Think • Vol 24 • No 69 • Spring 2025, 11–18

DOI: https://doi.org/10.1017/S1477175625000119

A Short Introduction to Wittgenstein's Private Language Argument

*Stephen Law**

Department of Continuing Education, University of Oxford, UK
*Corresponding author. Email: stephen.law@conted.ox.ac.uk

Keywords: Wittgenstein; private language argument; mind; language

Abstract

Wittgenstein's Private Language Argument is one of the most famous arguments in philosophy. It is a surprisingly tricky argument to understand. Some philosophers think it's a good argument. Others disagree. In fact they even disagree about what the argument actually is. This short essay gives three different interpretations of the argument and explains why I believe none succeed.

In his *Philosophical Investigations*, Wittgenstein presents a famous argument against the possibility of a private language. The argument appears in section §258 and surrounding text.

In this essay, I will first of all explain the main aim of this Private Language Argument (or *PLA*, for short), and then look at three different interpretations of the argument, which I call:

1. The No Check, No Rule Argument
2. The Stage-Setting Argument
3. The Circularity Argument.

The Aim of the Private Language Argument

Before we get to the aim of the PLA, let's take a step back and look at one of Wittgenstein's broader aims. One of the key targets of the *Philosophical Investigations* is what Wittgenstein considers to be a deeply confused picture of the mind as a sort of *private inner kingdom.*

It can be tempting, when we first start thinking about our own minds, to think of them as private spaces – as secret gardens populated by thoughts, feelings and other mental items to which we alone have direct access.

Just as I can gaze outwards at an external landscape of objects such as houses, trees and mountains, so, we might suppose, I can also focus my attention *inwards* at a sort of internal landscape that features my sensations, thoughts and feelings.

What I find within this inner landscape seems to be private in a very peculiar way. Normally, if I keep something private, it can *in principle* be revealed. For example, if I have a private stamp collection that I keep locked in a drawer and never show anyone, it remains *possible* for me to show others my stamps. But when it comes to my thoughts and feelings, it seems they are *necessarily* inaccessible to others. You cannot enter my mind and have the visual experience I am having right now. The most that you, as another mind, can have are experiences *just like* mine (that are, as philosophers like to put it, *qualitatively* the same as mine, but *numerically* distinct, much as you might own a car that is exactly *like* mine, but numerically

https://doi.org/10.1017/S1477175625000119 Published online by Cambridge University Press



distinct: there are still two cars: one in each garage).

Wittgenstein attacks this picture of the mind as sort of internally observed private realm in several ways. Wittgenstein doesn't deny that we have sensations, thoughts, feelings, and so on. But he does think that this way of thinking about them – as *things located within a necessarily private inner realm* – is confused.

So what role does the PLA play in this broader attack on the picture of the mind as a private inner kingdom? Its aim is to show something much narrower: *that it is not possible to introduce meaningful words or other signs that stand for such private inner things*.

https://doi.org/10.1017/S1477175625000119 Published online by Cambridge University Press

Private Meanings

The suggestion that it's not possible to introduce meaningful signs standing for such private inner things might seem just obviously false. After all, you might think, if my pain is a sort of private inner thing, why shouldn't I introduce a label for it? How? Well, I could focus my attention on it and say to myself: 'I shall use "S" to represent *this*.' Having inwardly defined 'S' in this way, I can go on to use it in my diary to record those days on which I have that experience. Then, later, I can flick back through my diary to check which days I had that sensation, and which days I didn't.

Here is the key thing to notice about this use of 'S': its meaning would be *necessarily* private. Because, *necessarily*, no one else can access my sensations, so, *necessarily*, no one else can know what I mean by 'S'.

Actually, it's tempting to think this is *already* how at least some of our words function. Take the word 'pain'. Surely, I might think, when I use the word 'pain', I use it to refer to *this* – a certain private inner experience on which I am now focusing my attention. So the word 'pain', in my vocabulary, *already* has such a private meaning.

'It can be tempting, when we first start thinking about our own minds, to think of them as private spaces – as secret gardens populated by thoughts, feelings and other mental items to which we alone have direct access.'

Of course, if that's how 'pain' is used, then, while I use it to refer to what I privately experience, you use it to refer to what you privately experience. But then we can never establish that we use 'pain' with the same meaning. We might perhaps correctly *guess* that we use the term with the same meaning, but, *necessarily*, we can't enter each other's minds to check.

In fact, some philosophers, such as John Locke (1632–1704), seem to go even further, committing themselves to the view that *all* language is private. For, according to Locke, the meaning of any word is an 'Idea' in the mind of the speaker:

> Words in their primary or immediate Signification, stand for nothing, but the Idea in the Mind of him that uses them . . .

Yet Locke also holds that our Ideas are 'hidden from others, nor of themselves can be made to appear'. But then this seems to entail that we each speak a private language: no one else can check what Ideas I use my words to signify.

Wittgenstein: No Private Meanings

Wittgenstein argues that while you might try to introduce a symbol like 'S', or a word like 'pain', to refer to such a private inner object, you can never succeed. The expression will always be meaningless. This is an important conclusion, because it entails, among other things, both that 'pain' cannot, after all, be the name of such a private inner experience, and also that Locke's account of how all words come by their meanings cannot be correct.

To be clear, Wittgenstein is not here targeting the suggestion that you could introduce a private code that keeps your meaning hidden. You certainly can do that. Famously, the seventeenth-century diarist Samuel Pepys sometimes wrote his diary using just such a private code so that others could not understand what he meant (he wanted certain things kept secret).

However, Pepys's private meanings could be, and eventually were, made public. What Wittgenstein is targeting is something else: the suggestion that you could introduce symbols the

https://doi.org/10.1017/S1477175625000119 Published online by Cambridge University Press

meaning of which are *necessarily* inaccessible and unteachable to others, because the symbols refer to what *necessarily* can only be known to the speaker.

The PLA

So what is Wittgenstein's argument against the possibility of such a *necessarily* private language? Here is the first half of section §258.

> Let us 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 will remark first of all that a definition of the sign cannot be formulated. – But still I can give myself 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.

Here Wittgenstein outlines how he thinks a symbol with a private meaning would need to be introduced. It could not be defined using other language, as we can define for example a triangle as a three-straight-sided plane figure, or a vixen as a female fox. The problem with a such a definition is it would then be parasitic upon your pre-existing language, the meaning of which is public. What your internal definition is supposed to do is introduce your first *private* meaning.

Still, you can provide another sort of definition of 'S' – what we call an *ostensive definition*. This is where we point to or focus our attention on something that we use as a *sample*. For example, I can define 'pencil' by saying 'This is a pencil' while pointing at an actual pencil.

Wittgenstein supposes our private linguist must introduce their first term by doing something similar: by saying or writing 'S' while concentrating their attention on their private sensation. The inner sensation is to function as the relevant sample. And, on the face of it, there doesn't seem to be anything particularly difficult about doing that.

And yet Wittgenstein thinks this internal naming ceremony achieves nothing. It fails to give 'S' a meaning. Why? Wittgenstein explains in the second half of section §258:

> But what is this ceremony for? For that is all it seems to be! A definition surely serves to establish the meaning of a sign. – Well, that is done precisely by the concentration of my attention; for in this way I impress on myself the connexion between the sign and the sensation. – But 'I impress it on myself' can only mean: this process brings it about that I remember the connexion *right* in 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 'rightness'.

This passage is not easy to understand. Wittgenstein's objection seems to be that our private linguist has failed to introduce a meaningful sign because they have failed to provide a 'criterion of correctness' – a way of *checking* that the sign is being applied correctly in the future.

Now it's certainly true that our private linguist doesn't have anything against which they can check whether they are later applying 'S' correctly. Suppose I define 'S' by reference to *this* sensation (I'm focusing my attention on it inwardly). Then, a bit later, I have another sensation, and it does *seem* like S again. However, I have no way of checking that it *really is* S again.

The situation is quite different if I define a term using a sample that exists in the public realm. Suppose I define 'puce' by reference to a piece of coloured card. I point at the colour on the card and say to myself 'By "puce", I mean *that* colour.' I now have something objective against which I can later check I'm applying 'puce' correctly. My sample of puce is objective in the sense that it exists mind-independently. I can put it in a drawer for safe keeping. If I later see something I *think* is puce, I can check by

https://doi.org/10.1017/S1477175625000119 Published online by Cambridge University Press

pulling out the card and comparing them. I possess a genuine criterion of correctness. But when it comes to 'S', no such check is possible. All I can say is that this *seems* like S again.

Wittgenstein now draws his conclusion. Because I have no criterion of correctness – nothing against which I can *check* I'm remembering correctly what S means – so we cannot talk of a 'correct' or 'incorrect' use of 'S'. But if there is no correct or incorrect use, then 'S' is just meaningless squiggle.

So runs Wittgenstein's PLA. He concludes that it is not possible to introduce a sign with a *necessarily* private meaning.

But if Wittgenstein is correct, and a sign cannot function as a label for a private inner 'something', then 'pain' can't function in that way. But then it seems that a tempting picture of how 'pain' functions cannot be right.

Why do I Need a 'Criterion of Correctness'?

I have just sketched out Wittgenstein's private language argument, but actually, as it stands, the argument might seem obviously incomplete. One question that may already have occurred to you is:

> Why, in order to have introduced a meaningful sign, do I need a 'criterion of correctness'?

Wittgenstein's argument, as presented in section §258, doesn't answer this question. In fact, Wittgenstein seems just to *presuppose* that if there's no criterion of correctness, then there is no such thing as correct. But then Wittgenstein's private language argument would be an obvious failure – it would rely on a claim that is not obviously true, and for which no argument is provided.

But what if Wittgenstein could show that in order to have a meaningful sign, you do need a criterion of correctness? Then the PLA would succeed.

It is at this point that interpretations of Wittgenstein's argument start to differ. Philosophers disagree about *why* Wittgenstein thinks that, in order for a sign to be meaningful, there must be a criterion of correctness, and *why* such a criterion is lacking in the private case. Here is a brief introduction to three main interpretations.

'Suppose I define 'S' by reference to *this* sensation (I'm focusing my attention on it inwardly). Then, a bit later, I have another sensation, and it does *seem* like S again. However, I have no way of checking that it *really* is S again.'

1. The No Check, No Rule Argument

I begin with Glock, who offers the following.

> [R]ules are standards of correctness … [T]here is no such thing as a non-operational standard of correctness, one which cannot even in principle be used to distinguish between correct and incorrect applications.

A rule, suggests Glock, is a standard of correctness. But a standard of correctness is, in turn, something you can *use to check* whether you are applying your sign correctly. If no check is possible, then there is no rule, and so there is no correct or incorrect.

How plausible is the suggestion that, in order for someone to follow a rule, they must be able to *check* they are doing so correctly? I

https://doi.org/10.1017/S1477175625000119 Published online by Cambridge University Press

don't find it plausible at all. Here's a simple counter-example.

Suppose a prisoner in solitary confinement begins to keep a diary in which he records the days he sees a rat. He writes the sign '!' in the top right corner of the page on those days when he sees a rat, and not otherwise. Suppose that, after several weeks of recording rats in this way, the prisoner has a moment of doubt. He wonders whether he has been using '!' to record days on which a rat is seen, or days when no rat is seen.

Notice our prisoner has no operational standard of correctness available to him to check whether he is using '!' correctly. He didn't write down '! = rat', for example. Nor can he appeal to someone else to check how '!' is used. Yet, surely, despite lacking an operational standard of correctness, the prisoner has been following a rule. Were he now to use '!' to record seeing a rat, he would continue to use his sign correctly. But then it is false that there is only a rule where there is an operational standard of correctness. Yes, rules determine what's correct and incorrect. However, they need not determine what's correct and incorrect in a way that we can check.

It seems that what matters, so far as following a rule is concerned, is that you possess a certain *skill* or *ability*. You just need to be a *reliable user*, applying the sign consistently in accordance with your original intention. You don't, in addition, need some way of *checking* you're reliable.

In short, I think Glock's version of the private language argument fails because it rests on a falsehood – it's simply not true that if the user of a sign can't *check* they are following a rule correctly, then there is no correct or incorrect.

2. The Stage-Setting Argument

A rather different interpretation of Wittgenstein's private language argument draws our attention to what immediately precedes section §258. According to Marie McGinn, for example, §258 is best read as a comment on §257, where Wittgenstein says:

> a great deal of stage-setting in the language is presupposed if the mere act of naming is to make sense. And when we speak of someone's having given a name to pain, what is presupposed is the existence of the grammar of the word pain: it shows the post where the new word is stationed.

What does Wittgenstein mean by 'stage-setting'? Elsewhere, he discusses the introduction of a new word, 'tove', defined by saying the word pointing at a pencil. Is this act of pointing a saying enough to give the word a meaning? No, insists Wittgenstein. For 'tove' could mean innumerable things. 'Tove' might mean pencil. Or it might mean *that* pencil (so no other pencil can be 'tove'). Or 'tove' might mean yellow, the colour of the pencil. Or 'tove' might mean wooden object, or writing implement, or human artefact, or object weighing less than 500 g, or object located in the Northern Hemisphere of the Earth. There are countless things 'tove' could mean. The mere act of pointing and saying 'tove' fails to fix which of these innumerable things it means.

According to Wittgenstein, in order to introduce 'tove' successfully as a meaningful sign, a certain amount of 'stage-setting' is required. It needs to be clear, for example, that I am picking out an example of *a certain sort of writing implement*, say. Or, if I am defining a colour term, say, it also needs to be clear how much of the colour spectrum is picked out (so that 'tove' means yellow, rather than, say, only that very specific shade of yellow – lemon yellow).

It seems that under 'stage-setting' Wittgenstein means to include such *pre-existing systems of classification* – systems that carve the world up into different colours, different kinds of writing implement, different kinds of artifact, and so on – so that the position into which 'tove' is to slot then becomes clear. In the absence of such stage-setting, pointing and saying 'tove' is a mere 'empty ceremony'.

However, such systems *are something language provides*. It's through language that we carve the world up into different sorts of colour, writing implement, and so on. But then, in the absence of some pre-existing language, the mere

https://doi.org/10.1017/S1477175625000119 Published online by Cambridge University Press

act of pointing and saying can't succeed in introducing a meaningful sign

If we now return to the definition of 'S' in §258, it seems this sort of stage-setting is missing. I just say 'S' and focus my attention inwardly. But what, exactly, have I named? That's entirely unclear. But then it's unclear what's to count as a 'correct' future application of 'S'. I have failed to lay down a criterion of correctness.

A Response to the Stage-Setting Argument

In reply, a critic might say: but the required stage-setting *is* present. When I define 'S', I make it clear that what I'm naming belongs to the category of *sensations* (under which we divide things up into different sorts of tickle, pain, and so on). That provides the necessary stage-setting.

The problem with this response is that our definition of 'S' is supposed to start a new language *from scratch*. But, just like writing implements, colours, and so on, sensations are also a category of thing carved out by our public language. So, by helping ourselves to the category of sensations, *we're piggy-backing on a prior, public system of linguistic rules.* But then we would not have succeeded in starting a new, *private* language entirely from scratch.

So runs the stage-setting version of the private language argument. I find it unconvincing.

What the Stage-Setting Argument shows, at best, is that language cannot get started with a bare act of ostensive definition, for that requires stage-setting, and stage-setting presupposes a language. But then how did *public* language get started? It did not begin with an act of ostensive definition – with for example a cave person sitting by a fire, pointing to a rock, and saying 'Ug!', with others around the fire nodding in agreement that that was how 'Ug!' would now be used. Public language must have started *in some other way.*

But if public language can be started *in some other way*, other than with an ostensive definition, then why couldn't this other way, whatever it is, be used by someone to develop a private language? That's not clear. But then this version of the PLA, as it stands, also fails.

3. The Circularity Argument

A particularly ingenious version of the private language argument suggests that there is something *fundamentally circular* about the explanation of how 'S' can be applied once it has been introduced.

Elsewhere, Wittgenstein points out that a tempting picture of how we apply for example colour terms cannot be correct. Suppose I am asked to pick the red flower out of a bunch of differently coloured flowers. How do I know which flower to pick? One suggestion is that, in order to recognize an external object (the flower) as red, I need to call up an inner object – a mental sample. I go to my memory bank, as it were, pull out the mental sample I previously learned to associate with 'red', and then compare it with the flowers until I get a match. I then pick that flower.

This explanation can seem plausible until we ask, 'And how did you know *which mental sample to pull up*? How did you know that *that* mental sample is the red sample, rather than, say, the yellow or blue sample?' If, as you just assumed, in order to identify something as being red, you must first compare it to a mental sample, then to identify which sample is red, you would need to compare *that* sample to a mental sample. And in order to identify that second sample as being red, you'd need to compare it to a third sample, and so on ad infinitum! In other words, this explanation of how you identify which flower is red is hopelessly circular.

So it seems, on reflection, that you must be able to identify *some* things as red *without comparing them to anything at all.* But if that's true, *why can't it be true of the red flower?* What is the point of introducing the mental sample, if it just presupposes what it's supposed to explain: the ability to recognize things as red?

This is a wonderful little argument. And the suggestion from some interpreters of §258 is

https://doi.org/10.1017/S1477175625000119 Published online by Cambridge University Press

that Wittgenstein is offering a similar criticism of how it is that the private linguist is supposed to identify something as being 'S' again. Suppose I define 'S' by focusing my attention on *this* mental item. If I want to know, later, whether some new mental item is 'S', I'll need to compare it with my original sample. But then I'll need to be able to identify, among all the previous experiences I've had, which one I previously labelled 'S'. But that requires that I *already* be able to identify what is and isn't 'S'.

The situation is quite different if I define 'S' by reference to a public sample, such as a piece of coloured card on which I have written 'S'. In order to know whether some new item is 'S' again, I don't need to remember what 'S' means. Instead, I can just rummage in my drawer of samples and find the one marked 'S'. But in the private linguist's case, their sample isn't something objective that they can physically mark and file away for future reference. They just have to *remember* what they previously labelled 'S'. And that requires that they already have the ability to identify which of their previous experiences is 'S'. So *they need already to know what 'S' means*. But then, as the proposed account of how the linguist knows what 'S' means just presupposes they know what 'S' means, it explains nothing.

As Peter Hacker puts it,

> [d]eliberately calling up the *memory* sample of *S* rather than some other sensation requires that one knows what '*S*' means; yet calling up *this* sample was meant to be what knowing the meaning of '*S*' consists in, not to presuppose it.

This is an ingenious version of the private language argument, but I still think it fails. Notice that Hacker assumes that, in order to know whether this is S again, the private linguist needs to *call up the right memory sample.* But remember, Wittgenstein denies that, in order to apply 'red' correctly, you need to call up a memory sample. He thinks you can just identify which flower is red *without comparing it to anything at all*. And if that's true of flowers, why can't it be true of private sensations?

So, to repeat, the problem with the Circularity Argument is that *it presupposes that applying 'S' again requires that the private linguist must engage in a mental looking-up process*: they must call up their original sample and compare it to what they have now. But as Wittgenstein elsewhere *denies* such a looking up process is required in order to identify things as being red, why assume such a process is required in order to identify things as being S? Why can't the private linguist identify a new experience as being S again, *without comparing it to anything at all?* Surely they can. But then there's no reason to deny that the private linguist can apply 'S' meaningfully.

Conclusion

We have looked at three different interpretations of Wittgenstein's PLA. On none of them does the argument succeed. Perhaps there is a good argument against the possibility of a private language, but, as yet, I've not found one.

Stephen Law

Stephen Law is Departmental Lecturer in Philosophy, Department of Continuing Education, University of Oxford.

Cite this article: Law S (2025) A Short Introduction to Wittgenstein's Private Language Argument. *Think* **24**, 11–18. https://doi.org/10.1017/S1477175625000119

© The Author(s), 2025. Published by Cambridge University Press on behalf of The Royal Institute of Philosophy. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/S1477175625000119 Published online by Cambridge University Press