What We Can Learn about Phenomenal Concepts from Wittgenstein’s Private Language

Abstract

This paper is both systematic and historical in nature. From a historical viewpoint, I aim to show that to establish Wittgenstein’s claim that “an ‘inner process’ stands in need of outward criteria” (PI §580) there is an enthymeme in Wittgenstein’s private language argument (henceforth PLA) overlooked in the literature, namely Wittgenstein’s suggestion that both perceptual and bodily experiences are transparent in the relevant sense that one cannot point to a mental state and wonder “What is that?” From a systematic viewpoint, I aim to show that Wittgenstein’s PLA teaches us that the prevailing picture of the nature of phenomenal concepts (henceforth PCs) is upside down: we can only introspectively know what is going on inside our heads, after we learn of what is going on outside (PI §580). In this regard, I aim to defend two associate claims against the prevailing view of PCs on the basis of PLA. First, by means of transparency, I aim to show that there is no de re awareness of our private sensation that could determine the meaning of sensation-words; for example, I am never aware of the phenomenal blueness of my experience of something blue. The second associated claim is that introspective self-knowledge of our private sensation is always de dicto. We can only know introspectively that phenomenal blueness is the phenomenal character of the experience we are undergoing after we have learned that (de dicto knowledge) blue is the color that usually causes in us that kind of experience. Likewise, we can only introspectively know that pain is the phenomenal character of the experience we are undergoing after we have learned that pain is what usually causes some typical pain behavior.
Introduction

Recently, David Papineau has formulated a direct criticism of Wittgenstein’s private language argument (henceforth PLA). The meaning of sensation-words (the linguistic analogue of phenomenal concepts, henceforth PCs) is determined a priori by an association between sensation-words and private sensation, and is only a posteriori communicated (either by indicating the objects that usually cause perceptual sensation, or by indicating the typical behavior associated with the bodily sensation). Marianne stares at a blue patch on the wall of her Technicolor vestibule (where she is imprisoned after her release from the original black-and-white room). Now, by turning her mind inwards, so to speak, she comes to learn what is like to experience blue and gives a name to it before she communicates it.

This paper is both systematic and historical in nature. From a historical viewpoint, I aim to show that to establish Wittgenstein’s claim that “an ‘inner process’ stands in need of outward criteria” (PI §580), there is the enthymeme in Wittgenstein’s private language argument (henceforth PLA) overlooked in the literature, namely Wittgenstein’s suggestion that both perceptual and bodily experiences are transparent in the relevant sense that one cannot point to a mental state and wonder “What is that?” From a systematic viewpoint, I aim to show that Wittgenstein’s PLA teaches us that the prevailing picture of the nature of phenomenal concepts is upside down: we can only introspectively know what is going on inside our heads, after we learn of what is going on outside (PI §580). In this regard, I aim to defend two associate claims against the prevailing view of PCs on the basis of PLA. First, by means of transparency, I aim to show that there is no de re awareness of our private sensation that could determine the meaning of sensation-words; for example, I am never aware of phenomenal blueness of my experience of something blue. The second associated claim is that introspective self-knowledge of our private sensation is always de dicto. We can only know introspectively that phenomenal blueness is the phenomenal

---

1 See Papineau 2011.
character of the experience we are undergoing after we have learned that blue is the color that usually causes in us that kind of experience. Likewise, we can only introspectively know that pain is the phenomenal character of the experience we are undergoing after we have learned that pain is what usually causes some typical pain behavior.2

The paper is conceived in five sections. After this brief introduction, the next section is dedicated to the historical reconstruction of PLA. In the third, I will explain the rationale for positing phenomenal concepts: the so-called phenomenal concept strategy. In the fourth section, I will expose the bone of contention between proponents of PCs and Wittgenstein’s PLA. The focus here is Papineau’s direct criticism of PLA. In the last section, I will argue in favor of the claim that we can only introspectively know what is going on inside our heads when we know what is going on outside.

1. Setting the Stage

The first task of this section is to clarify the major structure of PLA and to dispel a series of misunderstandings. In the opening remarks at PI §243, Wittgenstein announces what he means by private language:

But 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.

At PI §246, he adds:

In what sense are my sensations private? Well, only I can know whether I am really in pain; another person can only surmise it. — In one way this is false, and in the other is nonsense.

2 Wittgensteinians usually take issue with the assumption that mental states cause behavior rather than rationalize it. But I take that to be a minor problem. I prefer to here follow Papineau’s manner of phrasing the problem.
The first question we have to address is the meaning of a private language as an incommunicable language. According to Kenny, this incommunicability is a conjunction of two independent theses: (i) I can know that I am in pain; (ii) other people cannot know that I am in pain.\(^3\) In contrast, Tugendhat disputes Kenny’s first feature and emphasizes the modal “cannot” in the second. A private langue is one whose words immediately refer to the private experience of the individual, and hence one that cannot be understood by somebody else.\(^4\)

However, the crucial explanation is missing. What makes sensation-words incommunicable is not the fact that they refer to private experiences. Rather, the distinguishing feature of a private language is that the connection between sensation-words and whatever is going on outside the mind is purely contingent. The idea is that sensation-words are individuated independently of anything that is going on outside the mind. First, I name a sensation by turning my gaze inwards. Later I observe what is causing that peculiar sensation or, alternatively, which behavior that sensation is causing. Thus, I communicate indirectly, by means the speech, what I am feeling.\(^5\) Still, something remains essentially incommunicable, namely qualia, that is, what is like to in a sensory state.

Another source of misunderstanding arises from what Wittgenstein states at the end of the famous passage PI §258:

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.”

This passage gives rise to the misunderstanding of PLA as an argument about the reliability of memory and memory judgments. Ayer, for example, has tried to get from certain followers of Wittgenstein straight answers to the questions: “Why must we have

---


\(^4\) Tugendhat 1979: 81.

\(^5\) See LPE: 279. I will return to this point.
objective checks upon memory?” and “How can we have independent checks on memory as such?” However, the appeal to memory does not mean that PLA targets some form of skepticism about memory or memory judgment. The idea is that we cannot appeal to another mental state as a criterion for correction of the selection of a private sensation that has putatively given meaning to a sensation-word. Wittgenstein clarifies this point at PI §265:

– 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.)

Another source of misunderstanding is the confusion of private ownership of sensory states with private meaning:

The essential thing about private experience is really not that each person possesses his own exemplar, but that nobody knows whether other people also have this or something else. The assumption would

---

6 See Ayer 1954.
7 See Kenny: “This crucial passage [PI §258] is often misinterpreted. What would it be to remember the connection right? Many philosophers have taken ‘I remember the connection right’ to mean ‘I use “S” when and only when I really have S.’ They then take Wittgenstein’s argument to be based on skepticism about memory: how can you be sure that you have remembered aright when next you call a sensation ‘S’?” (2006: 151). See Hacker: “The point does not concern the fallibility of memory, but is rather that the putative mental ostensive definition was intended to provide a rule for the correct use of ‘S’ and now it transpires that in order to do so it presupposes the concept ‘S’. For to remember correctly can only be to remember that a certain sensation or mental image is an image of S” (1987: 267).
8 Kenny was one of the first to call attention to this distinction: “Using words that are not used by Wittgenstein, we may use for the first sense of ‘private’ the abbreviation ‘incommunicable’ and for the second sense of ‘private’ the abbreviation ‘inalienable’. The question ‘Are sensations private?’ breaks up into two questions: (i) ‘Are sensations incommunicable?’ (ii) ‘Are sensations inalienable?’ Crudely, Wittgenstein’s answer to (i) is ‘No’ and to (ii) is ‘Not in any way peculiar to sensations” (2006: 146). Hacker (1987: 263) emphasizes the same point.
thus be possible though unverifiable – that one section of mankind had one sensation of red and another section another. (PI §272)

In this sense of private ownership, Wittgenstein could certainly agree with Nagel (1974) when he claims that to feel what it is like to be a bat, you certainly have to be one. The problem is whether you need to be a bat in order to understand what is like to be one.

A further misunderstanding emerges with the quite interesting but unorthodox Kripkean reading.9 Fogelin has proposed a similar reading.10 The focus there is Wittgenstein’s previous remarks about following rules, practices, and the “form of life” (see PI §201 and 241). According to Kripke, what is in question in PLA is how one can know that I am following a rule correctly and the answer is: consulting the community in which I am embedded as a speaker. This requirement of community agreement for meaning renders the further development of PLA (PI §§256–271) senseless. Similarly, according to Fogelin, what is in question according to Fogelin is how to use an expression meaningfully according to a practice and his answer is training. Given this, Fogelin reads the famous diarist passage at PI §258 as if Wittgenstein was stating that such language is impossible because there is no trainer for the private diary keeper, and therefore no agreement or disagreement.11

Here we find the well-known chasm in the literature. According to the mainstream of Wittgensteinian scholars, the focus of PLA is Wittgenstein’s remarks §§256–271, rather than the previous discussion about following a rule. Again, what is in question in PLA is not the possibility of solitary language, but the possibility of a private language. Moreover, according to Backer and Hacker, the solution to the problem of how to follow a rule is not consultation with the community where the speaker is embedded, but rather the

---

10 See Fogelin 1976: ch. XII and Fogelin 1987. In his last paper about PLA, Fogelin changes his mind somewhat. He continues to claim that the core of PLA is Wittgenstein’s reflection at §202 on rule following. But now he denies that Wittgenstein ever attempted to provide an argument intended to show that private language is private (Fogelin 2012: 92).
understanding of the paradigmatic uses of linguistic expressions as *in some internal relations* with the rule. 12 In this regard, we must distinguish a private language from a solitary, but comprehensible language spoken by one person alone.

But let us return to the idea of a private language. The mainstream Wittgensteinian scholar agrees that what makes a language in principle incommunicable is the crucial assumption that this is a language whose words refers to the speaker’s immediate private sensation to which only she and nobody else has access. That is what we can read from PI §257: “Well, let’s assume that the child is a genius and invents a name for the sensation by himself! ... ‘He gave a name to his sensation’”. But the crucial passage is PI §258, where Wittgenstein explains how this could happen:

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.

The fundamental idea here is that if I cannot literally point to the private sensation to name it, I still can pick it out directly via introspection, namely by turning my mind inward and attending to the private sensation. The passage continues with Wittgenstein claiming that such ceremony is meaningless since we do not possess a criterion of correction:

But what is this ceremony for? 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)

The last fundamental step is a diary passage, at PI §270:

Let us now imagine a use for the entry of the sign “S” in my diary. I find out the following from experience: whenever I have a particular sensation, a manometer shows that my blood pressure is rising. This puts me in a position to report that my blood pressure is rising

12 See Baker and Hacker 1984: 98.
without using any apparatus. This is a useful result. And now it seems quite indifferent whether I’ve recognized the sensation correctly or not. Suppose that I regularly make a mistake in identifying it, this does not make any difference at all. And this alone shows that the supposition of this mistake was merely a sham. (We, as it were, turned a knob which looked as if it could be used to adjust something in the machine; but it was a mere ornament not connected with the mechanism at all.)

And what reason do we have here for calling “S” the name of a sensation? Perhaps the kind of way this sign is employed in this language-game. – And why a “particular sensation”: that is, the same one every time? Well, we’re supposing, aren’t we, that we write “S” every time.

Hacker interprets Wittgenstein as simply denying the possibility of a check on what is not mind-independent. The diary passage at §270 indicates that the subject has no means of checking whether the person is using “S” correctly or not.\textsuperscript{13} Along a different line, Kenny claims that Wittgenstein’s main argument at §258 takes the form of a classical infinite regress.\textsuperscript{14} That reading is focused on PI §265, Wittgenstein’s ironic remark that to appeal to memory is the same as to buy several copies of today’s morning paper to assure oneself that what it said was true. These two plausible readings of §§256–270 show that PLA is quite independent of Wittgenstein’s previous remarks at §201 about following a rule and about meaning as a practice.

However, at the end of his commentaries Hacker raises the crucial question. He formulates it as follows:

One is inclined to think that the use of public samples in ostensive definitions must have a mental analogue. Why cannot a sensation function as a sample for the correct use of a word? (Hacker 1987: 267)

Still, he dismisses it too quickly:

Against this beguiling fiction Wittgenstein directs a battery of arguments. There is no such thing as pointing at any sensation in the sense in which I point at a sample in giving an ostensive definition (PI §258). (Hacker 1987: 267)

\textsuperscript{14} See Kenny 2006: 151-2.
To be sure, we do all agree that, for Wittgenstein, an ostensive definition of sensation-words by acquaintance with the correspondent private sensation via introspection is out of the question. Still, if we take seriously the assumption that we can turn our gaze inwardly and attend to our private sensations, we cannot dismiss without further argument that we are able to pick it out as a sample for further uses of the same sensation-word. It is here that I identify an enthymeme, that is, an implicit premise in PLA.

2. Phenomenal Concept Strategy

“Qualia” as a technical term is usually defined in the literature as the intrinsic, nonphysical, ineffable properties of experience.15 Some philosophers16 use the term “qualia” in a still more restricted way, so that qualia are also “given” to their subjects incorrigibly (without the possibility of error). This is the reason why Dennett denies the very existence of qualia. It is also the reason why several philosophers avoid using the term “qualia,” preferring to use the metaphysical neutral binominals “phenomenal character,” “subjective character,” and “conscious character.”17 Now, whether qualia are nonphysical, intrinsic properties or physical, relational properties, etc., is a metaphysical question that does not concern us here.

PCs are defined as the concepts one deploys to pick out the qualia that one could only acquire by undergoing the relevant experience. For example, I cannot have a phenomenal concept of phenomenal blueness (of what it is like to see the blue color of the sky) unless I have previously had the experience in question. Using this common sense, it is undeniable that what we today call phenomenal concepts are mental analogues of what Wittgenstein called predicates or names of sensation, when he argued against the possibility of a private language. However, before confronting PLA with what philosophers of mind (and cognitive scientists) have to

15 See Tye 2015.
16 Dennett 1991 is the best example.
17 Tye 2015 is the best example that I know.
say about phenomenal concepts, I must present the rationale behind why those concepts have been recently postulated.

The rationale behind the postulation of PCs is to overcome Frank Jackson’s “knowledge argument” for neo-dualism. I believe that everyone reading this will be familiar with the story of Mary, but let us re-describe it. Let us suppose that Mary is a neuroscientist who has an exhaustive knowledge of color (as a physical property) and color vision (the physiological processes involved). However, she has been kept imprisoned in a black-and-white room since her birth. Thus, even knowing everything about what it is to see blue, she has never had the opportunity to see it. One day, she is released from her imprisonment and contemplates the color of the sky for the first time in her life. Thus, she thinks to herself “oh, that is what it is like to see something blue!” The neo-dualist assumption is that Mary obviously learns something; she makes a cognitive discovery, namely, she comes to know what it is like to experience blueness. Nevertheless, as she already had an exhaustive knowledge of color and color vision, the moral to be drawn is that what she learns cannot be a physical property.

There are two classical physicalist reactions to the knowledge argument. The first reaction is rejecting the key assumption that Mary makes a cognitive discovery. The physicalist might directly impugn the neo-dualist conclusion of the argument. There is no ontological chasm between physical and phenomenal properties because there is no genuine discovery or cognitive progress in the first place. Thus, from her exhaustive knowledge of colors and color vision Mary can know a priori what is like to experience blueness. Nevertheless, as she already had an exhaustive knowledge of color and color vision, the moral to be drawn is that what she learns cannot be a physical property.

The second reaction assumes as a fact that Mary makes cognitive progress after she finally sees the color of the sky. After all, we are told, it is a little odd to claim that no one needs to experience blueness to know exactly what it is like to experience it! This second reaction to the knowledge argument is what Chalmers

---

18 See Jackson 1982.

calls type-B materialism. It is here that phenomenal concepts are brought into the picture. Thus, phenomenal concepts are supposed to do two jobs. First, they must expose and account for the epistemic gap between the physical and the phenomenal domain. On the basis of her exhaustive knowledge of color and color vision alone, Mary cannot know a priori the phenomenal character of color experience. Only on her release can she learn what is like to experience blue. Yet, at the same time that phenomenal concepts explain the epistemic gap, they must also close the putative metaphysical gap between the physical and the phenomenal domain: the phenomenal character is either some intrinsic property of the brain (materialist qualia-realism) or some physical properties represented by the brain (what is known as representationalism).

By far the most popular version of this type-B materialism assumes that Mary’s cognitive progress can be accounted for by assuming that she acquires new, special phenomenal concepts of some physical property or fact she already knew by means of a physical concept in her predicament. Thus, instead of an ontological dualism between two kinds of properties, what we have is only an epistemic conceptual dualism: physical or physiological concepts on the one side, and phenomenological concepts on the other. Following Stoljar and the literature, we can call this the phenomenal concept strategy (henceforth PCS).

The *locus classicus* for PCS is Loar’s paper “Phenomenal States,” in which he claims that phenomenal concepts are recognitional concepts. A recognitional concept, unlike a theoretical concept, is applied directly on the basis of perceptual acquaintance with its instances, that is, when we recognize an object “as being one of those,” without relying on theoretical knowledge or other background knowledge. Carruthers, Tye, and Levine have endorsed similar accounts in the recent past. Thus, a recognitional concept

---

is a typical concept seeming to contradict the core claim of PLA that terms of private sensation must be connected to something publicly accessible.

In contrast, according to another trend, phenomenal concepts are indexical by nature. They are concepts that pick out brain states in an indexical mode of presentation. The suggestion here is that the epistemic gap between physical and phenomenal properties is similar to the familiar gaps between objective and indexical concepts. As I mentioned, phenomenal concepts are thought of here as flexible inner demonstratives that pick out the phenomenal character in the same way that demonstratives pick out objects in space. Thus, if the indexical account is sound, then Wittgenstein must be wrong when he claims that phenomenal concepts need to be primarily connected to something publicly accessible.

Finally, worthy of particular note is the view according to which phenomenal concepts are quotational concepts. That is, they are concepts that somehow contain the very mental or phenomenal states to which they refer. Again, the meaningfulness of phenomenal concepts does not depend on a reference to anything publicly accessible. A final group of philosophers worth mentioning defines phenomenal concepts by their conceptual role. Phenomenal concepts and physical concepts are associated with distinct faculties and modes of reasoning.

The general structure of PCS can be represented as follows. There are concepts (PCs) we employ, to introspectively pick out the phenomenal character of our experience, that we can only acquire by means of the experience in question. Mary finally sees something blue. Then she switches her attention from the sky she was contemplating to the phenomenal blueness that her experience presents to her, by turning her gaze inwards and picking out that character by means of a newly acquired phenomenal concept. Again, this is supposed to accomplish two tasks. The first is to explain the existence of an epistemic gap between physical and

25 Two names appear at the forefront of this view: Papineau (2006) and Block (2006).
phenomenal properties and, hence, Mary’s cognitive progress. The second is to close the ontological gap between those same properties. Thus, instead of an ontological dualism, PCS suggests a conceptual dualism. In Mary’s case, these are supposed to explain Mary’s cognitive progress, without assuming an ontological chasm between physical and phenomenal properties, that is, without assuming that what she now thinks by means of a new phenomenal concept is not the same physical property she already knew when she was confined to her black-and-white room.

4. PLA under attack

Wittgenstein’s PLA has come under considerable pressure. The challenge comes from two directions. First, in cognitive science such as cognitive psychology and neuroscience, methods for data collection in experimental psychology, as well as many of the inferences from image scanning, are implicitly based on the presupposed assumption that meaning is forged by a primitive association between sensation-words and sensation.27 However, the main attack comes from philosophy of mind with the emergence of the idea of phenomenal concepts (henceforth PCs).28 Again, even recognizing that PCs can be communicable, proponents of PCs still assume that what gives meaning to sensation-words is their primary association with private sensations. This is the prevailing view between qualia-realists, that is, the widespread view both in cognitive science and in philosophy of mind that the phenomenal character of experience is some intrinsic physical property of the brain.29


28 I remind the reader that I have assumed here that PCs are the mental analogues of Wittgenstein’s sensation-words.

29 The only noteworthy exceptions to this prevailing view are the so-called qualia-externalists: Dretske 1995; 1996 and Tye 1995; 2015 because they claim “qualia ain’t in the head,” they automatically reject the traditional object-perception model of introspective knowledge. See Shoemaker 1994. Still, someone could argue that qualia externalism is a quite implausible position.
Conee (and more recently Balog\textsuperscript{30}), for example, came up with the revival of Russell’s notion of knowledge by acquaintance: what fixes the reference and meaning of those phenomenal concepts is an introspective knowledge by acquaintance with those mental states with their phenomenal character. As Balog puts it, a person possesses a phenomenal concept “when she is acquainted with her own conscious states in introspection” (2012: 1).\textsuperscript{31} Even avoiding Russell’s controversial notion of acquaintance, qualia-realist versions of PCs suggest that phenomenal concepts are introduced via mental images as samples via introspection. Block, for example, claims that phenomenal concepts have the form of the _experience_. \textsuperscript{____}, in which the blank is filled by the phenomenal property that is the referent of the concept (2006: 48, n. 31). In the same vein, Papineau’s proposal is that phenomenal concepts have the structure _that experience_, where the demonstrative refers to the experience type exemplified in an associated image or copy of the experience in question.\textsuperscript{32} But, certainly, the most unusual revival of Russell’s acquaintance is Perry. He conceives of phenomenal concepts indexically as “flexible demonstratives” and “inner demonstratives,” as if we could look inwards and pick out the phenomenal character as “this,” just like we demonstrate objects in space.\textsuperscript{33}

In a recent paper, Papineau has presented a direct argument against this based on his conception of PCs.\textsuperscript{34} Papineau does not dispute whether PCs can possess a public language expression, or whether there are methods that ensure that other members of a

\textsuperscript{30} See Conee 1994 and Balog 2012.
\textsuperscript{31} Sundström defines phenomenal concepts along similar lines. We can think about conscious states under both phenomenal and nonphenomenal concepts. He states: “The former are concepts we employ to think about conscious states in ‘inner’ and ‘direct’ ways; the latter are concepts we employ to think about them in ‘outer’ and ‘indirect’ ways” (Sundström 2011: 267).
\textsuperscript{32} See Papineau 1993: 112.
\textsuperscript{33} See Perry 2001: 146. It is worth contrasting Perry’s flexible demonstratives with what Wittgenstein says in PI §258 already quoted above: “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”.
\textsuperscript{34} Papineau 2011.
linguistic community use PCs with the same meaning. He does not dispute one of the major tenets of PLA in which PCs must be communicable. His target is the claim that PCs already had a meaning prior to any a posteriori discovery of what causes them in a way that can be publicly identifiable. Following the mainstream of proponents of PCS, Papineau assumes that whenever we introspect, switching our gaze inwards, so to speak, we are able to pick out that conscious or phenomenal character before we learn how to communicate it. Papineau’s assumption is that the reference of phenomenal concepts to the phenomenal character is fixed independently of the reference to something publicly accessible:

So I myself think that the incommunicability of “senso” is superficial. But, as I said, that is not the crucial point. The question is whether Wittgenstein would agree. Now, there is a way of reading the private language argument that suggests that he would. According to this reading, the main point of the private language argument is to make it clear that private terms like Φ and “senso” must refer to something publicly accessible if they are to be genuinely meaningful. (Papineau 2011, 7)

To substantiate his claim, Papineau appeals to the variant case of Marianna, suggested by Nida-Rümelin, that parallels Wittgenstein’s diary passage (PI §258). Like Mary, Marianna is kept captive in a black-and-white room. Unlike Mary, however, when Marianna leaves her black-and-white room, she is led into a Technicolor vestibule in which there are various patches of different colors on the walls. At this point, she will have visual experiences that she has never had before, of red, yellow, blue, and

---

35 In the sense suggested by Papineau, Wittgenstein has never denied communicability of sensation-words. We first name directly the sensation, turning our gaze inwards, so to speak. Latter, we observe what causes that sensation in the outside world or what behavior that sensation causes. Now, by means of analogy, that is, by observing the same objects and behavior of other people we communicate by “means of the language” what we feel. The crucial point that Papineau misses is that the connection between sensation-words and what is going on outside is purely contingent. Therefore, what is incommunicable is what is essentially subjective, what it is like to be in such and such state. In LPE Wittgenstein says: “Our teaching connects the word ‘red’ (or is meant to connect it) with a particular impression of his (a private impression, an impression in him). He then communicates this impression- indirectly, of course through the medium of speech”. (LPE: 279).

36 See Nida-Rümelin 1996.
so forth. Yet, because Marianna has no clue which color is which, when she stares at a blue patch on the wall of the room she may reasonably wonder whether the character of the experience she is undergoing is actually phenomenal blueness. According to Papineau:

Consider how Marianna was originally supposed to attach a content to her concept $\Phi$ and the word “senso” she uses to express it. Presumably she turned her mind inwards, so to speak, when she first had the relevant kind of experience, and gave it these names. But this looks just like the kind of naming ceremony that Wittgenstein insists won’t work. He would object that at the point where Marianna coins these terms, she sets up no possibility of any public checks on her future usage – there is nothing in what she has done that will afford anybody else a potential way of checking whether she is using these terms with the same meaning in the future. (Papineau 2011: 8. Emphasis added)\(^{37}\)

In his original account of PCs,\(^{38}\) Papineau claims that we must be able to introspectively focus on qualia when we have them and to re-create them imaginatively at other times. Phenomenal concepts are mental demonstratives, and with them, we can form terms with the structure of the experience: $\ldots \ldots$, where the gap is filled either by a current token experience or by an imaginative recreation of an experience. Based on this so-called quotational-indexical proposal, the distinguishing phenomenal features of PCs are just the mental images, copies, or replicas of the experience housed in the sensory file. Therefore, exercising PCs involves recreating, simulating, remembering, and thinking of a phenomenal state or experience through introspection or memory.

Papineau later rejects his previous quotational-indexical account by saying that it ran a good idea together with a bad one. The bad idea is that phenomenal concepts are demonstrative-like concepts that pick out experiences indexically. Their distinguishing feature is not their putative indexical nature, but rather their cognitive function of accumulating information about experience. Since the

---

\(^{37}\) Indeed, the similarity between the target of PLA and Papineau’s account of PCs is striking. The passage quoted above stands in direct contradiction to Wittgenstein’s famous diary passage at PI §258.

\(^{38}\) See Papineau 2002: 56.
function of PCs is to carry information from one use to another, they cannot be modeled as mental demonstratives. In opposition to demonstratives, name-like files are suited to serve as repositories of information precisely because they refer to the same thing whenever they are exercised. When PCs are activated, the information hosted in them is activated too. Moreover, Papineau also assumes the possibility of a pure introspective classification:

Suppose that introspective classification depends on the existence of some kind of brain “template,” to use David Lewis’s phrase (1983). We don’t classify new experiences by seeing whether they have some phenomenal property with which we have previously been acquainted. Instead, we simply compare them with the “template” to see whether they correspond. This hypothesis too yields an obvious materialist explanation of why you should only be able to introspectively classify experiences of a kind that you have previously had. Again, the brain needs an original to form the mold. **In order to fix a neural pattern as a template against which to compare new inputs, we need some original experience to create the pattern.** (Papineau 2002: 58. Emphasis added)

Thus, the good idea, unencumbered by the bad one, is that phenomenal concepts are perceptual or sensory files whose function is to accumulate information about the relevant features of referents by storing copies or replicas of experience. The idea is that phenomenal concepts use the copies or replicas of the experience housed in the file in order to mention the experience.\(^{39}\) Papineau now claims, however, that PCs are simply special cases of ordinary perceptual concepts re-used in introspection to think about experiences themselves, rather than the objects of those experiences that normally cause the experiences.

The case of Wittgenstein’s opponent in his PLA is similar to Papineau’s account of PCs in all the relevant aspects. Let us take a closer look. Putting himself in the place of his opponent, Wittgenstein says, “I know what the word ‘toothache’ means…it produces one particular image in my mind” (LPE, 315. Emphasis added). The suggestion is that the words of the private language refer to elements of one’s experience: e.g., the blue patch in one’s

\(^{39}\) See Papineau 2007: 11.
visual field when one looks inward and contemplates the image of the phenomenal features of experience before his mind.

Strikingly, on several occasions, Wittgenstein uses exactly the same words as Papineau, suggesting that a meaningful term is associated with a replica of an element of one’s experience, an exemplar, which serves as a paradigm for the meaning of the term. Therefore, just as a color sample is employed as a part of the expression of a rule for the use of a color concept, so too a mental image or its replica is conceived as a sample (template) involved in the use of phenomenal concepts.

4. Transparency of Experience

Recall that the open question is why a sensation cannot function as a sample for the correct use of a word (Hacker 1987: 267). To answer this question, I assume that there is an enthymeme in PLA overlooked in the literature. This enthymeme is implicit in all of Wittgenstein’s writings about PLA, but has never received due attention from Wittgensteinians. The main argument comes from what I, following the recent literature, call here the transparency of perceptual and bodily experience.

To begin with, the very idea of the mind turning its attention onto itself has always struck Wittgenstein as the queerest thing that could be:

When speaking, one can refer to an object by pointing at it. Here pointing is a part of the language-game. And now it seems to us as if one spoke of a sensation by directing one’s attention to it. But where is the analogy? It evidently lies in the fact that one can point at a thing by looking or listening. (PI §669, original emphasis)

But the crucial aphorism is PI §275, where Wittgenstein states:

Look at the blue of the sky and say to yourself, “How blue the sky is!” – When you do it spontaneously – without philosophical purposes – the idea never crosses your mind that this impression of color belongs only to you. And you have no qualms about exclaiming thus to another. And if you point at anything as you say the words, it is at the sky. I mean: you don’t

---

40 See in particular NL, 341; PI, §272.
41 An exception worthy of note is Shoemaker’s (1994).
have the pointing-into-yourself feeling that often accompanies “naming sensations” when one is thinking about the “private language.” Nor do you think that really you ought to point at the color not with your hand, but with your attention. (Consider what “to point at something with one’s attention” means.) (Emphasis added)

At PI §275, Wittgenstein is making two complementary remarks. First, he is denying that phenomenal blueness is a property of the subject’s experience or of the subject’s brain. In other words, Wittgenstein is opposing what today is known as qualia-realism. Second, in support of his first claim, he adds that when you try to point to the phenomenal blueness of your visual experience of the sky you end up pointing (if at anything) to the blueness of the sky itself. Thus, when you try to pick out the putative phenomenal blueness of your experience of the sky, you end up pointing to the blueness of the sky itself.

That is the core of Wittgenstein’s criticism of Russell’s sense-data theory, according to which phenomenal properties are primarily connected to sense-data, and only indirectly connected with external objects, by means of some causal inference. According to Wittgenstein:

Does it follow from the sense impressions which I get that there is a chair over there? – How can a proposition follow from sense impressions? Well, does it follow from the propositions which describe the sense impressions? No. – But don’t I infer that a chair is there from impressions, from sense-data? – I make no inference! – and yet I sometimes do. I see a photograph, for example, and say “So there must have been a chair over there,” or again, “From what one can see here, I infer that there is a chair over there.” That is an inference; but not one belonging to logic. An inference is a transition to an assertion; and so also to the behaviour that corresponds to the assertion. “I draw the consequences” not only in words, but also in deeds. (PI §486)

On closer inspection, Wittgenstein’s remark of PI §275 echoes Moore’s famous thesis of transparency. Moore states as follows:

And, in general, that which makes the sensation of blue a mental fact seems to escape us: it seems, if I may use a metaphor, to be transparent — we look through it and see nothing but the blue. We may be convinced that there is something but what it is no philosopher, I think, has yet clearly recognized. (Moore 1903: 446)
When we try to introspect the sensation of blue, all we can see is the blue: the other element is as if it were diaphanous. (Moore 1903: 450)

Harman recently formulated the thesis even more sharply:

When Eloise sees a tree before her, the colors she experiences are all experienced as features of the tree and its surroundings. None of them are experienced as intrinsic features of her experience. Nor does she experience any features of anything as intrinsic features of her experiences. And that is true of you too. There is nothing special about Eloise’s visual experience. When you see a tree, you do not experience any features as intrinsic features of your experience. Look at a tree and try to turn your attention to intrinsic features of your visual experience. I predict you will find that the only features there to turn your attention to will be features of the presented tree. (Harman 1990: 667)

Dretske states the same claim in different words:

If one is asked to introspect one’s current gustatory experience — “Tell us, if you can, exactly how the wine tastes” — one finds oneself attending, not to one’s experience of the wine, but to the wine itself (or perhaps the tongue or palate). There seems to be no other relevant place to direct one’s attention. At least one does not have experiences other than the wine-experience one is asked to describe. (Dretske 1995: 62)

But Tye was certainly the philosopher who gave more emphasis to this thesis in contemporary philosophy of mind:

Intuitively, you are directly aware of blueness and squareness as ... features of an external surface. Now shift your gaze inward and try to become aware of your experience itself, inside you, apart from its objects. Try to focus your attention on some intrinsic feature of the experience that distinguishes it from other experiences, something other than what it is an experience of. The task seems impossible: one’s awareness seems always to slip through the experience to blueness and squareness, as instantiated together in an external object. In turning one’s mind inward to attend to the experience, one seems to end up concentrating on what is outside again, on external features or properties. (Tye 1995: 30)

The transparency thesis supports, of course, different readings. According to Papineau, for example, when you switch your attention from the sky, and instead turn your mind introspectively
to the nature of your visual experience “none of your conscious properties alter; that is, introspection makes no difference to the conscious nature of your experience itself” (2014: 22). Thus, there is no difference to the conscious nature of your visual experience of a blue sky, if you are contemplating the blueness of the sky, or introspecting the phenomenal blueness of your experience of the sky. That said, Papineau implicitly endorses the claim that when you introspect you are de re aware of the conscious blueness. This is what, I believe, PLA calls into question.

On closer look of what Wittgenstein says at PI §258, PI §275, and PI §669, it seems clear that what Wittgenstein is ruling out is exactly the idea that introspection is awareness of the phenomenal blueness, by excluding the possibility of pointing to the phenomenal blueness in the same sense that we point to the blueness of the sky. Wittgenstein says there is no “pointing-into-yourself feeling.” At PI § 401, he states:

You interpret the new conception as the seeing of a new object. You interpret a grammatical movement that you have made as a quasi-physical phenomenon which you are observing. (Remember, for example, the question “Are sense-data the stuff of which the universe is made?”)

Using Shoemaker’s distinction between “perception of” (de re awareness) and “perception that” (fact-awareness), what Wittgenstein is clearly denying is the possibility of a de re awareness of the phenomenal blueness of your visual experience as something before one’s mind (inner perception). Thus, the crux of the disagreement between Wittgenstein and Papineau and other proponents of PCs can be formulated as follows. Following the traditional object-perception model of introspection, Papineau endorses the claim that when you switch your attention from the blueness of the sky, and instead turn your mind inwards, you become de re aware of the phenomenal blueness of your visual experience as some image or replica you have before your mind’s eye. Now we can name that phenomenal blueness (senso) and take

---

42 See Papineau 2014: 23.
the image of it as a template for future use, as if the memory of the image would guide all future uses of the concept. In contrast, I want to suggest that, according to Wittgenstein, there is no such a thing as a de re awareness of phenomenal blueness that could fix the reference of a corresponding phenomenal concept to phenomenal blueness. As Wittgenstein says, that is an idle ceremony that actually achieves nothing.

Yet, Papineau can reasonably ask, what is the big deal in opposing seeing of to seeing that? He wonders, after all, what is it to be de re aware of something? I think that Tye has a straightforward answer: you are de re aware of something when you can wonder “what is that?” with respect to the entity you are aware of. Thus, you are de re aware of the blueness of the sky because you can wonder “what is that?” But if you switch your attention from the blueness of the sky to the phenomenal blueness of your visual experience of the sky, there is nothing you can be aware of. You cannot be de re aware of the phenomenal blueness in the same way you are de re aware of the blueness of the sky itself. For one thing, when you introspect the phenomenal blueness of your visual experience of the sky, you cannot ask yourself “what is that?” For, as Wittgenstein says, there is no “pointing-into-yourself feeling.” Wittgenstein is right when he insists that phenomenal concepts must be primarily connected to something publicly accessible.

However, to deny that you can be de re aware of phenomenal blueness in the same way that you can be de re aware of the blueness of the sky is, of course, not to deny the possibility of introspection in the non-etymological sense of knowing your own mental states or, in our case, knowing the phenomenal blueness. Wittgenstein suggests that we should understand introspective knowledge as a fact-awareness rather than an objectual awareness by acquaintance with phenomenal character:

---

45 Personally, he has told me – sincerely – that he has never understood (that is to say, accepted!) the difference between de re and de dicto awareness.

46 See Tye, 2009: 100; 2014: 44.
Do I observe myself, then, and perceive that I am seeing or conscious? And why talk about observation at all? Why not simply say “I perceive I am conscious?” – But what are the words “I perceive” here for? (PI, §417. Emphasis added)

Let us look at PI §412:

But what can it mean to speak of “turning my attention on to my own consciousness?” There is surely nothing more extraordinary than that there should be any such thing! What I described with these words (which are not used in this way in ordinary life) was an act of gazing. I gazed fixedly in front of me – but not at any particular point or object. My eyes were wide open, brows not contracted (as they mostly are when I am interested in a particular object). No such interest preceded this gazing. My glance was vacant; or again, like that of someone admiring the illumination of the sky and drinking in the light. (Emphasis partly added.)

What Wittgenstein is clearly suggesting is that we can only know introspectively that phenomenal blueness is the phenomenal character of the experience we are undergoing after we have learned (de dicto knowledge) blue is the color that usually causes in us that kind of experience. Likewise, we can only introspectively know that pain is the phenomenal character of the experience we are undergoing after we have learned that pain is what usually causes some typical pain behavior.

Let us return to Marianna’s case. On her release from the back and white room, Marianna enters the Technicolor vestibule. She stares a blue patch on the wall and experience what is like to see blue for the first time. Now, if we assume that she cannot point to her blueness sensation and name it, the only way available to her to pick out her sensation is by fixing her gaze to the patch in front of her. Thus, Papineau’s prevailing picture of phenomenal concepts is upside down: she can only introspectively know de dicto that blueness is the phenomenal character of the experience she is undergoing in her vestibule after she has learned that blue is the color of the patch that she is staring at on the wall of her vestibule.
References


Dretske, F. 1996. “Phenomenal externalism, or if meanings ain’t in the head, where are qualia?” Philosophical Issues 7, pp. 143-158.


**Biographical Note**

Roberto Horácio de Sá Pereira received degrees in philosophy from the Federal University of Rio de Janeiro and Freie Universität Berlin (PhD), and is now Professor of Philosophy at the Federal University of Rio de Janeiro. He is the author of *Aussenwelt-Skeptizismus. Eine Sprachanalythische Behandlung* (Konstanz: Hartung-Gorre Verlag 1993). His most recent papers are: “Combining the representational and the Relational view” (*Philosophical Studies* 2016), “A Non-conceptualist