The Private Language Argument and Mind-Body Dualism: A Reassessment

Álvaro R. G. Barredo
The Private Language Argument (PLA) turned Wittgenstein’s *Philosophical Investigations* into a source of heated debate due to the overarching implications it has for the way we have traditionally understood the endeavour of philosophy, casting doubt on our ability of introspection, and even threatening to render such a concept unintelligible. In this essay, I shall explore the bearing the PLA may have on the philosophy of mind, and, more precisely, whether it is a conclusive objection to mind-body dualism – a concern that has been raised before in the literature (Villanueva, 2: 30). The PLA, under a common interpretation, reduces all putative mental states to dispositional or behaviouristic states (Luckhardt, 1983: 319). Given the pervasiveness of this reading, Wittgenstein becomes a relevant force to reckon with when considering the mind-body debate.

The “core” of the PLA is stated in §§243ff. of the *Investigations*. As such, the theses put forth there will be the main focus of our investigation. I aim to explore whether the PLA, strictly speaking, is deleterious to mind-body dualism, not whether Wittgenstein would approve of a dualistic philosophy of mind. That being said, the PLA is notoriously ambiguous if read on its own, which means that references to other parts of the *Investigations* will be inevitable, if only to elucidate what is actually meant by it.

In order to explore, then, whether we can support any kind of body-mind dualism and accept, at the same time, the validity of the PLA, I will proceed as follows. I will begin by challenging some of the so-called “orthodox interpretations” (Stern, 2011: 331) of the PLA, and proposing what I find to be its most plausible characterization, namely, the PLA as a special case of a general problem with identity and ostension. After that, I will dedicate some space to addressing specifically behaviouristic concerns regarding the PLA. Finally, I will discuss the different ways in which we may understand
mind-body dualism and I will show that the PLA does have a bearing on some sorts of dualism, while not necessarily on others.

II. Elucidating the PLA

As I have mentioned, the PLA is notably obscure, and figuring out what it actually means has preoccupied much of the literature on the topic. Wittgenstein’s aphoristic style has not lent itself to easy formalization, so much so that some have argued that interpreting §§243ff. as an “argument” of any sorts is to misconstrue Wittgenstein’s point (Stern, 2011: 342-3). Wittgenstein, according to this line of reasoning, would not be interested in “proving” the impossibility of a private language by means of a reduction ad absurdum; rather, his aphoristic style would be warranted by the need to show, not prove, the unintelligibility of the thesis. This interpretation is not without its merits, and it is probably adequate to the latter Wittgenstein’s general anti-theorizing attitude (Pears, 1988: 214-215).

Nevertheless, I think it will be most appropriate for us to treat it as an argument, if, perhaps, not as a simple reductio ad absurdum, due to several reasons. First of all, some level of formalization is useful if we are to objectively assess the implications of Wittgenstein’s treatment of private languages; we will hardly be able to draw clear conclusions from vague aphorisms taken at face value. Aside from that, however, since our aim relates to the PLA, and to how it has been covered by the literature, it does not seem necessary to dialogue with Wittgenstein’s idiosyncrasies if they do not directly contribute to this particular academic debate. Thus, I will attempt to formalize the PLA, beginning by showing why I think two prevalent approaches – the “fallibility of memory” approach and the “verificationist” approach – are lacking. Subsequently, I will present my own interpretation.

a. The PLA as memory scepticism
Most commentators who take the PLA as an argument proper seem to agree, at least, that it implies what follows.106

(1) A private language stands in opposition to a public language insofar as the meaning of its terms is privately set.

(2) No meaning can be privately set.

(3) Therefore, there can be no private language.

Many of these terms are in dire need of definition, and that is so intentionally, since the crux of the dispute resides in how we come to understand them. Namely, we will see that what we mean by “privately set”, and what we base (2) on, will suppose the main source of disagreement among commentators. I will call those nodes of dissent the “privacy clause” (PC) and the “criterion-setting clause” (CC), which should be added as elided premises to the main argument.

The “fallibility of memory” approach, most famously defended by A. J. Ayer, interprets the PLA to entail this:

(PC): A language is private when the objects it refers to are, themselves, private. (Ayer, 1954: 64)107

(CC): No meaning can be privately set because, if we grant that we cannot immediately ascertain how to use a private term, then we cannot trust any of our private grounds for evidence. (ibid.: 68)108

---

106 For a general review of the bibliography around the PLA, vid. (Villanueva, 1975a), (Villanueva, 1975b) and (Stern, 2011).

107 “What philosophers usually seem to have in mind when they speak of a private language is one that is, in their view, necessarily private, in as much as it is used by some particular person to refer only to his own private experiences”.

108 “For if one cannot be trusted to recognize one [private sensation], neither can one be trusted to recognize the other”.

124
Once the PLA is set up like this, Ayer has strong reasons to dismiss it. If we cannot trust any of our private grounds of evidence (sight, memory, etc.), then it is not just private languages that cause trouble; it appears to be impossible to use any language (ibid.).

Much has been written against Ayer’s interpretation on both fronts. Regarding (PC), Ayer has been accused of misinterpreting what is relevant about the hypothetical private language that Wittgenstein discusses. It is not that it denotes private objects, but that it is a language that nobody but its user may, even potentially, come to learn.\(^{109}\) (Thomson, 1964: 20; Villanueva, 1975: 81; Luckhardt, 1983: 327; Stern, 2011: 333). We are detaching ourselves from the realm of the languages that we actually do use (Candlish, 1980: 86) since, as far as we know, our languages are inter-translatable, and, what is more, the entire point of the PLA is to show that any such languages are a logical impossibility. Ayer’s (CC) does not fare much better. I will expound more on this point, but Wittgenstein’s concern is not that we may be “fallible” when confined to our private fora; rather, that there cannot be anything like a criterion of correctness, fallible or not, that is entirely private (Pears, 1988: 333). In other words, Ayer interprets (CC) to mean something akin to: there is a process \(P\) by which I identify private objects and I name them. \(P\) has a non-zero chance of failing, therefore, \(P\) is not to be trusted. The actual clause in the PLA seems to be, on the contrary, that there can be no such process \(P\).\(^{110}\)

\[b. \] The PLA as verificationism

Some of our points merit further elucidation, and they will be subject to closer examination in the next section. Before doing that, however, it is necessary that we

---

\(^{109}\) This is evidenced, for instance, in PI §261. “‘Has’ and ‘something’ also belong to our common language”. What does he mean by this? Our putative private linguist is one which in no way depends on terms whose meaning may be publicly communicated; it is a matter of fact that we can publicly talk about “private objects” – as evidenced by this very paragraph. That does not mean, however, that we can “denote” private objects.

\(^{110}\) Wittgenstein does not merely say that we can fail; “in the present case, I have no criterion of correctness” (PI §258)
address the interpretation Judith Jarvis Thomson proposes for the PLA. She contends that we ought to interpret said clauses as follows:

(PC): A language is private when it is logically impossible for anyone but its sole user to understand it. (Thomson, 1964: 21)

(CC): No meaning can be privately set, because, for a sign to be a kind-name, it must be possible to find out (publicly) whether a thing is of that kind, which amounts to the principle of verification (ibid.: 29).

Thomson’s assessment of the (PC) seems entirely adequate, but I cannot agree with her understanding of the (CC) as a restatement of the verification principle. The reason why will be made clearer in the following section, but, as of now, we can consider an example, and compare Wittgenstein’s actual assessment to a “verificationist” one.

Compare a person undergoing tremendous pain to a great actor, who mimics “pain-behaviour” to such degree of perfection that there is no discernible difference between their acting and actual pain-behaviour. Consider this actor performing such moving scene on a stage. We can tell that this person is not “actually” in pain, they simply excel at their art, and, if pressed, we can mention other extraneous factors to support our judgment, like them not leaving the stage, the normal reactions of their colleagues, and so forth. But how does that amount to finding out, in a verificationist sense, that they are not, in fact, in pain? Is it essential to pain-behaviour that it transpires outside a stage? Or that witnesses react in a given manner? It does not appear so, yet we do not think it misjudged to say that the actor was not actually in pain, and, more

111 The main question here revolves around the possibility of mimicry, deception or acting. There are two theses that seem to conform to our common understanding. (1) Somebody may convincingly fake the behaviour associated to a mental state; (2) faking a behaviour implies not being in the mental state typically associated to it. Given that, there are contexts where there is an expectation for faking, and where it seems we can be said to, despite perceiving the exact same behaviour that would make us think that somebody is experiencing some mental state, tell it apart from the “actual thing”. If this is so, a purely verificationist stance does not adequately portray the language games at stake here, which go “beyond” denotation. (Cf. Putnam, 1980: 29)
importantly, neither does Wittgenstein.\textsuperscript{112} The verificationist approach does not seem to be compatible with Wittgenstein’s actual theses.

c. \textit{The general problem with ostension}

In order to present our interpretation of the PLA, we need to characterize a general issue present in Wittgenstein’s philosophy, of which the PLA would be a particular case. This issue is brought up by Thomson in the previously cited article:

The question, “How do I identify a kind of sensation?” is a very respectable philosophical question. But of course it is only a special case of the very respectable philosophical question, “How do I identify a kind of \textit{thing}?" (\textit{ibid.}: 26)

And I cannot but agree with her, since therein lies the question. The \textit{Investigations} may be primarily concerned with so-called “inner states”, but it starts off as a general discussion about language use. How do we come to use any word whatsoever, if, for any rule on word-using we may encounter, we would need yet another rule on rule-following, falling into a \textit{regressus ad infinitum}? (\textit{vid. PI} §86; Kripke, 1984: 62)

This problem has been, perhaps, most famously exposed by Saul Kripke in his work on the PLA. I will skim over his very suggestive interpretation, due to space constraints. Shortly, Kripke posits that the \textit{Investigations} are chiefly concerned with a so-called “sceptic” objection to all rule-following (\textit{ibid.}: 8). It seems as though we act in certain ways – for instance, giving the “correct” answers to arithmetic problems – because we follow specific rules in doing so. Nonetheless, we can never be confident that we are justified in following a rule, because any amount of past instances that, putatively,

\footnote{\textsuperscript{112}“‘But you will surely admit that there is a difference between pain-behaviour accompanied by pain and pain-behaviour without any pain?’–Admit it? What greater difference could there be?’” (\textit{PI} §304)}
express the rule, could be made fit into arbitrarily many other rules. A finite set of past instances is never “enough evidence” for a particular general rule, since there can always be counterexamples not covered by that finite set that we add to construct an apropos alternative rule. (ibid.: 18).

I have no objections to this interpretation; I am, in fact, highly indebted to it. Nonetheless, we need not operate at such a level of abstraction to make our case. I will apply Kripke’s problem of rule-following to the more concrete problems of ostension and identity, which will be shown to be necessarily intertwined.

In order for us to “point at things” and use these crude denotations as the building blocks of a language, as does Wittgenstein in PI §2, we surely need to have a criterion to distinguish what is “equal” from what is not. That is, in order for me to be proficient at bringing bricks, I need to know which two things are equal qua bricks, or, what “about the brick” is being pointed at when I am taught what a brick is. If I do not, I will not be able to follow the task at all. Let us move away from bricks, and consider the Greek alphabet. Alpha, delta, and lambda are three similar-looking yet completely different letters; my proficiency at decoding the name: ΑΛΚΙΒΙΑΔΗΣ

depends on my being able to understand which differences between letters are “significant” to this task, namely, reading a Greek name. The first and second alphas are not “the same”, insofar as they are two distinct “tokens”, but they are of the same type. Now consider the following diagram:

![Diagram](image)

113 Cf. PI §33. Am I pointing to the rectangular shape of the brick’s face? To its colour? Ostension is always ambiguous on its own.
It is not clear now whether we are before three tokens of the same type or not. They seem to be slightly differing arrows, which point at circles nevertheless. They have been deprived from their original context, and what constitutes “equalness” in this new environment becomes ambiguous.

All of this is to illustrate Wittgenstein’s note in PI §215: “Then are two things the same when they are what one thing is? And how am I to apply what the one thing shews me to the case of two things?” Equalness itself, which seems to be at the heart of all possible ostension, requires a criterion, because it is not self-evident (Pears, 1988: 386). Let us turn back to the experiment described in §258. Our private linguist has some sensation on day 1, let us call it $S_1$, accordingly, and she denotes it with an $S$ on her diary. It is day 2 now, and she experiences $S_2$. How is she to proceed now, to know whether it is appropriate to write down an $S$ or not? Contra Ayer, we can endow our linguist with a perfect memory, she can mentally reconstruct her every past state at will. Yet, she will have to compare $S_1$ and $S_2$ and emit a judgment on whether they are tokens of the same type or not. How will she do it? Trying to appeal to higher-order disambiguation criteria just moves the problem a step backwards, since she will have to wonder whether this situation is “equal” to a past situation where disambiguation rule $R$ applied, running again into the regressus problem.

We seem to have reached an impasse, since this objection would apparently hold equal to private and public languages. There is, nonetheless, a crucial distinction. The public forum establishes what we may call semi-rigid grounds of significance, which are to be regarded as “brute facts” given that we can use language to communicate (Kripke, 1984: 98). (1) There needs to be some regularity in the world in order for us to be able to assign meaning to our terms. If things could never reliably be said to have a colour, and our visual perceptions were chimeras, we could hardly be expected to come
to apprehend a colour-language.\textsuperscript{114} (Rhees, 1954: 93). The “stubbornness” of reality makes quaddition-like formulas ultimately unusable. But, of course, what counts as “regularity”, as we have already discussed, cannot be established \textit{a priori} (Kripke, 1984: 105); we also need (2) some regularity in the linguistic uses of the community of speakers, to whom the appropriate use of terms is of some significance, and are thus able to enforce it and teach it. (Kripke, 1984: 96; Pears, 1988: 370).

Why cannot we disambiguate ostension by purely private means? A private context runs into regression problems because there is never a “last” ground of justification that is not “simply chosen” to be so, and thus amounts to not distinguishing between “being right” and “feeling right”, which is the entire point (PI §258). Meanwhile, the tree you run into, or the teacher that corrects you, do not admit further appeals, they are “coercive” in their disambiguation.\textsuperscript{115} This does not mean that, analysed in the abstract, public practices are without ambiguity, but the coerciveness of use overrides the need for an “indubitable” grounding.

The world and our language colleagues conform the necessary context in which we can disambiguate ostension, they are “what happened before and after the pointing” (PI §35). They are only \textit{semi-rigid} grounds, because the meaning of our words does change, and there is room for idiolectal variation, but this has to be ultimately constrained within the bounds of usability. We only get to disambiguate our terms if there is any consequence to getting them wrong. Therefore, after this laborious exercise at elucidation, I can give my proposed interpretations of (PC) and (CC):

\textsuperscript{114} I interpret PI §80 to serve a double function. On the one hand, at face value, it is a reflection about how our rule-following does not depend on our effectively being able to know how to use the rule under outlandish circumstances. But, additionally, it points out how, for a rule to be meaningful, there needs to be some regularity to the cases where it applies. Our language about chairs is not equipped to talk about flickering and disappearing objects, because it does not need to be. If chairs did flicker and disappear, however, our language would not be appropriate.

\textsuperscript{115} Cf. PI §303. “Just try—in a real case—to doubt someone else’s fear or pain”.

130
(PC): A language is private when it is logically impossible for anyone but its sole user to understand it.

(CC): No meaning can be privately set because any criterion of identity, public or private, requires disambiguation through semi-rigid grounds of significance. Any purely private context fails in this regard because it is always subject to a regressus ad infinitum where no ultimate ground of justification is to be found. Public contexts find a way to halt the regression by coercing the speaker through the necessity to act.

III. Dualism and the PLA

Having elucidated an operational form of the PLA, we can now move on to analysing our main concern, whether it has any bearing on mind-body dualism. I will begin by addressing a possible way to interpret the PLA that would immediately discard any kind of dualism.

a. The behaviourist challenge

We have established that, for any term to have a set meaning, it requires a public context of disambiguation, making all attempts at constructing a private language meaningless. What does this entail, however, for the terms we allegedly use to denote private objects like “pain”? A behaviourist interpretation seems simple enough:

(1) Terms like “pain” have meaning for us.

(2) Per the PLA, no meaning can be privately set.

(3) Therefore, the meaning of “pain” is publicly set.

(4) Pain-behaviour is public, pain-sensations are private.

(5) Therefore, the meaning of “pain” cannot be grounded on pain-sensations.

(6) Therefore, “pain” denotes pain-behaviour.
If this is so, we do not use “pain” to denote anything about our inner experience, but about some behaviour through which we can be taught to talk about “pain”, and through which a certain use of the word can be enforced. Wittgenstein, in fact, is adamant about how the word “pain” is not used to denote or describe any hidden mental state (PI §290). Is all talk about dualism linguistic nonsense then, dissolved by the PLA?

Not quite. There is a clearly unjustified leap from (5) to (6). Up until then, the argument holds, it is true that under the provisos of the PLA we cannot ground the meaning of “pain” on pain-sensations (Pears, 1988: 350). But that does not mean that “pain” needs to have a denotative function of any sort, let alone that it needs to denote behaviour. Recall our earlier example about the actor. How do we explain it under our current interpretation of the PLA? As members of a community, we come to disambiguate our references to pain on contextual bases. We can only tell whether certain behaviour is “pain”, or “acting”, or anything at all based on how our linguistic community has acted regarding certain scenarios and how they have enforced the use of certain rules. We can imagine a child going for the first time to a theatre and telling his parents that the actor needs help, an assessment that the parents would correct by noting how “pain” – the “pain-language game” – does not apply in that situation.\footnote{116 Cf. PI §584.}

There is a difference between assertability conditions (Kripke, 1984: 111) and the meaning of a term. A given behaviour is necessary for us to come to learn the meaning of a sensation-term, but the meaning of a sensation-term is not exhausted by behaviour (vid. Putnam, 1967: 57-8; Luckhardt, 1983: 328).

\textit{b. What do we understand by dualism?}

What we can derive from the past discussion is that, even if we grant that there are such things as “private objects”, we would not be able to define them by ostension without the
concurrence of publicly enforced criteria. Our discussion about the private linguist
distinguishing between $S_1$ and $S_2$ tacitly implied their persistence as “objects” of some
sort, and their ontological status played no role in our argument. Strictly speaking, then,
the PLA says nothing about the ontology of the private forum, and, in a trivial sense, it is
compatible with any manner of dualism.

But this answer is not satisfactory, because, as J.J.C. Smart puts it, even though
a state of affairs about our mental reality may be compatible with several explanations,
mere compatibility is not enough to merit accepting any one of them (Smart, 1959: 155-6). The question we should be asking is, do we have any reasons to maintain dualism
given the PLA?

Mind-body dualism comes in many different shapes. A tripartite distinction that
has enjoyed some popularity, regarding the different ontological presuppositions that
dualism may have, is that of substance, property, and predicate dualism (Robinson,
2020), in decreasing order of ontological commitment. We may characterize them as
follows:\textsuperscript{117}

Substance dualism, which would be a thesis such as the one espoused by René
Descartes, holds that:

(1) There are mental states, different from physical states.

(2) The mental states of a subject $S$ correspond to mental properties of $S$.

(3) These mental properties belong to a distinct mental substance.

\textsuperscript{117} This is a decidedly simplistic account of what are deeply complex theories about the mind. Nonetheless, focusing on these three particular theses seems to (a) show the main sources of disagreement between the three positions, (b) establish some points whose contention against the PLA seems most relevant. Property dualism, for instance, may say much more than the very vague theses (1) and (2) would have it, but it is of the utmost importance whether the PLA posits serious problems to those theses; more so than other, perhaps, less central stances within such theory.
Property dualism does not commit to thesis (3), being compatible with the idea that there are only physical substances; emergentism being an example of it (Mitchell, 2010: 172), and predicate dualism does not commit to either theses (2) or (3), basing itself, for instance, on multiple realizability to argue that mental states, while reducible as tokens to physical states, are not so reducible as types; an example of this being Davidson’s anomalous monism (Davidson, 1970: 99-100).

By characterizing these three classes of dualism, we can see that the more ontologically compromised theories necessarily entail the less compromised ones. Substance dualism, for instance, as characterized, would entail both property and predicate dualism. There could be other stances, but these seem to be the most useful for our discussion.

Let us begin, then, by assessing the plausibility of thesis (3) vis-à-vis the PLA. The PLA says nothing about ontology, but what reasons could we have to support the existence of a mental substance? Descartes argues that, since we first come to be certain of our being mental, and we can have a clear and distinct notion of it (2011 [1641]: 76-7), the union between our mental and physical states is to be held as contingent, and, thus, said states correspond to different substances.

If we accept the PLA, however, we cannot sustain that there be any privilege in acquiring knowledge about our mental states; Cartesian introspection is required to make this argument work, but the PLA forces mental concepts to be set in the very same public forum as physical concepts, as we have already seen. Thus, the source of distinctiveness that Descartes alleges as sufficient reason to defend the existence of a res cogitans is lost. The PLA equalizes the ground for all states, mental and physical, there is no priority

---

118 Mitchell tackles emergence tout court, but that is, naturally, applicable to our case.
other than that established by the linguistic uses of the community. Therefore, multiplying the substances, if we are to accept the PLA, seems capricious.\textsuperscript{119}

What about thesis (2)? The argument for characterizing mental states as properties of their own can take many shapes. We can consider emergent states, says Mitchell, as relevant entities subject to natural selection, for example, or as causally efficient (2010: 179-80). This seems like the sort of hypostasising that Wittgenstein would forbid (\textit{vid.} PI II§76), but we must recall that we are not accepting or assessing all of Wittgenstein’s psychology, we are merely addressing the relation between the PLA and dualism. Does the public setting of meaning affect in any way the assessment that mental states may be considered as properties of physical substances, insofar as they are causally efficient, or insofar as they are subjected, by their own, to natural selection? There may be other arguments against them, but it does not seem that what is posited by the PLA alone does anything to problematize them. We learn publicly, for example, to refer to some mental state of ours that precedes our \textit{acting} as “determination”. We cannot learn to use the word by ourselves, but this does not preclude that, once we learn to use it, we think it best to analyse it as a property that instantiates onto us.

If this is so, and the PLA does not pose serious problems to thesis (2), \textit{a fortiori} it will not be problematic for thesis (1). Thesis (1) is not directly implied by thesis (2), since it is an assertion about there \textit{being} mental states. But this one has been already tackled by our previous discussions regarding verificationism and behaviourism. Privacy is not discarded from our language games. Our using it and talking about it cannot be fully independent from the physical and public (Pears, 1988: 350), but that does not entail that we may \textit{reduce} mental states to physical states (Luckhardt, 1983: 329). Terms about

\textsuperscript{119} Cf. PI §293; the example of the beetle in the box is particularly relevant to this effect. Once again, it is impossible to tackle all forms of substance dualism. The main argument, in any case, is that a particularly strong form of first person privilege seems necessary to deem substance dualism a reasonable position to sustain. If someone were to convincingly sustain the necessity for \textit{two substances} even without said privilege, then it would not fail the PLA test either.
mental states are, as a matter of fact, present all throughout our common linguistic experience, and they do not necessarily seem to be reducible.

IV. Conclusion

It may be legitimately objected that the sort of dualism that withstands the PLA is so far removed from Cartesian dualism that it is not appropriate to even consider it at the same level. However, it does not appear that the discussion has been fruitless. I have attempted to present a non-behaviourist, non-verificationist view of the PLA, which allows for far more flexibility in the status we may attribute to mental states. If I have succeeded in my argument, the PLA does not reduce mental terms to their behavioural counterparts, it simply establishes general conditions for criteria-setting, which then may apply beyond the strict scope of what is publicly verifiable. This is a notable shift from the starting point, and, if it is not to be called dualism – although emergentist stances are typically called dualist (Gregory and Zangwill, 1987: 204) – it certainly is not pure physicalism.

There are many interesting topics relating to the general PLA discussion we have not been able to tackle here, and which may warrant further research. Is reductive physicalism even intelligible from a Wittgensteinian point of view since it ignores the problems regarding rule-following and criteria of identity? What are the links between neutral monism and language games grounded on an indeterminate sort of world regularity? Can there be an ontologically uncompromised Wittgensteinian functionalism? How does the PLA fare with qualia? These, among many others, are questions we will have to leave unanswered for the time being.
Works cited


Álvaro is a last year Philosophy student at Universidad Nacional de Educación a Distancia (Spain), and he recently received his BA in Philosophy, Politics and Economics at Universidad Carlos III de Madrid. His research interests include Kant's philosophy and the philosophies of mind, action and biology. Starting in October 2022, as part of "La Caixa" fellowship for postgraduate studies abroad, he will be studying a Masters by Research in Philosophy at Durham University (UK). Aside from philosophy, Álvaro likes to participate at radio shows, gardening, and (poorly) playing chess.