

Orly Shenker

Denialism: What Do the so-called Consciousness Deniers Deny?

Abstract. Some philosophers consider that some of their colleagues deny that consciousness exists. We shall call the latter ‘deniers’, adopting a term that was initially meant pejoratively. What do the deniers deny? In order to answer this question, we shall examine arguments, both of some deniers and of their critics, and present denialism as a systematic highly non-trivial position that has had some interesting achievements. We will show that the denialist project concerns the epistemology of the mind and specifically of consciousness: what can be known about it, and how it can be known. The main argument of denialism is that first-person reports about the mental realm are not always the best source of information about that realm, and are certainly not reliable. This leads the deniers to realize that the reference and meaning of mental terms as used in standard philosophical literature are vague and call for clarification, and that what many take to be empirical data in the study of the mind are actually heavily laden with theory. Denialism thus makes scientific and philosophical research clearer and more fruitful.

Preface

Mark Steiner was an anti-naturalist; he believed that human beings have a special place and role in the cosmos. The important conclusion he drew from studying *The Applicability of Mathematics as a Philosophical Problem*

Acknowledgements. I am grateful to Yakir Levin, Meir Hemmo and Aya Evron for their strong objections and insightful criticisms on previous drafts, to Daniel Dennett for suggesting some further references, and to Gedalyahu Wittow and Gal Vishne for their remarks. All of them helped me clarify my thinking about denialism and realize, and hopefully demonstrate, how important this position is for understanding the mind.

(Harvard University Press, 1998) was that the universe is not indifferent to the aims and values of human beings. In the introduction Mark recently wrote for the Hebrew translation of Hume's *A Treatise of Human Nature*, of which he was scientific editor, he wrote that, according to Hume, passion rather than reason is the basis of scientific knowledge and ethics, and argued that in this way Hume inadvertently brought back to center stage the element he had made huge efforts to remove from metaphysics: the self. An idea that underlies Mark's anti-naturalist philosophy, either explicitly or implicitly, is that people can have something useful, even something true, to say about their own beliefs, aims, values, and the other facets of their mental reality. That assumption is challenged by *denialism*, which is the view I characterize and discuss in this paper. Ever since my earliest days as Mark's student our debates concerning our radically dissenting views were a constant source of inspiration, and his inimitable humor always added a special flavor to his sharp objections and analyses. This paper continues this exchange.

1. Introduction

The operation of the mind, conscious and¹ unconscious, free and unfree, in perception, action, and thought, in feeling, emotions, reflection, and memory, and in all its other features, is not so much an aspect of our lives, but in a sense, it is our life. (Searle 2004, 6)

While to many this sounds a truism, some people think that others deny it. For example, Galen Strawson, in a paper titled "The Consciousness Deniers," writes:

What is the silliest claim ever made? The competition is fierce, but I think the answer is easy. Some people have denied the existence of consciousness: conscious experience, the subjective character of experience, the 'whatitislike' of experience. Next to this denial – I'll call it 'the Denial' – every known religious belief is only a little less sensible than the belief that grass is green. (Strawson 2018)

What exactly is denied by endorsers of the Denial? Why do critics of the Denial bother to discuss it, and yet treat it with unusual terms like "silly" (Strawson 2018) or an "intellectual pathology" (Searle 1997, chap. 5)? This

¹ "And" here is probably not meant as the logical conjunction but rather as the exclusive disjunction.

view, so harshly attacked, is endorsed by prominent philosophers: the two I study here are Daniel Dennett and Paul Churchland. I suspect that the strong attack is a hint that deeply entrenched intuitions are being threatened by the Denial. Such a case is worthy of philosophical investigation. What are the challenged intuitions, and more importantly: what are the arguments on both sides?

Those accused of denying consciousness deny that they deny it. Dennett (2018) writes: “I don’t deny the existence of consciousness; of course, consciousness exists; it just isn’t what most people think it is.” So *why* does Strawson insist that Dennett denies the existence of consciousness? What exactly is the debate about? In particular, what exactly do those accused of denying consciousness actually deny? And what do they propose instead? The debate between the so-called deniers and their critics is not about definitions, nor terminology: it is about ontology and epistemology. Dennett writes about this debate: “One of us is dead wrong, and the stakes are high” (Dennett 1995).

I begin (in *section 2*) by describing the position that I call ‘denialism,’ and distinguishing it from other positions, especially ‘illusionism,’ which was discussed in recent literature (see Frankish, ed. 2016). Denialism is often too quickly dismissed as being either incoherent or empirically inadequate, on the basis of unidentified intuitions rather than serious argument (as I point out very briefly in *section 2.1*). I will explain (throughout the paper but especially in *section 2.2*) why I think certain thinkers (main examples are Dennett and Churchland) are denialists, and why ascribing this position to them fits their overall views and enhances their contribution to science and philosophy. For denialists, even more than for their critics, it is crucial to clarify the meaning and reference of mental terms, and some of them developed an argumentative schema in order to carry this out (this topic is discussed in *section 3*). Additionally, they have pointed out that in order to make some progress in philosophy as well as science, it is important to improve our understanding of what, in discussion of the mental realm, is taken as data and what is understood as theory, for which alternatives can be offered (this topic is discussed in *section 4*). Finally (in *section 5*), to illustrate the fruitfulness of the distinctions I make in this paper, I examine a case study: a recent scientific theory by Michael Graziano, in which there are denialist as well as illusionist elements, and show that distinguishing between them is helpful in understanding his theory.

2. Denialism versus Illusionism

Denialism is a term that I coin in this paper,² in order to refer to the following simple but non-trivial idea: *first-person reports about the mental realm are not reliable*.³ As is sometimes done in other contexts, I take a term that was meant to be pejorative and adopt it to denote a view that I take to be worth considering. Denialism is greatly misunderstood in the literature. Here is an example of such misunderstanding:

One of the strangest things the Deniers say is that although it seems that there is conscious experience, there isn't really any conscious experience: the seeming is, in fact, an illusion. (Strawson 2018)

Indeed, that “there isn't really any conscious experience” may be a “strange” and “silly” idea: but it is not the view of “the Deniers” that Strawson attacks. Their main claim is the one mentioned above, and what Strawson describes is not even one of its implications. My first task in this paper, which I undertake in the present section, is to *clarify* the denialist position. Let me stress that I do not *defend* in this paper any of the denialist theories that I present; my aim is to point out that there *is* such a position, that it is endorsed with good reasons by some prominent thinkers, that it is *significantly* different from some other positions mistakenly ascribed to these thinkers (e.g., illusionism), and that clarifying this position and correcting this mistake is important because the denialist position is worthy of consideration in both philosophy and science.

Since the term ‘denialism’ (with the above characterization) is coined here, and the position is not systematically studied as yet in the philosophical or scientific literature, it is helpful to begin presenting it by distinguishing it from other approaches. The distinctions that I now discuss are summarized in table 1. The table distinguishes between two questions, and this distinction clarifies the difference between the two approaches of *denialism* and

² Terminological disambiguation: the term “science denialism” has been used in the literature to denote forms of pseudo-science that are driven by a desire to fight down some scientific theory or branch of science. Examples are climate change denialism, HIV denialism, and vaccination denialism. This term is used for example in Slater et al. (2018) and Diethelm and McKee (2009).

³ Chirumuuta (2014) and Bronfman, Jacobson, and Usher (2019) are examples of discussing recent methodological arguments concerning the degree to which first-person reports provide reliable or conclusive data in experimental psychology.

illusionism, that are sometimes conflated (e.g., in Dennett 2016).⁴ When presented in this way it is hard to see how one might conflate between them, but vagueness sometimes has this consequence, and my main goal here is one of clarification. One question concerns metaphysics: What is the mental reality? Are its features as described in the contents of first-person reports? Another question is one of epistemology: How can we know the features of our mental realm? For example, are first-person reports reliable evidence for this? Two families of approaches that differ on the metaphysical question are *realism* and *illusionism*; and two families of approaches that differ on the epistemological question are *reliabilism* and *denialism*. Let me discuss briefly what these four families of approaches are.

The distinction between realism and illusionism is this.

Theories of consciousness typically address the hard problem. They accept that phenomenal consciousness is real and aim to explain how it comes to exist. There is, however, another approach, which holds that phenomenal consciousness is an illusion and aims to explain why it *seems* to exist. ... Illusionists agree that we have introspective subjectivity, though they hold that it is radically misleading. ... [They] deny that phenomenal properties exist in the real world, as properties of brain states. (Frankish 2016a, 14, 38)

Realists and illusionists about phenomenal consciousness *disagree* about the nature of reality: realists think that “phenomenal consciousness is real,” whereas illusionists say it isn’t; it only “seems to exist.” What do the realists and illusionists *agree* on? According to Frankish, both *agree* that “it is a datum that phenomenal properties exist as intentional objects ..., that we have introspective subjectivity” (Frankish 2016a, section 3.1).

According to the distinction I presented above, Frankish thinks that realists and illusionists agree on the epistemological question: they are both *reliabilists*. They agree (on the epistemological question, namely,) that we are reliable reporters concerning our mental realm, and they disagree (on the metaphysical question, namely,) on what gives rise to this mentality or what explains it. But according to the distinction between the metaphysical question and the epistemological one these are two distinct questions, so that in principle realists as well as illusionists are not bound to agree on the epistemology: in particular, they can opt for either reliabilism or denialism (see table 1, p. 313 below).

⁴ Dennett also conflates denialism with “Fictionalism,” see Dennett 2020.

When it comes to explaining knowledge of our own thoughts the only three options are that this knowledge is:

1. Based on inference or reasoning.
2. Based on inner observation.
3. Based on nothing. (Cassam 2014, chap. 10)⁵

Reliabilists can opt for *any* of these options to support the claim that first-person reports on the mental realm are reliable. Of course, the degree of support for the reliability claim can differ between these options, since they differ with respect to the gap between the knower and the known. Whereas Cassam opts for option 1 and, being aware of this gap, acknowledges that we can be mistaken about our beliefs,⁶ Strawson – in the context of attacking denialism – makes a very strong claim in this respect:

Suppose you're hypnotized to feel intense pain. Someone may say that you're not really in pain, that the pain is illusory, because you haven't really suffered any bodily damage. But to seem to feel pain is to be in pain. It's not possible here to open up a gap between appearance and reality, between what is and what seems. (Strawson 2018)

Strawson's point is epistemic, not metaphysical; he attacks denialism, not illusionism (see table 1). If one endorses reliabilism, as for example Strawson does, one may still opt for either realism – according to which the best explanation for how things feel is that the content of the feeling is true about the world; or for illusionism – which rejects this explanation, and says that the best explanation for the way that things feel is that the content of this feeling is not true about the world, for example: the properties ascribed to the mental realm are properties of matter (possibly of the brain).⁷ It seems to me that Strawson's metaphysical theory of pan-psychism is *realist*, and that metaphysics could turn out to be part of future science (as Strawson emphasizes). But although Strawson is an example of a *reliabilist & realist*, this nature of his metaphysics does not follow from his preference of extreme reliabilist epistemology over denialism. This is easily seen by considering the alternative, to which I now turn.

⁵ Cassam (2014) distinguishes between the “inward-directed” question “Do I believe that P?” and the “outward-directed” question “Is it the case that P?”

⁶ Cassam (2014) focuses on beliefs. Our discussion in this paper concerns all aspects of the mental realm. I return to intentional states in secs. 3.2 and 4.2 when discussing Churchland's version of denialism.

⁷ It is not always clear whether one is a realist or an illusionist in this respect. For example, approaches that endorse multiple realizability are usually taken to be physicalist, but are shown to be dualist in Shenker 2017, Hemmo and Shenker 2020.

		<p>Metaphysics What is the mental reality? Are its features as described in first-person reports?</p>	
		<p>Realism. Things are as described in first-person reports. E.g., if the features of mental reality seem non-physical, then this is the case.</p>	<p>Illusionism. Things are not necessarily as described in first-person reports. E.g., even though the features of mental reality may seem non-physical, the facts may be physical.</p>
<p>Epistemology. How can we know the features of our mental realm? In particular, are first-person reports a reliable source about it?</p>	<p>Reliabilism. First-person reports are reliable evidence concerning the mental realm. E.g., Introspection is a form of observation, and reports about it reliably reveal the true contents of the mental realm.</p>	<p>Position 1 Reliabilist Realism. The true matters of fact about the mental realm are <i>reflected</i> in the contents of first-person reports, and they <i>correspond</i> to the ontology.</p>	<p>Position 2 Reliabilist Illusionism. The true matters of fact about the mental realm are <i>reflected</i> in the contents of first-person reports, but they <i>do not correspond</i> to the ontology.</p>
	<p>Denialism. First-person reports are not reliable evidence concerning the mental realm. E.g., Introspection is not observation, or its content is not correctly reported. The true nature of mental states and processes is best revealed by third-person observations.</p>	<p>Position 3 Denialist Realism. The true matters of fact about the mental realm are <i>different</i> from the contents of first-person reports. The <i>real</i> mental experience corresponds to the ontology.</p>	<p>Position 4 Denialist Illusionism. The true matters of fact about the mental realm are <i>different</i> from the contents of first-person reports. The <i>real</i> mental experience <i>does not</i> correspond to the ontology.</p>

Table 1

Frankish's view is one example of a *reliabilist & illusionist*.⁸ An example of a *scientific* implication of the combination of reliabilism and illusionism is this. It is well known that we report having a psychological arrow of time and feeling a qualitative difference between past and future. This is manifested, for example, in the difference between memories and hopes. According to reliabilism, this is indeed how we *feel*. It is also well known that contemporary fundamental physics is time symmetric, so that the above difference cannot even be translated to physical terms. One way to explain the psychological arrow is to reduce it to a non-temporal asymmetry in the brain, e.g., the right-left asymmetry. In this case we take the report on the feeling of the psychological arrow as reliable, but we explain it as an illusion. (See Hemmo and Shenker 2019c.)

We have seen examples of the combinations of reliabilism & realism and reliabilism & illusionism. I bring examples of denialism in the next section.

Finally, perhaps a way to present the difference between illusionism and denialism, and the difference between their significance in the philosophy of mind, is via the 'meta-problem of consciousness' recently introduced by Chalmers as follows: "The meta-problem of consciousness is (to a first approximation) the problem of explaining why we think that there is a problem of consciousness" (2018, 6; see also Frankish 2019).

Regarding the meta-problem, illusionists might say as follows. (a) People report that their mental realm has certain features, namely, those that give rise to the hard problem of consciousness. (b) These reports are reliable, that is, they are evidence that those things do happen in the reporters' mental reality. (c) The reason why there is this mental reality is that the reporters are physically wired in a certain way.

⁸ Frankish distinguishes between "strong illusionism" and "weak illusionism," explaining that "Illusionism makes a very strong claim: it claims that phenomenal consciousness is illusory; experiences do not really have qualitative, 'what-it's-like' properties, whether physical or non-physical. This should be distinguished from a weaker view according to which some of the supposed *features* of phenomenal consciousness are illusory" (Frankish 2016a, sec. 1.2; see also Chalmers 2018, sec. 7). To understand his position it is helpful to consider the examples that Frankish brings in 2016a, sec. 1.3. I think denialism is also *not* Frankish's (2016a) "quasi phenomenal properties," nor Frankish's (2016a) "eliminativism" (which is not Churchland's 1996 eliminativism). But in the end, some of this may be a matter of terminology, and I am willing to accept any name that refers to essentially the same position.

Denialists endorse (a) but reject (b). According to denialists, the report of having a mental reality with certain features is not evidence that these elements indeed obtain in the mental reality of the reporters. The reports can be mistaken and should not be taken at face value. Instead of (b) it might be that something else happens in reality, call it (b'), which is different from (b). On the denialist view, the explanation of (a) by (b') and the corresponding (c') will allegedly be better than explaining (a) by (b) and (c).

Finally, as I said, I am not going to argue in the paper for or against denialism in any depth. Let me just mention very briefly some main arguments, just to recall the context of discussion, since this may be useful for the following sections.

2.1 Some Objections to Denialism

Denialism strikes many people as an absurdity. Searle (1997, 112) thinks that it is “a form of intellectual pathology,” and Strawson (2018) thinks that it is “the silliest claim ever made.” Dennett (1995) writes in return that “the feeling is mutual.” But what are the arguments for rejecting denialism? One kind of objection is that this position is *incoherent*, and is illustrated by Strawson’s following argument: “One of the strangest things the Deniers say is that although it seems that there is conscious experience, there isn’t really any conscious experience: the seeming is, in fact, an illusion. The trouble with this is that any such illusion is already and necessarily an actual instance of the thing said to be an illusion” (Strawson 2018).

This line of argument might strike some people as reminiscent of the famous Cartesian line of thinking for the *cogito ergo sum* conclusion. Notice, however, first, that Strawson’s claims notwithstanding, the denialists do not make a metaphysical point, but an epistemological one. The denialist claim is not a denial (nor the affirmation) of the existence of anything; rather, the point is epistemological-skeptical, concerning the reliability of first-person reports concerning the nature of the mental realm. Denialists may differ considerably on metaphysics, for example: they may be either realists or illusionists concerning the question of whether the actual mental reality (as opposed to the reported one) reflects that nature of reality, including the reality of what gives rise to the actual mental reality. Second, while there may be differences between Descartes and any of the denialists concerning what the features of the mental realm are and what the reasons are for ascribing

those features to it, they seem to be closer to Descartes's skeptical spirit than some of their critics.

Others agree that denialism is coherent, but think it is empirically false and, moreover, that its falsity is "obvious" (e.g., Chalmers 2018, sec. 7). As usual, when claims about the "obvious" are made, philosophical investigation is called for. What is the argument? Chalmers (2018, sec. 7) mentions "a simple Moorean argument, reminiscent of Moore's pointing to his hands to demonstrate that there is an external world."

Presumably, Moore's argument, which is about the external world, is brought *mutatis mutandis*; and the objections to Moore's argument apply, *mutatis mutandis* as well. But since such arguments are not brought as part of a general common-sense philosophy,⁹ it seems that they are simply based on strong intuitions. Philosophers often turn to intuitions in their arguments: Sometimes those are given the status of conjectures that are subject to further examination, and at other times they are seen as some sort of evidence for the truth of their contents (see overview in Pust 2019). Kripke wrote: "I think having intuitive content is very heavy evidence in favor of anything, myself. I really don't know, in a way, what more conclusive evidence one can have about anything, ultimately speaking" (Kripke 1980, 42).¹⁰

Suppose that intuitions are the final arbiter. Still the questions arise: *which* intuitions should we accept? And which should we take as brute facts or as supporting brute facts, and which call for explanation?¹¹ Is finding them

⁹ See recent papers on common-sense philosophy in de Ridder, van Woudenberg, and Peels 2020.

¹⁰ Ayer seems to disagree, when he writes that "Philosophers who fill their books with assertions that they intuitively 'know' this or that moral or religious 'truth' are merely providing material for the psycho-analyst" (Ayer 1936, chap. 6). I leave open here the question of whether the intuitions that Kripke talks about are "moral or religious" (in Ayer's sense) or other. Ayer writes this remark in a chapter dedicated to the critique of ethics and theology from the perspective of his logical-positivist worldview. In Ayer's view one can take certain sense data to be 'mistaken' only relative to other sense data, and hence "anyone who condemns the sensible world as a world of mere appearance, as opposed to reality, is saying something which, according to our criterion of significance, is literally nonsensical" (1936, chap. 1). This line of argument seems to me quite close to Kripke's (but I will not expand on this point).

¹¹ See more on what calls for explanation in Dan Baras, *Calling for Explanation* (ms. in preparation). In my view, Fodor's 1974, 1997 acceptance of multiple realizability is a case of unacceptable mysterianism (see Shenker 2017, Hemmo and Shenker 2019b and 2020).

the job of “experimental philosophy” (Knobe and Nichols 2017)? Or should only the intuitions of some sort of experts in ‘inner study’ be considered (see, e.g., Varela 1996)? I leave these questions open. My point of concern in this is to identify some intuitions that are in play, whether those of the denialists or of their critics.

In this paper I do not address the heavy arguments that could be mounted for or against denialism, which involve conceptual, epistemological, and metaphysical issues concerning varieties of skepticism. The point is merely that intuitions need to be identified, clarified, and then seriously discussed.

2.2. Are there Denialists?

Denialism is held by some philosophers, although not by many. Of course, this is not a matter to be decided by majority vote, but by arguments. So in characterizing certain philosophers as denialists, and distinguishing between denialism and other positions, I am not interested in “just definitions defended” (Dennett 2016), nor about intellectual biographies. As I said above, denialism is based on the idea that self-knowledge is based on some form of inference or reasoning. With respect to inferentialism about intentional states, Cassam writes:

Inferentialism about intentional self-knowledge, the view that knowledge of our own beliefs, desires, hopes, and other ‘intentional’ states is first and foremost a form of inferential knowledge ... hasn’t exactly been a popular approach to self-knowledge in recent years. Many currently influential discussions of self-knowledge start out with a statement to the effect that intentional self-knowledge is normally ‘immediate’, that is, non-inferential. (1) Philosophers who defend inferentialism – Ryle is usually mentioned in this context – are then berated for defending a patently absurd view. (2) The assumption that intentional self-knowledge is normally immediate ... is rarely defended; it’s just seen as obviously correct. In contrast, I think that the immediacy premise is not obviously correct, and that inferentialism is a live option. (Cassam 2014, chap. 11)

I am not sure whether or not Cassam is a denialist and I leave this question open here.¹² One prominent example of a contemporary denialist, who focuses

¹² On the one hand he writes that “self-knowledge is fallible” (Cassam 2014, 189), but on the other hand he “builds on the idea that self-knowledge is cognitively substantial by giving examples of how human beings come to know their own attitudes by inference.” Moreover, his inferentialism is only about beliefs, and the knowledge claims about those are based on taking other statements about the mental realm as

on the same kind of mental states that Cassam addresses, is Churchland, who writes explicitly about his theory of eliminative materialism (discussed later in this paper):

Eliminative materialism is the thesis that our common-sense conception of psychological phenomena constitutes a radically false theory, a theory so fundamentally defective that both the principles and the ontology of that theory will eventually be displaced, rather than smoothly reduced, by completed neuroscience. Our mutual understanding and even our introspection may then be reconstituted within the conceptual framework of completed neuroscience, a theory we may expect to be more powerful by far than the common-sense psychology it displaces, and more substantially integrated within physical science generally. (Churchland 1981, 67)

Churchland's focus is on intentional states (I expand on this below in sections 3.2 and 4.2). As is well known (following Quine's 1960 *Word and Object*), it is hard to naturalize intentionality by making this notion compatible with contemporary physics (see overview in Jacob 2019). Churchland gives priority to physics, and assumes that the same difficulty will persist with respect to future physics. His conclusion is that intentionality cannot be real, despite the impression given by first-person reports; a clear case of denialism. (In this paper I do not criticize nor defend Churchland's argument for this idea: my aim is to present denialism, characterize it, and show some of its implications; and Churchland is an example of a philosopher who holds it.)

I find it interesting to think of Churchland's view as endorsing a version of the *denialism & realism* combination, position 3 in table 1, in the following sense. He denies that the contents of first-person reports – which are the core of what he calls “common-sense psychology” – are reliable evidence of the true matters of fact in the mental realm, and conjectures that in the future, when we shall have a theory of neuroscience that will provide a complete and true account of the mental realm, people will not only realize the merits of this theory, but will endorse it in such a way that their own first-person *experiences* will change in a way that will reflect the true matters of fact concerning their mental realm. So, he is a denialist with respect to *contemporary* first-person reports, but is a realist concerning the *conjectured* contents of *future* first-person reports. (But other interpretations

evidence. “Beliefs and desires aren't feelings but what you feel can sometimes tell you what you believe or desire.”

of Churchland's metaphysical views are possible, as long as they are compatible with his denialism.)

Another prominent denialist is Dennett. Above I mentioned his saying "I don't deny the existence of consciousness; of course, consciousness exists; it just isn't what most people think it is" (Dennett 2016). Similarly, concerning freedom of will, Dennett insists that it really exists, it just is not what people intuitively think it is (see, for example, Dennett 2003a, 2020). He extends this general idea to intentional states as well (in his 2020) and I will discuss below in section 3 his denialism concerning qualia. And here I will present a fourth example of the philosophical and scientific consequences of Dennett's denialism.

Dennett and Kinsbourne (1992) describe the well-known 'phi' phenomenon: If a row of nearby small spots (call them 1, 2, 3, 4 for example) are lit one after another slowly, the observers *report* experiencing them one after the other; but if they are lit in rapid succession, observers *report* seeing a moving spot – that is, they experience the spot appearing also in the area between the actual spots. It is further known that if, when all four spots are lit one after the other and the subjects *report* having seen a spot between, say, 3 and 4, and then spot 4 is removed and only a succession of 1, 2, and 3 is activated, the subjects *report* that they no longer see a spot in the position between 3 and 4: the spot 'stops' at 3. Dennett and Kinsbourne (1992) express denialism, when they conjecture that – the abovementioned reports notwithstanding – the subjects did not have the reported experience of a spot between 3 and 4: the reported spot appeared *only in the report*, provided after the visual information has already arrived at the relevant brain region. In this sense the first-person report concerning one's own mental perception is flawed, and the experimenter who activated the lighted spots, and can check the observer's brain, etc., is in a better position than the reporter to know that it is indeed flawed. Let me emphasize that whether or not the conjecture of Dennett and Kinsbourne is true is a question of fact, and the challenge is to find a way to test their factual conjecture. The point that Dennett and Kinsbourne make in "Time and the Observer" (1992) is precisely that whatever the method for testing this fact is, it cannot be just taking the reports on face value. (They illustrate this point by their distinction between the Orwellian and Stalinesque scenarios.) So epistemologically, Dennett and Kinbourne (1992) express denialism.

Dennett and Kinsbourne, and Dennett in other arguments for denialism, emphasize the gap between the mental reality and the first-person reports about it, between the report and the reported, due to which, in general, one is not in a better position to report one's own mental states than the position of others who investigate those states:

anyone who wants to appeal to private, subjective properties has to prove first that in so doing they are not making a mistake. ... Unless he seeks outside help, the state of his own qualia must be as unknowable to him as the state of anyone else's qualia: hardly the privileged access or immediate acquaintance or direct apprehension the friends of qualia had supposed 'phenomenal features' to enjoy!... Indeed, a subject's 'introspective' convictions will generally be worse evidence than what outside observers can gather. (Dennett 1988, 382, 389, and 395)

Dennett is aware of the proximity between his denialist views and those of Churchland:

[Churchland and myself] agree that there is no good motivation for shoehorning these folk [psychology] categories into neuroscientific pigeonholes via a strict type identity theory, and even a strict functionalism would require some Procrustean labors that might better be postponed indefinitely, since the domain on the left hand side of the equation – the folk categories – is composed of items that are just not up to the task. (Dennett 2005, 193)

The first-person reports on which folk psychology is based are so mistaken that they are not even candidates for physicalist reduction. Does this support a realist metaphysics or an illusionist one? More generally, on the metaphysical front, is Dennett's view realist or illusionist? In Dennett's case the question is a bit tricky (one might even say, ill posed), since according to Dennett's multiple drafts theory there wasn't any conscious mental state prior to the trigger that elicited the state in which the observer mis-reports a past mental state. However Dennett recently described his own views as "illusionist" (Dennett 2016, in a volume on illusionism), or as "fictionalist" (Dennett 2020, in a volume on fictionalism). Frankish (2016a, 2016b), himself a reliabilist illusionist (I think), writes that Dennett is an illusionist, and Dennett (2016) says that Frankish's analysis of his own views is flawless. Dennett writes: "Illusionism, I am saying, should not be seen as a lame attempt to deny the obvious, but as the leading contender, the default view that should be assumed true until proven otherwise."

As I said above, it is a bit tricky to classify Dennett's metaphysics and its relation to his account of first-person reports. Dennett is emphatically not

a reliabilist: so *if* he is an illusionist, then he is a denialist one, endorsing position 4 in table 1. (Let me however remark that Dennett endorses a version of computational functionalism, which allows for multiple realizability of the mental by the physical. In my view all approaches that allow for multiple realization are forms of dualism, not of physicalism. If that is the case then whether Dennett's position is 4 or 3 depends on the details of his dualist view. The very possibility that Dennett may be holding position 3 rather than 4 strongly supports the need for the distinctions made in table 1.)

Finally, let us recall here – just to be clear about what we say here and what is said in other places – that Dennett has been *mistakenly* 'accused' of consciousness denial, which is *not* denialism in our sense. Searle writes:

[Dennett] thinks there are no such things as ... the feeling of pain. He thinks there are no such things as qualia, subjective experiences, first-person phenomena, or any of the rest of it. Dennett agrees that it *seems to us* that there are things as qualia, but this is a matter of a mistaken judgement we are making about what really happens. ... The main point of Dennett's book is to deny the existence of inner mental states and offer an alternative account of consciousness, or rather what *he* calls "consciousness." The net effect is a performance of *Hamlet* without the Prince of Denmark. (Searle 1997, 99–100)

Searle is wrong: Dennett does not deny consciousness in this sense, but is only a denialist in the sense presented here, that is, he denies that first-person reports concerning the mental realm are reliable. (Searle's misunderstanding of Dennett's view is similar to Strawson's misunderstanding mentioned above. Dennett addresses similar mistaken accusations concerning his conception of freedom in 2003a, chap. 8.)

A third thinker who may be seen as a denialist is Michael Graziano. I will describe his views later in the paper, using them in order to illustrate some of the features of the denialist view that I will present in section 5.

As I said, I will not address in this paper arguments for denialism or against it. Denialism is not addressed in the literature as such, and I think the main arguments that can be understood as criticizing denialism are much too quick and turn to highly non-trivial ideas as if they were obvious. Here my aim is only to present this position and describe some of its implications.

3. Denialism on the Reference or Meaning of Mental Terms

Denying the reliability of first-person reports about the mental realm does not entail that all the contents of such reports are false; rather, it requires that we

take nothing for granted, and that we provide arguments for accepting some parts of these reports as evidence and rejecting others. The need to make this distinction, which is part of denialism but not of reliabilism, forces denialists to be very attentive to what precisely they talk about, that is, about the reference of terms that appear in first-person reports or in their philosophical analyses, and about the distinction between data and theory. Needless to say, these are topics to which reliabilists should pay attention as well, but for denialists they are more crucial and therefore they received special attention from some denialists. In this section I describe why clarifying the reference or meaning of mental terms is crucial for denialist arguments, and what role this clarification plays in them. The distinction between theory and data is addressed in the next section.

Remark: I shall not argue for any theory of reference or meaning nor about their relations, therefore I use the phrase ‘reference or meaning’ generally, and readers are invited to see whether my arguments apply to their favorite theories.¹³

3.1 *Dennett on the Reference or Meaning of Mental Terms*

‘Qualia’ is an unfamiliar term for something that could not be more familiar to each of us: the ways things seem to us. (Dennett 1988, 381).

But what *is* it? Block writes:

You ask: What is it that philosophers have called qualitative states? I answer, only half in jest: As Louis Armstrong said when asked what jazz is, ‘If you got to ask, you ain’t never gonna get to know’. (Block 1978, 281)

Nobody is born knowing what jazz is (not even Armstrong), and yet people (like Armstrong) become able to distinguish between paradigmatic cases of jazz and paradigmatic cases of non-jazz.¹⁴ In this sense Armstrong was to a

¹³ Quine and Davidson (for example) put forward (different) theses saying that there is no matter of fact concerning the “absolute” reference or meaning of a given expression or (correspondingly) of a given belief, and instead reference and meaning are relative to a framework. What we said above concerning the fixing of reference or meaning of terms like “qualia” or “conscious experience” should be understood within one’s favorite theory of reference or meaning, and may be fixed relative to a suitable framework.

¹⁴ Maybe ‘jazz’ is a somewhat open term, although arguably not every new case is admitted to the disjunction (even if Armstrong declares it to be a case of ‘jazz’). Maybe

large extent *wrong*, and the analogy that Block draws is misleading. But the gist of Block's saying is prevalent: some people think that you don't need to be taught what qualitative states are or, more generally, what conscious experience is; you are born knowing it. And this is a mistake: *nobody is born knowing what the terms 'qualitative states' or 'conscious experience' refer to or what they mean (not even Block)*. Strawson expresses the erroneous view even more explicitly:

If someone asks what conscious experience is, you say, 'You know what it is from your own case'. (You can add, 'Here's an example', and give them a sharp kick.) When it comes to conscious experience, there's a rock-bottom sense in which we're fully acquainted with it just in having it. *The having is the knowing*. So when people say that consciousness is a mystery, they're wrong – because we know what it is. It's the most familiar thing there is – *however hard it is to put into words*. (Strawson 2018; final italics added)

Suppose that one hears the phrase “conscious experience” for the first time (and there always is a first time about this) and asks what it means, or what it refers to. Strawson's response (as well as Block's) is of no help at all. What precisely (the novice might ask) about “your own case” is “conscious experience”? The having of *what* is the knowing of it? When filling-in the blank in “your own case of ___” or “The having of ___ is the knowing of it,” why is “conscious experience” better than ‘money’ or ‘jabberwocky’? The term that fills the blank should be a suitable one: for example, it should be something ‘to be had’ by a suitable owner, and should give rise to ‘knowledge’ by a suitable knower. The line of thinking of Strawson and Block is, then, not very useful to either science or philosophy.

How are we, then, to answer questions like “what conscious experience is” or “what qualitative states are”? Dennett illustrates how to carry out this project by addressing ‘qualia’; he is interested in this clarification as part of his denialist project: denying that first-person reports, which purport to be about qualia, are reliable evidence concerning the mental realm. He proposes: “You have to break up the given, and the taking of the given, into more modest parts whose operation we can actually begin to understand” (Dennett 2005, 194).

‘jazz’ is a disjunction of jazz1, jazz2, etc., each of which has a different reference or meaning, but importantly this is not ‘multiple realization’ since ‘jazz’ does not exist beyond the convention of calling all these pieces ‘jazz’ (that might be explained by the history of music, for example).

Dennett begins this project by analyzing arguments in the standard philosophical literature, in which it is claimed that the mental realm contains somethings called ‘qualia’. Dennett finds, possibly following a literature survey, that many take ‘qualia’ to be somethings with the following properties: being

- (1) ineffable
- (2) intrinsic
- (3) private
- (4) directly or immediately apprehensible in consciousness. (Dennett 1988, 385)

Let me stress that I (and possibly Dennett) *don't* insist that this is a correct description of the literature: the reader may object to the list's adequacy, on any grounds. In that case we shall *replace* the above four features by others, but the outline of Dennett's argument – which is the main point here – will apply equally to the new list. To be general, and to avoid unnecessary debates concerning the details of this list, let us call the list of features that characterize ‘qualia’ (whatever those features might be) Q1...Qn.

But a feature of the list Q1...Qn which *is* significant for our discussion of denialism is that, even if the specification of Q1...Qn involves conceptual (or other a priori) analysis, it *also* involves empirical claims about the features of the human mind. And according to standard reliabilism, the evidence supporting these empirical claims is the contents of first-person reports.

The schema of Dennett's argument that I now present combines partial arguments that Dennett brings explicitly in several places and implicitly in others. For example, “Quining Qualia” (1988) is more about (a), (b), and (c), whereas his discussions of free will in (2003a) has also elements (d), (e), and (f).

- (a) Claim: People report having qualia that are characterized as things that have features Q1...Qn. (This is learned from standard literature.)
- (b) Claim: Nothing has features Q1...Qn. (This claim is based on conceptual analysis, thought experiments, intuition pumps, experimental results).
- (c) From (a) and (b): If qualia are things that have features Q1...Qn, then qualia do not exist.
- (d) Claim: There are things that have features C1...Cn, which differ from Q1...Qn. (This is based on conceptual analysis, thought experiments, intuition pumps, experimental results).

- (e) Claim: $C_1 \dots C_n$ explain or predict (etc.) the same data that ‘qualia’ are supposed to explain or predict (etc.). (The notion of ‘data’ in this context is discussed in the next section.)¹⁵
- (f) From (d) and (e): If we choose to call the thing that has features $C_1 \dots C_n$ ‘ C ’, then C exists. (I discuss different options for the name C below.)

Concerning (a), (b), and (c), he writes:

My claim, then, is not just that the various technical or theoretical concepts of qualia are vague or equivocal, but that the source concept, the ‘theoretical’ notion of which the former are to be presumed refinements, is so thoroughly confused that, even if we undertook to salvage some ‘lowest common denominator’ from the theoreticians’ proposals, any acceptable version would have to be so radically unlike the ill-formed notions that are commonly appealed to that it would be tactically obtuse – not to say Pickwickian – to cling to the term. Far better, tactically, to declare that there simply are no qualia at all ... ‘qualia’ is a philosophers’ term which ... refers in the end to no properties or features at all. (Dennett 1988, 382 and 387)

Since the standard claim that there are qualia in the sense of the thing that has features $Q_1 \dots Q_n$ is – as I stressed – an empirical claim based on first-person reports, and not merely a conceptual exercise, the first three elements of the argument form a case of *denialism*.

As I said, Dennett is a denialist not only about qualia, but also about free will, the self, and other elements of the mental realm (he mentions this generalization in his recent 2020). He puts forward similar arguments in all of these cases, although these are not always easy to distill. In my view, one of the more important contributions of these discussions is that they bring to our attention the need to say more clearly what we talk about when we use terms like ‘qualia’, ‘conscious experience’, and the like, if we hope to make philosophical, let alone scientific, progress about it.

¹⁵ Assuming that we have some theory about $C_1 \dots C_n$ (and Dennett does have such a theory), since some people (especially philosophers) report having $Q_1 \dots Q_n$, it turns out that having $C_1 \dots C_n$ disposes those people to report that they have $Q_1 \dots Q_n$, and to insist that their reports are reliable about it. $C_1 \dots C_n$ are perhaps, in this sense, what Frankish (2012) calls *zero qualia*: “Zero qualia: The properties of experiences that dispose us to judge that experiences have introspectable qualitative properties that are intrinsic, ineffable, and subjective” (Frankish 2012, sec. 1). But Frankish (2012, sec. 1) thinks that “My putative zombie twin judges that its experiences have classic qualia, even though it has no phenomenal consciousness at all.” I disagree: Frankish’s putative zombie twin will report having whatever that is, without having it.

Statements (d), (e), and (f) are the idea behind Dennett's saying, another version of which I quoted in the introduction above:

Dennett's Alternative: X's are perfectly real; they just aren't what you think they are! (Dennett 2020)

Statement (f) has been the origin of some confusion and misunderstanding concerning Dennett's arguments. When applying this schema for the case of free will (in 2003), to make his point, Dennett insists that the thing that has features C1...Cn ought to be called 'free will', although people usually use this term to refer to whatever has features Q1...Qn; and one of his reasons for this is the claim that nothing has features Q1...Qn, so the term is empty unless we fill it in with features C1...Cn. To clarify the argument one (including Dennett) should be aware that if a given term X is understood as referring to whatever has features X1...Xn, one should examine whether anything has these features, if not – what features do exist, and if those are Y1...Yn, then decide whether to still use the term X or replace it with the term Y, so as to be clear about the meaning and reference and avoid confusion.

Dennett is aware of this kind of confusion (see his recent 2020) but unfortunately does not offer the simple clarification offered above. In this, although his motivation is one of clarification, he added to the confusion.

3.2 Churchland on the Reference or Meaning of Mental Terms

Churchland holds views that are very close to those of Dennett on this matter. While Dennett closely examines what the reference or meaning of qualia might be, Churchland is concerned with 'intentionality' statements: he is not interested in the question of whether there is a natural relation called "intentionality" between mental states and the external world, but rather, in whether these mental states, in which one feels as if such a relation obtains, exist at all (on the former problem see Jacob 2019).

Churchland is a denialist, and his argument has the same schematic structure as Dennett's. For him, despite the fact that first-person reports state that our mental realm has features A1...An that we call 'intentional states', the third-person evidence gives us good reason to suspect that nothing has features A1...An: intentional states do not exist. And the same holds for all the terms that folk psychology talks about, and in this sense it is false. Churchland conjectures that future science will give us good reasons to think

that features $B_1 \dots B_n$, which are different from $A_1 \dots A_n$, exist, and explain the same observations that $A_1 \dots A_n$ were meant to explain. Churchland conjectures that $B_1 \dots B_n$ will be so radically different from $A_1 \dots A_n$, that the two sets of features will not even be related by any kind of reduction. Moreover, according to his conjecture, once the new theory is in place and its explanatory and predictive merits will be realized, no one in their right mind will continue to use the old theory and it will disappear, to become a relic with only historical interest. People will internalize the new theory and will describe themselves in its terms – in which case the first-person reports will become true.

4. Denialism on the Theory Ladenness of First-Person Reports

As we saw, Dennett and Churchland deny that there is something (called in philosophical literature qualia or conscious experience or propositional attitude) that has features $Q_1 \dots Q_n$. Instead, they say, what exists is something else, which has features $C_1 \dots C_n$. Their critics object. What are the criteria for deciding whether or not there is something that has features $Q_1 \dots Q_n$? In addition to seeing this debate through the prism of the problem of fixing the reference it may be helpful to view it as also concerning the following question: What are the data? In philosophy of science it is well known that the distinction between theory and data is not clear, since propositions that allegedly describe data are in general theory laden. We shall see that there is a debate between the denialists and their critics concerning the question of what is to be taken as data, even if the general problem of the theory ladenness of data is acknowledged by all. Of course, once something is characterized as theory rather than data, the door is open to proposing alternative theories, and another question arises, namely, what is the best theory to explain the data; but I shall not expand on this last question here.

4.1 Dennett on Data and Theory

Here is a prevalent view:

Consciousness is not an explanatory construct, postulated to help explain behavior or events in the world. Rather, it is a brute explanandum, a phenomenon in its own right that is in need of explanation. ... To take the line that explaining our judgments about consciousness is enough (just as explaining our judgments about God is enough) is

most naturally understood as an eliminativist position about consciousness (as one analogously takes an eliminativist position about God). As such it suffers from all the problems that eliminativism naturally faces. In particular, it denies the evidence of our own experience. This is the sort of thing that can only be done by a philosopher – or by someone else tying themselves in intellectual knots. Our experiences of red do not go away upon making such a denial. It is still like something to be us, and that is still something that needs explanation. (Chalmers 1996, 188)

Phenomenal properties are not theoretical posits introduced to explain other data, but are themselves core data. (Frankish 2016a, sec. 3.1)

Where consciousness is concerned the existence of the appearance is the reality. If it seems to me exactly as if I am having conscious experiences, then I am having conscious experiences. (Searle 1997, 112)

Chalmers, Searle, and Frankish are all reliabilists: they agree that the data about the mental realm are, first and foremost, the contents of first-person reports. They only disagree about the way in which these data are to be explained: for example, Chalmers opts for metaphysical dualism whereas Frankish opts for a version of physicalism.¹⁶ Dennett, as a denialist, disagrees with all three of them about what the data are. For him, the contents of first-person reports are *not* reliable data concerning the subject's mental states and processes. The data, according to the denialists that we examine here, are third-person observations. Dennett describes his methodology, which goes under the name of 'heterophenomenology', thus:

Heterophenomenology is a scientific methodology for gathering and interpreting data on human consciousness, described by Dennett. ... Dennett defends the hypothesis that there is a straightforward, conservative extension of objective science that handsomely covers all the ground of human consciousness, doing justice to all the data without ever having to abandon the rules and constraints of the methods that have worked so well in the rest of science. This third-person methodology, heterophenomenology (phenomenology of another not oneself), is, he claims, the sound way to take the first-person point of view as seriously as it can legitimately be taken. ... That is the core of heterophenomenology: it exploits our capacity to perform and interpret speech acts, yielding a catalogue of what the subject believes to be true about his or her conscious experience. This catalogue of beliefs fleshes out the subject's *heterophenomenological world*, the world according to S, the subjective world of one subject. The total set of details of heterophenomenology, plus all the data we can gather about concurrent events in the brains of subjects and in the surrounding environment, comprise the total data set a theory of human consciousness must explain. Dennett claims that it leaves out no objective phenomena and no subjective

¹⁶ I think, as many do, that Searle is a dualist, but he disagrees, see Searle (2002).

phenomena of consciousness. The interpretation required to turn data about speech sounds and button pressings into reports and expressions of beliefs involves adopting what Dennett calls the *intentional stance* (Dennett 1987): adopting the working hypothesis that the subject is an agent whose actions are rationally guided by beliefs and desires that are themselves rational, given the subject's perceptual history and needs. (Dennett 2009)

Since, according to Dennett, what his critics claim to be data is in fact theory, the door is open to proposing alternative theories, and the question of which is the best theory remains to be decided according to the criteria for theory choice:

For Searle, it is all really quite simple. There are these bedrock, time-tested intuitions we all have about consciousness, and any theory that challenges them is just preposterous. I, on the contrary, think that the persistent problem of consciousness is going to remain a mystery until we find some such dead obvious intuition and show that, in spite of first appearances, it is false! (Dennett 1995)

It is not something too obvious to need support, however appealing it may at first seem. It is a first shaky step toward a philosophical theory. I offer a rival theory. (Dennett 2018)

It is well known in the philosophy of science that, in general, observation statements are theory laden, and this applies of course also to the kind of observations that Dennett accepts as data. However, Dennett is in a better position on this front than his critics. The reason is that in the critics' view we have direct and unmediated access to our experience, that is, they think that the data obtained this way are not theory laden. Dennett, by contrast, points out that this access is indirect and mediated, despite intuitions on this matter (see Cassam's three options, in section 1 above). For example, in "Time and the Observer" they write:

Although this is a 'theoretical' paper, it is addressed especially to those who think, mistakenly, that they have no theories and no need for theories. We show how uncontroversial facts about the spatial and temporal properties of information-bearing events in the brain require us to abandon a family of entrenched intuitions about 'the stream of consciousness' and its relation to events occurring in the brain. (Dennett and Kinsbourne 1992, 183)

Of course, once a theory of the mental is endorsed, we remain with the general problem of scientific realism; that is, with the question of whether we take the theoretical entities and laws to be real. This, in Dennett's words, is the question about "whether beliefs and other mental states are, shall we say, as real as electrons or centers of gravity" (Dennett 1991, 30).

What kind of things are beliefs? We may stay maximally non-committal about this – pending the confirmation of theory – by treating beliefs and their contents or objects as theories, fictions, or abstractions, similar to centers of mass, the Equator, and parallelograms of forces. (Dennett 2009)

4.2 *Churchland on Data and Theory*

Churchland, too,¹⁷ thinks that what are often considered data are heavily theory laden. Specifically, while many people take the contents of the folk-psychological statements to be data (and hence as true), in his view these are elements of a theory, and moreover, this theory is not a good one: it is false. He writes:

The semantics of the terms in our familiar mentalistic vocabulary is to be understood in the same manner as the semantics of theoretical terms generally: the meaning of any theoretical term is fixed or constituted by the network of laws in which it figures. (Churchland 1996, 69)

And to emphasize that folk psychology is a theory and not a collection of data Churchland adds: “A Martian could justly ascribe to us the familiar run of mental states, even though his own psychology were very different from ours. He would not, therefore, be generalizing from his own case.” His skepticism concerning folk psychology – which is based first and foremost on the contents of first-person reports – is a *consequence* of Churchland’s denialism. He writes:

Introspective judgments about one’s own case turn out not to have any special status or integrity anyway. On the present view, an introspective judgment is just an instance of an acquired habit of conceptual response to one’s internal states, and the integrity of any particular response is always contingent on the integrity of the acquired conceptual framework (theory) in which the response is framed. Accordingly, one’s introspective certainty that one’s mind is the seat of beliefs and desires may be as badly misplaced as was the classical man’s visual certainty that the star-flecked sphere of the heavens turns daily. (1996, 70)

¹⁷ Churchland and Dennett disagree on important aspects of their views; see, for example, Churchland 1996, 77. Dennett (1991, 50–51) thinks that folk psychology is not such a bad theory, since it enables us to make not too bad predictions. Its ontology is not worse than the ontology described in terms of the physical constituents, which would be perhaps more accurate, but less useful and more energy consuming. It is all a matter of tradeoff between theoretical merits: ease of use and immunity from error.

By contrast, third-person reports concerning data such as behavior or brain states and processes are to be given the status of data; those may be as theory laden as all the data used in science.

4.3 *Data vs. Theory: A Remark on the Indeterminacy of Reference Theses*

To see the far-reaching implications of the disagreement between reliabilists and denialists, it may be helpful to consider, briefly, another context in which a similar debate arises. Searle (2002) criticizes Quine's thesis concerning the indeterminacy of reference or meaning, and its implications for the philosophy of mind. Whereas Quine aims to illustrate his thesis using the famous example that there is underdetermination of radical translation of 'gavagai' between, for example, 'a rabbit' and 'a stage in the life of a rabbit', Searle takes this very example to be a *reductio ad absurdum* argument against that thesis. Searle thinks that there is a universal fact concerning the meaning of 'a rabbit' and 'a stage in the life of a rabbit', while Quine denies that such a universal fact obtains. For Searle, the alleged universal fact is one concerning data obtained by introspection:

If [Quine's] argument is valid, then it must have the result that there isn't any difference *for me* between *meaning* rabbit or rabbit stage, and that has the further result that there isn't any difference for me between *referring to* a rabbit and referring to a rabbit stage, and there isn't any difference for me between something's being a *rabbit* and its *being* a rabbit stage. ... [But] ... we know from our own case that we do mean by 'rabbit' something different from 'rabbit stage'. ... If somebody has a theory according to which there isn't any difference between my meaning rabbit and my meaning rabbit part, then I know that his theory is simply mistaken. (Searle 2002, 233–34, 229–30)

There are various perspectives from which to discuss this matter, and here I focus on one that is relevant for the subject of this paper. Searle, as we saw above, believes that what he thinks are his mental states and processes is indeed true; for him, reports on introspections should be taken as evidence for the truth of their contents. He seems to conclude that if one person (call her J) reports her state of mind (concerning rabbit, etc.), then another person (call him W) may, upon reading these reports, arrive at the same state of mind. Quine denies this in the following sense: for him, as a principle we can never know whether or not W has obtained a state of mind identical to that of J, and in this sense there is no fact concerning whether or not W has obtained such a state of mind. No amount of reports of W about his state of

mind will suffice for that, including the case where W provides the same reports provided by J; Searle adds: “We know from our own case, from the first-person case, that behaviorism is wrong, because we know that our own mental phenomena are not equivalent to dispositions to behavior” (Searle 2002, 240).

I think he misses the point, because he takes the identity of reports by J and W as implying that their content is everything that there is. For Quine, what Searle takes to be the definite data is only one theory, and there is unavoidable underdetermination between all the theories on this matter.

5. Another Case Study: Graziano’s Attention Schema

In this section I will try to apply the lessons learned from analyzing the views of Dennett and Churchland in order to understand a recent theory of consciousness, Michael Graziano’s ‘attention schema’ which is, extremely briefly, as follows (see overview in Graziano 2019). Fundamentally, a human being is nothing but a physical system that interacts with its environment. In order for the human system to interact with only those features of its environment that are useful for its survival, the internal states of the human brain represent only those salient features of the environment, as well as the aspects of the human body that take part in interacting with the environment. This focusing on certain features of the environment, and ignoring the rest, is enabled by the mechanism of ‘attention’ (as defined by Graziano), which magnifies the information from certain features of the environment and the body, and reduces others. The outcome of the focus on certain features of the environment and the body and the ignoring of others is a simplified model, a caricature, of the environment and the body. In addition to this ‘overt attention’ – representation of salient features of the environment and body – there is ‘covert attention’; representation of some processes of the overt attention, possibly in a way that is not immediately reflected in behavior. The overt attention results in simplified models of the environment and body, and the covert attention results in simplified models of the internal states and processes of the overt attention. Finally, in humans (and possibly other creatures) there is yet a higher level of covert attention, which is directed at covert attention. This level of attention and abstraction is ‘consciousness’. Here, since the object of attention is covert attention itself, it is not directly explicitly given in physical terms, and does not explicitly

have a physical object, such as the environment or the body. Thus, when reporting such covert attention, the system (in particular, the person) may find itself having to use terminology in which the reference is not directly to physical reality, and so it may seem to be non-physical. This is a rough outline of Graziano's proposal (for more details see Graziano 2016 and Graziano 2019 and references therein).

Let us examine the nature of this theory in terms of the positions developed above (summarized in table 1). As I now show, his view has elements from each of these positions, and understanding this may contribute to further understanding of this matter. (Since the conceptual and terminological distinction between denialism and illusionism, as asking different questions, was not available to Graziano, he uses the latter term where the former is more appropriate, see, e.g., Graziano 2019, chap. 7, and 2016, sec. 4.)

At the initial stages of his argument Graziano appears to be a *reliabilist*, but later explanations bring out his denialism. Let's begin with the expressions of reliabilism. Graziano begins with the *working hypothesis* (that he later re-examines) that the reported feeling that our consciousness is "physically weightless and substanceless" and is "an ethereal essence and an inherent property dwelling inside you" (2019, chap. 7) are true statements about our mental experience. In terms of section 4, these are statements that the features of the mental are Q1...Qn, and if we take these reports to be reliable, the task of brain science is to explain how Q1...Qn come about. This, at this initial stage of Graziano's argument, is the task of the 'attention schema' account. This is an expression of *reliabilism*: the attention schema is a physical account of Q1...Qn, that is, of the features of the mental as given in first-person reports. And since these mental features are explained by physics, he endorses here position 2 in table 1: *reliabilist illusionism*.

But later in his argument Graziano returns to examine his reliabilist assumption and rejects it. Going deeper into the implications of his theory, he expresses a *denialist* view, and claims that the contents of the first-person reports are not reliable, after all: in the terms of Section 4 above, the features of the mental are not Q1...Qn, but C1...Cn – first- person reports notwithstanding. He writes: "We *don't* understand consciousness as being physically substanceless. Instead, we understand it as something for which physical attributes are *irrelevant*. And those two intuitions are very different" (2019, chap. 7).

And so, after all, our first-person reports in which people say that they have a feeling of something ethereal, etc., are not to be taken at face value. At this point in his argument it turns out that his task is to explain why we make the false reports: in the terms of section 4 above, why our physics is such that having $C1...Cn$ makes us prone to report, mistakenly, that we have $Q1...Qn$. This seems to me to be an expression of *denialist illusionism*, position 4 in table 1.

We see that the double distinction, between denialism and reliabilism on the epistemological front and between realism and illusionism on the metaphysical front, enables us to understand Graziano's claims and arguments.

Professor Orly Shenker
The Department of Philosophy
The Hebrew University of Jerusalem
orly.shenker@mail.huji.ac.il

Bibliography

- Ayer, A. J. 1936. 1971. *Language, Truth and Logic*. London: Penguin.
- Balog, Katalin. 2016. "Illusionism's Discontent." In Frankish, ed. 2016, 40–51.
- Baras, Dan and Orly Shenker. 2020. "Calling for Explanation: The Case of the Thermodynamic Past State." *European Journal for Philosophy of Science* 10:1–20. <https://doi.org/10.1007/s13194-020-00297-7>
- Bermudez, Jose Luis. 2005. "Arguing for Eliminativism." In *Paul Churchland*, ed. Brian L. Keeley, 32–65. Cambridge University Press.
- Blackmore, Susan. 2016. "Delusions of Consciousness." In Frankish, ed. 2016, 52–64.
- Block, Ned. 1978. "Troubles with Functionalism." *Minnesota Studies in the Philosophy of Science* 9:261–325.
- Boghossian, Paul. 1990. "The Status of Content." *The Philosophical Review* 99:157–84.
- Bronfman, Zohar, Hilla Jacobson, Marius Usher. 2019. "Impoverished or Rich Consciousness outside Attentional Focus: Recent Data Tip the Balance for Overflow." *Mind & Language* 34:423–44.
- Cassam, Quassim. 2014. *Self-knowledge for Humans*. Oxford University Press.
- Chalmers, David. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.
- Chalmers, David. 2018. "The Meta-Problem of Consciousness." *Journal of Consciousness Studies* 25:6–61.
- Chirimuuta, Mazviita. 2014. "Psychophysical Methods and the Evasion of Introspection." *Philosophy of Science* 81(5): 914–926.

- Churchland, Paul. 1981. "Eliminative Materialism and the Propositional Attitudes." *Journal of Philosophy* 78:67–90.
- Churchland, Paul. 1996. *The Engine of Reason, the Seat of the Soul*. Cambridge, Ma.: Bradford and MIT Press.
- Dennett, Daniel. 1988. "Quining Qualia." In *Consciousness in Contemporary Science*. Ed. A. J. Marcel and E. Bisiach. Oxford U.P.
- Dennett, Daniel. 1991. "Real Patterns." *Journal of Philosophy* 88:27–51.
- Dennett, Daniel. 1995. "The Mystery of Consciousness: An Exchange." *The New York Review of Books*, December 21, 1995 Issue.
- Dennett, Daniel. 2003a. *Freedom Evolves*. London: Penguin.
- Dennett, Daniel. 2003b. "Explaining the 'Magic' of Consciousness." *Journal of Cultural and Evolutionary Psychology* 1:7–19.
- Dennett, Daniel. 2005. "Two Steps Closer to Consciousness." In *Paul Churchland*, ed. Brian L. Keeley, 193–225. Cambridge University Press.
- Dennett, Daniel. 2009. "Heterophenomenology." In *The Oxford Companion to Consciousness*, ed. Tim Bayne, Axel Cleeremans, and Patrick Wilken. Oxford University Press.
- Dennett, Daniel 2015 "Why and How Does Consciousness Seem the Way it Seems?" In *Open MIND*. Ed. T. Metzinger & J. Windt. Frankfurt am Main: MIND Group. doi: 10.15502/9783958570245
- Dennett, Daniel. 2016. "Illusionism as the Obvious Default Theory of Consciousness." In Frankish, ed. 2016, 65–72.
- Dennett, Daniel. 2018. "Magic, Illusions, and Zombies: An Exchange." *The New York Review of Books*, March 13, 2018 issue.
- Dennett, Daniel. 2020. "Am I a Fictionalist?" In *Mental Fictionalism*, ed. Adam Toon. Routledge, forthcoming.
- Dennett, Daniel and Marcel Kinsbourne. 1992. "Time and the Observer." *Behavioral and Brain Sciences* 15:183–247.
- de Ridder, Jeroen, René van Woudenberg, and Rik Peels, eds. 2020. *Cambridge Companion to Common Sense Philosophy*. Cambridge University Press. Forthcoming.
- Diethelm, Pascal and Martin McKee 2009. "Denialism: What Is It and How Should Scientists Respond?" *European Journal of Public Health* 19:2–4. doi:10.1093/eurpub/ckn139
- Firt, Erez, Meir Hemmo, and Orly Shenker. 2020. "Extending Hempel's Dilemma." Manuscript in preparation.
- Fodor, Jerry. 1974. "Special Sciences: Or the Disunity of Science as a Working Hypothesis." *Synthese* 28:97–115.
- Fodor, Jerry. 1997. "Special Sciences: Still Autonomous After All These Years." *Nous* 31:149–163.
- Frankish, Keith. 2012. "Quining Diet Qualia." *Consciousness and Cognition* 21:667–676.

- Frankish, Keith, ed. 2016. *Illusionism*. Special issue of *Journal of Consciousness Studies* 23(11–12). Reprinted as a volume by Imprint Academic in 2017.
- Frankish, Keith. 2016a. “Illusionism as a Theory of Consciousness.” In Frankish, ed. 2016, 11–39.
- Frankish, Keith. 2016b. “Introduction.” In Frankish, ed. 2016.
- Frankish, Keith. 2016c. “Not Disillusioned: Replies to Commentators.” In Frankish, ed. 2016, 256–289.
- Frankish, Keith. 2019. “The Meta-Problem is The Problem of Consciousness.” *Journal of Consciousness Studies* 26(9–10): 83–94.
- Graziano, Michael. 2016. “Consciousness Engineered.” In Frankish, ed. 2016, 98–115.
- Graziano, Michael. 2019. *Rethinking Consciousness: A Scientific Theory of Subjective Experience*. New York: Norton.
- Hemmo, Meir and Orly Shenker. 2015. “The Emergence of Macroscopic Regularity.” In *Mind and Society* 14:221–244. <https://doi.org/10.1007/s11299-015-0176-x>
- Hemmo, Meir and Orly Shenker. 2019. “The Physics of Implementing Logic: Landauer’s Principle and the Multiple-computations Theorem.” *Studies in History and Philosophy of Modern Physics* 68:90–105. <https://doi.org/10.1016/j.shpsb.2019.07.001>
- Hemmo, Meir and Orly Shenker. 2019b. “Two Kinds of High-Level Probability.” *The Monist* 102(4): 458–477.
- Hemmo, Meir and Orly Shenker. 2019c. “The Second Law of Thermodynamics and the Psychological Arrow of Time.” *British Journal for Philosophy of Science*. <https://doi.org/10.1093/bjps/axz038>
- Hemmo, Meir and Orly Shenker. 2020. “Is the Mind in the Brain in Contemporary Neuroscience?” Manuscript in preparation.
- Jacob, Pierre. 2019. “Intentionality.” *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), ed. Edward N. Zalta. URL = <<https://plato.stanford.edu/archives/win2019/entries/intentionality/>>.
- Knobe, Joshua and Shaun Nichols. 2017. “Experimental Philosophy.” *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), ed. Edward N. Zalta. URL = <<https://plato.stanford.edu/archives/win2017/entries/experimental-philosophy/>>.
- Kripke, Saul. 1980. *Naming and Necessity*. Oxford, Basil Blackwell.
- Kripke, Saul. 2019. Lecture at <https://www.youtube.com/watch?v=3zazonG6zBk>
- Lycan, William and George Pappas. 1972. “What is Eliminative Materialism?” *Australasian Journal of Philosophy* 50:149–159.
- Moore G. E. 1939. “Proof of an External World.” *Proceedings of the British Academy* 25:273–300.
- Neurath, Otto. 1932–33. “Protocol Statements.” In Robert S. Cohen and Marie Neurath, eds. *Otto Neurath: Philosophical Papers* 1913–1946. Springer.
- Papineau, David. 2002. *Thinking about Consciousness*. Oxford: Clarendon Press.

- Pust, Joel. 2019. "Intuition." *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition), ed. Edward N. Zalta. URL = <https://plato.stanford.edu/archives/sum2019/entries/intuition/>.
- Searle, John. 1997. *The Mystery of Consciousness*. New York: A New York Review Book.
- Searle, John. 2002. "Why I Am Not a Property Dualist." *Journal of Consciousness Studies* 912:57–64.
- Searle, John. 2004. *Mind: A Brief Introduction*. Oxford University Press.
- Shenker, Orly. 2017. "Flat Physicalism: Some Consequences." *Iyyun: The Jerusalem Philosophical Quarterly* 66:211–225.
- Shenker, Orly. 2020. "Common sense and scientific realism." In de Ridder, van Woudenberg, and Peels 2020. Forthcoming.
- Slater, Matthew H., Joanna K. Huxster, Julia E. Bresticker, Victor LoPiccolo. 2018. "Denialism as Applied Skepticism: Philosophical and Empirical Considerations," *Erkenntnis* <https://doi.org/10.1007/s10670-018-0054-0>
- Strawson, Galen. 2018 "The Silliest Claim." In *Things That Bother Me*. New York: A New York Review Book.
- Taylor, Richard. 1992. *Metaphysics*. Englewood Cliffs: Prentice Hall.
- Varela, Francisco. 1996. "Neurophenomenology: A Methodological Remedy for the Hard Problem." *Journal of Consciousness Studies* 3(4): 330–49.

