# Scepticism About Scepticism or the Very Idea of a Global 'Vat-Language'

Genia Schönbaumsfeld[1]

## Abstract

This paper aims to motivate a scepticism about scepticism in contemporary epistemology. I present the sceptic with a dilemma: On one parsing of the BIV (brain-in-a-vat) scenario, the second premise in a closure-based sceptical argument will turn out false, because the scenario is refutable; on another parsing, the scenario collapses into incoherence, because the sceptic cannot even save the appearances. I discuss three different ways of cashing out the BIV scenario: 'Recent Envatment' (RE), 'Lifelong Envatment' (LE) and 'Nothing But Envatment' (NBE). I show that RE scenarios are a kind of 'local' sceptical scenario that does not pose a significant threat to the possibility of perceptual knowledge as such. I then go on to consider the more radical (or global) LE and NBE scenarios, which do undermine the possibility of perceptual knowledge of an 'external' world by positing that it is conceivable that one has always been envatted and, hence, trapped in a 'global' illusion. I start by assuming that we could be in such a scenario (LE or NBE) and then spell out what we would need to presuppose for such scenarios to be capable of being actual. Drawing on some central insights from Wittgenstein's anti-private language considerations, I show that the truth of a global scepticism would presuppose the possibility of a private 'vat-language', a notion that cannot be rendered coherent. But, if so, then neither can the sceptical scenarios that presuppose such a conception.

**Keywords** Scepticism · Closure-based sceptical argument · Local sceptical scenario · Global sceptical scenario · Recent envatment · Lifelong envatment · Private language · Wittgenstein

## 1 Introduction

Scepticism, in the contemporary literature, is most commonly motivated by way of appeal to a 'closure'-based sceptical argument. Such arguments run as follows:

(BIV1) If I know I have two hands, then I know I'm not a brain-in-a-vat (BIV).

((BIV1) is motivated by the closure principle, 'If S knows that p, and S competently deduces q from p thereby coming to believe that q on this basis while retaining her knowledge that p, then S knows that q.,[1]). But, it seems,

(BIV2) I don't know I'm not a BIV.

So,

(BIV3) I don't know I have two hands.

✉ Genia Schönbaumsfeld
  gmes@soton.ac.uk

1  Department of Philosophy, Faculty of Arts and Humanities, University of Southampton, Southampton, UK

In other words, if knowledge is 'closed' under known entailment, it appears that being able to rule out that one is a BIV has epistemic priority: the onus seems to be on the anti-sceptic to show that we do have perceptual knowledge of the world around us. In previous work (Schönbaumsfeld 2016), I have extensively criticized the notion that the closure principle actually does any sceptical (or anti-sceptical) work, and that it, therefore, cannot be used to motivate scepticism at all. I will not revisit this terrain here. Rather, I will focus on (BIV2), in order to determine what the 'scope' of the commonly invoked BIV scenario is actually supposed to be—a question frequently not addressed in the literature. My overall purpose is to present the sceptic (or our sceptical *alter ego*) with a dilemma: On one parsing of the BIV scenario, (BIV2) will turn out false, because the scenario is refutable, on another parsing, the scenario collapses into incoherence, because the sceptic cannot even save the appearances. So, whichever horn of this dilemma the sceptic

---

[1]  I am here following Pritchard (2012). See also Hawthorne (2005).

grasps, the closure-based sceptical argument will fail, as it can be blocked at (BIV2).

My strategy, in this paper, is as follows. In the next two sections, I discuss three different ways of cashing out the BIV scenario: 'Recent Envatment' (RE), 'Lifelong Envatment' (LE) and 'Nothing But Envatment' (NBE).[2] I will show that RE scenarios are a kind of 'local' sceptical scenario that does not pose a significant threat to the possibility of perceptual knowledge as such. I then go on to consider the more radical (or global) LE and NBE scenarios, which do undermine the possibility of perceptual knowledge of an 'external' world by positing that it is conceivable that one has always been envatted and, hence, trapped in a 'global' illusion. I start by assuming that we could be in such a scenario (LE or NBE) and then spell out what we would need to presuppose for such scenarios to be capable of being actual. Drawing on some central insights from Wittgenstein's anti-private language considerations, I show that the truth of a global scepticism would presuppose the possibility of a private 'vat-language', a notion that cannot be rendered coherent. But, if so, then neither can the sceptical scenarios that presuppose such a conception. In the penultimate section, I consider some objections before offering an overall assessment.

## 2 Recent Envatment Scenarios

Let's start by asking, what are recent envatment scenarios? Well, similarly, to other local sceptical scenarios, such as Dretske's famous zebra example (Dretske 1969), a recent envatment scenario casts doubt on particular, 'local' perceptual knowledge claims, rather than attacking the very idea that perceptual knowledge is ever possible (which is what LE and NBE scenarios do). Thus, local sceptical scenarios, including recent envatment scenarios, give one grounds for doubt about particular perceptual cases (or larger swathes of them), while leaving intact the thought that perceptual knowledge is sometimes available. In a 'lifelong envatment' scenario (LE), on the other hand, I have always been a BIV and have, therefore, never been in touch with an 'external' world, but only with electrode stimulations, whereas in a 'nothing but envatment' scenario (NBE), we are supposing that there may be nothing at all except me and my vat. This is a kind of metaphorical BIV scenario, analogous to Descartes' Evil Demon argument, where we are discounting the notion that someone must have set up the vat and is continuing to service it.

In order to get RE scenarios off the ground, we need to make a set of quite determinate assumptions, which undermine the sceptical scenario's virulence (as the sceptic should not have to make claims that we need to accept). For example, we must suppose that there is a world, broadly like our own, in which the vat containing my brain exists; that there are evil scientists (or robots or aliens or what-have-you) who have abducted me, harvested my brain and subsequently envatted it; that it is scientifically possible to separate brains from bodies without killing off the brain; that I might find out about my previous envatted state by, for instance, having my brain 'reinserted' into my body; that I 'am' my brain, and so on. Such a literal fleshing-out of the sceptical scenario is precisely what turns it into a 'local' case: if I imagine that I might be the victim of such a predicament, this does not imply that anyone else is, or, indeed, that the 'external' world as such does not exist (since, as we have just seen, a world containing vats, brains and evil scientists is in fact presupposed by this form of scepticism). Neither does it imply that perceptual knowledge of the world as such is impossible, since we are presupposing that I did, in fact, possess such knowledge *before* I was envatted. All it entails is that, for as long as I am a BIV, most of what I believe about the world is false.[3]

It is important to note here, however, that not *all* of my beliefs about the world will be false. Since I have formerly interacted with the world in the normal way, many beliefs I hold about the state of this world will survive envatment. For example, it will remain true that London is the capital of England, that Great Britain is a small island off the Continent, that cats are mammals etc. RE will, therefore, mainly pose a threat to my current perceptual beliefs—e.g. about there being a computer in front of me, or a tree, or a person etc. This is worth noting, as this fact poses a major stumbling block to any attempt at generalizing local sceptical arguments to get one the 'global' conclusion that if one can sometimes—or very often—be wrong, it is conceivable that one could always be wrong and, hence, know nothing about the world around one.

In this respect, there is a similarity between RE scenarios and arguments from perceptual illusion. For, just as RE scenarios are insufficient to motivate the possibility of 'global' illusion, so the possibility of 'local' perceptual error does not imply systematic or 'global' perceptual unreliability. I am only able to determine, for instance, that when I look at a square tower from a distance, it will appear round, because I can trust my perception that from close-by it looks square, and there is a scientific explanation available that can tell

---

[2] I first drew these distinctions in Schönbaumsfeld (2019).

[3] In this respect, a RE scenario is like Descartes' dreaming argument: for as long as we are asleep, we cannot know anything about the world (even if it seems otherwise to us), but it does not automatically follow from this that, therefore, I cannot know anything when I am *not* asleep. For more on the dreaming argument, see Schönbaumsfeld (2019).

me why it nevertheless appears round from some way off. If perception in general were deceptive, I could not make the judgment that perceptual appearances are sometimes misleading. All I could do would be to report, for example, that at time t1 I have the impression that 'thing 1' that I see is square, while at time t2 I have the impression that 'thing 2' that I see is round. And since 'thing 1' and 'thing 2' might, for all I know, be different things, I could not even conclude that one perceptual experience might be an accurate representation of the way things are, while the other might not. But if I cannot make this judgement, I am similarly unable to conclude that at time t2 I am being misled, for my perceptual experience at t2 would only be misleading if it were an experience of the same thing that I encountered at t1, so that my reports at t1 and t2 would turn out to be in conflict with each other. As long as I have no reason for assuming that my perceptual experiences of 'thing 1' and 'thing 2' are in fact experiences of the same physical object, however—which, of course, I would not if perception were generally defective (for then I would have no grounds for trusting one report more than the other; I would rather have to assume they are equally misleading)—there is no way of determining that a perceptual 'illusion' has in fact occurred. All I could say is that I am having different perceptions at different times, but this, of course, does not suffice to allow me to infer that at time t2 I was misled, and, hence, that perceptual errors are possible.[4] Consequently, far from showing that perception, in general, is defective, the possibility of perceptual error actually presupposes that perception is generally in good working order. Hence, 'aggregate arguments' based on the possibility of perceptual error fail: I am not entitled to infer that because I can sometimes be wrong ('local case'), it is possible that I could always be wrong ('global case'). In other words, 'local' error cannot be 'aggregated' to give us 'global' error. For this reason, global sceptical scenarios cannot be constructed out of 'local' ones.

Consequently, just as we need to presuppose the existence of an 'external' world that is broadly like our own, in order to get a RE scenario off the ground, so perception must generally be taken to be reliable, if an 'argument from illusion' is to be constructed.[5] What makes 'local' sceptical scenarios possible, therefore, is the very thing that they are, ironically, drafted in to undermine: a background of generally veridical perceptual experience. In this respect, the RE scenario is just a souped-up, 'sci-fi' version of the more pedestrian arguments from illusion and cannot, for this reason, achieve more than they can: to show that perceptual error is possible (that human beings are fallible).

Now, one might, perhaps, wish to object here that even if the foregoing is correct, it is at least *logically* possible that I am currently envatted and only vainly imagining that I am writing a paper (or, to speak with Descartes, that I am currently dreaming that I am doing this). Although this is true, it is important to bear in mind that the mere fact that I am able to conceive of such a state of 'envatment' does not give me a real reason to suppose that I might *actually* be the victim of such a situation. Why not? Because, in the absence of a general argument designed to undermine the possibility of perceptual knowledge per se, I have no real grounds for thinking that the imagined scenario might be the actual one. For example, the state of current science is not yet advanced enough to make such scenarios even empirically possible; there is no evidence of alien or robot activity etc. In short, the thought that, despite all the evidence to the contrary, I might nevertheless be a BIV, purely because this is logically conceivable, is not yet to offer a reason *in favour* of this scenario: logical conceivability alone is not a form of evidence.[6]

But if this is right, then contrary to what is often assumed, there is no 'local' route to a 'global' scepticism. For as we have just seen, neither arguments from perceptual illusion, nor RE scenarios, are by themselves sufficient to entail the conclusion that perceptual experience in general is never to be trusted. So, if the scenario in (BIV2) is taken to be a RE scenario, we can know that (BIV2) is false by relying on ordinary empirical evidence.[7]

---

[4] What if we suppose that, at t1, I have the impression that thing 1 that I see is square, at t2 I have the impression that thing 2 that I see is round, and between t1 and t2 I have the impression that thing 1 hasn't changed shape or been replaced—wouldn't we then have to conclude that a perceptual illusion has in fact occurred? We might indeed conclude in such a case that a perceptual illusion has occurred (as we might also in the other scenarios), but the relevant point is that we can only draw this conclusion because we are taking our impression that thing 1 has not changed between t1 and t2 (or been replaced) to be veridical, which confirms what I was trying to show: namely, that we cannot conclude that a perceptual error has occurred unless we are willing to grant that some of our sense-perceptions can be veridical. If we don't, then we certainly cannot grant that we can take our impression that thing 1 hasn't changed between t1 and t2 at face value (and hence we could not conclude that a perceptual error has occurred).

[5] This has nothing to do with the question of whether or not 'local' error is always detectable. It may not be detectable in any given case. The point is rather that if I start with the assumption that *all* perceptions are unreliable, then I can *never* formulate an argument from illusion, as such an argument presupposes that some of our perceptions are reliable (e.g. that the tower has not changed shape and now looks square). Hence, I cannot use an argument from illusion in order to undermine the very thing the argument itself presupposes: the reliability of some of our sense-perceptions. So I am not endorsing the false principle that if it is not possible to tell that p, it is not possible that p.

[6] Compare Pritchard (2012) and his contention that sceptical hypotheses must be motivated. RE is not motivated. In fact, all the available empirical evidence speaks against it.

[7] This conclusion is strengthened, if we accept a form of epistemological disjunctivism that allows us to hold that while it is not possible to know that one has been recently envatted if one has, it does not follow that it is not possible to know that one has not been envatted, if one has not.

## 3 Global Envatment Scenarios

If the argument in section 2 is correct, what is it, then, that makes us worry about global sceptical scenarios, such as LE and NBE, if local sceptical concerns give us no real reason to do so? For, clearly, most people don't entertain such global sceptical scenarios because they strike one as probable, or, in any way, as evidentially motivated. So, perhaps such scenarios appear compelling, because they articulate a deep-seated anxiety human beings appear to have about their whole relation to the world: How do we know that we are ever in touch with something genuinely objective and 'external', and not just locked into the 'cage' of our own perspective? In other words, rather than being, in any way, evidentially or argumentatively motivated, LE and NBE scenarios give expression to the concern that there might be a systematic mismatch between everything we collectively believe to be the case about the world, and the way the world actually is. And, if this were in fact so, then, even if it seemed otherwise, we would never have knowledge of anything; nor could we ever find out (even in principle) whether such a scenario obtained, as any form of evidence one could appeal to would itself be part of the 'grand illusion'. Stroud gives potent expression to this fear:

> We are confined at best to what Descartes calls 'ideas' of things around us, representations of things or states of affairs which, for all we can know, might or might not have something corresponding to them in reality. We are in a sense imprisoned within those representations, at least with respect to our knowledge. Any attempt to go beyond them to try and tell whether the world really is as they represent it to be can yield only more representations, more deliverances of sense experience which themselves are compatible with reality's being very different from the way we take it to be on the basis of our sensory experiences. There is a gap, then, between the most that we can ever find out on the basis of our sensory experience and the way things really are. In knowing the one we do not thereby know the other. (Stroud 1984: 32)

It is this angst that fuels what I'm going to call the Cartesian Picture of our epistemic situation, which conceives of the 'external' world not as something that is, in principle, directly present to the senses, but rather as something that we can only know about by dint of making inferences from our mental states or apparent 'perceptual' experiences.[8] And

once such a picture is in place, it indeed becomes hard to see how we could so much as know about this world beyond the 'veil of appearances', including whether it exists. Consequently, LE and NBE scenarios are merely ways of making vivid the anxiety that we are somehow not in direct touch with the objects that make up the physical world, as we might be locked into a global 'vat-world', where all is illusion.

Although this picture continues to be at work in much contemporary epistemology,[9] it is neither presupposition-free, nor compulsory. In the remaining sections of this paper, I will explore one serious problem that undermines it (but there are others[10]), namely, that this picture just takes for granted that it is possible to have knowledge of appearances (of how things seem to one perceptually), even though these appearances may never be appearances *of* anything (as would be the case if one were in a LE or NBE scenario).[11] Not only does such a view imply that appearances can, in some sense, 'subsist' on their own (i.e., without the presence of the objects they are putatively appearances of) and be something one has introspective access to, it also means that these appearances are able to possess all the content that they do, regardless of whether there is an 'external' world 'out there' to supply it or not. Hence, on this conception, 'appearances' must be conceived as having a determinate content which one can be presented with in sense-experience—for example, an appearance 'as of there being a mug in front of one' as opposed to an appearance 'as of there being a computer on the table'—even though there might be no physical objects (such as mugs and computers) at all (since we might be trapped in a simulation).

If such a thesis makes sense, then it entails that one's putative knowledge-claims about the world can be derived solely from knowledge of how things appear to one, since, as per the closure-based sceptical argument, unless one can rule out that one is in the 'global' 'vat-world' (LE or NBE), one can never have knowledge of anything else. Given that it must be possible to express these alleged 'knowledge-claims' in a language (if they are to be expressed at all), this conception implies that a language must be possible that does not presuppose any acquaintance with an 'external' world, or with 'other minds', as if we were in the global vat-world (LE or NBE), we would have access to no such things. Consequently, the words of this language would have

---

to derive their meaning purely from being linked to episodes of one's own consciousness (since nothing else would be available), which means that it must be possible to identify and describe these episodes independently of whether they have 'the world in view' or not (independently of whether they ever inform one of 'reality' or of 'how things are'). So, the intelligibility of such a language requires that one's inner experiences (the episodes of one's consciousness) come with a 'readymade' content that one can 'grasp' and make reference to, even though there is, if one were in the global vat-world, no connection between them and anything 'outer', 'objective' and 'external'.

Such a 'vat-language' would, therefore, be an intrinsically unshareable 'private' language,[12] in the sense that its signs would refer to something that no one, apart from oneself, can have any direct access to—the episodes of one's consciousness (one's private inner experiences). Consequently, if it can be shown that no language is possible that can be intelligible to me *alone*, then this would also serve to undermine the notion that knowledge of the content of appearances can be preserved, while knowledge of the facts (the way things are) must be jettisoned: If one radically divorces the 'subjective' from the 'objective'—'appearances' from 'reality' (or from 'the facts')—then everything, to speak with McDowell, 'goes dark in the interior as we picture it' (McDowell 1998: 409).

The thought that a 'private' language is possible, famously comes under attack in Wittgenstein's *Philosophical Investigations* (PI). Since we have just seen that the possibility of such a language is presupposed by global envatment scenarios—the notion that I and my envatted counterpart could share the same perceptual experiences even though the experiences of my envatted twin are never experiences of an 'external' world—if Wittgenstein is right that such a 'language' is not possible, then such a result has devastating consequences for the tenability of global sceptical scenarios. For this reason, we need to subject Wittgenstein's case against such a notion to careful scrutiny, in order to ascertain whether it is correct that there is no such thing as a language whose words could refer to what *only* the speaker can experience and know.

## 4 Wittgenstein's Anti-Private Language Considerations

Wittgenstein's arguments[13] against the possibility of a 'private' language are commonly thought to start at PI §243:

> Is it also conceivable that there be a language in which a person could write down or give voice to his inner experiences—his feelings, moods, and so on—for his own use?—Well, can't we do so in our ordinary language?—But that is not what I mean. The words of this language are to refer to what only the speaker can know—to his immediate private sensations. So another person cannot understand the language.

In this passage, Wittgenstein seems to be drawing a contrast between what he calls 'giving voice to one's inner experiences in our ordinary language', and the possibility of a language whose words refer to 'what only the speaker can know'—one's 'immediate private sensations'. Filling in some of the blanks, this contrast appears to imply the following:

(1) Our ordinary sensation language is not a private language, because other speakers can come to know and understand what one means. Naturally, one can choose to conceal one's feelings and record them only in a coded diary, for example, but were one to teach another the code—or were another to steal one's diary and break the code—then they could come to know about one's feelings. So, in this respect, one's 'inner experiences' would only be 'contingently' private in the sense that one could, in principle, reveal them if one wanted to, even though, in actual fact, one might choose not to do so.

(2) In a 'logically' (or necessarily) private language, by contrast—the notion that comes under indictment in PI—the words refer to what *only* the speaker can know, and their meaning is, in this much, necessarily incommunicable. That is to say, even if one wanted to, one would not be able be to teach someone else what the words in this language mean, since no one, apart from oneself, can have access to the 'immediate private sensations' that one's words are allegedly referring to.

The thought that no one can have access to one's 'immediate private sensations' is, of course, part and parcel of the traditional Cartesian conception according to which 'inner experience' is conceived as an internal equivalent of 'outer experience'—i.e. as a class of

---

[12] An intrinsically unshareable private language is a 'logically private' language. See discussion below.

[13] I am not, here, going to pursue the question whether Wittgenstein advances any 'arguments'. It seems to me obvious that he does and have, indeed, argued this elsewhere.

appearances that one is presented with in introspection (rather than in sense-perception[14]), but which no one else can have any direct access to in virtue of not being me. With this conception in place, any language whose words refer exclusively to what goes on in this 'private arena', access to which is necessarily denied anyone apart from oneself, will be radically incommunicable and, in this sense, 'logically' private. Given that nobody can, even in principle, become acquainted with the entities that the words of this language refer to, it is impossible for anyone other than oneself to acquire this tongue. Let's call this view the 'Private Inner Object Picture'. It implies the following:

(3)  If the words of our ordinary language referred to 'Private Inner Objects', our ordinary language would be radically incommunicable and, hence, unintelligible to anyone apart from the 'speaker'.

But (3) is inconsistent with (4), and (4) appears obviously true:

(4)  Our ordinary language is not unintelligible to anyone apart from the speaker.

So, it seems that we must accept (5) instead:

(5)  The words of our ordinary language do not refer to Private Inner Objects.

This is, indeed, the conclusion that Wittgenstein wants us to espouse. But, as it stands, the argument would presumably fail to convince someone who is enticed by the idea that even if our ordinary language cannot be a 'private' language in the relevant sense, it might, nevertheless, be possible to construct one. What is more, what if it is the case that everyone speaks the ordinary language, but, on top of that, also has a 'private' language for their own private use[15]? Rather than immediately offering the reader a proof that the Private Inner Object Picture is mistaken, Wittgenstein, in the next passage, raises the question of how words are supposed to refer to sensations—or 'inner experiences'—in the first place. This is a good move to make, since, if we have a skewed conception of what it means to employ sensation-language in ordinary, 'public' circumstances, we are inevitably going to miscast what goes on when we try to 'refer' to something in a logically private context as well.

How, then, Wittgenstein asks, is the connection between the name and the thing named set up (PI §244)? In other words, how does a human being learn the *meaning* of the names of sensations? Wittgenstein suggests one possibility:

Words are connected with the primitive, natural, expressions of sensation and used in their place. A child has hurt himself and he cries; then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain-behaviour (*ibid.*).

The interesting thing to note about this remark is not that it immediately prompts Wittgenstein's *alter ego* to proffer charges of behaviourism,[16] but rather that in it no mention is made of the 'Augustinian' approach to language[17]—that is to say, of the idea, discussed at the beginning of PI, that the essence of language consists in naming things and, hence, that the primary way of setting up a connection between word and object is by ostensive definition.

The reason why Wittgenstein eschews this route is no coincidence, however, since the Augustinian and Private Inner Object pictures are companions in guilt: they work in tandem to produce the notion of a private language. For with both pictures in play, one is naturally going to assume that the way one learns sensation words is by assigning names to 'internal', 'introspectible' items; that is to say, to Private Inner Objects. For example, one is going to find it plausible to think that just as one might first learn what 'cat' means by repeatedly being told by someone, pointing to cats, that these are 'cats', one might learn what the names for sensations are by, as it were, 'pointing inwards' and naming a particular sensation 'pain' say.

The problem with the latter idea is of course that while, in the cat case, someone else can teach one what the word 'cat' means, in the case of sensations, it appears that one must do all of the hard work oneself. For how does one know *what* to name 'pain' in the first place? Perhaps, if one already has a language, one might be told that 'pain' names an unpleasant sensation. But how does one first learn what a *sensation* is? If one is supposed to learn all this by ostensive definition, one is in a quandary, since, while I can see the thing you are pointing to when you say that this is a cat, I have no idea what you are 'pointing to' when you ask, for example, whether I am in pain after having fallen off a chair, say. So, it seems that when it comes to naming 'internal' items, one must guess the essential thing—i.e., one must guess which of one's private, introspectible items is the referent of the word 'pain'—and this implies that the possibility of

---

[14]  Of course, if one is in the 'global' 'vat-world' (in a LE or NBE scenario) all apparently 'outer experience' collapses into 'inner experience' anyway.

[15]  This is a possibility that Wittgenstein considers at PI §273.

[16]  '"So you are saying that the word 'pain' really means crying?"— On the contrary: the verbal expression of pain replaces crying, it does not describe it' (PI §244).

[17]  'These words [Augustine, *Confessions*, I.8], it seems to me, give us a particular picture of the essence of human language. It is this: the words in language name objects—sentences are combinations of such names.—In this picture of language we find the roots of the following idea: Every word has a meaning. This meaning is correlated with the word. It is the object for which the word stands' (PI §1).

misidentification can arise (I might name something different 'pain' than you have done).

But things are worse than they seem, for the proponent of the Augustinian picture of language is over-confident even when it comes to learning by ostension in the cat case. For just as one needs to know what a sensation is before there can be any question of 'naming sensations', unless one already has prior knowledge of the 'post where the new word is to be stationed' (PI §257), one will not understand the definition, 'this is called "tove"',[18] uttered in the presence of a pencil, for example, either. Given that 'tove' might mean a whole host of different things—for instance, 'writing utensil', 'number', 'physical object in general', 'position on the table', 'colour', 'sharp', 'blunt' etc.—unless one already knows that someone means ostensively to define the *name of a particular writing utensil*, say (which in turn presupposes that one has already acquired the concept 'writing utensil'), one is not going to understand the ostensive definition, as one will have no idea what the 'this' in 'this is called "tove"' is supposed to refer to. In short, ostensive definition underdetermines the *definiendum* unless some prior linguistic competence is already present.

In the light of the foregoing discussion, the proponent of the Private Inner Object picture is faced with a major problem: If she continues to insist that the connection between the name and the thing named is to be set up by ostensive definition (which, it seems, she must, given that she is committed to the idea of naming private, introspectible items), then, in the case of 'inner experience', this will have to be *private* ostensive definition (since no 'public' ostensive definition is possible). In what follows, I argue that such a notion is fraught with difficulties and, hence, that the private linguist cannot meet this challenge.

Wittgenstein's argument against the possibility of private ostensive definition begins at PI §257, which is worth quoting in full:

"What would it be like if human beings did not manifest their pains (did not groan, grimace, etc.)? Then it would be impossible to teach a child the use of the word 'toothache'."—Well, let's assume that the child is a genius and invents a name for the sensation by himself!—But then, of course, he couldn't make himself understood when he used the word.—So does he understand the name, without being able to explain its meaning to anyone?—But what does it mean to say that he has 'named his pain'?—How has he managed this naming of pain? And whatever he did, what was its purpose?—When one says "He gave a name to his sensation", one forgets that much must be prepared in the

language for mere naming to make sense. And if we speak of someone's giving a name to a pain, the grammar of the word 'pain' is what has been prepared here; it indicates the post where the new word is stationed.

Contrary to his initial suggestion at PI §244 that the word 'pain' is connected to the primitive, natural expressions of pain and is learnt in conjunction with them, Wittgenstein is, in this passage, humouring the interlocutor (his 'sceptical' *alter ego*) by contemplating a scenario—perhaps as a radical (and ironic) means of countering charges of behaviourism by getting rid of behaviour altogether—where no pain-behaviour is *ever* manifested by anyone. He then goes on to ask whether it would be possible, under such conditions, to teach someone what the word 'toothache', for instance, means.

*Prima facie*, it is not easy to see what we are supposed to imagine here. Of course human beings can choose to conceal their toothache or, indeed, to simulate toothache when they are not, in fact, feeling any. But that cannot be the completely 'pain-behaviour-free' environment that Wittgenstein's interlocutor is hankering after, since being able to simulate or conceal toothache presupposes that human beings generally express their toothache, otherwise there would be nothing *to* simulate or conceal in the first place. That is to say, if people didn't generally complain, moan, grimace, express fear of the dentist etc. when they had toothache, it would not be possible to choose *not* to express these things either. Similarly, one cannot wonder whether someone is hiding their pain, if there weren't an established practice of manifesting pain, i.e. if people never showed any signs of pain. And this entails that certain forms of scepticism just don't make sense:

Are we perhaps over-hasty in our assumption that the smile of a baby is not pretence?—And on what experience is our assumption based? (Lying is a language-game that needs to be learned like any other one) (PI §249).

In other words, it would be senseless to wonder whether a baby might not be simulating, as a baby does not yet know what it *means* to simulate. It has not yet been inducted into this practice; it does not yet have the linguistic and social competence to pretend that things are other than they are; it, in fact, has no conception of *how things are* yet. Consequently, the following 'aggregate argument', constructed by Wittgenstein's *alter ego*, fails to go through:

"If it is possible for someone to make a false move in some game, then it could be that everybody made nothing but false moves in every game" (PI §345).

For, as Wittgenstein goes on to point out, although orders, for example, are sometimes not obeyed, if *no* orders were *ever* obeyed, 'the concept of an order would have lost its purpose' (*ibid.*). Similarly, just as the possibility of perceptual

illusion presupposes the general veridicality of sense-perception, so the possibility of simulating (or concealing) pain presupposes that pain is generally expressed. Just as it does not follow from the fact that someone can sometimes be wrong that everyone could always be wrong, so one cannot get from the thought that someone might sometimes simulate (or conceal) pain to the notion that everyone could always be simulating (or concealing) pain. For in a society where no one *ever* presented *any* pain-behaviour, the concept of 'pain' would lose its point and so would the notions parasitic on this concept, such as 'simulation' or 'concealment'. In other words, just as it is logically impossible for something both to *be* a game *and* to consist of nothing but 'false moves', so it is impossible to 'conceal' pain in a context where 'pain manifestation' is *ex hypothesi* impossible.

If this is right, then it is not surprising that at §247 Wittgenstein says that in contexts where no one ever manifested any pain-behaviour (showed any natural signs of pain), it would not be possible to introduce the term 'pain'. But let's suppose, for the sake of argument, that the child in question is a genius and just invents a name for the sensation himself. In that case, Wittgenstein responds, the child could not explain the meaning of this word to anyone, and this is problematic because we might legitimately wonder whether a word that cannot (even in principle) be understood by anyone else, can nevertheless be understood by its 'speaker'. For, as we have already seen, in order for something to qualify as an act of 'naming', for example, the relevant linguistic background conditions that make this practice possible, must already be in place: 'One forgets that much must be prepared in the language for mere naming to make sense. And if we speak of someone giving a name to a pain, the grammar of the word "pain" is what has been prepared' (§247). But given that no such grammar can be presupposed in the completely pain-behaviour-free environment we are still being asked to consider—for if it were, there would be no reason to suppose that someone else could not learn the meaning of the word in question—it seems that the child genius can do nothing but rely on bare ostension. So, we are now approaching the very heart of Wittgenstein's polemic against the idea of a private language:

> Let's imagine the following case. I want to keep a diary about the recurrence of a certain sensation. To this end I associate it with the sign 'S' and write this sign in a calendar for every day on which I have the sensation.—I first want to observe that a definition of the sign cannot be formulated.—But all the same, I can give one to myself as a kind of ostensive definition!—How? Can I point to the sensation?—Not in the ordinary sense. But I speak, or write the sign down, and at the same time I concentrate my attention on the sensation—and so, as it were, point to it inwardly.—

> But what is this ceremony for? For that is all it seems to be! A definition serves to lay down the meaning of a sign, doesn't it?—Well, that is done precisely by concentrating my attention; for in this way I commit to memory the connection between the sign and the sensation.—But "I commit it to memory" can only mean: this process brings it about that I remember the connection *correctly* in the future. But in the present case, I have no criterion of correctness. One would like to say: whatever is going to seem correct to me is correct. And that only means that here we can't talk about "correct" (PI §258).

It is probably safe to say that no other passage from Wittgenstein's later *corpus* has been subjected to more misrepresentation than the present one. Given how compressed Wittgenstein's argument is, most commentators[19] tend to assume that it is supposed to attack the reliability of memory in the private context, when its actual target is the very idea of private ostensive definition as such, something that has nothing to do with a general scepticism about memory.

Philosophers who lean towards the 'reliability of memory' interpretation are prone to argue in the following manner.[20] In the above passage, Wittgenstein is claiming that there is no such thing as 'private' rule-following, since in the private scenario no distinction can be drawn between what seems right to me, and what is *actually* right.[21] They then go on to link this idea to the thought that, in the private case, where I cannot appeal to the judgement of others, I have no way of checking whether my present sensation, which I take to be 'S' again, actually corresponds to the sample I originally concentrated my attention on and thus labelled 'S' in the first place. In other words, it may be that I misremember which sensation is supposed to be 'S', and, since I have nothing outside of myself to appeal to—as it were no 'external checks' to corroborate what I believe to be the case—whatever is going to seem right to me, is right, and that just means that one can't talk about right (PI §258).

If this is supposed to be Wittgenstein's argument, then it is vulnerable to the following objection: It would undermine public language just as much as 'private' language. For, if the problem with private language is supposed to be that it is vulnerable to a scepticism about the reliability of memory,

---

[19] See for example, A.J. Ayer (1971), P.F. Strawson (1971), Helen Hervey (1971), C.W.K. Mundle (1971), Judith Jarvis Thomson (1971), Saul Kripke (1982), Robert Fogelin (1987).

[20] I don't here intend to reconstruct any particular argument, just to give the general gist of this type of view.

[21] 'That's why "following a rule" is a practice. And to *think* one is following a rule is not to follow a rule. And that's why it's not possible to follow a rule "privately"; otherwise thinking one was following a rule would be the same thing as following it' (PI §202).

then, naturally, the same would be true of ordinary, 'public' language as well. As Saul Kripke, for example, has forcefully argued, how do I know that by '$+$' I meant *plus* in the past and not, for instance, *quus* (where 'quus' is defined as $x*y = x + y$, if x, $y < 57; = 5$ otherwise) (Kripke 1982: 9)?[22] If it is possible that the private linguist misremembers what he meant by S, then the 'public' linguist surely faces exactly the same predicament. Indeed, it is precisely because Kripke thinks that Wittgenstein is putting forward a 'sceptical argument' in PI that he believes that an appeal to the community at large is our only safeguard against an all-consuming 'meaning scepticism'. This is Kripke's 'sceptical solution': I am only able to fend off the 'quus' interpretation of 'plus', because I can check the answers that I am disposed to give to addition problems against what the community is doing. The private linguist lacks this resource, and, hence, on the Kripkean reading of Wittgenstein, a private language is impossible for this reason.[23]

The problems with Kripke's 'sceptical solution' are notorious. The most serious one is a self-undermining relativism: whatever seems right to the community is right, and, one might like to add, using Wittgenstein's own words against Kripke, that just means that one can't talk about right. That is to say, if Kripke's 'community response' is to provide any kind of answer to the sceptical problem, it must already presuppose that we know what it means 'to follow according to the rule'—as opposed to doing whatever pops into one's head—otherwise what the community at large is doing is going to be as ad hoc and as arbitrary as anything that the individual might do by herself (and therefore just as vulnerable to 'meaning scepticism').

In other words, I won't be able to recognize whether or not there is general community agreement about what 'addition', for example, means, unless I already have a way of determining that two responses to a sum are in fact the *same* response. But I can only do that if I already have criteria available that allow me to recognize that two utterances are tokens of the same type, i.e. if I already know what 'plus' means! Consequently, Kripke's 'sceptical solution' fails, because it either has to assume what is to be proved (i.e. prior knowledge of what 'plus' means) or it collapses back into the 'meaning scepticism' it was supposed to save us from. So, the community is in no better position, if Kripke is right, than the individual members by themselves.

But if community agreement does not solve the 'sceptical paradox', then the thought—essential to Kripke's 'solution'—that the sum is the result of whatever the community

calls 'plus' or 'addition', has to be rejected as well.[24],[25] Of course, a distinction needs to be drawn between what the individual and what the community are doing, but it is a mistake to think that this lends any support to Kripke's thesis. For, although, as Wittgenstein emphasizes, 'a person goes by a signpost only in so far as there is an established usage, a custom' (PI §198) of using signposts, the 'established custom' is not, as it were, the result of a community-wide 'agreement in *opinions*' (PI §241). Rather, the 'established custom' or practice provides the necessary background conditions that make agreement and disagreement possible.

In other words, it would be a misconception to think that since a word can change its meaning as a result of the community adopting a different use of a word over time—such as happened, for example, in the case of the word 'gay'—it is the 'community' that 'determines' or 'decides' which *particular* applications of words can be correct or incorrect.[26] It is our shared linguistic practices that make meaning, and, hence, disagreements in opinion possible,[27] but to describe

---

[22] Compare also Fogelin (1987: 179–83).

[23] Fogelin also believes that Wittgenstein is putting forward a 'sceptical solution' (1987: 181).

[24] Strangely enough, Kripke himself seems to reject this 'community-wide version of dispositionalism' (1982: 111), but that just makes his conception even *more* implausible, for it now turns out that the 'sceptical solution' is no solution at all, but merely a reiteration of the platitude that 'if everyone agrees upon a certain answer, then no one will feel justified in calling the answer wrong' (1982: 112).

[25] Kusch (2006) defends a more sophisticated version of the 'community agreement thesis' that is compatible with what he calls a 'minimal factualism' about word meaning. Nevertheless, Kusch's position does not seem entirely consistent, since his notion that a community actually needs 'to confer' the status of being a 'rule-follower' on an individual before they can be accepted into the community of 'practitioners of addition' say, seems incompatible with even a 'minimal' factualism about what a word means. For, unless *all* forms of 'meaning factualism' are in fact rejected, what would such 'conferral' accomplish over and above someone's just satisfying the ordinary criteria for rule-following (or adding)? That is to say, if someone counts as a practitioner of addition merely because 'the community', in virtue of having 'bestowed' this status, 'says so', then Kusch's response is not relevantly different from Kripke's. If, on the other hand, there is supposed to be room for some 'minimal factualism' about meaning, then, whether someone satisfies the conditions for having performed an addition problem correctly, must be independent of 'community agreement'. One cannot have it both ways here.

[26] Compare Stroud: 'Similarly, it cannot be said that plus or addition is whatever a community uses the plus-sign or the word "plus" to stand for. That we all use it as we do is what gives that sign its particular meaning. But it could have had a quite different use. Just as I could easily recognize that another community use that same sign differently so I could even recognize that my own community had changed its use of it.' Stroud (2000: 92).

[27] In this sense, and as Baker and Hacker (1984: 44–5) point out, 'agreement' is a 'framework condition' for the possibility of rule-following and *not*, as Kripke thinks, 'constitutive' of it. Kusch (2006: 258) disagrees with this assessment, claiming that Kripke does appreciate the 'framework' role of agreement. But the passage Kusch cites in favour of his reading (Kripke (1982: 96–7)), in fact lends no support to this claim at all, since, even though Kripke talks of 'agreement in form of life', he regards this agreement as a 'brute fact', rather than as something emerging out of a common understanding

those practices—to give a description of the grammar of our language (our concepts)—is not yet to make any particular truth-claim (not yet to voice any particular assenting or dissenting opinion). Rather, truth-claims can be made because we already have at our disposal a language that provides us with the necessary network of concepts that makes agreement and disagreement possible. So, for example, without the rules of chess in the background, one could not have a dispute about whether a particular chess move is legitimate or not, for without those rules, it would not be a *chess* move at all—not even the *appearance* of one. This is the significance of Wittgenstein's remark that 'what is true or false is what human beings *say*; and it is in their *language* that human beings agree. This is agreement not in *opinions* [emphasis added], but rather in form of life' (PI §241).

If this is right, then there is, in the end, no such thing as making sense of the 'sceptical' alternative that, for all the individual taken in isolation knows, one might have meant 'quus' by ' + ', which Kripke is proposing as a challenge to the thought that a rule or a word has a determinate meaning. For it is only against the background of a mistaken conception of what rule-following involves that the so-called 'rule-following' paradox arises in the first place:

> This was our paradox: no course of action could be determined by a rule, because every course of action can be brought into accord with the rule. The answer was: if every course of action can be brought into accord with the rule, then it can also be brought into conflict with it. And so there would be neither accord nor conflict here.
>
> That there is a misunderstanding here is shown by the mere fact that in this chain of reasoning we place one interpretation behind another, as if each contented us at least for a moment, until we thought of yet another lying behind it. For what we thereby show is that there is a way of grasping a rule which is *not* an interpretation, but which, from case to case of application, is exhibited in what we call 'following the rule' and 'going against it'.

That's why there is an inclination to say: every action according to a rule is an interpretation. But one should speak of interpretation only when one expression of a rule is substituted for another (PI §201).

It ought, nowadays, almost to be commonplace to say that Kripke, by presenting his 'sceptical' reading of Wittgenstein, has simply chosen to disregard the second paragraph of this passage. For, there, Wittgenstein says very clearly that 'that there is a *misunderstanding* [my emphasis] here is shown by the mere fact that in this chain of reasoning we place one interpretation behind another, as if each contented us at least for a moment, until we thought of yet another lying behind it. For what we thereby show is that there is a way of grasping a rule which is *not* an interpretation, but which, from case to case of application, is exhibited in what we call "following the rule" and "going against it".' In other words, Wittgenstein is indicating that what gives rise to the 'sceptical paradox' is the view that what enables us to 'follow according to the rule' is an *interpretation* of the rule which tells us how to apply the rule correctly in any given case. Not only does such a conception lead to an infinite regress—for, how do we know how to interpret the interpretation of the rule and so on?—it also paves the way for the sceptical paradox. If it is granted that an interpretation of a rule needs to be given before we can know how to apply it, then, depending on the 'method of projection', any action can, in principle, be viewed, on some interpretation, to be 'in accordance with the rule', and so, also, to be 'in conflict' with it. This is the significance of Wittgenstein's discussion of the 'bent-rule follower' at PI §187, who 'naturally reacts' to the order 'add 2' by adding 2 up to 1000, and then adding 4.

To say that such responses are possible, however, is not, *pace* Kripke, to *endorse* the sceptical paradox. Rather, Wittgenstein is trying to undermine the idea that there is an 'epistemic gap' between rule and application that needs to be bridged by an interpretation of the rule.[28] Kripke's own proposal, by contrast, never challenges the sceptic's terms of engagement—never properly acknowledges that 'there is a misunderstanding here'. Instead, it is the community's dispositions, on Kripke's account, which, as it were, provide the 'last' and 'final' interpretation of the rule (the 'interpretation' that, as it were, needs no interpretation). So, we never, on his reading, actually move beyond the conception that gives rise to the paradox in the first place.

In order to recognize this, however, we need to understand that the relation between a rule and its application is an *internal* one and not, as Kripke's conception would have us believe, an 'external' relation between two unconnected items (since it is this idea that leads to the

---

Footnote 27 (continued)

of the same concept: 'We cannot say that we all respond as we do to '68 + 57' *because* we all grasp the concept of addition in the same way, that we share common responses to particular addition problems *because* we share a common concept of addition…Rather our license to say of each other that we mean addition by ' + ' is part of a 'language-game' that sustains itself only because of the brute fact that we generally agree' (Kripke (1982: 97)). This makes it clear, it seems to me, that Kripke *does* believe that 'agreement' is constitutive of rule-following rather than a 'framework condition' for it. Indeed, it is the basis for the 'sceptical solution' (a view that Fogelin also shares).

---

28 Compare Baker and Hacker (1984: 19–21), McDowell (2002).

underdetermination problem). For it is a mistake to think that one can separate 'grasping' a rule from knowing how to apply it, since the former is actually constitutive of the latter.[29] If I have understood an order, for example, then I also know what counts as complying with it. So, there are not *two* things going on here: understanding the order and knowing what counts as complying with it. In this sense, there is no 'gap' between 'order' and 'execution' that needs to be 'bridged' by invoking 'intermediary' items such as mental acts, dispositions and so forth.

Of course, that does not mean that I can never make a mistake. It is possible sometimes to misunderstand an order. But this has to be the exception rather than the rule, since, as we have already seen, one is not licensed to infer that one could *always* be wrong from the fact that one can sometimes be (no 'aggregate arguments'). Therefore, it would not be possible *always* to misunderstand *every* order, since otherwise the practice of giving orders would just lose its point. In other words, the possibility of misunderstanding orders presupposes that a general practice of obeying orders is already in place in just the same way as the possibility of 'shamming' pain is parasitic on the general practice of expressing pain.

Consequently, once the idea that a rule always needs an interpretation is no longer in play, it becomes hard to see what Kripkean 'meaning scepticism' can really amount to. Let us imagine, for example, that we are radical interpreters trying to make sense of an alien tribe's practice of writing signs down on pieces of paper, and we are shown the 'equation' '68 + 57 = 5' by the chief of the tribe. Would this give us reason to think that the tribe can add, but disagrees with us about the value of particular sums? No. For we could only judge that the tribe is disagreeing with us about *addition*, if it actually engaged in the practice of *arithmetic* (and were not, say, doing something else). And whether the tribe is engaged in arithmetic or not can only be determined by attempting to make sense of their signs in conjunction with their behaviour, which proves none too easy a thing to do.

One might, perhaps, start off by thinking that given that the tribe's signs look like our numbers '68', 57' and '5', they also symbolize in the same way. If they do mean the same as our numbers, however, then translation of the signs ' + ' and ' = ' becomes difficult. For if '68', '57' and '5' mean what they ordinarily mean, then it seems that ' + ' and ' = ' cannot mean what they ordinarily mean. For no one who is operating with the concept of *addition* can intelligibly judge that '68' and '57' *added* together equals '5'. In other words, if, *per impossibile*, all the signs in the equation '68 + 57 = 5' meant the same as our symbols, then this 'equation' would not even qualify as a *mistaken* judgement (since it is not,

in the relevant sense, a 'judgement' at all[30]). For someone who has learnt how to add can make all sorts of mistakes (e.g. '68 + 57 = 124'), but '68 + 57 = 5' is *not* among them (as making this move would make it questionable whether I know what 'addition' means; whether I have so much as acquired the concept). As Wittgenstein remarks in *On Certainty* §156: 'In order to make a mistake, a man must already judge in conformity with mankind.' That is to say, making a mistake presupposes that one has acquired the concept that one is, in a particular instance, misapplying, but regarding the 'equation' in question, it is hard to see what the relevant concept is so much as supposed to be. Consequently, we would have to conclude that either we cannot make sense of the tribe's signs at all, or a different translation of them is required. Either way, it would be impossible to come to the conclusion that the tribe is disagreeing with us about addition. For, we would either be unable to understand its signs or it would be speaking of something else (i.e. not about 'addition'). But whichever horn of this dilemma we grasp, we cannot conclude that they mean addition by 'plus'.

We are now in a position to return to the discussion of the 'memory sceptical' reading of the 'private language' sections that we started to consider earlier. On Kripke's view, the only thing wrong with the idea of a 'private' language is that in the 'private' context no community exists to ratify my responses, and, consequently, that I can never be sure that I ever really mean anything by my signs (since my memory might deceive me about what I thought they meant in the past). Hence, on Kripke's interpretation, both 'public' and 'private' language is vulnerable to the same sceptical concerns, and there is, to this extent, no real asymmetry between them: Wittgenstein has, according to Kripke, shown *all* language to be unintelligible (Kripke 1982: 62)! Given that we have just seen, however, that Kripke's 'sceptical' reading of the PI is off target, we need to be wary of 'memory sceptical' interpretations of §258.

As previously discussed, PI §258 is concerned with the problem of setting up a connection between the sign 'S' and some internal episode in the absence of any public grammar for the sign—any already available concept—or any connection to natural reactions, so that no definition of the sign can be formulated. Nevertheless, Wittgenstein's interlocutor is confident that he can give himself a kind of ostensive definition by concentrating his attention on the episode, and at the same time writing the sign down. In this way, Wittgenstein's putative 'private linguist' believes, he can 'commit to memory' the connection between the sign and the episode.[31]

---

[29] Compare Baker and Hacker (1984: 95–97).

[30] As Wittgenstein said in the *Tractatus*, I cannot judge a nonsense (TLP 5.5422), since, in such a case, there would be nothing *to* judge.
[31] I specifically do not say 'sensation', since, so far, the interlocutor has no right to call whatever it is that he thinks he is 'naming' a 'sensation'.

It is this idea that Wittgenstein goes on to challenge in the following passage:

> But "I commit it to memory" can only mean: this process brings it about that I remember the connection *correctly* in the future. But in the present case, I have no criterion of correctness. One would like to say: whatever is going to seem correct to me is correct. And that only means that we can't talk about 'correct' (PI §258).

In other words, what Wittgenstein is challenging in this passage is not the notion that I will not be able correctly to identify *future* instances of 'S'; rather, he is challenging the idea that *any* connection has so far been set up between the putative internal episode and the sign 'S'. For the question of future correct identification presupposes that the sign 'S' has already been given a meaning, and that is just what is at issue. That is to say, 'memory sceptical' readings of PI §258 already grant to the private linguist that a connection between 'S' and a 'sensation' has been set up, when it is that very notion that Wittgenstein is setting out to challenge. But, if so, then it cannot be the case that what is problematic in the private linguist's situation is that it is not possible to verify whether a future application of S accords with one's previous usage, as the very point at issue is the question whether 'private ostensive definition'—the 'concentrating of one's attention' on a putative private particular—can serve to endow a sign with a use in the first place. Consider the following passage:

> What reason have we for calling 'S' the sign for a *sensation*? For 'sensation' is a word of our common language, which is not a language intelligible only to me. So the use of this word stands in need of a justification which everybody understands.—And it would not help either to say that it need not be a *sensation*; that when he writes 'S' he has *Something*—and that is all that can be said. But 'has' and 'something' also belong to our common language.—So in the end, when one is doing philosophy, one gets to the point where one would like just to emit an inarticulate sound.—But such a sound is an expression only in a particular language-game, which now has to be described (PI §261).

If one already grants to the private linguist that he intends to name only a new kind of *sensation*, then the sign he introduces could, in principle, be learnt by others. For example, 'S' might denote, in this case, the sensation of having 'pins and needles', or a very peculiar kind of vertigo. But, presumably, the whole point of the 'private' language is supposed to be that no one else can so much as be taught it (since no one can become directly acquainted with the internal particular that one is allegedly naming for oneself by private ostensive definition), in which case the private linguist is not entitled

to help himself to the concept of a 'sensation' (as this is a 'public' one). In other words, Wittgenstein is really setting up a dilemma for the private linguist: Either the sign 'S' is logically private—in which case it will fail to mean anything (as private ostensive definition cannot give it a use)– or 'S' names a sensation, in which case others could learn what it means. Whichever horn of this dilemma the private linguist grasps, there can be no such thing as a (logically) private language.

If this is a correct representation of the dialectical structure of Wittgenstein's discussion, then 'memory sceptical' readings of PI §258 not only attribute bad arguments to Wittgenstein, they also concede *everything* to the private linguist, since they grant that he has already been able to set up a connection between 'S' and the 'internal episode', when the coherence of that notion in the 'logically private' context is the very thing that is in question. For, if I already have the *concept* of a sensation, then, naturally, I can introduce new words for as-yet-unheard-of sensations, but these will then turn out to be only 'contingently' private: I can, if I want to, share these words with others and their use might 'take on'. If I don't yet have the concept, however, but rather try to derive it from 'concentrating my attention' on a private, internal phenomenon, then appeals to a 'memory of S' won't help me, since to have a 'memory of S' presupposes that I already know what 'S' means. Consequently, *pace* Ayer et al., it is actually irrelevant whether or not my memory deceives me, since if I don't yet know what 'S' means, there is nothing for my memory to deceive me about. Wittgenstein tries to make this clear in the following passage:

> Let us imagine a table, something like a dictionary, that exists only in our imagination. A dictionary can be used to justify the translation of a word X by a word Y. But are we also to call it a justification if such a table is to be looked up only in the imagination?—'Well, yes; then it's a subjective justification.'—But justification consists in appealing to an independent authority— 'But surely I can appeal from one memory to another. For example, I don't know if I have remembered the time of departure of a train correctly, and to check it I call to mind how a page of the timetable looked. Isn't this the same sort of case?' No; for this procedure must now actually call forth the *correct* memory. If the mental image of the timetable could not itself be *tested* for correctness, how could it confirm the correctness of the first memory? (As if someone were to buy several copies of today's morning paper to assure himself that what it said was true) (PI §265).

Far from making any kind of verificationist point here, Wittgenstein is really accusing the private linguist of begging the question. For to call up a memory, in order to confirm something, presupposes that the memory in question is

indeed a memory of the sort of thing that will, if the memory is accurate, confirm the thing in question. That is to say, it is only if my 'memory' is a memory of a *train timetable* (and is accurate) that it will confirm the time of departure of the train. If I called up the memory of a page in my Gordon Ramsey recipe book instead, say, then the memory, even if accurate, would not confirm the train departure time. In other words, the private linguist is only entitled to appeal to his 'memory of S' as a way of confirming that his present internal episode is S, if the 'memory of S' is a memory of *S* (and not, say of 'T', 'U' or 'XYZ').[32] But, again (as in PI §258), the private linguist can only know that it is a memory of *S,* if he already knows what 'S' means (this parallels the problem faced by Kripke's 'sceptical solution'). It is the appeal to the 'memory of S', however, that is supposed to provide the criterion for what 'S' means, so the private linguist is, in effect, using the memory of the meaning of S to confirm itself. It is for this reason that Wittgenstein says that what the private linguist is doing is like buying several copies of the (same) morning paper, in order to assure himself that what it says is true.[33]

Ayer, Strawson et al. fail to read this passage correctly, because they just take for granted that one can give a sign a meaning simply by uttering a sound in the presence of a certain phenomenon. This parallels the attempt to reduce rule-following to dispositions and backfires in a similar way: One may be disposed to say, 'this is S again', but that is not sufficient to give 'S' any normative force (to supply rules for the correct use of 'S').

What the 'private language' dialectic is, therefore, supposed to show is that in the 'logically private' context, 'S' will remain a meaningless sign, since, *pace* the Private Inner Object Picture, introspection alone cannot give it a use. That is to say, Wittgenstein is trying to undermine the old myth that we can somehow 'read off' from the 'object' itself the way that its 'name' is to be applied[34]—a myth that is constitutive of the 'Augustinian picture of language', according to which 'naming' is the primary function of language and 'ostensive definition' the paradigmatic way of teaching word-meaning. This conception is helped on its way by what Putnam calls a 'pictorial semantics' (Putnam 2002: 15)— the empiricist notion that words refer to ideas which are mental 'copies' of the objects we perceive. Even if none of these notions are explicitly endorsed by most contemporary philosophers, they form part of the inherited (and, perhaps,

'intuitive-seeming') background that gives sustenance to the idea that a 'logically private' language must be possible. For with both pictures in place, it is going to seem natural to think that just as we learn what 'cow' means by looking at cows, so we learn what 'sensation' means by, as it were, 'looking at' our sensations. So, a mistaken conception of how we acquire concepts in the ordinary, public sphere, when applied to the 'internal realm' conceived as a kind of inner analogue of an 'outer' 'external' world (the Private Inner Object Picture), is naturally going to suggest to us that a 'logically private' language must be possible. But, if Wittgenstein is right, and my reconstruction of his argument correct, this turns out to be an illusion.

## 5 Implications of Wittgenstein's Argument

The consequences of Wittgenstein's argument for global scepticism are far-reaching, but insufficiently recognized in contemporary epistemology. For if the Private Inner Object Picture goes out of the window, then the thought—essential to global scepticism—that I and my envatted counterpart could share the same experiences even though my envatted twin's experiences *never* have the world in view—is undermined as well. And if the latter is no longer thought to constitute a coherent idea, then the notion that the way things appear to us could be the same, whether we are systematically deluded or not, can no longer be upheld either.

Hence, to undermine the possibility of a 'private' language is also to undermine the coherence of the idea that we could be victims of global BIV scenarios. For, if, for example, I do not learn what 'pain' or 'red' (or any other word) means by being presented with a self-subsistent, 'internal' 'intermediary item' (a Private Inner Object), thought to form part of an 'interface' of 'appearances' which 'intervenes' between myself and the 'external' world, and from which one can somehow 'magically' 'read off' how the word 'pain' or 'red' is to be used, then there is no longer any room for the thought that what lies beyond the 'interface' might be radically different from what we think it is, or for the notion that what lies on the 'inside' of it might vary radically between observers.[35]

---

[32] Compare Kenny (2006: 152).

[33] Compare also Rhees (1971): 64).

[34] This idea is similar to what Putnam (1981: 51) calls a 'Magical Theory of Reference'—the view that some occult 'noetic' rays connect a word with its referent; David Wiggins (1980) calls this notion a belief in 'Self-Identifying Objects'.

[35] It should be obvious by now that Wittgenstein's critique of 'private ostensive definition' undermines the coherence of the notorious 'inverted spectra' problem—the thought that you might have named a 'green' Experience 'red', say. For what is a 'green' Experience? Unless I again appeal to occult, 'self-identifying' objects, an 'experience' cannot be described as 'green' unless I already know what 'green' means (i.e. unless I already have the concept). But if I know what 'green' means, I will have learnt this from having been shown green samples (and had I called them 'red', I would have been corrected). So there is no room here for an 'intermediate step', where I, as it were, 'name something' that 'intervenes' between my mind and the physical object that I am looking at. For further discussion of this problem and its relevance to the 'qualia' debate, see Horwich (2012).

That is to say, if we were in a LE or NBE scenario, then the 'experiences' that we would allegedly be being 'fed' could only have the *same* content as our real, 'external' world experiences, if it is taken for granted that the putatively phenomenologically indistinguishable BIV-experiences possessed some kind of 'intrinsic' content that could be 'accessed' merely by 'inspecting' these 'inner experiences' themselves. Since such 'private inspection' could provide us with the relevant concepts only if private ostensive definition (and 'magical' acts of reference) were possible, Wittgenstein's arguments against this notion show that there is, in the end, no making sense of the idea that we could be wrong about everything all of the time.

## 6  Some Objections

The objection is sometimes made that even if one grants that Wittgenstein has shown that there can be no such thing as private ostensive definition, this is not sufficient to undermine 'global' scepticism, since the sceptic does not have to rely on this notion. Irwin Goldstein (1996), for example, is of this view.

Goldstein's exposition of Wittgenstein's case against private ostensive definition starts on the wrong foot, however, as he begins by attributing the 'public check' (memory sceptical) reading to Wittgenstein. Like Ayer et al. before him, Goldstein grants that the private linguist is trying to 'name a sensation'. For this reason, the only problem that Goldstein thinks that Wittgenstein is raising for his sceptical interlocutor is that since the original sensation is 'irretrievable' and not connected to 'public, bodily signs', she is in no position to reidentify her sensation, even were she to have it again in the future, as there are no 'public checks' for her to appeal to (1996: 140). As we have already seen, by reading Wittgenstein's arguments in this manner, 'the decisive movement in the conjuring trick has been made, and it was the very one that seemed to us quite innocent' (PI §308).

The apparently 'innocent' movement, in this context, is of course the Augustinian one of assuming that it is possible to identify 'a sensation'—as opposed to a 'red flash', say, or, indeed, nothing at all—in the absence of a public language that provides the necessary concepts or any link to natural reactions and expressions. If one concedes that it is possible to 'name' an 'internal episode' without the necessary background being in place that indicates the post where the new word is to be stationed, then one is simply assuming the intelligibility of the Private Inner Object Picture without argument. For, on this picture, 'sensations' are, as it were, 'self-identifying objects' that 'make themselves known to one' through introspection, regardless of whether one is in possession of the necessary concepts enabling one to identify them or not. And this, in turn, is just to espouse

a version of the Myth of the Given[36]: One is assuming that 'Experience' confronts one with something 'ready-made'—an 'internal' 'pre-conceptual' *this*—upon which one subsequently goes on to impose a 'public' conceptual structure.[37] But, of course, if one allows *that* move, one has already conceded *everything* to the sceptic. Consequently, it is not surprising that on Goldstein's reading, Wittgenstein's arguments will have no force against the sceptic, for he has already defanged them by *petitio principii*.

Finally, let's consider a sceptical move of last resort: Could the sceptic try to bite the bullet and grant that I have successfully shown that she is only under the illusion of having established conditions of correct use for her 'words', but nevertheless maintain that this could *be* our situation? No. If the private linguist/sceptic is under the illusion of having established conditions of correct use—which, of course, is precisely Wittgenstein's point—then her 'words' mean nothing, for they have no conditions of correct use. And if her words mean nothing, then the private linguist is not 'saying' anything. Consequently, the sceptic/private linguist cannot even 'save the appearances' or make any kind of sceptical point. It is for this reason that Wittgenstein says that, in the end, 'when one is doing philosophy, one gets to the point where one would like just to emit an inarticulate sound' (PI §261). Even such a sound, however, is only an expression, Wittgenstein notes, in a particular language-game, which now has to be described (*ibid.*). In the absence of such a sense-giving context, even the sound means nothing. So, if it is true that the private linguist is under the illusion of having established conditions of correct use for her 'words', the game is over.

## 7  Conclusions

If what I have argued in this paper is correct, the dilemma I presented the sceptic (or our sceptical *alter ego*) with in the introduction is fatal. Neither RE, nor LE or NBE, scenarios make scepticism compulsory. RE scenarios can be ruled out by way of appeal to ordinary empirical evidence, while LE or NBE scenarios can be discarded, since they presuppose the possibility of a global vat-language and there can be no such thing. But, if so, then (BIV2) in the closure-based sceptical argument can be denied: I can know that I have not recently been envatted, as all empirical evidence speaks against such a notion, and I no longer have to

---

[36] See Sellars (1956/1997).

[37] It is equally implausible to assume that a computer programme can just feed us these 'Givens'.

take LE or NBE scenarios seriously, for they collapse into incoherence.[38]

## Declarations

## References

Ayer AJ (1971) Could language be invented by a Robinson Crusoe? In: Jones OR (ed) The private language argument. Macmillan, London, pp 50–61

Baker G, Hacker PMS (1984) Scepticism, rules and language. Blackwell, Oxford

Coliva A (2014) Moderatism, transmission failures, closure and humean scepticism. In: Dodd D, Zardini E (eds) Scepticism and perceptual justification. Oxford University Press, Oxford, pp 248–276

Coliva A (2015) Extended rationality: a hinge epistemology. Palgrave Macmillan, London

Dretske F (1969) Seeing and knowing. Routledge and Kegan Paul, London

Dretske F (2010) Epistemological self-portrait. In: Dancy J, Sosa E, Steup M (eds) A companion to epistemology. Blackwell, Oxford, pp 130–133

Fogelin R (1987) Wittgenstein. Routledge and Kegan Paul, London

Goldstein I (1996) Ontology, epistemology, and private ostensive definition. Philos Phenomenol Res 56(1):137–147

Hawthorne J (2005) The case for closure. In: Sosa E, Steup M (eds) A companion to epistemology. Blackwell, Oxford, pp 26–42

Hervey H (1971) Private language and private sensations. In: Jones OR (ed) The private language argument. Macmillan, London, pp 76–95

Horwich P (2012) Wittgenstein's Metaphilosophy. Oxford University Press, Oxford

Jarvis Thomson J (1971) The verification principle and the private language argument. In: Jones OR (ed) The private language argument. Macmillan, London, pp 183–204

Kenny A (2006) Wittgenstein. Blackwell, Oxford

Kripke S (1982) Wittgenstein on rules and private language. Harvard University Press, Cambridge

Kusch M (2006) A sceptical guide to meaning and rules. Acumen, Chesham

McDowell J (1998) Knowledge and the Internal. Meaning, knowledge and reality. Harvard University Press, pp 395–413

McDowell J (2002) Wittgenstein on following a rule. In: Miller A, Wright C (eds) Rule-following and meaning. Routledge, London, pp 45–80

Mundle CWK (1971) Behaviourism and the private language argument. In: Jones OR (ed) The private language argument. Macmillan, London, pp 103–117

Neta R, Pritchard D (2007) McDowell and the new evil genius. Philos Phenomenol Res 74:381–396

Pritchard D (2012) Epistemological disjunctivism. Oxford University Press, Oxford

Putnam H (1981) Reason, truth and history. Cambridge University Press, Cambridge

Putnam H (2002) The collapse of the fact/value dichotomy and other essays. Harvard University Press, Cambridge MA

Rhees R (1971) "Private experience" and "sense data." In: Jones OR (ed) The private language argument. Macmillan, London, pp 229–232

Schönbaumsfeld G (2016) The illusion of doubt. Oxford University Press, Oxford

Schönbaumsfeld G (2019) Epistemic angst intellectual courage and radical scepticism. Int J Study Skept 9(3):206–222. https://doi.org/10.1163/22105700-20191388

Sellars W (1956/1997) Empiricism and the philosophy of mind. Harvard University Press, Cambridge, MA.

Strawson PF (1971) Exposition and criticism of Wittgenstein's *investigations*. In: Jones OR (ed) The private language argument. Macmillan, London, pp 27–33

Stroud B (1984) The significance of philosophical scepticism. Oxford University Press, Oxford

Stroud B (2000) Wittgenstein on meaning, understanding and community. In: Stroud B (ed) Meaning, understanding and practice. Oxford University Press, Oxford, pp 80–94

White R (2014) What is my evidence that here is a hand? In: Dodd D, Zardini E (eds) Scepticism and perceptual justification. Oxford University Press, Oxford, pp 298–321

Wiggins D (1980) Sameness and substance. Blackwell, Oxford

Wright C (2008) 'Comment on John McDowell. In: Haddock A, Macpherson F (eds) Disjunctivism: perception, action, knowledge. Oxford University Press, pp 390–404

Wright C (2014) On epistemic entitlement (II): welfare state epistemology. In: Dodd D, Zardini E (eds) Scepticism and perceptual justification. Oxford University Press, Oxford, pp 213–247

Wittgenstein L (1965) The blue and brown books. Harper and Row, New York

Wittgenstein L (1969) On certainty, Ed. GEM Anscombe and GH von Wright, trans. Anscombe and D Paul. Oxford: Blackwell

Wittgenstein L (1975) Tractatus Logico-Philosophicus, trans. D.F. Pears and B.F. McGuinness. Routledge and Kegan Paul, London

Wittgenstein L (2009) Philosophical investigations, Ed. PMS Hacker and J Schulte, trans. Anscombe, Hacker, Schulte. Blackwell, Oxford

---