-pain, except where you are *A*. For *A*-pain is defined in terms of the phenomenal character of *A*'s *F*-pain, and given [3] only *A* is acquainted with this. However I do not think this follows. For unless one has to be acquainted with what a sensation is like, in order to understand a sensation term, there is no reason why one should not understand and apply the term '*A*-pain' without being acquainted with what *A*-pain is like.

So, even if it *is* impossible to ‘imagine’ A-pain on the basis of one’s own pain, it is unclear why this imaginative deficiency should prevent one from understanding and applying the term ‘A-pain’. This can be backed up by noting that in other cases imaginative understanding seems superfluous for understanding and application of a term: for example, having never sampled *tsampa*, I cannot imagine what it tastes like. But I have no trouble understanding the term ‘the flavour of *tsampa*’, and I have empirical knowledge that many Tibetans are currently enjoying the flavour of *tsampa*.

A further objection to the move I have suggested is voiced by Schroeder. He proposes that it is not open to the ‘inner-object theorist’ to propose that sensations have behavioural criteria for the following reason:

With that criterion in place... the idea that in spite of that behaviour—by reference to which the word ‘pain’ has been given its meaning (PI § 244)—there might never be any pain becomes inconsistent. (Schroeder 2006, p. 209)

If this is correct then it looks as though by defining *F*-pain and (hence, indirectly) A-pain in terms of behaviour, proponents of [1]–[3] must give up [2], the claim that sensations are logically independent of behaviour. For if sensations are logically independent of behaviour then there is a possible world that is the same as the actual world with respect to behaviour, but where there is no pain.

However, it is not clear to me where the inconsistency to which Schroeder refers is supposed to lie. Schroeder defines Wittgenstein’s notion of a ‘criterion’ as follows: ‘if it is in the very meaning of a term ‘*F*’ that some phenomenon is (good though not infallible) evidence for the presence of *F*, then that phenomenon is a criterion of *F*.’ On *that* definitions, A’s pain-behaviour is a criterion of both A’s *F*-pain and A’s A-pain. But the fact that *G*s are in

fact good but fallible evidence for *F*s does not make it *inconsistent* to suppose that there might have been *G*s but no *F*s.

Perhaps Schroeder is using a more stringent notion of criteria than his definition entails, such that if *G*s are a criterion of *F*s there could be no world with some *G*s but no *F*s. But proponents of [1]–[3] do not require a criterion of this sort in order to explain our empirical knowledge of other's sensations. For the weaker evidential relationship described in this section is sufficient—insofar as anything is—for empirical knowledge.

6. Private Language—PI § 256 ff

In the preceding sections I have defended [1]–[3] against several arguments from *Investigations* § 243–315 concerning our publicly intelligible sensation terms. I now consider the passages concerning the possibility of sensation terms that are *not publicly intelligible*.¹⁰ The idea of a private language is introduced as follows:

But could we imagine a language in which a person could write down or give vocal expression to his inner experiences...for his private use?—Well, can't we do so in our ordinary language?—But that is not what I mean. The individual words of the language are to refer to what can only be known to the person speaking; to his immediate private sensations. So another person cannot understand the language. (PI § 243)

¹⁰ Canfield (2001, p. 38) distinguishes a now unpopular memory scepticism interpretation of these passages; Kripke's rule-following interpretation; and the 'new orthodoxy' originating with Kenny (1976, pp. 151–2). My discussion is based on interpretations of the third variety.

At § 256 Wittgenstein considers whether someone could introduce such a language by associating a certain sensation with the sign ‘S’ and keeping a diary of its recurrences. He emphasises again that the imagined language is one that ‘only I myself can understand’. This requirement is crucial to three objections that Wittgenstein raises to such a the scenario.

Wittgenstein’s first objection is that ‘S’ cannot really name a *sensation*. ‘For “sensation” is a word of our common language, not one intelligible to me alone’ (PI § 261). Priest explains:

Wittgenstein’s point is that this would be impossible for someone who did not already have a command of a public language within which ‘S’ could be allocated a role as the name of a sensation. (Priest 1991, p. 62)

Likewise, Hacker suggests that the point here is that an ostensive definition ‘presupposes the grammatical category of the definiendum’ and that such a category could only be supplied by a term of our publicly intelligible language (Hacker 1990, p. 99). This is not permitted:

The private language theorist is not at liberty to help himself to expressions which are linked to behavioural manifestations of the mental, on pain of relinquishing his claims to the privacy of his putative private language. (Hacker 1990, p. 103)

Wittgenstein’s second objection is that we should not ‘consider it a matter of course that a person is making a note of something when he makes a mark... for a note has a function, and this “S” so far has none’ (PI § 260). According to Schroeder the private diarist cannot give the term ‘S’ a function because:

As soon as we can see the point of registering certain occurrences, we will have some notion of what kind of occurrences they are... So the meaning of 'S' will no longer be entirely incommunicable. (Schroeder 2006, p. 216)

With no grammatical category and no function, it seems that the diarist must merely 'impress' on himself the connection between 'S' and the relevant sensation. Wittgenstein's third objection is that this would be insufficient to define a name:

But what is this ceremony for?... "I impress it on myself" can only mean: This process brings it about that I remember the connexion right in the future. But in the present case I have no criterion of correctness. One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about 'right'. (PI § 258)

It is crucial to these arguments that the diarist be committed to the incommunicability of the meaning of 'S'. For otherwise, there should be no worry about relinquishing this incommunicability by using publicly intelligible words to determine the grammatical category of the definiendum; no worry that in acquiring a function, 'S' ceases to be 'entirely incommunicable'; and no need for the ostensive definition to rely merely on 'impressing' upon oneself a connection between 'S' and the sensation it is to name. If this is correct, then supposing that proponents of [1]–[3] are not committed to the possibility of giving publicly unintelligible names to sensations, Wittgenstein's objections at §§ 258–61 will pose no threat to their position.

The only direct reason proffered for the incommunicability of the diarist's language is that its words 'are to refer to what can only be known to the person speaking' (PI § 243). But this seems once again to rely on the assumption that acquaintance with the referent of a sensation term is necessary in order to understand the meaning of that term. I have already suggested that proponents of [1]–[3] reject this assumption. Proponents of [1]–[3] will then be free to say that people other than the private diarist could understand the meaning of 'S', despite the fact that they will be unacquainted with what *S* is *like*, just as I understand the meaning of 'the flavour of *tsampa*' even though I am unacquainted with the flavour of *tsampa*.

Another way *Investigations* § 243–315 might show that proponents of [1]–[3] are committed to the possibility of an incommunicable sensation language would be by adducing arguments to the effect that any term of which [1]–[3] holds will itself be publicly unintelligible. The arguments discussed in previous sections of this essay, the *Beetles* argument of § 293 in particular, might be thought to support this conclusion. However, I have maintained that these arguments are unsuccessful.

7. Priest and the Semantics of Sensation Terms

It seems that without further argument Wittgenstein's remarks concerning private language at §256 ff are not relevant to [1]–[3]. This is good news for proponents of the picture of the mind championed by Priest. It also good news for Priest's project of using phenomenology as a starting-point for metaphysics, insofar as the communicability of insights concerning phenomenal states is a necessary condition for that project.

In defending [1]–[3], however, I have found it useful to reject the further claim that to understand the meaning of a sensation term, like 'pain', it is necessary to have undergone pain. Perhaps Priest would be hesitant to reject this claim. For it might be thought to follow

from his suggestion that ‘the language of exteriority is inadequate to the inner life, because experience is necessary for understanding first-person singular psychological ascriptions’ (Priest 2012, p. 312).

If so, I would repeat that when we say that one can understand the meaning of ‘pain’ only by being acquainted with pain, the fundamental insight that we are expressing is that one can only understand what pain is *like*—one can only know the *nature* of pain—by being acquainted with pain. This fundamental insight is not at issue. I would add that for this reason it is in fact true, after all, that there is a sense in which somebody who has never undergone pain cannot *fully* understand of the meaning of ‘pain’. I briefly explain.

The meaning of a referring term is (at least in part) its intension. An intension is a function from the way a world is to an extension. So one grasps the meaning of a term *T* to the degree that one is able to determine the extension of *T* at a world. The information one needs to determine the extension of a term at the actual world is distinct from the information one needs to determine the extension of that term across possible worlds. For to determine the extension of a term *T* at the actual world, you only need to know of some *accidental* property that uniquely picks *T*s in the actual world, whereas to determine the extension of *T* across possible worlds, you need to know of some *essential* property that uniquely out *T*s across possible worlds. For example, to pick out water at the actual world, all you need to know is that water is the clear drinkable liquid in lakes, rivers and so on. But to pick out water across possible worlds, you need to know that water is H₂O.

In the same way, to pick out pain in the actual world—with some competence—all you need to know is that pain is a state that is typically caused by illness and injury and typically causes grimacing and howling. But to pick out pain across possible worlds, you need to be acquainted with phenomenal essence of pain: you need to have undergone pain. To

that extent, somebody who has not undergone pain can only have a relatively shallow understanding of the intension, and hence the meaning, of 'pain'.

There is more that could be said here. It would be possible, I believe, to construct a detailed theory of the meanings of sensation terms, using the resources of two-dimensional semantics, that would fully elucidate the mistakes that underlie the private language argument. But I put off that project for a future occasion.¹¹

References

- Canfield, J. (2001). Private Language: The Diary Case. *Australasian Journal of Philosophy*, 79, 377–394.
- Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Glock, H. (2007). Perspectives on Wittgenstein: an intermittently opinionated survey. In G. Kahane, E. Kanterian, and K. Oskari (Eds.), *Wittgenstein and His Interpreters* (pp. 37–65). Wiley-Blackwell.
- Gomes, G. (2005). Is Consciousness Epiphenomenal? *Journal of Consciousness Studies*, 12, 77–79.
- Hacker, P. (1986). *Insight and Illusion* (Rev. ed.). Clarendon.
- Hacker, P. (1990). *An Analytical Commentary on the Philosophy of Wittgenstein* (Vol. 3). Blackwell.
- Hacker, P. (2006). Of knowledge and of knowing that someone is in pain. In S. Pichler and S. Säätelä (Eds.), *Wittgenstein: The Philosopher and His Works* (pp. 244–76). Ontos Verlag.

¹¹ I am grateful to John Hyman for providing comments on an early draft of this chapter.

- Hanfling, O. (2001). Consciousness: the last mystery. In S. Schroeder (Ed.), *Wittgenstein and Contemporary Philosophy of Mind* (pp. 36–59). Palgrave.
- Hyman, J. (2001). Knowledge and self-knowledge. In S. Schroeder (Ed.), *Wittgenstein and Contemporary Philosophy of Mind* (pp. 171–93). Palgrave.
- Kenny, A. (2006). *Wittgenstein* (Rev. ed.). Blackwell.
- Lewis, D. (1973). Causation. *Journal of Philosophy*, 70, 556–67.
- Pagin, P. (2000). Sensation terms. *Dialectica*, 54, 177–199.
- Priest, S. (1991). *Theories of the Mind*. Penguin.
- Priest, S. (2012). The unconditioned soul. In B. P. Göcke (Ed.), *After Physicalism* (pp. 295–334). University of Notre Dame Press.
- Schroeder, S. (2001). Private language and private experience. In H. Glock (Ed.), *Wittgenstein: A Critical Reader* (pp. 174–98). Blackwell.
- Schroeder, S. (2006). *Wittgenstein: The Way Out of the Fly Bottle*. Polity Press.
- Stern, D. (2007). The uses of Wittgenstein’s beetle: philosophical investigations §293 and its interpreters. In G. Kahane, E. Kanterian, and K. Oskari (Eds.), *Wittgenstein and His Interpreters* (pp. 248–68). Wiley-Blackwell.
- Wittgenstein, L. (1958). *Philosophical Investigations* (2nd ed., G. E. M. Anscombe trans.). Blackwell.