

Experiential Physicalism versus Biological Naturalism

Mario Martin



Master of Science
Cognitive Science
School of Informatics
University of Edinburgh

2016

Abstract

Several theories within the philosophy of mind are addressing the phenomenal aspects of consciousness, i.e. focus on the question how the subjective character of our experiences can arise from a physical substrate. This paper introduces and compares *Biological Naturalism* by John Searle (1992) and *Experiential Physicalism* by Galen Strawson (2010), two theories that reject any supernatural explanation of consciousness, and yet insist on the irreducibility of the mental to the physical. *Biological Naturalism* defines consciousness as an emerging biological feature of the brain, while *Experiential Physicalism* explains its phenomenal aspects via panpsychism, i.e. the view that the most basic building blocks of physical reality themselves have experiential properties. The analysis shows that Searle's account has to be considered as inconsistent while Strawson's view – although logically consistent – might not qualify as a foundation for a scientific framework. It is concluded that the problem of the phenomenal aspects of consciousness might turn out to be unsolvable because the human mind may simply be unable to form the required concepts for a sufficient theory.

Keywords: consciousness, biological naturalism, experiential physicalism, emergentism, panpsychism;

Table of Content

1	Introduction	1
2	Background	2
2.1	Historical origins of the philosophy of mind	2
2.2	Classification of contemporary theories in the philosophy of mind.....	3
3	Searle's Biological Naturalism	6
3.1	The claim	6
3.2	The justification.....	8
3.3	The critique	12
4	Strawson's Experiential Physicalism	17
4.1	The claim	17
4.2	The justification.....	19
4.3	The critique	22
5	The two Theories in Comparison	26
5.1	Similarities	26
5.2	Differences	27
6	Conclusion	28
7	Acknowledgments.....	31
8	Bibliography	32

1 Introduction

The term *consciousness* is a rather fuzzy one. It can refer to functional aspects of the mind primarily dealing with representational content, i.e. content that is accessible “for use as a premise in reasoning, poised for rational control of action, and poised for rational control of speech” (Block 1998), as well as to phenomenal aspects of the mind, i.e. to the subjective experience of “what it’s like” to be in a certain mental state (cf. Chalmers 2010, p. 4; cf. Nagel 1974).

Although we do not yet have a full description of all the functional aspects of the mind (i.e. a detailed explanation for a cognitive system’s ability to discriminate, categorize, and react to environmental stimuli or for its capacity to learn and reason, etc.) these aspects could – in principle – be tackled via scientific means. It is conceivable that someday we will have fine-grained cognitive models of the mind respectively neurophysiological models of the brain that allow us to understand the underlying functional mechanisms (cf. Chalmers 2010, p. 3). Therefore, the *functional aspects* of consciousness can be regarded as the *easy problem* (ibid.).

However, for a conscious organism there is also “something that it is like to *be* that organism – something it is like *for* the organism. We may call this the subjective character of experience” (Nagel 1974; [emphasis in original]). Since we do not have an explanation of how or why experience arises from its physical basis (cf. Chalmers 2010, p. 4) and do not even know how this explanation could be achieved within the current boundaries of our scientific framework (cf. Nagel 1974), the subjective character of experience, the *phenomenal aspect* of consciousness, forms the *really hard problem* (cf. Chalmers 2010, p. 4):

“How it is that anything so remarkable as a [subjective] state of consciousness comes about as a result of irritating nervous tissue, is just as unaccountable as the appearance of the Djinn when Aladdin rubbed his lamp” (Huxley 1886, p. 170).

Two theories that address the *hard problem* of consciousness (among other theories) are:

- *Biological Naturalism* by John Searle (1992) and
- *Experiential Physicalism* by Galen Strawson (2010).

The aim of this dissertation lies in introducing these two theories as well as in providing a critical evaluation and comparison of their relevant arguments.

Therefore this paper:

- will give a brief overview of how the *philosophy of mind* has developed and will introduce Chalmers' framework for classifying theories within the *philosophy of mind* (which serves as a background for analyzing and comparing Searle's and Strawson's theories throughout the entire dissertation),
- will introduce Searle's *Biological Naturalism* as well as Strawson's *Experiential Physicalism*, while focusing on their main claims and justifications, and finally
- will compare these two positions in detail by pointing out and discussing their individual strengths and weaknesses.

2 Background

2.1 Historical origins of the philosophy of mind

By stating that the *act of perceiving* itself is something that does not "escape our notice" Aristotle became one of the first philosophers that addressed problems associated with *phenomenal consciousness* (cf. Caston 2002). Nevertheless, it is still a topic of discussion whether Aristotle had a clear concept of *consciousness* (ibid.) and due to the fact that he considered the brain merely as a cooling device for the blood (Gross 1995), while assuming that the center of all mental processes is located in the heart, it is rather difficult to align his view on this matter with the current discourse.

The first explicit theory about the nature of the mind and its relation to the body (respectively to the physical world in general) was provided in the 17th century by the French philosopher René Descartes. By distinguishing the human mind as a thinking substance (*res cogitans*) from the human body as an extended substance (*res extensa*) Descartes founded a line of thought known as *substance dualism*, which is still referenced and discussed by more recent approaches to consciousness (cf. Strawson

2010, p. 123; cf. Searle 1992, p. 28)¹. Yet, apart from the definition of body and mind as two ontologically different substances that interact via the pineal gland (cf. Shapiro 2006), Descartes' investigation focuses rather on epistemological questions (e.g. on the question whether we can trust our senses) than on the phenomenal aspect of experience (cf. Jorgensen 2014).

One of the earliest appearances of the term *consciousness* – close to its modern sense – can be found in Locke's definition of *personal identity*:

"[...] it being impossible for any one to perceive without perceiving that he does perceive. When we see, hear, smell, taste, feel, meditate, or will anything, we know that we do so. Thus it is always as to our present sensations and perceptions [...]. For, since consciousness always accompanies thinking, [...] in this alone consists personal identity [...]" (Essay II, xxvii, 9).

In this passage the term *consciousness*² can be treated synonymously with *being aware of having sensations and perceptions* and therefore is directly connected with the phenomenal aspect of experience. Furthermore, by differentiating between *primary qualities* that are intrinsic to the object (e.g. extension, figure, motion, solidity, number, etc.) and *secondary qualities* that depend on the mind of the observer (e.g. color, taste, smell, sound, etc.) (Essay II, viii, 9), Locke laid the foundation for the previously mentioned *hard problem* of consciousness: the subjective nature of our experiences.

2.2 Classification of contemporary theories in the philosophy of mind

In the 20th century, influenced by progress of *natural sciences* and the program of the *Wiener Kreis*, the philosophy of mind largely refrained from making bold

¹ „To each substance there belongs one principal attribute; in the case of mind, this is thought, and in the case of body it is extension" (Descartes 1644, p. 20 [transl. from Latin]; cf. Descartes 2008, p. 55).

² For an alternative interpretation of the term *consciousness* in the work of Locke considering the context – i.e. defining *personal identity* – see Noonan (2003, pp. 43).

metaphysical statements (like “body and mind are two ontologically different substances”) (cf. Morton 1997, pp. 143) and started the attempt to develop theories that are based on (or are at least compatible with) scientific principles and facts (e.g. *behaviorism* (e.g. Skinner 1974), *functionalism* (e.g. Putnam 1960) or different kinds of *identity-theories* (e.g. Smart 1959)).

According to Chalmers all these theories as well as their counterparts and successors can be classified via the following categories – depending on how they try to tackle respectively avoid the *hard problem* of consciousness (Chalmers 2010, pp. 103-140); hence they will be referenced throughout the entire course of this dissertation:

Type-A Materialism: This type of materialism denies the existence of an *epistemic gap*, i.e. claims that physical truth entails phenomenal truth: the *how-it-is-done* explains the *what-it-is-like*. In its most radical form (*eliminativism*) *type-A materialists* avoid the hard problem simply by denying the existence of the phenomenal aspects of consciousness. Other forms of *type-A materialism* (e.g. *functionalism* and *behaviorism*) try to define the *concept* of consciousness entirely through input-output relations or behavioral terms.

Type-B Materialism: This kind of materialism accepts the existence of an *epistemic gap* but rejects the view that an *epistemic gap* entails an *ontological gap*: although the *concept* of consciousness cannot – in principle – be explained via functional or physical concepts, we still might discover empirically that all these concepts refer to the same ontological entity/thing in nature. *Type-B materialism* closes the *epistemic gap* via the assumption of a fundamental law-like connection between physical and phenomenal states that cannot be expressed in more basic terms.

Type-C Materialism: For *type-C materialists* there is a deep *epistemic gap* but this gap is considered as closable if we were able to overcome our own limitations. A conceptual revolution of physics (that goes beyond functional and structural descriptions) might be able to explain all aspects of consciousness, i.e. even give an explanation for the subjective nature of our experiences via physical terms (cf. Nagel 1974).

Type-Q Materialism: This kind of materialism arises after rejecting the taxonomy that causes the distinction between other types of materialism (i.e. the distinction between empirical and conceptual truth that separates *type-B* from *type-C materialism* since all concepts may be revised in the presence of new empirical data). Quine for example states that concepts like *physical objects*, *forces* or *abstract entities* are “myths on the same footing with [...] gods”, serving merely to simplify our treatment of experience (1951). However, also *type-Q materialism* faces the same problems as materialism in general and does not contribute anything particularly new to the analysis of the question in focus.

Type-D Dualism: This type of dualism claims that phenomenal properties may play a causal role in the physical world and hence is known as *interactionism*. Prominent examples for *type-D dualism* are Descartes’ *substance dualism* and certain forms of *property dualism* as long as their “phenomenal properties will play an irreducible role in affecting the physical properties” (Chalmers 2010, p. 126).

Type-E Dualism: For *type-E dualists* there is an *ontological gap* between phenomenal and physical properties (i.e. phenomenal states might be caused by physical states but still cannot be reduced on a mere physical account). Furthermore, phenomenal properties are considered as *epiphenomenal* (i.e. without effect on the physical) which makes *type-E dualism* compatible with scientific theories that insist on causal closure of the physical world (cf. Jackson 1982).

Type-F Monism: This type of monism suggests that *protophenomenal* properties, i.e. properties that are not phenomenal but that – when arranged/combined in the right way – constitute phenomenal properties, are “located at the fundamental level of physical reality and in a certain sense underlie physical reality itself” (Chalmers 2010, p. 133; cf. Chalmers 2015, pp. 247). Depending on the point of view this type of monism can be seen as *neutral monism*, *idealism* or *panpsychism* – or in its most general form: *panprotopsychism*.

The following two chapters will introduce Searle's *biological naturalism* as well as Strawson's *experiential physicalism* and classify these theories (respectively delimit them from other theories) according to the presented categories of this framework.

3 Searle's Biological Naturalism

This chapter will:

- lay out the claim (i.e. the core content) of Searle's *biological naturalism*,
- address Searle's justification for his conception of consciousness and
- examine the argumentation for (respectively against) Searle's claims.

3.1 The claim

John Searle presents in *The Rediscovery of the Mind* his conception of consciousness that basically claims:

"Mental phenomena are caused by neurophysiological processes in the brain and are themselves features of the brain. [...] Consciousness, in short, is a biological feature of human and certain animal brains. It is caused by neurobiological processes and is as much a part of the natural biological order as any other biological features such as photosynthesis, digestion, or mitosis" (1992, p. 1 and p. 90).

Furthermore, Searle specifies consciousness explicitly as an emergent property:

"The brain causes certain 'mental' phenomena, such as conscious mental states, and these conscious states are simply higher-level features of the brain. Consciousness is a higher-level or emergent property of the brain in the utterly harmless sense of 'higher-level' or 'emergent' in which solidity is a higher-level emergent property of H₂O molecules when they are in a lattice structure (ice), and liquidity is similarly a higher-level emergent property of H₂O molecules when they are, roughly speaking, rolling around on each other (water). Consciousness is a mental, and therefore physical, property of the brain in the sense in which liquidity is a property of systems of molecules" (1992, p. 14).

At first glance – due to their close connection to the physical – these definitions would suggest that Searle advocates some kind of materialistic conception of

consciousness (*type-A*, *-B*, *-C* or *-Q materialism*; see chapter 2.2). Yet, although Searle claims that mental states are *caused* by neurophysiological processes (which allows a *causal reduction* of consciousness to physical properties of the brain) he explicitly denies the possibility of an *ontological reduction* (cf. 1992, p. 116) since experience – according to Searle – has a *first-person ontology* (cf. 1992, p. 16), i.e. a subjective character that would get lost if we tried to reduce it to the physical.

Searle illustrates the ontological distinction between mental states and brain states via an example (cf. 1992, p. 117): pain is experienced from a subjective *first-person point of view*. Pain is also caused by neurons firing in the brain. We can reduce the experienced pain *causally* to the firing of neurons. However, if we tried to reduce the *first-person* experience of pain *ontologically* to the firing of neurons (i.e. identify the *subjective aspects* of pain with *objective aspects* that are accessible from a *third-person point of view*) we would miss out essential features of the pain:

“No description of the third-person, objective, physiological facts would convey the subjective, first-person character of the pain, simply because the first-person features are different from the third-person features” (Searle 1992, p. 117).

The crucial point in Searle’s conception of consciousness is that the gap between the *what-it-is-like* and the *how-it-is-done* is not just an *epistemic one* (i.e. is not just a gap between our experience and our means of explanation): the *what-it-is-like* and the *how-it-is-done* happen on two entirely different levels. How this could be possible will be addressed in chapter 3.2. and why – in Searle’s case – it is self-contradictory will be shown in chapter 3.3.

The mere inclusion of a subjective character of experience into the conception of consciousness distinguishes *biological naturalism* from *type-A materialism* and if there are two ontologically different entities (a *subjective* one and a *physiological* one) then *biological naturalism* is also different from *type-B* and *-C materialism* which only allow an *epistemic gap* but no *ontological gap* (i.e. *type-B* and *-C materialism* just as *materialism* in general ultimately insist on the *ontological reducibility* of the mental to the physical).

But does the fact that Searle distinguishes a *first-person ontology* from *third-person features* while denying *ontological reducibility* of the former to the latter make *biological naturalism* a dualist theory? – In *Mind a Brief Introduction* Searle states that “[b]ecause conscious states are real features of the real world, they function causally” (2004, p. 114). By granting consciousness a causal role in the physical world *biological naturalism* cannot be classified as *type-E dualism* but still seems to be compatible with *type-D dualism*. Nevertheless, Searle would disagree and the justification for this objection will be given in the next chapter.

3.2 The justification

Searle claims that consciousness “is part of the ordinary physical world and is not something over and above it” (2004, p. 127). Furthermore, he stipulates the existence of a *first-person ontology* that is different from a *third-person ontology* (2004, p. 209). Yet he still rejects *materialism*, which claims “that there are no ontologically irreducible mental phenomena” (2004, p. 126), as well as (*substance*) *dualism*, which considers the mental as something “over and above” the physical (ibid.). – Why?

Searle considers Descartes’ dualist dichotomy (i.e. the mental vs. the physical) as rather unfavorable because along with Descartes’ vocabulary we inherited “a certain set of categories, within which we are historically conditioned to think about these problems” (1992, p. 14). Oppositions like “physical” vs. “mental”, “body” vs. “mind”, “materialism” vs. “mentalism” or “matter” vs. “spirit” are immanent to this vocabulary and implicitly contain the thesis “that the same phenomenon under the same aspects cannot literally satisfy both terms” (ibid.). Due to this fact every theory about consciousness that addresses the phenomenal character of experience appears – according to Searle – automatically as *suspicious* (1992, p. 13) and *antiscientific* (1992, p. 4), just as if acknowledging the existence of consciousness would have to entail the affirmation of something *nonphysical* (1992, p. 13). However, an essential claim of Searle’s view is: “The fact that a feature is mental does not imply that it is not physical; the fact that a feature is physical does not imply that it is not mental” (1992, p. 14).

How can Searle insist on an ontologically irreducible difference between the *first-person view* and a *third-person view* (i.e. in old Cartesian terms: between something *mental* and something *physical*) if the mental (according to the quote above) has *not* to be considered as entirely different from the physical? How should this be possible? Searle's answer is that the mental is *caused* by the physical but the subjective features of the mental are *irreducible* to the objective features of the physical, since the *third-person view* only delivers an incomplete description:

“mental phenomena have a first-person ontology, in the sense that they exist only insofar as they are experienced by some human or animal subject, some ‘I’ that has the experience. And this makes them irreducible to any third-person ontology, any mode of existence that is independent of any experiencing agent” (Searle 2004, p. 98).

I am sympathetic with Searle's view but given the task of explaining it I feel like Augustine when asked to define time: “If no one ask of me, I know; if I wish to explain to him who asks, I know not” (Augustine 1900; xi, 14 [p. 244]). Nevertheless, I will try to illustrate *how consciousness can be part of the physical world but yet be irreducible to the physical* via an analogon – inspired by Leibniz' 5th letter about the nature of space (Leibniz & Clarke 1717, p. 199):

Imagine four identical objects (e.g. four ideal white billiard balls that are qualitatively identical down to the amount and arrangement of their atoms). We place these balls on an (invisible) billiard table according to Figure 1.

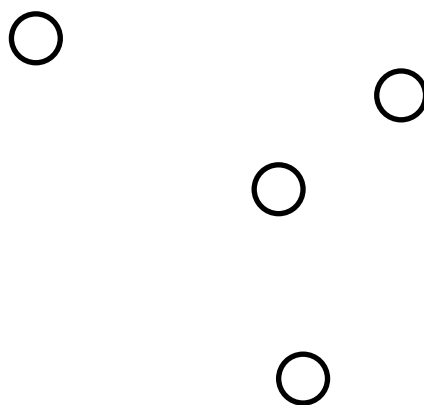


Figure 1: Four identical billiard balls on a pool table

Since the billiard balls are absolutely identical the only way to distinguish them is via their geometrical relations. The relations $\{R\}$ of ball A (to the remaining balls) are different from the relations of ball B (see Figure 2); the relations of ball B are different from the relations of ball C and so forth:

$$R(A,B), R(A,C), R(A,D) \neq R(B,A), R(B,C), R(B,D) \neq \dots$$

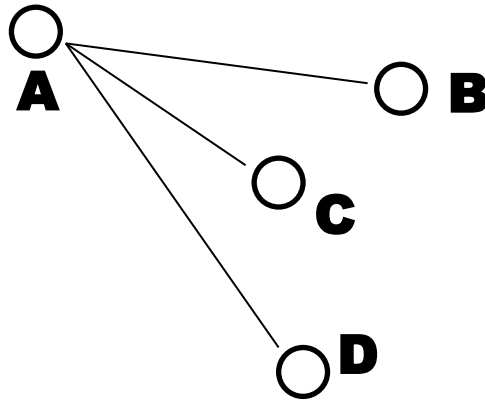


Figure 2: Relations of ball A to the remaining balls

Now imagine that ball A and ball D switch positions. Since the identity and the distinguishability of each billiard ball are solely defined through its geometrical relations to the remaining balls: ball A turns into ball D and vice versa. – But how is this possible? Neither the amount nor the arrangement of the balls' atoms changed when the billiard balls moved from their former position to their new position and yet their identity changed: A turned into D and D turned into A .

It is possible because the identity change didn't happen in the *physical* reality. The billiard ball D (formerly known as A) is still physically the same as it was before the move. The identity change from A to D happened on a *different ontological layer*: it happened in our mind (cf. Leibniz & Clarke 1717, p. 201). – Now, what exactly has this to do with the discussion of *biological naturalism* and how can it be used to understand that *consciousness is part of the physical world but irreducible to the physical*?

In the described scenario we have four physical objects that are identified through their geometrical relations. (Instead of geometrical relations we could also use a set of gravitational force vectors that each billiard ball exerts on the remaining balls –

which would make the scenario even “more physical”.) A physical change within the physical world (the movement of billiard balls) led to a change in our mind (the reassignment of identities). If we replace the billiard balls with neurons and (instead of geometrical relations) use electrochemical relations that exist among these neurons and furthermore generously allow (for the sake of argument) that this configuration (i.e. neurons and their electrochemical relations) constitutes³ a mind, we receive

- an entirely physical system with
- a mind which is as well a product as a part of the system.

Just as a physical change in the original scenario led to a change (the reassignment of identities) in *our* mind, a physical change in the new scenario would lead to a change in the mind that is constituted by this *system*. In the former case we experience the change from our *first-person view* – in the latter case it is experienced from the system’s own *first-person view* which is tied to the experiencing agent and is therefore inaccessible to us. If we observed the new system’s physical change we might also have an experience but it would be different from the experience of the system’s mind: our experience, which is just an observation, would happen on a new (3rd) ontological layer and therefore would be different from the experience that the system’s physical change (on the 1st ontological layer) causes in the system’s mind (on the 2nd ontological layer).

Although this analogon⁴ (hopefully) was able to illustrate how consciousness and the *what-it-is-like* can be seen as part of the physical world and yet be – due to their *first-person ontology* – irreducible to the physical, it is (to my surprise) not what Searle had in mind, which makes *biological naturalism* implausible as the next chapter will show.

³ The difference between *constitution* and *causation* will be discussed in chapter 3.3.

⁴ The Austro-Hungarian writer Franz Kafka wrote 1922 in his prose work *Von den Gleichnissen* [On Analogies]: „Alle diese Gleichnisse wollen eigentlich nur sagen, daß das Unfaßbare unfaßbar ist, und das haben wir gewußt“ [“All these analogies really intend to say is only that the incomprehensible is incomprehensible, and this we knew already”].

3.3 The critique

Most of the facts about *biological naturalism* or consciousness in general seem for Searle to be *obvious*⁵ which might explain some of his bold statements like: “If you are tempted to functionalism, I believe you do not need refutation, you need help” (Searle 1992, p. 9). – However, there are philosophers that consider Searle’s positive account of consciousness merely as “sufficiently contentful to be self-contradictory” (Honderich 2000) or think of his *first-person ontology* as a “metaphysical extravagance (at best just peculiarity or at worse incoherence)” (Dennett 1993). Furthermore, Chalmers (1996, p. 164) and others (e.g. Burton, 1995) criticize that *biological naturalism* would just be another form of *property dualism* (i.e. *type-D dualism*). Searle’s claim that consciousness is an emergent property with causal powers is also a matter of discussion (e.g. Kim 1995). All of these aspects will be addressed in this chapter.

3.3.1 The First-Person Ontology Objection

Dennett is criticizing that Searle’s use of the term *first person ontology* is “unprecedented and unobvious” (1993). According to Dennett, Searle never explains what is meant by the term *ontology* and only uses it to deny that subjectivity of the mental is just an epistemological fact (ibid.)⁶.

I, on the other hand, think that Searle was quite clear about the content of this term by stating that it “is just a fancy way of saying that every mental state has to be somebody’s mental state. Mental states only exist as subjective, first-person phenomena” (1992, p. 70), which – in my view – can only be interpreted as clear commitment to dualism. However, by doing so Searle’s account becomes self-

⁵ I counted 77 occurrences of the term *obvious* (respectively *obviously*) in (Searle 1992) and additional 30 occurrences in (Searle 2004), e.g.: “How is it that so many philosophers and cognitive scientists can say so many things that, to me at least, seem obviously false?” (Searle 1992, p. 3). See also (Dennett 1993).

⁶ For a critique of Dennett’s alternative to the *first-person ontology*, which he calls *heterophenomenology* (Dennett 1991, p. 72), see Zahavi (2007).

contradictory/inconsistent when he states that “there are not two different metaphysical realms in your skull, one ‘physical’ and one ‘mental’” (2004, p. 128).

The term *first-person ontology* only makes sense if we are precisely talking about two different realms in form of two different realities⁷: a subjective *reality* and an objective/physical *reality* – each coming with its *own* ontology (just not necessarily with different substrates). In my view also the *irreducibility* of the *first-person ontology* necessarily entails *different* ontology levels (see chapter 3.2): if an experience only has *what-it-is-like* features on the mental level and if these features (i.e. the experience itself) ceased to exist if we tried to reduce them down to the physical level (cf. Searle 1992, p. 117), then there must be a fundamental difference between these two levels.

For example, the “experience of a broken heart” happens on two *different* ontology levels: on a mental level and on a (neuro-) physical level (as depicted in Figure 3).

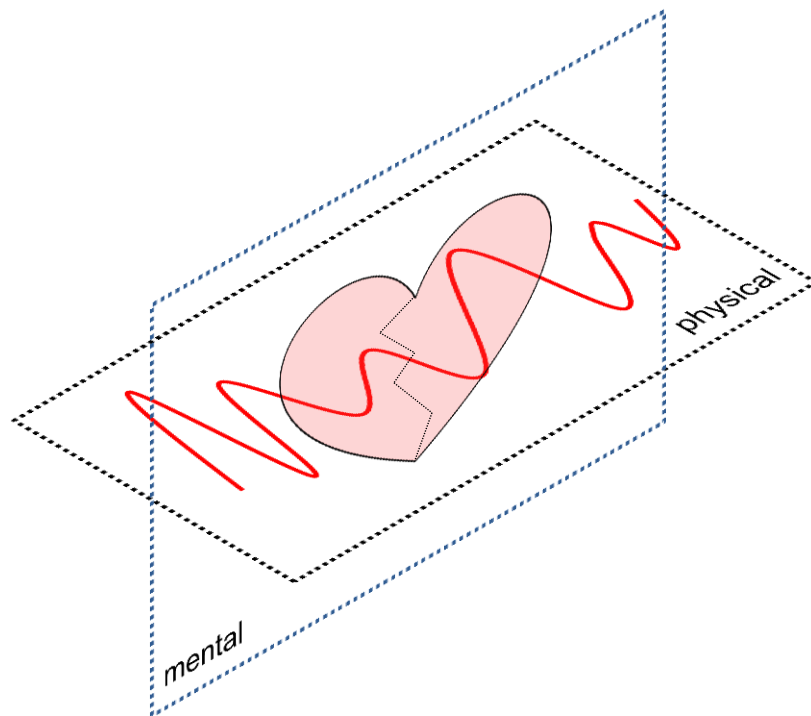


Figure 3: The two different ontology levels of a broken heart.

⁷ The term *reality* means a domain (or ontology level) in which certain things exist. Each ontology is tied to a certain domain: you won't find hope in a toolbox and you won't be able to tighten a loose nut with an imaginary wrench, no matter how detailed your imagination of this wrench is.

On each level (the mental and the physical) a broken heart has its own *essence*, i.e. some intrinsic quality that is specific for (and limited to) the respective level/reality:

“No one has ever directly perceived the neuronal states in her or his own brain as neuronal states. [...] While we are able to experience mental states from the first-person perspective, we remain unable to experience our neuronal states as such” (Northoff & Musholt 2006).

The phenomenal qualities of a broken heart are exclusive to the mental and its electrochemical properties are exclusive to the physical, since the respective qualities only exist in their associated domain. We do not (and cannot) experience a change of electrochemical properties; we experience *agony*, which – in my view – *requires* a claim of two *distinguished incommensurable ontologies* (cf. Kim 1995) each characterized by a set of qualities exclusive to the respective domain.

After all Searle contradicts himself by explicitly stating that the point of the argument showing the irreducibility of the phenomenal aspects of pain (see chapter 3.1) is “ontological and not epistemic” (1992, p. 117) while rejecting the existence of two different ontological realms (2004, p. 128).

3.3.2 The Property Dualism Objection

Searle rejects an interpretation of *biological naturalism* as some form of *property dualism*, which considers unique *mental* properties as *emergent* from an underlying physical substrate (cf. Chalmers 1996, p. 125), by denying the existence of different metaphysical realms (2004, p. 128; see chapter 3.3.1 above) and by explicitly stating:

“The property dualist wants to say that consciousness is a mental and therefore not physical feature of the brain. I want to say consciousness is a mental and therefore biological and therefore physical feature of the brain” (Searle 2002, p. 61).

At this point it seems to me as if we are going around in circles: the mental has the same ontological status as the physical and is yet irreducible to the physical. This appears to me as entirely unintelligible.

The mere claim of Searle's biological naturalism that consciousness is

“caused by neurobiological processes and is as much a part of the natural biological order as any other biological features such as photosynthesis, digestion, or mitosis” (Searle 1992, p. 90)

gives no *epistemic satisfactory explanation* of consciousness (cf. Rowlands 2001, p. 59).

Furthermore, as sympathetic as I am with Searle's attempt to naturalize consciousness, I cannot understand what is meant when he describes consciousness and its irreducible properties as “ordinary higher-level biological properties of neurophysiological systems” (1992, p. 28; cf. p. 13). Since all ordinary biological properties are reducible to the physical, nothing about consciousness seems to be *ordinary* – at least not in my view. Even Searle himself clearly assigns consciousness and its irreducible properties a special status among all other natural phenomena by writing:

“Conscious mental states and processes have a special feature not possessed by other natural phenomena, namely, subjectivity” (1992, p. 93).

3.3.3 The Epiphenomenalism / Overdetermination Objection

Searle gives consciousness the same ontological status as that of physical features of H₂O (cf. Chalmers 1996, p. 130) by claiming that consciousness is just an emergent higher-level feature of the brain, in the same sense as liquidity is a higher-level emergent property of H₂O molecules (cf. Searle 1992, p. 14). When Searle states consciousness is *caused* by the brain (ibid.; cf. Searle 2000; cf. Searle 2004, pp. 208) he seems to ignore that liquidity is not *caused* by the physical properties of H₂O – the physical properties of H₂O rather *constitute* the liquidity of water (cf. Chalmers 1996, p. 130). The difference between *causing* and *constituting* is an important one and can be illustrated by following example: three straight lines arranged in a way that their angular sum amounts to 180° *constitute* a triangle – the triangle is *logically entailed* in the specific arrangement of the lines.

However, the triangle is not *caused* by the specific arrangement of the lines. The triangle is *caused* by someone who is drawing / arranging the lines in a specific way so that they *constitute* a triangle (cf. Kant AA II, p. 202).

If consciousness is *constituted* by the physical then it is *logically entailed* in the physical, which would bring consciousness much closer to a materialistic account than Searle might intend. Just as triangularity is not an *emerging* feature of a triangle, consciousness would no longer be an emerging property if it is *logically entailed* in the physical, which means that its *irreducibility* would have to be questioned.

If, on the other hand, consciousness is indeed *caused* by the brain questions about the interaction between consciousness (as a higher-level feature) and the brain (as the lower-level substrate) would arise: if the world is causally closed (i.e. if all events can be explained by natural laws) and if consciousness is not epiphenomenal (i.e. if it has an effect in the physical world) – how can Searle’s account avoid the problem of overdetermination (cf. Kim 1995)? What causal power could consciousness (as a higher-level feature) have in addition to the causal power of its lower-level substrate?

Searle’s answer to these questions is once more that the mental is biological and at the end physical, yet irreducible:

“[T]he ontological irreducibility of consciousness comes not from the fact that it has some separate causal role to play; rather, it comes from the fact that consciousness has a first-person ontology and is thus not reducible to something that has a third-person ontology, even though there is no causal efficacy to consciousness that is not reducible to the causal efficacy of its neuronal basis” (Searle 2004, p. 209).

Since there are not two different ontologies in *biological naturalism* (as stated in chapter 3.3.1 and chapter 3.3.2) Searle’s account of consciousness collapses finally into plain *type-B materialism* (see chapter 2.2), which allows an epistemic gap between the mental and the physical but ultimately identifies the former with the latter. This means that Searle’s arguments against physicalism might at the end successfully refute his own theory.

4 Strawson's Experiential Physicalism

This chapter will:

- lay out the claim (i.e. the core content) of Strawson's experiential physicalism,
- address Strawson's justification for his conception of consciousness and
- examine the argumentation for (respectively against) Strawson's claims.

4.1 The claim

Strawson himself does not use the term *experiential physicalism* to describe his own conception of consciousness. In fact this term does not even appear in Strawson's *Mental Reality* (2010). It appears only once in *Consciousness and its place in nature* where its content is defined as: "all experiential being is physical being" (Strawson 2006, p. 222). Nevertheless, I think it is very well suited to subsume Strawson's overall account of consciousness, because if one would want to put one of Strawson's own labels on his view, one would have to choose among terms like *naturalized Cartesianism* (Strawson 2010, p. xi), *realistic monism* (2006, p. 3), *materialist intentional realism* (2010, p. 324), *agnostic materialism* (2010, p. 98), *real materialism* (2003a) or *realistic materialist monism* (1999).

The wide variety of names results from the circumstance that Strawson's own view is not following a single strain of thought leading to a single theory. His positive account of consciousness is much more derived from his thorough critique of already established theories within the philosophy of mind and their associated vocabularies (e.g., physicalism, materialism, monism, idealism etc.). Instead of claiming "the solution for the hard problem is X" Strawson asks questions like: If we take consciousness seriously, i.e. if we claim that consciousness really exists, and if we, furthermore, take physicalism seriously, i.e. if we claim that everything that exists is at the very bottom something entirely physical, what would have to be the case for consciousness and physicalism to make both claims compatible with each other? What do these claims entail when they are combined? – a technique that Strawson describes as "dialectically *ad hominem*, in the non-aggressive sense of the term" (2006, p. 186), since it is directly addressed to specific positions that a person might hold.

Despite this variety of names for Strawson's account of consciousness, a set of core assumptions can be extracted from his work:

1. there is a physical world, although we may be very wrong about the nature of the physical (2010, p. 1);
2. experience exists, is part of reality and is not a mere illusion (2010, p. 52);
3. experience is realized in / realized by physical goings-on (2010, p. 1; cf. p. 105);
4. experience is not all there is to reality (there are entities as part of the physical world with non-experiential properties) (2010, p. 105);
5. experiential properties are not emergent from non-experiential properties (2010, p. 105; 2006, p.12);
6. in a fundamental sense there is only one kind of stuff (i.e. some variety of monism is true) (2010, p. 1; 2006, p. 7).

Strawson's second core assumption makes his view incompatible with *type-A materialism*. Furthermore, his fifth assumption implicitly holds that experiential properties are irreducible to physical properties, which makes his view incompatible with *type-B* and *-C materialism* since *materialism* in general relies on the reducibility of the mental to the physical (see chapter 2.2). Since the third assumption clearly ties consciousness to the physical world *experiential physicalism* is also incompatible with *type-E dualism*. Depending on whether Strawson's view allows both, substances with and without *phenomenal* properties, or just substances with *phenomenal* respectively *protophenomenal* properties (see chapter 2.2), his view can either be classified as *type-D dualism*, i.e. *property dualism* (as done by Macpherson 2006), or as *type-F monism* (as done by Chalmers (2010) in footnote 30 on p. 133).

In *Consciousness and its Place in Nature* Strawson explicitly rejects a classification of his position as *property dualism* (2006, p. 193; see chapter 2.2) and furthermore claims that

“[...] physicalism, i.e. real physicalism, entails panexperientialism or panpsychism [...] and whatever problems are raised by this fact are problems a real physicalist must face” (2006, pp. 25-26),

which leaves *type-F monism* as the only suitable category for Strawson's view.

How Strawson arrives to this astonishing claim that physicalism entails panpsychism, i.e. the view that everything that exists has some experiential properties, will be shown in the following chapter.

4.2 The justification

Strawson asks: “what is the first (or at least equal first) natural fact with which we are acquainted?” (2010, p. xviii) and gives as answer:

“Experience – and necessarily so. And experience is not only the first (or at least equal first) natural fact with which we are acquainted; it’s also the most certainly known fact” (ibid.).

By taking experience as a *fact* and equating experience with the *what-its-likeness* of a being (2010, p. 3) Strawson rejects eliminativism as solution for the hard problem and claims that acknowledging the (physical) reality of consciousness has to be the starting point for “any remotely realistic version of physicalism” (2006, p. 4).

Furthermore, Strawson states that “physics is one thing, the physical is another” (Strawson 2003a, p. 49) and therefore wants the term *physicalism* to be distinguished from *physicSalism* (cf. 2006, p. 4). The former term labels the view that everything concrete (i.e. none abstract) that exists has to be physical (i.e. not supernatural) (ibid.; cf. Strawson 2010, p. 2) while the latter is referring to the view that *physics* is able to capture the true nature or essence of all concrete reality (cf. Strawson 2006, p. 4). Given this differentiation, what can *physicSalism* tell us about the true nature of the physical that would justify a fundamental distinction between mental phenomena and physical phenomena?

Strawson reminds us that Descartes' distinction between the *physical* and the *mental* was a plausible response to the mind-body problem in a time which was dominated by classical mechanistic materialism, i.e. dominated by the view that the physical world consists entirely of “small, solid, intrinsically inert particles in motion” (Strawson 2003a, p. 66). But this view is not valid anymore: meanwhile we have discovered that atoms consist of even smaller particles, which themselves are not

particles in the classical sense, i.e. solid entities, but rather fields and repulsive forces (ibid.). However, the ultimate nature of the physical is still unknown and the notion of the physical is therefore not less problematic than the notion of the mental (cf. Strawson 2010, p. 32). By quoting Eddington (1929, p. 258) Strawson asks:

“what knowledge have we of the nature of atoms which renders it at all incongruous that they should constitute a thinking [Experiencing] object?” (Strawson 2003a, p. 71; [annotation added by Strawson]).

According to Strawson it doesn't make sense to draw a fundamental distinction between the mental and the physical, because doing so would be similar to drawing a distinction between cows and animals (cf. 2010, p. 58). Cows *are* animals – just as experience *is* something physical. Instead of using the traditional mental/physical dichotomy Strawson advocates the usage of the terms *experiential* and *non-experiential* which would allow addressing different aspects of the same physical reality. By doing so, the question whether the *mental* is reducible to the *physical* turns into the question whether the *experiential* is reducible to the *non-experiential*.

If we assumed the reducibility of the *experiential* to the *non-experiential* we would have to accept an asymmetry in the status of the *experiential* and the *non-experiential* (Strawson 2010, p. 56). Why should the former be dependent on the latter? Given our limited understanding of the true nature of the physical, what could justify this asymmetry and – maybe even more important – why should the *experiential* be reducible to the *non-experiential* and not the other way around?

If we accept an asymmetry, idealism (i.e. the claim that – to use Strawson's terminology – the *non-experiential* is at the very bottom something *experiential*) might become as plausible as reductive materialism (cf. Strawson 2010, p. 107). But apart from the fact that such an asymmetry in the status of the *experiential* and the *non-experiential* seems unjustified: how should it even be possible?

Since “emergence can’t be brute” (Strawson 2006, p. 18) an emergent property must arise in an essentially non-arbitrary, non-miraculous way:

“For any feature Y of anything that is correctly considered to be emergent from X, there must be something about X and X alone in virtue of which Y emerges, and which is sufficient for Y” (Strawson 2006, p. 18).

If we declare, for example, the liquidity of water as a property that emerges from the non-liquid properties of H₂O molecules, we are explaining a *non-experiential* property via other *non-experiential* properties, which is unproblematic. However, if we try to explain *experiential* properties via *non-experiential* properties we are comparing properties that seem to lack a common ground. In order to do so we would have to assume *protophenomenal* properties of the *non-experiential*, where the term *protophenomenal* either means:

- not actually experiential, but somehow necessary for experience, or
- already intrinsically experiential, although qualitatively different from “the experience whose realizing ground we are supposing it be” (Strawson 2006, p. 22).

In the first case the problem is not solved because the gap between the *experiential* and the *non-experiential* is still unbridged; in the second case it has to be admitted that there is no significant difference between the *experiential* and the *non-experiential* left, since there are “experiential properties all along” (Strawson 2006, p. 23), which leaves physicalism finally two options:

- admitting that at least *some* types of ultimates (i.e. fundamental building blocks of physical reality) are intrinsically experiential or
- admitting that *all* physical stuff has experiential properties.

Since option one would lead us to some kind of (substance) dualism (cf. Strawson 2006, p. 25) and would leave us wondering why there is “such a radical heteronomy at the very bottom of things” (ibid.), option two, i.e. *panpsychism*, seems simpler and therefore more plausible, which results in the previously mentioned claim that physicalism entails panexperientialism or panpsychism (cf. Strawson 2006, pp. 25-26).

4.3 The critique

Considering Strawson's argumentation from the previous chapter and his core assumptions (as listed in chapter 4.1), the claim that physicalism entails panpsychism is consistent and hence seems to be justified. Nevertheless, Strawson's conception of consciousness raises at least two serious questions:

- How can micro-experiences result in macro-experiences?
- How does experiential physicalism differ from property dualism?

Both of these questions will be addressed in this chapter.

4.3.1 The Composition Objection

Frege postulated that "an experience is impossible without an experient" (1956). Strawson is following this claim by stating that "[t]here cannot be experience without a subject of experience, because experience is necessarily *for* someone or something" (2010, p. 129; [emphasis added by Strawson]). It follows from this claim that in case of human consciousness (explained via panpsychism) a myriad of experience-involving ultimates (i.e. fundamental building blocks of physical reality, each with its own micro-experiences) come together to constitute somehow a single thing that is the subject of its own (macro-)experience (cf. Goff 2006 and 2009).

Goff (2006 and 2009) as well as Carruthers and Schechter (2006) are criticizing that it is unintelligible how and why the composition of numerous micro-experiences could add up to a single macro-experience – just as it is unintelligible how and why the composition of numerous individual subjects (an experience-involving ultimate would have to be considered as a subject) could constitute a new single subject. Via a variation of Chalmers zombie argument⁸ (cf. Chalmers 1996, p. 94) Goff concludes that panpsychism is not a satisfactory explanation of consciousness (cf. Goff 2009):

⁸ Chalmers argues that if the existence of zombies (i.e. beings that are physically identically to us but lack any form of conscious experience) are conceivable, physical facts cannot a priori entail consciousness. Consciousness therefore must be something additional to the pure physical facts.

- P1: It is conceivable that a composition of experience-involving ultimates (each with its own micro-experiences) does not result in a single macro-experience, i.e. does not result in consciousness.
- P2: If P1 is conceivable, P1 is possible.
- P3: If P1 is possible, micro-experiences do not logically entail macro-experience.
- P4: If micro-experiences do not logically entail macro-experience, panpsychism does not explain consciousness.
-

C: Panpsychism does not explain consciousness.

In order to explain macro-experience via micro-experiences, one would have to assume an additional unknown law-like relationship of micro-experiential properties that constitutes macro-experience (cf. Goff, 2009). However, such an assumption would make panpsychism not superior to ordinary physicalism (*type-B materialism*). According to Goff (2009) this would lead us to following two options:

- accepting that consciousness is the result of a law-like relationship of micro-experiential properties;
- accepting that consciousness is the result of a law-like relationship of standard physical properties (i.e. classical non-experiential properties).

In Goff's view (2009), the second option is simpler than the first option because it does not require the introduction of micro-experiential properties. Furthermore, Goff considers the second option also as more plausible because it does not stipulate that physical ultimates are conscious (*ibid.*).

I can understand why declaring physical ultimates as being conscious does not appear to Goff as very plausible, since the hypotheses of conscious ultimates is undisputably very counterintuitive. Nevertheless, I think that a reduction of macro-experiential phenomena to micro-experiential phenomena seems far less problematic than a reduction of macro-experiential phenomena to non-experiential phenomena (cf. Strawson 2006, p. 250) as supposed by *type-B materialism*. Therefore, I still consider panpsychism (as a form of *type-F monism*) more plausible than ordinary physicalism (*type-B materialism*).

4.3.2 The Property Dualism Objection

Macpherson (2006) claims that Strawson's account of consciousness is in fact just a version of *substance monism*⁹ combined with *property dualism*. While Strawson himself agrees with the label *monism* (cf. 2010, p. 1; 2003a; 2006, p. 8, 186) he explicitly declares *property dualism* as incoherent (cf. 2006, p. 28) what would suggest that he wants his own view clearly distinguished from *property dualism*.

Macpherson starts her argument by distinguishing two types of *property dualism*:

- 1) a form of dualism holding that "two different types of properties cannot exist in the same type of substance, or cannot exist in collections of the right kind that constitute a substance" (2006);
- 2) a form of dualism holding that "two different types of properties can exist in the same type of substance, or can exist in collections of the right kind that constitute a substance" (ibid).

Given these two types, Strawson's account of consciousness (consisting of a differentiation between *experiential* and *non-experiential* features of the physical in combination with the claim of an irreducible nature of the *experiential* plus a clear commitment to *monism*) seems indeed to correlate with Macpherson's second type of *property dualism*. However, from this does not follow that Strawson's view has to be considered as incoherent¹⁰.

Strawson considers *property dualism* as incoherent because he rejects the ontological distinction between the object and its properties:

"[...] one has already gone wrong if one thinks that there is any sort of ontologically weighty distinction to be drawn according to which there is the object, on the one (ontological) hand, and the properties of the object, on the other: according to which one can distinguish between the existence of the object, at any given time, and its nature, at that time" (Strawson 2003b).

⁹ A view holding that the physical and the mental are finally realized in / by the same substance.

¹⁰ Macpherson herself doesn't state that Strawson's view is incoherent, but it would be the logical consequence if Strawson were a property dualist and property dualism should turn out as incoherent.

According to Strawson (2003b), objects without properties are impossible. In his view objects are “nothing but collections or ‘bundles’ of properties — property-concretions” (ibid.). There are no objects that are in themselves entirely independent of their properties, since the latter constitute the former. The distinction between the object and its properties is just a conceptual distinction rather than an ontological one (ibid.). Therefore, Strawson has to consider the first type of Macpherson’s *property dualism* as incoherent because, according to Strawson’s view, this type of dualism is rather a *substance dualism* than a *property dualism*. However, Macpherson’s second type of *property dualism* does not show such an incoherency and therefore doesn’t pose a threat to the coherency of Strawson’s view. Nevertheless, the question that remains is: Is there a difference between Strawson’s view (panpsychism) and Macpherson’s second type of *property dualism*? – Yes, I think there is (at least) one because panpsychism requires, as indicated by the prefix *pan*, that *all* fundamental building blocks of reality have experiential properties, whereas Macpherson’s second type of *property dualism* doesn’t make this explicit claim. Furthermore, Strawson’s panpsychism declares both the *experiential* and the *non-experiential* as irreducible parts of the physical (which I, following his previously shown justification, consider as fully intelligible), while *property dualism* is based on the classical mental-physical dichotomy. Finally, even Macpherson herself states:

“Suppose one thought that panpsychism is the claim that all the fundamental constituents of reality are experiential. If, on the one hand, ‘the fundamental constituents of reality’ can refer to the fundamental properties then Strawson is not a panpsychist as, as I have been arguing, he thinks that there are both fundamental experiential and fundamental non-experiential properties. [...] If, on the other hand, ‘the fundamental constituents of reality’ only refers to fundamental objects, substances, or collections of properties of the right kind that comprise objects, then because Strawson holds that these things always involve at least one fundamental experiential property he can reasonably be classified as some type of panpsychist” (Macpherson 2006).

Considering the fact that Strawson rejects the object-property distinction by clearly stating that objects are “nothing but collections or ‘bundles’ of properties” (2003b) and considering the fact that Strawson requires *all* fundamental building blocks of reality to have experiential properties, I don’t see why Strawson’s view should be taken as a version of *property dualism* according to Macpherson’s own definition. All that follows from Macpherson’s argument is that (in order to be a panpsychist) experiential and non-experiential properties always have to co-occur which is entirely unproblematic since it is compatible with all of Strawson’s core assumptions (as listed in chapter 4.1).

5 The two Theories in Comparison

Based on the explications of Searle’s *Biological Naturalism* and Strawson’s *Experiential Physicalism* in the two previous chapters, this chapter briefly summarizes their most important similarities and differences.

5.1 Similarities

When Searle uses the term *first-person ontology* and Strawson uses the term *experience*, they are referring to the same phenomenon: the *what-it-is-like*, i.e. the phenomenal aspects of consciousness, and for Searle as well as for Strawson these aspects are part of reality and not just mere illusions.

“Conscious states, with their subjective, first-person ontology, are real phenomena in the real world” (Searle 2004, p. 113).

“For if one thing is clear, it is that experience is as real as rabbits and rocks” (Strawson 2010, p. 103).

Both philosophers naturalize phenomenal aspects of consciousness, i.e. consider them as part of the *physical* rather than the *supernatural* world, and yet insist on the irreducibility of the mental to the (ordinary) physical (Searle 2004, p. 113; Strawson 2010, p. 73). According to both theories human consciousness has to be considered as a result of biological evolution (cf. Searle 1992, p. 90; Strawson 2010, p. 36).

Furthermore, Searle (cf. 1992, p. 14) as well as Strawson (cf. 2003a, p. 66) consider the classical Cartesian mental/physical dichotomy as rather unfavorable, since it comes with a set of obfuscating preconceptions. However, the justifications for these claims and their individual consequences are quite different, resulting in an incompatibility of *biological naturalism* and *experiential physicalism*.

5.2 Differences

Although *biological naturalism* as well as *experiential physicalism* are both theories that naturalize the phenomena aspects of consciousness, they are supposing quite different roots of these phenomena. According to Searle's view consciousness is an emergent higher level feature of the brain (cf. 1992, p. 14), caused by neurobiological, hence physical, processes (cf. 1992, p. 90), and yet cannot be reduced "to its neurobiological basis, because such a third-person reduction would leave out the first-person ontology of consciousness" (cf. 2004, p. 113). Strawson, on the other hand, rejects the idea of consciousness as an emergent property of the non-experiential physical (cf. 2006, p. 18) and claims that "the experiential and non-experiential coexist in such a way that neither can be said to be based in or realized by or in any way asymmetrically dependent on the other" (cf. 2010, p. 73). Strawson supposes "experiential properties all along" (2006, p. 23) and declares the experiential as a fundamental feature of *all* ultimates, i.e. the most basic building blocks of physical reality (cf. 2006, p. 25).

Both theories are rejecting the classical mental/physical dichotomy but the resulting consequences of this rejection are quite different. In Searle's case the rejection leads to the bold attempt of declaring the mental as a mere biological feature that "is as much a part of the natural biological order as any other biological features such as photosynthesis, digestion, or mitosis" (Searle 1992, p. 90) which (as shown in chapter 3.3) is at odds with his claim of irreducibility (cf. Burton, 1995). In Strawson's case the rejection of the classical mental/physical dichotomy leads to a strong skepticism regarding our actual knowledge about the true nature of the physical (cf. Strawson

2003a, p. 71). Strawson therefore advocates a differentiation between *experiential* and *non-experiential* properties of the physical which (in combination with his understanding of emergentism) finally results in a commitment to *panpsychism* (cf. Strawson 2006, pp. 26).

Apparat from their specific content, the two theories also differ in their explanatory power. Searle's stipulation that consciousness is just a biological feature of the brain does not equate to a satisfactory explanation of the phenomenon in question (cf. McGinn 1999), because – as Searle himself admits – “the complexity of the structure itself, and the precise nature of the brain processes involved remains unanalyzed by this characterization” (Searle 2004, p. 158). Moreover, I cannot see how current physicalism, i.e. scientism, could ever be able to derive such an explanation. Maybe we might find correlations between certain brain states and certain mental states but “mere correlations cannot provide any sort of explanation or understanding of the existence of experience or what-it's-likeness or ‘qualia’ in physical-science terms” (Strawson 2010, p. 86). Strawson's account of consciousness lacks such an explanation too, but it is at least pointing us in a new direction: in order to understand consciousness we need a revolution of physics (cf. Strawson 2010, p. 92 & 102) and accepting panpsychism would clearly qualify as such.

6 Conclusion

Since the literature in the field of consciousness is quite extensive, secondary literature, focusing on thorough analyses and comparisons between existing theories, builds an important contribution to the philosophical discourse. Therefore, the aim of this dissertation lied in a comparative analysis of *biological naturalism* (Searle 1992) and *experiential physicalism* (Strawson 2010) – two positions within the philosophy of mind, interested in the phenomenal aspects of consciousness, i.e. interested in the hard problem (cf. Chalmers 2010, p. 4).

This dissertation showed that both theories try to naturalize the phenomenal aspects of consciousness, i.e. consider them as part of the *physical* rather than the *supernatural*

world, and yet insist on the irreducibility of the mental to the physical. Both theories are rejecting pure reductive materialism as well as any form of dualism as solution for the hard problem. Despite these similarities, a direct comparison of the two positions exposed fundamental differences in their core content.

Although I agree with Searle when he states that “[c]onsciousness is not like some fluid squirted out by the brain. A conscious state is rather a state that the brain is in” (Searle 2000), his conception of consciousness turned out to be inconsistent. By rejecting the existence of two different ontological realms, Searle is denying an ontological difference between the *mental* and the *physical* which makes the gap between the phenomenal subjective character of our experience and its physiological realization just an epistemic one. By explicitly calling the *first-person ontology* an *ontology*, while denying an ontological difference between the *mental* and the *physical* (cf. 2004, p. 128) and yet insisting on the irreducibility of the former to the latter, Searle’s account of consciousness becomes self-contradictory. If there is no ontological difference between the *mental* and the *physical*, Searle’s claim of irreducibility is no longer tenable, which means that the whole concept of *biological naturalism* collapses into plain *type-B materialism* (see chapter 2.2).

Strawson, on the other hand, accepts an ontological difference between the *non-experiential* and the *experiential*. If we deny the existence of the supernatural and acknowledge that experience (i.e. consciousness) is something real, consciousness has to be part of the natural world. Therefore, if the aim of physicalism lies in explaining the natural world, consciousness must be part of every realistic version of physicalism (cf. Strawson 2006, p. 4). Strawson’s conclusion that *all* ultimates, i.e. the most basic building blocks of physical reality, have experiential properties (cf. 2006, p. 25), might appear as counterintuitive but his claims represent a consistent theory.

Although I consider Strawson’s *panpsychism* superior to Searle’s self-contradicting *biological naturalism*, I doubt that *panpsychism* will help us to solve the hard problem. Strawson himself states that whatever problems are raised by the fact that physicalism entails panpsychism “are problems a real physicalist must face” (2006,

pp. 26). I see at least one problem of panpsychism that appears to me as unsolvable for physicalism: the problem of other minds, i.e. the question whether respectively how we can know about the existence of other minds.

If we suppose consciousness to be realized in the physical brain we might be justified to think that organisms with a similar brain are conscious too. However, we are not able to make such an inference – based on “same-causes-same-effects, and relevantly-similar-causes-relevantly-similar-effects” (Searle 1992, p. 22) – when we are looking at the question whether ultimates have experiential properties, since a physical ultimate can hardly be compared with a human brain.

If we have no way to check whether physical ultimates have experiential properties, panpsychism is as good in explaining consciousness as an account which stipulates that we are conscious *sola gratia* (i.e. through the grace of god). The unsolvability of the problem of other minds disqualifies panpsychism as a scientific theory, since “[t]he criterion of the scientific status of a theory is its falsifiability, or refutability, or testability” (Popper 2002, p. 48) which might turn experiential physicalism (in form of panpsychism) ultimately into a theory that has

“the faintly sickening odor of something put together in the metaphysical laboratory” (Nagel 1986, p. 49).

However, I think Strawson is right when he states that in order to understand consciousness we need a revolution of physics (cf. Strawson 2010, p. 92 & 102). It is just questionable whether such a revolution is possible since the human mind is only able to create theories according to its capacities¹¹. These capacities are the product of evolution and were formed by selective pressure to maximize the chance of successful reproduction. Unlike other (mental) capacities which finally made us the dominant species on this planet, being able to *fully* understand a person’s mind does

¹¹ „The physical atom is, like everything else in physics, a schedule of pointer readings” (Eddington 1929, p. 258), i.e. a concept, which itself is just a product of our mental capacities.

not seem necessary for a successful reproduction¹². Therefore, there is a chance that the hard problem of consciousness will remain unsolvable because the human mind simply lacks the necessary capacities to form the required concepts for a sufficient theory.

ignoramus et ignorabimus!

[du Bois-Reymond 1874]

7 Acknowledgments

I would like to thank my supervisor Dr. Paul Schweizer for accepting my thesis proposal as well as for his advice and helpful remarks throughout the course of this thesis. Furthermore, I would like to express my sincere gratitude to Ms. Helena Veum whose (short) acquaintance enabled me to gain profound insights regarding the ontological status of certain phenomenal aspects of consciousness, which definitely had an influence on the execution and the content of this thesis – probably even beyond.

¹² Based on my own experience the mere attempt to fully understand another person's mind usually turns out to be rather counterproductive from an evolutionary standpoint.

8 Bibliography

- AUGUSTINE, St. (1900): *The Confessions of St. Augustine*. [Digital Edition, <https://archive.org/details/confessionsofsta00augu> (Last accessed on 14th of June 2016)]. Garden City, 1900.
- BLOCK, N. (1998): "On a Confusion about a Function of Consciousness". In: *The Nature of Consciousness: Philosophical Debates*. pp. 375–415. MIT Press, 1998.
- BURTON, R. (1995): "Searle on Rediscovering the Mind". In: *Man and World*. Vol. 28(2), pp. 163-174, 1995.
- CARRUTHERS, P., SCHECHTER, E. (2006): "Can Panpsychism bridge the Gap?". In: *Consciousness and its Place in Nature*. Imprint Academic, 2006.
- CASTON, V. (2002): "Aristotle on Consciousness". In: *Mind*. Vol. 111(444), pp. 751-813, 2002.
- CHALMERS, D. (1996): *The Conscious Mind : In Search of a Fundamental Theory*. Oxford University Press, 1995.
- CHALMERS, D. (2010): *The Character of Consciousness*. Oxford University Press, 2010.
- CHALMERS, D. (2015): "Panpsychism and Panprotopsyism". In: *Consciousness in the Physical World: Perspectives on Russellian Monism*. Oxford University Press, 2015.
- DENNETT, D. (1991): *Consciousness explained*. Back Bay Books/Little, Brown & Co., 1991.
- DENNETT, D. (1993): "The Rediscovery of the Mind" [Review]. In: *The Journal of Philosophy*. Vol. 90(4), pp. 193-205, 1993.
- DESCARTES, R. (1644): *Renati Des-Cartes Principia Philosophiae*. [Digital Edition, <http://hdl.loc.gov/loc.rbc/Rosenwald.1431.2> (Last accessed on 8th of June 2016)]. Amstelodami, apud Ludovicum Elzevirium, 1644.
- DESCARTES, R. (2008): *Meditations on First Philosophy with Selections from the Objections and Replies* [Translated with an Introduction and Notes by Michael Moriarty]. Oxford University Press, 2008.
- DU BOIS-REYMOND, E. (1874): "The Limits of our Knowledge of Nature". In: *Popular Science Monthly*. Vol. 5, pp. 17-32, 1874.
- EDDINGTON, A. (1929): *The Nature of the Physical World*. [Digital Edition, <http://www.archive.org/details/natureofphysical00eddi> (Last accessed on 3rd of July 2016)] Macmillan, 1929.
- FREGE, G. (1956): "The Thought: A Logical Inquiry". In: *Mind*. Vol. 65(259), pp. 289-311, 1956.

- GROSS, Ch. (1995): "Aristotle on the Brain". In: *The Neuroscientist*. Vol. 1(4), pp. 245-250, 1995.
- GOFF, P. (2006): "Experiences don't sum". In: *Consciousness and its Place in Nature*. Imprint Academic, 2006.
- GOFF, P. (2009): "Why Panpsychism doesn't help us explain Consciousness". In: *dialectica*. Vol. 63(3), pp. 289-311, 2009.
- HUXLEY, Th. (1868): *The Elements of Physiology and Hygiene: A Textbook for Educational Institutions*. Appleton & Co., 1868.
- HONDERICH, T. (2001): "Mind the Guff". In: *Journal of Consciousness Studies*. Vol. 8(4), pp. 62-78, 2001.
- JORGENSEN, L. (2014): "Seventeenth-Century Theories of Consciousness". In: *The Stanford Encyclopedia of Philosophy*. [Digital Edition, <http://plato.stanford.edu/archives/win2014/entries/consciousness-17th> (Last accessed on 9th of June 2016)]. 2014.
- JACKSON, F. (1982): "Epiphenomenal Qualia". In: *The Philosophical Quarterly*. Vol. 32 (127), pp.127-136, 1982.
- KANT, I. (1902): "Versuch den Begriff der negativen Größen in die Weltweisheit einzuführen". In: *Gesammelte Schriften*. [AA II] [29 vols., publ. by Königliche Preußische Akademie der Wissenschaften zu Berlin.] Reiner / de Gruyter, 1902.
- KIM, J. (1995): "Mental Causation in Searle's 'Biological Naturalism'". In: *Philosophy and Phenomenological Research*. Vol. 55(1), pp.189-194, 1995.
- LEIBNIZ, G. W., CLARKE, S. (1717): *A Collection of Papers, which passed between the late learned Mr. Leibnitz and Dr. Clarke, in the Years 1715 and 1716*. [Digital Edition, <https://archive.org/details/acollectionpape00leibgoog> (Last accessed on 15th of June 2016)]. Crown, 1717.
- LOCKE, J. (1825): *Essay Concerning Human Understanding* [Essay]. [Digital Edition, <https://archive.org/details/humanunderstandi00lockuoft> (Last accessed on 9th of June 2016)]. Tegg, 1825.
- MCGINN, C. (1999): "Can we ever understand Consciousness". In: *The New York Review of Books*. Vol. 46(10), pp. 44-48, 1999.
- MACPHERSON, F. (2006): "Property Dualism and the Merits of Solutions to the Mind-Body Problem". In: *Consciousness and its Place in Nature* [pp. 72-89]. Imprint Academic, 2006.
- MORTON, P. (1997): *A Historical Introduction to the Philosophy of Mind: Readings With Commentary*. Broadview Press, 1997.
- NAGEL, Th. (1974): "What is it like to be a Bat?". In: *The Philosophical Review*. Vol. 83 (4), pp. 435-450. Duke University Press, 1974.

- NAGEL, Th. (1986): *The View From Nowhere*. Oxford University Press, 1986.
- NOONAN, H. (2003): *Personal Identity*. Routledge, 2003.
- NORTHOFF, G., MUSHOLT, K. (2006): "How Can Searle Avoid Property Dualism? Epistemic-Ontological Inference and Autoepistemic Limitation". In: *Philosophical Psychology*. Vol. 19(5), pp. 589-605, 2006.
- POPPER, K. (2002): *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge, 2002.
- PUTNAM, H. (1960): "Minds and Machines". In: *Dimensions of Mind*. [Digital Edition, <http://www.archive.org/details/dimensionsofmind013653mbp> (Last accessed on 30th of July 2016)]. New York University Press, 1960.
- QUINE, W. (1951): "Two Dogmas of Empiricism Symposium". In: *The Philosophical Review*. Vol. 60(1), pp. 20-43, 1951.
- ROWLANDS, R. (2001): *The Nature of Consciousness*. Cambridge University Press, 2001.
- SHAPIRO, L. (2006): "Descartes's Pineal Gland Reconsidered". In: *Midwest Studies in Philosophy*. Vol. 35(1), pp. 259-286, 2006.
- SEARLE, J. (1992): *The Rediscovery of the Mind*. MIT Press, 1992.
- SEARLE, J. (2000): "Consciousness". In: *Annual Review of Neuroscience*. Vol. 23, pp. 557-78, 2000.
- SEARLE, J. (2002): „Why I Am Not a Property Dualist". In: *Journal of Consciousness Studies*. Vol. 9(12), pp. 57-64, 2002.
- SEARLE, J. (2004): *Mind - A Brief Introduction*. Oxford University Press, 2004.
- SKINNER, B.F. (1974): *About Behaviourism*. Cape, 1974.
- SMART, J.J. (1959): "Sensations and Brain Processes". In: *Philosophical Review*. Vol. 68 (2), pp. 141-156, 1959.
- STRAWSON, G. (1999): "Realistic Materialist Monism". In: *Towards a Science of Consciousness III*. MIT Press, 1999.
- STRAWSON, G. (2003a): "Real Materialism". In: *Chomsky and his Critics*. Blackwell, 2003.
- STRAWSON, G. (2003b): "What Is the Relation Between an Experience, the Subject of the Experience, and the Content of the Experience?". In: *Philosophical Issues*, Vol. 13(1), pp. 279-315, 2003.
- STRAWSON, G. (2006): *Consciousness and its Place in Nature*. Imprint Academic, 2006.
- STRAWSON, G. (2010): *Mental Reality*. [Second Edition]. MIT Press, 2010.
- ZAHAVI, D. (2007): "Killing the Straw Man: Dennett and Phenomenology". In: *Phenomenology and the Cognitive Sciences*. Vol. 6(1), pp. 21-43, 2007.