



ΠΡΟΜΕΘΕΥΣ

Journal of Philosophy

n. 33 maio - agosto de 2020

Dossiê Linguagem e Cognição

Editores: Marcus Souza, Marcos Silva & Maxwell Lima Filho



**PSYCHOLOGICAL CONCEPTS IN COGNITIVE
NEUROSCIENCE: SOME REMARKS ON BENNETT &
HACKER'S *PHILOSOPHICAL FOUNDATIONS OF
NEUROSCIENCE***

Marcelo Carvalho
Universidade Federal de São Paulo

ABSTRACT: The use of psychological concepts in cognitive neuroscience is heavily criticized by Bennett & Hacker's *Philosophical Foundations of Neuroscience*. The central objection points to neuroscience's attribution to the brain of psychological concepts that are meaningful only when applied to the entire being. That is supposedly the case of "seeing," "communicating," and "reading." Bennett & Hacker identify in such attributions what they call a *mereological fallacy*. The critical revision of Bennett & Hacker's argument is an opportunity to present the debate about philosophy and psychological neuroscience and outline a Wittgensteinian perspective about the meaning of psychological concepts, its interest, and its relevance to scientific research.

KEYWORDS: Cognitive Neuroscience. Philosophy of Mind. Wittgenstein.

RESUMO: O uso de conceitos psicológicos na neurociência cognitiva é fortemente criticado por Bennett & Hacker em *Philosophical Foundations of Neuroscience*. Sua objeção central dirige-se à atribuição ao cérebro pela neurociência de conceitos psicológicos que são significativos apenas quando aplicados a todo o ser. Esse é supostamente o caso de "ver", "comunicar" e "ler". Bennett & Hacker identificam em tais atribuições o que eles chamam de falácia mereológica. A revisão crítica do argumento de Bennett & Hacker é uma oportunidade para apresentar o debate sobre filosofia e neurociência psicológica e delinear uma perspectiva wittgensteiniana sobre o significado dos conceitos psicológicos, seu interesse e sua relevância para a pesquisa científica.

PALAVRAS-CHAVE: Neurociência Cognitiva. Filosofia da Mente. Wittgenstein.

1.

For most right-handed people only the left hemisphere can speak or communicate through writing. It also rules most of the capacity to deal with language, although the right hemisphere may understand spoken words to a limited extent and probably deals with the music of speech. When the callosum is cut, the left hemisphere sees only the right half of the visual field; the right hemisphere, the left half. Each hand is controlled by the opposite hemisphere, although the other hemisphere can produce some of the coarser movements of the hand and arm. Except under special conditions both hemispheres can hear what is being said...¹

With this description of the effect of a commissurotomy, Francis Crick intends to show that understanding and communicating through writing are functions of the left part of the brain. “What different brain regions do,” and how can they do it, are among the main questions in neuroscience.² These problems are not strange to philosophy, Descartes being the more evident of the authors associated with it, in a tradition that goes back to Plato and Aristotle.³ M. Bennett & P. Hacker’s *Philosophical Foundations of Neuroscience* have, in its core, the idea that descriptions like the one presented by F. Crick make a problematic use of *psychological concepts* and entangle cognitive neuroscience in error and confusions. The problem is what Bennett & Hacker characterize as the mistaken attribution of understanding, seeing, expecting, etc., to the brain. This kind of use of psychological concepts is, they say, “a source of much further confusion.”⁴ Crick’s description is only one example of what supposedly is a vast problem that pervades cognitive neuroscience:

the characteristic form of explanation in contemporary cognitive neuroscience consists in ascribing psychological attributes [perceiving, thinking, guessing and believing] to the brain and its parts *in order to explain* the possession of

¹ CRICK, 1995, p. 169f.; after BENNETT & HACKER, 2014, p. 388-9.

² BECHTEL, MANDIK & MUNDALE, 2001, p. 4-7.

³ Cf. e.g., DESCARTES, 1996, VOL 11 (*Les passions de l’ame*), Art. 31-31.

⁴ BENNETT & HACKER, 2014, p. 3.

psychological attributes and the exercise of cognitive power by human beings.⁵

Revising the use of the traditional psychological vocabulary is a relevant philosophical project, not only to identify mistakes and problems in it, as Bennett & Hacker propose, but also to make clear their complex relationship with our ordinary language games, and the way they revise and change it. Variation in the way we use words is not a problem. It is the very reason for Wittgenstein introducing the concept of “language games.” According to him, concepts do not have clear limits or follow fixed rules.⁶ Therefore, F. Crick’s ascription of vision to the brain is not an exception or a problem in itself. But it should be asked, for example, the consequences of saying, for example, that “I have a picture of it in my mind” is synonymous with “I remember it.” In the same way, we should ask about the relationship between Crick’s use of “understanding” and our ordinary use. Beyond that, we should be cautious about the philosophical consequences, if any, of such uses of “understanding” and “mind.” Do they have a place in our practices? Are they connected with our experiences?

Bennett & Hacker have a very critical position about the use of psychological concepts by cognitive neuroscience. Instead of presenting Wittgenstein as the primary source of their investigations⁷, this criticism resorts to philosophical conceptions that are closer to the heritage of Plato and Kant. They describe their work, for instance, in an uncomfortable vocabulary, as a “conceptual hygiene” that results in “[liberating] neuroscience from various forms of incoherence and from bogus mysteries, which are generated by current conceptual confusions.” The consequence of that is “elucidating the conceptual scheme that determines the psychological concepts with which neuroscience actually works, but which they inadvertently misuse.”⁸

The complex set of ideas and presuppositions outlined in *Philosophical Foundations of Neuroscience* may be organized as follows:

- A1) It may be established a clear boundary between philosophy and science;
- A2) Philosophical investigations precede anything of empirical;

⁵ BENNETT & HACKER, 2014, p. 3.

⁶ Wittgenstein says that “we can draw a boundary [for a concept] – for a special purpose. Does it take this to make the concept usable? Not at all! Except perhaps for that special purpose.” WITTGENSTEIN, 2009, 69.

⁷ BENNETT & HACKER, 2014, p. 71 and p. 97-103.

⁸ BENNETT & HACKER, 2014, p. 114.

A3) Consequently, philosophy is responsible for the foundations of any science, particularly of neuroscience;

A4) But the conceptual revision proposed by philosophy leaves the results of neuroscientists untouched;

B5) There is a correct and meaningful use of psychological terms;

B6) Neuroscience does not use psychological terms correctly;

B7) The correct use of psychological concepts is that of ordinary language, and it is the result of a (behavioral) criteria for the attribution of psychological terms;

C8) The mistakes of neuroscience in its use of psychological concepts are the consequence of a mereological fallacy;

C9) This fallacy comes from cognitive neuroscience having its origin in the Cartesian tradition;

C10) Eliminating the mereological fallacy results in hygienising neuroscience's vocabulary.

These suppositions are problematic in themselves and in their relationship with each other. In what follows, we will review some of them to understand Bennett & Hacker's project and its place in the philosophical debate. The result is the outline of an alternative, Wittgensteinian perspective on the relevance of philosophical investigations to psychology and neuroscience. To do that, we will survey Bennett & Hacker's conceptions about the relationship between philosophy and (neuro)science, their criterium for the correct use of a concept, and their description of the "mereological fallacy," confronting all that with Wittgenstein's late philosophy.

2.

Central to Bennett & Hacker's description of their project is that the conceptual investigation they propose, notwithstanding its relevance, "leaves" neuroscientist's achievements "as they are."

Abandoning the style of thought that is enmeshed in the mereological fallacy in neuroscience (...) will not deprive neuroscientists of any of their hard-won and remarkable achievements. On the contrary, it will enable those achievements to be seen clearly (...).⁹

⁹ BENNETT & HACKER, 2014, p. 114.

It is not easy to understand how it is possible. Bennett & Hacker's criticism is vast and seems to imply that there are confusion and fallacy in every relevant achievement of neuroscience. It seems to demand a revision of all its descriptions and conclusions, undermining the very existence of this science.

The idea that neuroscience is immune to the philosophical debate results from a particular appropriation of paragraph 124 of the *Philosophical Investigations*, where Wittgenstein says that:

Philosophy must not interfere in any way with the actual use of language, so it can in the end only describe it. / For it cannot justify it either. / It leaves everything as it is. / It also leaves mathematics as it is, and no mathematical discovery can advance it.¹⁰

However, *Philosophical Foundations of Neuroscience* presents an interpretation of this remark that is closer to Kant than to Wittgenstein. According to it, empirical science has a *conceptual schema* at its foundation. This conceptual schema *determines* (like a *transcendental schema*) how we understand and describe our experiences. As a consequence, instead of not being part of our experiences and having no empirical impact, a philosophical analysis puts concepts in order. It makes adjustments in the conceptual basis that formally precedes any empirical investigation, within which empirical science works. The relationship between the conceptual schema at the foundation of neuroscience and the descriptions of its results are presented by Bennett & Hacker in the following terms:

These categories of concepts [use in ordinary psychological descriptions], and the more specific concepts that they subsume, are the spectacles through which psychological phenomena are viewed and understood. If these spectacles are askew, then neuroscientists cannot but see the phenomena awry.¹¹

They propose to elucidate “the conceptual scheme that *determines* the

¹⁰ WITTGENSTEIN, 2009, 124.

¹¹ BENNETT & HACKER, 2014, 115; this also implicitly refers to WITTGENSTEIN, 2009, 103.

psychological concepts, with which neuroscientists actually work, but which they inadvertently misuse.”¹² The assertion that a conceptual scheme stands at the foundation of neuroscientific research implies (in this context) that it *determines* the description of the results of this research and of experience in general. Any such result is to be viewed only through the glasses of a particular conceptual scheme. It follows that, “meaning precedes truth,” conceptual investigation precedes empirical science, and that Bennett & Hacker’s philosophical criticism precedes (neuro)science.¹³

Instead of comparing conceptual schemes to glasses, Bennett & Hacker refuse and criticize Helmholtz’s supposition that perception involves the development of an unconscious hypothesis based on inductive inferences from the senses. According to Bennett & Hacker, this theory is incorrect, among other reasons, because

Perceptions cannot be conclusions of unconscious inferences the premises of which are unconscious and more or less indescribable sensations and (unconscious) generalizations about the correlations between past sensations and objects perceived. (...) To perceive something is not to form a hypothesis.¹⁴

The concept of “hypotheses” in Helmholtz’s theory seems very close to the spectacle metaphor used by Bennett & Hacker. Both are succedaneums to the Kantian distinction between *phenomenon* and *thing-in-itself*, and the correlated supposition that what we recognize as “our experience” results from an active interference from a *transcendental schema*. The spectacle metaphor has its origin in a Kantian context. It is part of a famous description of Kant’s philosophy made by Heinrich von Kleist in a letter from 1801.¹⁵ In *The Big Typescript*, Wittgenstein uses Kleist’s metaphor in a sense

¹² BENNETT & HACKER, 2014, 144 (italics added); “It [the crypto-Cartesianism of contemporary neuroscience] leads to incoherent descriptions of the results of neuroscientific research, and sometimes to the formation of incoherent research programs.”

¹³ “Whether a hypothesis makes sense must be settled in advance of determining its or its evidential support” (BENNETT & HACKER, 2014, 382). This idea finds explicit support in Wittgenstein’s earlier works, the *Tractatus* and in the *Philosophical Remarks*.

¹⁴ BENNETT & HACKER, 2013, p. 9.

¹⁵ “I recently became familiar with the more recent, so-called Kantian philosophy, and I may impart one of its leading ideas to you without fear of its shattering you as deeply, as painfully, as it has me. For, after all, you are not versed enough in the whole matter to grasp its import completely. I shall therefore speak as clearly as possible. If everyone saw the world through green glasses, they would be forced to judge that everything they saw *was* green, and could never be sure whether their eyes saw things as they really are, or did not add something of

that is close to Kant's philosophy. In the *Philosophical Investigations*¹⁶, however, the core of his mature work, it appears in a new context and with a different meaning that does not support a Kantian interpretation.¹⁷ The idea of a *conceptual schema* is part of an idealist perspective that the *Philosophical Investigations* explicitly refuses.¹⁸

The widely stressed independence between neuroscience's foundations (its conceptual schema) and its results¹⁹ raises some questions that do not find easy answers in Bennett & Hacker's book. If it is true, it implies the *irrelevance* of their work, at least to empirical scientists. If all the "hard-won and remarkable achievements" of cognitive neuroscience are left untouched and if these achievements are obtained without the hygienizing proposed by Bennett & Hacker, this entire job seems to be fruitless, or at least dispensable (from its own perspective). The relevance of such an investigation seems to be in the opposite way: it results in an adequate foundation for scientific research. How can that be without any empirical consequence? The authors only say that they intend to remove confusion, leaving the achievements untouched. But where are the consequences of such confusion, however, if not in the results of scientific investigations? Bennett & Hacker say, in a positivistic-tractarian vocabulary, that neuroscientific theories "transgress the bounds of sense" and that their philosophical criticism intends to prevent us from "talking nonsense."²⁰ But if nonsense is not in neuroscience's achievements, and if all the results remain as they are, why is the entire project worth at all?

That the *Philosophical Investigations* do not give clear support to assert the immunity of psychological neuroscience to philosophical criticism is, paradoxically, a conclusion found in G. Baker & P. Hacker's commentary of the same paragraph of

their own to what they saw. And so it is with our intellect. We can never be certain that what we call Truth is really Truth, or whether it does not merely appear so to us. If the latter, then the Truth that we acquire here is *not* Truth." KLEIST, 1982, p. 95.

¹⁶ WITTGENSTEIN, 2009, 103.

¹⁷ For the critique of the orthodox interpretation about the relationship between Wittgenstein's middle period works (*The Big Typescript*, *Philosophical Grammar*) and the *Philosophical Investigations*, cf. STERN, 2004; ENGELMANN, 2013; MEDINA, 2002; CARVALHO, 2014.

¹⁸ For Wittgenstein's use and criticism of the idea of a "schema", cf. WITTGENSTEIN, 2009, 73, 86, and 134-6.

¹⁹ BENNETT & HACKER, 2014, p. 114; "[It] will not deprive neuroscience of any of their hard-won and remarkable achievements. On the contrary, it will enable those achievements to be seeing clearly, striped of the conceptual confusion in which they are currently all too often wrapped."

²⁰ BENNETT & HACKER, 2014, p. 116.

Wittgenstein. Reversely, *Understanding and Meaning* implies a revision of the assured independence of neuroscience's results.

If, 'in psychology there are experimental methods and conceptual confusion' (PI p. 232/197), philosophical investigation into psychological concepts will affect empirical psychology, for it may show that some questions are senseless, that some experiments rest on incoherent presuppositions, and that some experimental results do not prove what they are held to demonstrate.²¹

From another, more uncomfortable perspective, the independence between conceptual schema and the findings of psychological neuroscience seems to imply that the discoveries of neuroscience are "hard facts" that remain the same whichever description we give of them, even after revising its fundamental conceptual schema. This idea is not easily maintained or conciliated with the centrality that Bennett & Hacker attribute to (psychological) concepts in their (Kantian) approach.

Following one example of Bennett & Hacker's criticism of neuroscience's vocabulary makes more explicit what they suppose that remains the same after a conceptual hygienizing. The result of a commissurotomy is, in Francis Crick's words, that "one half of the brain appears to be almost totally ignorant of what the other half saw."²² Roger Sperry describes the relationship between the two hemispheres of the brain in the following terms:

both the left and the right hemispheres may be conscious simultaneously in different, even in mutually conflicting, mental experiences that run along in parallel.²³

The experiment that supports this conclusion goes as follows. After the commissurotomy, the patient is presented to different pictures in his right and left visual fields. He describes what he sees with his right side as a "normal person" (a person without a commissurotomy) does. If, however, a picture is shown to his left side (to his right, "non-speaking" hemisphere), "the results are quite different."

²¹ BAKER & HACKER, 2009, p. 266.

²² After BENNETT & HACKER, 2014, p. 389.

²³ After BENNETT & HACKER, 2014, p. 389-90.

The left hand (largely controlled by this non-speaking hemisphere) can point to and identify unseen objects by touch (...). But when the patient is asked to explain why his left hand behaves in that particular way, he will invent explanations on the basis of what his left (speaking) hemisphere saw, not on what his right hemisphere knew.²⁴

Bennett & Hacker refuse attributions of consciousness and speech to the brain or its parts, like the ones that are used in this description. As an alternative, they propose to present the experiment's result in the following way:

The light stimulus (...) affected the right hemisphere, the severance of which from the left hemisphere deprived the patient of the ability to describe or be visually aware of what was present to him on the left of his visual field, although, remarkably, he was, by pointing, able to associate correctly what was there (...). Nevertheless, he did not know *why* he made that association (not being aware of the snow scene presented to him [to the right hemisphere]), and confabulated a tale to explain why he has done so (a confabulation comparable to those produced by subjects to explain their post-hypnotically suggested behavior). This, in turn, is (crudely) explicable by reference to the fact that the visual stimulation of the right hemisphere is disconnected from the left hemisphere, so that the patient is deprived of his normal cognitive capacity to be visually aware of what is presented to him and to recognize and describe familiar objects that are thus presented. It does not, however, deprive him of the ability to associate what was visually presented to him on the screen with an appropriate object (...) – but without knowing why he is doing so. / It is senseless to speak of the right hemisphere observing the action of the left hand, for neither hemispheres of the brain can observe or fail to observe anything at all.²⁵

This alternative description substitutes expressions like “what the right hemisphere saw” by “the affection of the right hemisphere by the light stimulus,”

²⁴ After BENNETT & HACKER, 2014, p. 389.

²⁵ BENNETT & HACKER, 2014, p. 392.

something that does not seem central to the argument. It also avoids talking about what the right or the left hemispheres do know since hemispheres “cannot be either aware or unaware of anything,” only human beings can do that.²⁶ But as a result, the description proposed by Bennett & Hacker is not able to explain why there are different reactions to what is presented to different sides of the brain (a consequence, in Crick’s description, of their different “abilities” and the lack of “communication” between them after the commissurotomy). To Bennett & Hacker, only human beings can communicate. How is it possible to explain that the patient in the example does not present a plausible verbal justification of his choice without asserting that the linguistic ability resides mainly in the left side of the brain, and that, as a consequence, the split-brain is not capable of talking about what is presented only to its right hemisphere, instead of being, in a clear sense of the word, *aware* of it? Bennett & Hacker’s alternative description does not give us a full and plausible explanation about the experiment’s results. The vocabulary’s revision they propose seems to be either irrelevant or unacceptable.

We may also ask what are the *facts* in the experiment? What remains the same after changing the description? Is the statement that the light stimulus affects the right side of the brain (made by Bennett & Hacker) a fact? Is it independent of any conceptual schema? Doesn’t it imply attributing awareness or sensibility to parts of the brain? Will this *fact* be the same in any different conceptual schema? What is the relevance, then, of a conceptual schema? How can we draw the line between what is a fact and what changes with the conceptual schema? It is not clear how these questions can be answered within Bennett & Hacker’s framework. In the end, if all our descriptions depend, at least partially, on the concepts we use, how is it possible to say that the philosophical investigation proposed in *Philosophical Foundations of Neuroscience* does not undermine neuroscientist’s empirical results?

There is a remaining question about the relationship between the book’s perspective and Wittgensteinian philosophy. That the ideas of conceptual “hygienizing” and “determination by a *conceptual schema*” do not find support in Wittgenstein’s *Philosophical Investigations*, as we assert, do not have relevance only as part of a debate in the history of philosophy, about how to read Wittgenstein. The conceptions presented by Bennett & Hacker are, in fact, close to some ideas from Wittgenstein’s work in his middle period, particularly to the *Philosophical Remarks* (1930) and *The Big Typescript*

²⁶ BENNETT & HACKER, 2014, p. 391.

(1932-3). But they are sharply criticized in the *Philosophical Investigations*. Wittgenstein's mature work presents a critique of his earlier concept of grammar (that is, of his particular version of a "conceptual schema") that make clear the mistakes and consequences of Bennett & Hacker's ideas about the role of concepts in science and experience and present an alternative perspective to this debate. Doing so, Wittgenstein offers a new and relevant alternative to the Kantian ideas about the relationship between concepts and practices, and, in this particular case, between philosophy and psychological neuroscience.

3.

Bennett & Hacker's conceptions of "philosophical *foundation* of science" and "conceptual hygiene,"²⁷ do not have a place in the *Philosophical Investigations*. The *Investigations* rarely refers to foundations, and never in Bennett & Hacker's sense. The unique exception to that is the very last remark of the *PPF*²⁸, where Wittgenstein says that the investigations about psychology that he presents are also possible in mathematics and that they are not *psychological* or *mathematical* investigations. He says that "it might deserve the name of an investigation of the 'foundations of mathematics'" (echoing the same expression that appears in the "Preface"). Instead of it, Wittgenstein criticizes foundationalism in many other places, particularly in his *Remarks on the Foundations of Mathematics* (a book that received from its editors a title that contradicts its content). Wittgenstein does not present his philosophical work as an analysis or a search for foundations, but as *clarification*.

What does mathematics needs a foundation for? It no more needs one, I believe, than propositions about physical objects – or about sense impressions, need an *analysis*. What mathematical propositions do stand in need is a clarification of their grammar [*Klarlegung ihre Grammatik*], just as do those other propositions.²⁹

That mathematics does not need a foundation does not imply that there is no

²⁷ Cf. BENNETT & HACKER, 2014, p. 116.

²⁸ WITTGENSTEIN, 2009, PPF, xiv, 372. Another relevant reference can be found in Wittgenstein's "Preface".

²⁹ WITTGENSTEIN, 1978, VII, 16; Cf. also MÜHLHÖLZER, 2010, I, 1.

foundation for our knowledge or our use of language. In *On Certainty* Wittgenstein presents a very interesting conception about this subject, concluding that what lies at the foundation of “all our operating with thoughts” and language are not the propositions of logic, but “propositions of the form of empirical propositions,” which “do not serve as foundations in the same way as hypotheses which, if they turn out to be false, are replaced by others.”³⁰ The distinction between what is an “empirical” proposition and what is a “grammatical” one does not result from their form, but of the way we use them in our practices. The entire work of philosophy consists, then, in clarifying our use of language, our practices. Therefore, there is no conceptual system at the foundation of our knowledge claims; there is only an “ungrounded way of acting.”³¹ The clarification that Wittgenstein proposes is not something that may be described as *conceptual hygiene*.

The *Philosophical Foundations of Neuroscience* has, in its core, the supposition that neuroscience is full of conceptual errors, confusions, and nonsense. What is the context that makes it possible to talk about errors this way? Why saying that “the brain feels...”, like in F. Crick’s description, is an “error”? How is it possible to determine what is right and wrong in neuroscience’s use of these words?³²

For Wittgenstein in the *Philosophical Investigations*, understanding a word is to understand its use in particular contexts, as part of our practices. As a result, the

³⁰ WITTGENSTEIN, 1974, 401-2.

³¹ WITTGENSTEIN, 1974, 110; cf. CARVALHO, 2016.

³² Prescribing a *correct* use of language is not easily adjustable to Wittgenstein’s later philosophy. It becomes clear to him in the early 1930s that he could not talk about a connection between language and something outside it, and that his analysis remains “inside language.” Even in an ostensive definition, “there is no confrontation of a sign with reality.” According to this intermediary conception, the meaning is relative to grammar. In the *Philosophical Investigations*, however, the concept of “grammar” does not have the same use that it had in the middle period, and it is not as relevant as it was before. From this latter perspective, the previous use of “grammar” was idealist, formalist, and still shared a residual referentialist perspective. It is particularly clear in the intermediary conception of “rule,” strongly criticized by Wittgenstein in the *Philosophical Investigations*. From this new perspective, a word has meaning only in the context of forms of life, of a culture. The first thing to do when we are presented to a new word, in a new context, is asking how people *use* it. The ordinary mistakes in the use of concepts are relative to these contexts. Philosophy may only show, for example, that we suppose to attribute use to some expressions when, in fact, we do not. This seems to be the case of the word “know” in expressions like “I know I have a pain”. The problem here is not a particular use of “to know” that is absolutely forbidden and nonsense, but that *we*, in *our* practices, do not attribute a use, a meaning, to it, and, as a consequence, that we cannot point to an effective case in which we mean it, we cannot identify any difference between saying “I know I have a pain” and saying “I have a pain.”

meaning is not given as a hierarchical system, with foundations and fixed determinations. If meaning arises only in the use of a word, it is not possible to be prescriptive about it. Saying that philosophy leaves “everything as it is”³³ points to an absolute lack of ground to *prescribe* the use of a word. The *determination* of meaning by a rule or the reference to an object is not even possible, as we learn in the debate about “private language.” There is no absolute standard of right and wrong to which we must adjust our language games.

But there are things that we recognize as mistakes and confusion. We may argue against particular uses of a word and certain perspectives. But that has nothing to do with “conceptual hygiene.” In his *Lectures on Aesthetics*, Wittgenstein tries to make clear the nature of this work as a fight to *persuade* about how to understand what is before us:

What I’m doing is also persuasion. If someone says: “There is not a difference”, and I say: “There is a difference” I am persuading, I am saying “I don’t want you to look at it like that.”³⁴

I am in a sense making propaganda for one style of thinking as opposed to another. I am honestly disgusted with the other. Also I’m trying to state what I think. Nevertheless I’m saying: “For God’s sake don’t do this.” E.g. I pulled Ursell’s proof to bits. But after I had done, he said that the proof had a charm for him. Here I could only say: “It has no charm for me. I loathe it.”³⁵

There are different descriptions in dispute, and there is no independent or absolute instance to settle it. Our everyday use of language does not remain unchanged face to anything that changes our lives – science, for example. It is open and changes all the time. We may evaluate new uses and their context, and we may conclude, for example, that saying that “the brain sees” something or that its left side “communicates” with its right side may lead to confusion and may be mystifying. We may try to persuade neuroscientists to use different concepts. But Bennett & Hacker’s conception about the foundation of cognitive science is something quite different and is far from what we find in Wittgenstein’s later philosophy.

³³ WITTGENSTEIN, 2009, 124.

³⁴ WITTGENSTEIN, 1966, p. 35.

³⁵ WITTGENSTEIN, 1966, p. 37.

4.

Bennett & Hacker seems to conceive that the “mereological fallacy” is a pervasive problem in psychological neuroscience.

A cardinal conceptual error of so much current cognitive neuroscience is that it ascribes to the brain attributes that it makes sense to ascribe only to the animal as a whole. In so doing, contemporary neuroscience commits what we called ‘a mereological fallacy.’ Strikingly, neuroscience ascribes to the brain much of the same range of properties that Cartesians ascribed to the mind. It thus operates with a conceptual scheme that is roughly isomorphic with Cartesian dualism.³⁶

The debate about the mereological fallacy refers to an argument from Aristotle’s *De Anima*: “to say that the soul [*psyché*] is angry is as if one were to say that the soul weaves or builds. For it is surely better not to say that the soul pities, learns or thinks, but that a man does this with his soul”.³⁷ Aristotle’s assertion that it is “better not to” attribute anger to the soul is part of his argument in support of a particular conception about the *soul*. It is not a general contraposition to any such attribution. Bennett & Hacker’s argument, by its turn, is not circumscribed to particular language games with the concepts of “soul” and “brain.” It is a logical objection to any attribution of psychological concepts to the brain. It also presents a general demand for a specific kind of *criterion* to the use of psychological concepts. A strikingly result of accepting their argument seems to be that we are not allowed to say about the soul that it knows, feels, dies, etc., since all our criteria for attributing these concepts refer to body actions, not to the soul.

Bennett & Hacker relate the mereological fallacy also to Descartes’ *Homunculus*

³⁶ BENNETT & HACKER, 2014, p. 111. “Like Cartesianism, it [cognitive neuroscience] ascribes psychological attributes to a part of a human being. Furthermore, it explains the possession of psychological attributes by a human being by reference to the psychological attributes allegedly ascribable to a part of the human being, namely, to the brain. This (...) is not an error of fact, but a logical or conceptual error.” (BENNETT & HACKER, 2014, 111); “It comes to this: Only of a human being and what resembles (behaves like) a living human being can one say: it has sensations; it sees, is blind; hears, is deaf; is conscious or unconscious.” (WITTGENSTEIN, 2009, 281)

³⁷ ARISTOTLE, 2016, 408b, 12-15.

argument. In the *Dioptrics*, Descartes rejects the supposition that the eye produces an image that is sent to the brain.³⁸ If there is an image *in the brain*, there must be another eye to see it (hence the *homunculus*), which, by its turn, produces another image, leading to an infinite regress. Descartes's intention is to refuse the demand for an explanation about how the "inverted image" produced in the eye is "corrected" in the brain. The *Homunculus* argument is interesting and useful but is different from that of Aristotle. It is useful to contrapose the supposition that there are images in the brain (or in the mind). However, it does not imply or support that there is a mereological fallacy. There is no reason in Descartes' argument to suppose that a mereological *fallacy* is a problem in the attribution of seeing to the brain, instead of being possible to reformulate his argument as a mereological fallacy.

In a reformulation of their argument about the mereological fallacy, Bennett & Hacker say that we do not have a criterion for the attribution of thinking to the brain, only to an entire animal. This is a general argument against any kind of "reductionism." The problem with it is to know how we should understand the demand for a *criterion* for the use of a word. Bennett & Hacker's conception about it seems close to that of N. Malcolm in his presentation of what he supposes to be Wittgenstein's "private language argument." Malcolm demands a public or behavioral criterion to determine the reference of a word. "Only by being associated with criteria could a word mean something."³⁹ It follows that sensation terms are "conceptually tied to criteria with behavioral conditions of satisfaction"⁴⁰. These criteria are, in the end, behaviors, and Malcolm's demand is "essentially a form of the old verificationist dogma of the Vienna Circle."⁴¹

To run together one use of 'criterion' which is oriented towards verification with another whose orientation is toward meaning is to make the notion of 'criteria' carry essentially the burden of a verificationist theory of meaning.⁴²

This seems to be precisely the case of Bennett & Hacker's use.

³⁸ Cf. DESCARTES, AT VI, 130: "as if there were yet other eyes within our brain by mean of which we could perceive [an image resembling an external object];" cf. also COTTINGHAM, 2008, p. 19-20, and KENNY, 2008, p. 172-3.

³⁹ NIELSEN, 2008, p. 77.

⁴⁰ NIELSEN, 2008, p. 92.

⁴¹ NIELSEN, 2008, p. 92.

⁴² BUCK, 1962, p. 196, *apud* NIELSEN, 2008, p. 92.

We ascribe pain to a person or an animal on the basis of their behavior (including their verbal behavior). If a person injures himself and screams or groans, (...) we take such pain-behavior in these circumstances to be justifying grounds or evidence for ascribing pain to the person. / Pain-behavior is a criterion – that is, *logically* good evidence for being in pain – and perceptual behavior (...) is a criterion for the animal’s perceiving.⁴³

There is well-established literature showing that behind the kind of demand for criteria presented by Malcolm lies the positivist concept of verification.⁴⁴ But even if we accept such a request, a neuroscientist could still answer that they may present clear criteria for their attribution of psychological concepts, clearer than human behavior: thinking and seeing is identified and situated in the brain by an MRI while the person answers some question or sees an image. It can even be shown that the same process happens in every instance that a perception or memory is accessed. Neuroscientists are not simply attributing to parts of the body what may be assigned only to the whole body. Otherwise, they would also conclude that feet also think and see. They have reasons for saying that thinking happens in the brain, that it is a brain process – the MRI, for instance. Beyond that, if we apply the mereological argument to other contexts, we will also reject, together with saying that the brain thinks, that Earth moves and that we know somebody is sick because we have a lab exam, and not because we can identify specific symptoms.

From a Wittgensteinian point of view, there is no clear-cut and stable use or meaning of a word: uses constitute families. Wittgenstein’s “private language argument,” for instance, is not an investigation about the criteria for applying some psychological words, but about the way we use them, with the intention of making clear that some uses of psychological concepts are not what they seem to be. Meaning is understood as the *effective use* we make of words. How can that conciliate with Bennett & Hacker’s supposition that they shall “sketch out the conceptual framework for neuroscience”?⁴⁵ Transposing ordinary psychological vocabulary to descriptions about the brain presents real problems. It is not evident that the vocabulary may be changed

⁴³ BENNETT & HACKER, 2014, p. 81-82.

⁴⁴ MALCOLM, 1954, p. 530-559.

⁴⁵ BENNETT & HACKER, 2014, p. 114.

this way without originating obscurity and confusion. Neuroscience's uses of psychological concepts may be problematic, but not because they result from a fallacy. The question we may present is not about the *logical correctness* of the use of a concept. We do not have any criteria for that. According to Wittgenstein, our task is to describe new contexts of use, identifying its peculiarities, ambiguities, and limits, aware that any word is used with blurred edges and that there is no absolute clarity.

Cognitive Neuroscience has an extensive superposition with traditional fields of philosophical research. Descartes' *homunculus* fallacy, for example, is a relevant argument in the contemporary debates about vision and perception.⁴⁶ The interest of research on philosophy and neuroscience is both in clarifying the use of cognitive and psychological vocabulary and in revising philosophical conceptions that are under the direct impact of these studies. This kind of interplay is strongly absent in Bennett & Hacker's work. Their idea that philosophy precedes and is independent of scientific research implies that it is not open to the fruitful results of psychological neuroscience, to their new and relevant use of concepts. Their position is strongly prescriptive and conservative in their attitude about concepts (instead of their explicit denial of that). A different conception about the meaning of concepts, a different perspective on Wittgenstein's mature work, may result in a more relevant intervention in the debate about philosophy and psychological neuroscience.

5.

In an exemplary fragment of the *Philosophical Investigations*, Wittgenstein evaluates the translation of psychological concepts into brain processes. It is part of the debate about the criteria for saying that a student is *reading* a word (instead of "accidentally" saying it). After a certain point, repetition in the student's behavior may lead to that conclusion. But the question "about that first word" remains: "the change when the pupil began to read was a change in his *behavior*; and it makes no sense here to speak of 'a first word in his new state'."⁴⁷

But isn't that only because of our too slight acquaintance with what goes on in the brain and the nervous system? If we had a more accurate knowledge of

⁴⁶ NÖE, 2006, p. 39-44.

⁴⁷ WITTGENSTEIN, 2009, 157.

these things, we would see what connections were established by the training, and then when we looked into his brain, we would be able to say: “Now he has *read* this word, now the reading connection has been set up.” — And it presumably *must* be like that – for otherwise how could we be so sure that there was such a connection? That it is so is presumably a priori – or is it only probable? And how probable is it? Now, ask yourself: what do you *know* about these things? — But if it is a priori, that means that it is a form of representation which is very appealing to us.⁴⁸

If more accurate knowledge about the brain could unveil that psychological concepts correspond to brain processes, one could look at the student’s brain and search for the first event of “reading.” It would allow us to identify the first word the student read, the first time “the reading connection has been set up.” Things being so, one could suppose that a brain process, not the student’s behavior, is what is called “reading.” The discomfort with not being possible to point to the “one thing” that we call “reading” is what leads, in Wittgenstein’s example, to the search for something hidden in the brain. A brain process is, supposedly, the *thing* that “is the meaning” of “reading.” The situation is similar to the one Wittgenstein describes ironically in his commentaries about ostensive definitions:

And we do here what we do in a host of similar cases: because we cannot specify any *one* bodily action which we call pointing at the shape (as opposed to the color, for example), we say that a *mental, spiritual* activity corresponds to these words. / Where our language suggests a body and there is none: there, we should like to say, is a *spirit*.⁴⁹

The brain process plays the role of a hidden thing meant by “reading.” Against this conception, Wittgenstein contraposes the central idea of the *Philosophical Investigations*, that meaning does not presuppose or imply reference. We play language-games with words, they have a place in our actions, in our forms of life. The use of a words presupposes training, not definitions. That a word or sentence is meaningful results from our practices, not from a fixed relationship between sounds and “things” (or

⁴⁸ WITTGENSTEIN, 2009, 158.

⁴⁹ WITTGENSTEIN, 2009, 36.

processes) in the world. We say the student is reading as a reaction to his behavior. That is the game we were trained to play with this word. As a consequence, we cannot determine the first word and sentence he read. The idea that the meaning of psychological concepts is a brain process insists on finding “something” that is *reading*, hidden behind our use of that word. It is accessory in supporting the conception about meaning that Wittgenstein is criticizing in the *Philosophical Investigations*.

The problem Wittgenstein sees in presenting brain processes as the meaning of psychological concepts (or, at least, the criterium for ascribing such concepts) is not the mereological fallacy that could be identified, for example, in saying that “the mind reads.” The problem is that moving the reference of a word to inside the brain does not solve the initial problem. How is it to be identified the particular brain processes correlated to specific psychological concepts? It is determined either *a priori* (as a necessary connection) or as something inferred from experience (and only probable). Saying that it is a conclusion from experience means that the assertion that the student is reading at a certain time is what allows us to associate *reading* with what is simultaneously happening in his brain. This is an empirical inference, and the possibility of misidentifying the proper mental process is always present. Any conclusion about such a connection must be confronted with our preliminary recognition that the student was reading. Our final criterium for ascribe “reading” is, therefore, still the student’s behavior. The *translation* of psychological concepts to brain processes seems to be a paraphrase of the habitual use of these concepts.

If we say, on the other hand, that the connection between “reading” and a particular process in the brain is necessary and *a priori*, Wittgenstein concludes that “it is a form of representation [*Darstellungsform*] which is very appealing to us.” To be *a priori*, the correlation should be set as a definition. We do not discover brain processes behind the use of such concepts, we assert them as the meaning of psychological concepts – even when it is not congruent with our everyday use of “reading,” for instance. There is no problem with such definitions, but the result is that we have a different concept, and, in the end, there is no reason to suppose that there is a *correlation* between brain processes and ordinary psychological concepts.

Wittgenstein’s remarks about “reading” and brain processes are exemplary of his perspective on the use of psychological concepts in neuroscience. It potentially involves a mistaken conception about meaning as a reference relationship. The reason for searching the brain processes correlated to “reading” is often to assure referentiality to

this concept. Beyond that, to assert that there is a brain process we call “reading,” either we presuppose the ordinary use of this concept and try to find something recurrent in the brain when the student is reading, or we say it by definition, obtaining a different concept (what is much less appealing and useful than saying that what we ordinarily call “reading” means a brain process). Wittgenstein does not object, however, to the supposition that reading is a brain process, or to attributing reading to the brain. If there is a place for that in our practices (and it can easily be supposed), if there is a game we play with it, the word is as meaningful as any other. Philosophy is not prescriptive about meaning, but it makes clear that the result of our experiments and ideas may not be what we expected: the development of neuroscience does not assure that psychological concepts mean brain processes. Even when the meaning of these concepts is the use neuroscience makes of it.

REFERENCES

- ARISTOTLE. *De Anima*. Oxford: Oxford University Press, 2016.
- BAKER, G. P. & HACKER, P. M. S. *Wittgenstein: understanding and meaning – volume i, part ii of an analytical commentary on the Philosophical Investigations: Exegesis §§1-184*. Oxford: Blackwell Publishing, 2009.
- BECHTEL, W., MANDIK, P. & MUNDALE, J. *Philosophy and the neurosciences*. Oxford: Blackwell, 2001.
- BENNETT M. *Neuroscience and philosophy*. New York: Columbia Univ. Press, 2010.
- BENNETT, M. & HACKER, P. *History of cognitive neuroscience*. Malden, MA: Wiley-Blackwell, 2013.
- _____. *Philosophical foundations of neuroscience*. Malden: Blackwell Publishing, 2014.
- BUCK, R. Non-other Minds. In: R. J. Butler (Ed.). *Analytical philosophy*. Oxford: Basil Blackwell, 1962.
- CARVALHO, M. The groundlessness of our practices - Wittgenstein on Moore’s Propositions. In: Moreno, A. (Org.). *Wittgenstein: conhecimento e ceticismo*. Campinas: Coleção CLE, vol. 77, 2016.
- _____. Linguagem sem regras fixas: a normatividade da linguagem e a recusa do modelo de cálculo nas *Investigações Filosóficas*. In: Smith, P.; Silva Filho, W. & Santos, P. (Orgs.). *Crença, verdade e racionalidade*. Salvador: EDUFBA, 2014.

- COTTINGHAM, J. *Cartesian reflections*. Oxford: Oxford University Press, 2008
- CRICK, F. *The Astonishing hypothesis, after Philosophical foundations of neuroscience*. New York, NY: Simon & Schuster, 1995.
- DESCARTES, R., Adam, C., Tannery, P. *Oeuvres De Descartes*. Paris: Librairie philosophique J. Vrin, 1996.
- ENGELMANN, M. *Wittgenstein's philosophical development*. Hampshire: Palgrave Macmillan, 2013
- KENNY, A. *From Empedocles to Wittgenstein*. Oxford: Oxford University Press, 2008.
- KLEIST, H. *The letters of Heinrich von Kleist*. Edited and translated by Philip Miller, New York: Dutton, 1982.
- MALCOLM, N. Wittgenstein's *Philosophical investigations*. *The Philosophical Review*, Vol. LXIII, 1954.
- MEDINA, J. *The unity of Wittgenstein's philosophy*. Albany: SUNY Press, 2002.
- MÜHLHÖLZER, F. *Braucht die Mathematik eine Grundlegung?* Frankfurt am Main: V. Klostermann, 2010.
- NIELSEN, K. *The evolution of the private language argument*. UK: Ashgate Publishing, 2008.
- NOË, A. *Action in perception*. Cambridge, MA: Bradford Books, 2006.
- STERN, D. *Wittgenstein on mind and language*. New York: Oxford University Press, 1995.
- _____. *Wittgenstein's Philosophical Investigation*. Cambridge University Press, 2004.
- WITTGENSTEIN, L. *Bemerkungen über die Grundlagen der Mathematik, Remarks on the foundation of mathematics*. Eds. G. E. M. Anscombe, Rush Rhees & G. H. von Wright, Oxford: Blackwell, 1956; 3rd. ed. 1978.
- _____. *Lectures and conversations on aesthetics, psychology and religious belief*. Compiled from notes taken by Yorick Smythies, Rush Rhees & James Taylor, ed. by Cyril Barrett. Oxford: Blackwell, 1966.
- _____. *On certainty*. Ed. G. E. M. Anscombe & G. H. von Wright, trans. D. Paul & G. E. M. Anscombe. Oxford: Blackwell, 1974.
- _____. *Philosophical grammar*. Ed. Rush Rhees, tr. Anthony Kenny. Oxford: Basil Blackwell, 1974.
- _____. *Philosophical remarks*, ed. R. Rhees, trans. R. Hargreaves & R. White. Oxford: Blackwell, 1975.

- _____. *Philosophische Bemerkungen*. Suhrkamp Verlag: 1984. Philosophical Remarks. Wiley-Blackwell: 1991.
- _____. *Philosophische Grammatik*. Ed. Rush Rhees. Oxford: Basil Blackwell, 1969.
- _____. *Philosophische Untersuchungen, Philosophical investigations*. 4th ed., Oxford: Blackwell, 2009.
- _____. *Philosophische Untersuchungen: Kritisch-genetische Edition*. Herausgegeben von J. Schulte, H. Nyman, E. v. Savigny und G. H. von Wright. Frankfurt am Main: Suhrkamp Verlag, 2001.
- _____. *The big typescript: TS-213*, ed. C. G. Luckhardt & M. A. E. Aue (Oxford: Blackwell, 2005).
- _____. *Tractatus logico-philosophicus*. trad. Luiz H. L dos Santos. São Paulo: Edusp, 1995.
- _____. *Über Gewissheit, On certainty*. G.E.M. Anscombe & G.H. von Wright, eds. Oxford: Basil Blackwell, 1969.