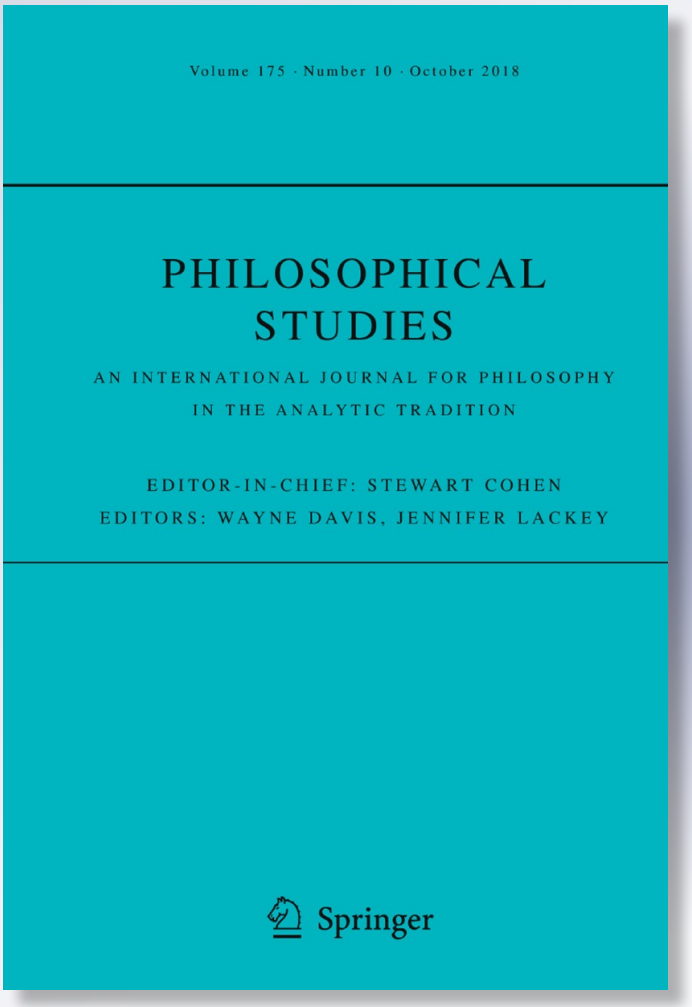
Experiencing organisms: from mineness to subject of experience

Tobias Schlicht

Philosophical Studies
An International Journal for Philosophy in the Analytic Tradition

ISSN 0031-8116
Volume 175
Number 10

Philos Stud (2018) 175:2447-2474
DOI 10.1007/s11098-017-0968-4



Springer

Philos Stud (2018) 175:2447–2474
https://doi.org/10.1007/s11098-017-0968-4

Experiencing organisms: from mineness to subject of experience

Tobias Schlicht[1]

Published online: 21 August 2017

Abstract Many philosophers hold that phenomenally conscious experiences involve a sense of mineness, since experiences like pain or hunger are immediately presented as mine. What can be said about this mineness, and does acceptance of this feature commit us to the existence of a subject or self? If yes, how should we characterize this subject? This paper considers the possibility that, (1) to the extent that we accept this feature, it provides us with a minimal notion of a subject of experience, and that (2) the *phenomenological* subject of experience, as it is represented in conscious experience, is the organism. While many philosophers agree that the *metaphysical* subject of experience is the animal, this claim is much less widespread, maybe even counterintuitive. The argument for this claim alludes to the structure of phenomenal consciousness and to recent work in cognitive science concerning the embodied character of consciousness and cognition. To illustrate the problems of current controversies, not only several recent rejections of a subject of experience are critically discussed, but also Hume's famous rejection of a subject is criticized making use of epistemological aspects from Kant's philosophy of mind. The final section situates the present discussion in the context of recently popular predictive coding accounts of perception and perceptual experience.

Keywords Consciousness · Subjective character · Self · Embodied cognitive science · Kant · Hume · Predictive coding

✉ Tobias Schlicht
tobias.schlicht@rub.de;
http://www.tobiasschlicht.com

[1] Institut für Philosophie II, Ruhr-Universität Bochum, Universitätsstr. 150, GA3/29, 44780 Bochum, Germany

Phenomenal consciousness is typically characterized in terms of 'what it is like' to be in a given mental state (Nagel 1974). For example, there is something that it is like for me to see something red, and this differs from the way it feels like for me to have a toothache. Phenomenally conscious experiences differ with respect to their quality.[1] Philosophers from different camps have attempted to explain conscious experiences solely by explaining this feature. For example, representationalists like Tye (1995) provide an explanation for the quality of seeing red in terms of surface features of the object perceived, and an explanation for the quality of feeling a toothache in terms of features of the tooth. Tye alludes to the alleged transparency of experience to support the notion that his explanation is exhaustive and that there is no remaining explanatory gap. Similarly, proponents of sensorimotor theories of consciousness like Noë (2004) and O'Regan and Noë (2001) provide an explanation for the quality of seeing red in terms of laws governing the contingencies between the sensorimotor capacities of the experiencing agent and the environment; such laws could be different with respect to the different senses. Again, the claim is that this provides an exhaustive explanation of consciousness. Other philosophers, not being satisfied with these ways of characterizing and explaining phenomenal consciousness, provide a more differentiated description. While seeing something red and suffering from a toothache may *differ* in qualitative character, these experiences *share* an important feature, namely, that they are both *subjectively* experienced by me. Sharing this feature means that the experiences present themselves as *mine* from my own (first-person) perspective. Thus, Kriegel (2009) distinguishes between the 'qualitative character' and the 'subjective character' of a conscious experience, where the latter is supposed to be an invariant aspect across different experiences.

This way of cashing out Nagel's phrase seems close to what Nagel himself had in mind when he writes: "an organism has conscious mental states if and only if there is something 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, 436). Nagel did not use the term 'qualia' at all—which was already familiar at the time (see Crane 2000). Neither was he interested in the *differences* among individual experiences like sensations of red and toothaches. Rather, he was concerned with the contrast between the objectivity of science and the subjectivity of consciousness since "facts about what it is like to be a human being… appear to be facts that embody a particular point of view" (Nagel 1974, 441), whereas scientific investigations of a phenomenon usually make progress by *ignoring* the subjective differences

[1] Some philosophers prefer to use the notion of "qualia" in this context, but since this term is more problematic than useful it will be ignored in this paper. See Dennett (1991) for a criticism of the coherence of this notion in contrast to the acceptance of the subjective impression that it *feels* like something to experience a sound or taste or smell. While there may be widespread disagreement about how to explain this subjective feeling, most philosophers of mind at least accept that there is something to be explained with respect to phenomenal consciousness. This minimal agreement of the majority is the starting point of this paper—Work on this paper was funded by the Volkswagen Foundation for the research project "Situated cognition. Perceiving the world and understanding other minds". I am grateful for valuable discussions of earlier drafts with Krys Dolega, Luke Roelofs, and Rob Rupert, and for the helpful and constructive comments by an anonymous reviewer.

associated with points of view: "If the subjective character of experience is fully comprehensible only from one point of view, then any shift to greater objectivity—that is, less attachment to a specific viewpoint—does not take us nearer to the real nature of the phenomenon: it takes us farther away from it" (Nagel 1974, 444f). Even though qualitative and subjective character may usually be present or absent together, it is useful to draw the distinction if only because some theories may provide good explanations for one of them but not for the other. Thus, for the purposes of this paper, I take it as a feature of *my* phenomenally conscious experiences that they present themselves as *mine*.

Other terms for subjective character are 'mineness' (Zahavi 2005) or 'me-ishness' (Block 1997).[2] Like subjective character, mineness applies to all *my* conscious experiences in the same way. There may be other aspects of consciousness (such as its unity or cognitive access to information etc.) but mineness clearly pertains to phenomenal consciousness. This means, in accordance with Nagel's formulation, that all *my* phenomenally conscious experiences present themselves *as mine* or are something *for me*. Among the central questions arising for all those who accept that conscious experiences have this feature are: With respect to experiences we may ask, what are the conditions that any given representation has to meet in order to be experienced as *mine*? And with respect to mineness we may ask whether accepting this feature commits us to the postulation of a phenomenal 'subject of experience' or even a 'self'? And if yes, how should we characterize this subject?

In this paper I address these questions, arguing that although accepting mineness commits us to the postulation of a subject of experience, this notion can be characterized in a way that is unproblematic for a naturalistic approach to consciousness. In a nutshell, the best way to conceive of the subject of experience is by identifying it with the organism as a whole. While this claim may sound plausible as far as the *metaphysical* subject is concerned, it is much less intuitively plausible as far as the *phenomenal* subject is concerned, i.e. the subject as it appears in experience. Yet, this is the claim that this paper aims to defend. We proceed as follows: First, some recent (and traditional) criticisms of the notion of a subject of experience (identified as a 'minimal self' by many authors) will be discussed in Sect. 1, which leads—in Sect. 2—to a critical discussion of Hume's famous incapacity to "find" an entity in his stream of consciousness to which the notion of "subject" could refer. Relying on important points from Kant's philosophy of mind it will be objected that this "Humean intuition", as it is often called, need not be taken too seriously as an argument against either the mineness of experience or the notion of a subject of experience. Then, Sect. 3 presents a case for the claim that *if* we accept the mineness of experience, *then* we are also committed to positing a subject of experience. This case rests on considerations about the structure of consciousness. It is well known that for various philosophical reasons, we have to

[2] The history of philosophy is filled with variations on this theme and with multiple ways of cashing out the notion of mineness, ranging from Kant's treatment of the unity of apperception to Sartre's pre-reflective self-awareness. Although we will return to Kant, this is not the place to reconsider all accounts from the history of philosophy in great detail.

distinguish between what Wittgenstein (1958) called the object-use and the subject-use of 'I', or between a consciousness of self-as-subject and a consciousness of self-as-object in order to capture the complex double life that the subject of experience plays in experience. Merleau-Ponty (1964) put the distinction in slightly different terms with respect to the double life of the body, distinguishing between the body as object and the body as experienced. This has important ramifications for a discussion of the sense in which the subject is represented in experience. On the basis of a proposal for what it is for a mental representation to be a *conscious* representation and of recent work on the embodied character of consciousness and cognition, it will then be argued that the most parsimonious way to think about the subject is in terms of the whole organism. This way of thinking about the subject of experience is both philosophically sound and empirically informed. Section 4 relates the results of this discussion to the recently popular framework of predictive coding and situates the subject (in both its roles) within such an account.

1 The cartesian curse: shying away from the subject

There is a clear division between those philosophers who think that consciousness has an egological structure and those who don't. The latter typically seem to think that acknowledging such a structure commits us to the notion of a substantial self or subject which then in turn leads to unnecessary problems for naturalism. Of course, accepting a notion of self along Cartesian lines, according to which it is a non-extended thinking thing that can exist independently of a body and is thus a substance in the Aristotelian sense, is devastating for any theory of consciousness that purports to be naturalistic. Therefore, it is typically rejected. Dennett (1991) even rejects materialist versions of this view, in his vehement rejection of a 'Cartesian Theatre', i.e. the assumption of a center or 'headquarters' of consciousness in the brain in which all information from the senses comes together and from which commands are send to the muscles of the limbs when actions are executed by the body. He suspects this assumption to be still widespread among materialists who have given up Cartesian dualism, although no serious philosopher nowadays believes that the theatre metaphor is useful for an explanation of consciousness (but some neuroscientists apparently do, see Hobson and Friston (2014) and Dolega and Dewhurst (2015) for criticism); in fact, Kant has already criticized this conception in the 1770s. Yet, it seems that many philosophers today find it still necessary to *reject* such a *substantial* notion of self which then leads them into accepting the equally extreme claim that the notion of self (and the notion of a subject as well) should be rejected altogether, as if there was no conceptual space in between (see Metzinger (2011) for an overview). It should be emphasized that as far as mineness is concerned, we are considering the 'phenomenal subject', as it appears in conscious experience.

With his rejection of the notion of a self, Dennett (1991) also attempts to deflate the notions of a first-person perspective and mineness of experience, arguing for a thoroughly third-person approach to consciousness. But since it is possible to take the first-person perspective and subjective point of view seriously and accept it

without positing Cartesian selves, the discussion does not stop here. Philosophers who accept the egological structure of consciousness have found different ways of characterizing the subject of experience, because this notion does not yield the association of a substantial self.[3] One example among many is Zahavi's (2005) elaborate defense of a 'minimal' notion of selfhood on the basis of considerations from phenomenological philosophers such as Husserl, Sartre, Henry and others. Zahavi emphasizes the mineness of experience and characterizes it in terms of a pre-reflective self-consciousness "in the weak sense that there is necessarily something it is like for the subject to have or live through the experience" (Zahavi 2014, 88).

What does this amount to? In short, he argues that phenomenal consciousness *implies* a basic form of self-consciousness. This self-consciousness is pre-reflective since my conscious experiences are typically intentional in the sense of being directed towards some object or other—the blue sky, say. But my conscious experience of seeing the blue sky is not exhausted by the presentation of the blue sky, Zahavi argues. It also involves my awareness of *seeing the blue sky* (among other experiences that comprise my current total state of consciousness). By seeing it I am also pre-reflectively aware that it is *me* who is seeing the blue sky. Being aware pre-reflectively amounts to an immediate awareness which is not the result of a further cognitive process (reflection). Therefore, such pre-reflective self-consciousness does not amount to or presuppose the possession of the *concept* of self. Neither does it, consequently, amount to self-knowledge. It is therefore not the kind of sophisticated self-consciousness that is often denied of young infants and nonhuman animals. One motivation for drawing this connection is that in order for me to be able to report *what* I am conscious *of* I have to be intimately acquainted with the experience itself. And this amounts to an acquaintance with myself having or undergoing the experience. Typically, this pre-reflective self-consciousness is not at the center of attention, since I am focused on the blue sky, say, not on myself. But it is simply a phenomenal fact that we can *reflect* on our experiences by focusing our attention away from the object of consciousness—the blue sky, say—to ourselves as subject of the experience. It would be problematic to suppose that only such an act of reflection *makes* the experience *mine*, or brings mineness into existence so to speak; the experience has been presented as *mine* already, it only exists as *mine*. But the pre-reflective self-consciousness is not a consciousness of an *object* akin to the blue sky; it is rather a consciousness of oneself as *subject*. This is to be contrasted with a consciousness of oneself as *object*. (We return to this important distinction below.) Consequently, mineness and pre-reflective self-consciousness do *not* imply that either the experience or the subject are experienced as objects. What Zahavi calls the minimal self is *identified* with the first-person givenness of phenomenally conscious experiences. Based on what has been said so far, this minimal self enjoys experiential reality and is an integral part of the structure of phenomenal consciousness. I admit that I find much of what Zahavi has to say about

[3] In his discussion of the nature of the self against the background of the Buddhist tradition, Matt MacKenzie is equally looking for a "middle way" between the "reductionist fictionalism" about selves held by the Abhidharma school of Buddhism on the one hand and "the substantialism or ātman and Cartesian ego theories" on the other (MacKenzie 2010, 84).

phenomenal consciousness, mineness and minimal selfhood very plausible, even though the position I will defend in this paper differs from his account in various respects.

Given Zahavi's triad of mineness, subject of experience and minimal self, it is interesting to note that Metzinger, despite his rejection of the self, also leaves room for what he and neuroscientist Olaf Blanke call "minimal phenomenal selfhood". But although one of their starting points is "mineness", they arrive at this notion of selfhood via research on body perception. The central notion in their identification of the "minimally sufficient conditions for the appearance of a phenomenal self" is embodiment in the sense of "global ownership" of one's whole body (Blanke and Metzinger 2009, 7).[4] In addition to identification with one's body, this *simplest* form of the phenomenal experience of being a self involves spatiotemporal self-location and a first-person perspective. Given what I am arguing in later sections, it is clear that organisms are the prime candidates for exhibiting the property of minimal phenomenal selfhood. But Blanke and Metzinger make clear that they understand the notion of a 'subject of experience' to be stronger than minimal selfhood, differentiating four different degrees of a first-person perspective, culminating in self-knowledge, involving the concept of a self. What makes consciousness subjective in their view is "the fact that a system represents itself not only as a self but also", quoting Damasio (1999), as "a self in the act of knowing" (Blanke and Metzinger 2009, 8). Although their notion of minimal phenomenal selfhood (MPS) is considerably weaker than the way of conceiving of a subject of experience which is at issue here, Blanke and Metzinger take MPS to be a necessary condition for all higher manifestations of a first-person perspective, and they agree with the present position that they are real phenomena. We will return to their view later in the context of a discussion of the body and bodily self-consciousness.

However, not all philosophers share this view. In a recent paper, Slors and Jongepier (2014) criticize Zahavi's account with the aim of defending what they call "mineness without minimal selves"—which is the title of their paper. That is, they want to take the 'mineness' of conscious experiences seriously but explain it without postulating a 'minimal self', as Zahavi does; they also take Zahavi's theory as a paradigmatic reference point for this family of views. Echoing Dainton's criticism of such an approach they argue that it is sufficient to account for the mineness of experience in terms of the co-occurrence of a given experience. Dainton, who accepts 'mineness' as a phenomenological datum, writes:

> Any sense I have that a typical experience is experienced by a subject when it occurs is due to the fact that this experience is co-conscious with certain other experiences, namely those comprising the inner component of the phenomenal background. The inner background largely constitutes what it feels like to be

[4] Blanke and Metzinger are primarily interested in exploring the differences between such global ownership and its possible breakdowns (in autoscopic phenomena like out-of-body-experiences: seeing a second own body in extracorporeal space) and ownership merely of a body part and its possible breakdowns (in cases of somatoparaphrenia or the rubber hand illusion). I won't discuss these phenomena any further. The upshot of their argument is that minimal phenomenal selfhood does not require agency; a passive experience of owning a body is sufficient.

> me… If so, then when the inner background is present, so too am I, phenomenologically speaking. Consequently, any experience which is co-conscious with the inner background will seem as though it is occurring to a subject (=me). (Dainton 2008, 243).

The first thing to note about the notion of co-occurrence is to acknowledge that Dainton as well as Slors and Jongepier seem to accept that the question of mineness thereby must be explained in the context of considering both the unity of consciousness and the notion of creature consciousness (Bayne 2007). But co-occurrence is not sufficient by itself. In order for this to be a sufficient condition one must make sure that the two (or more) co-occurring states are states *one and the same subject* is in. Therefore, the authors must explain what it means for the *same* subject to be in two or more states at the same time or what it means for different mental states to be united and therefore experienced by the same subject. That's why the unity of consciousness is important if co-occurrence is considered as an explanans of subjectivity.

Such co-occurrence, Dainton argues, suffices for saying that this given experience is also one of *my* experiences. What Dainton calls the 'phenomenal background' is constituted by outer and inner elements, where the former are unattended intentional objects and the latter include unattended mental states like moods, bodily sensations, but also thoughts, imaginations, memories, and also the sense of self which remains typically an unnoticed and even elusive aspects of one's overall conscious state, according to Dainton. Thus, what counts as a *conscious* experience in this framework, is singled out in terms of—and is therefore restricted by—the focus of attention (Dainton 2008, 240). Whereas what we focus on is supposed to constitute the center of consciousness, the whole background is typically on the fringe of it. It is not unconscious.

Slors and Jongepier find it necessary to enrich Dainton's notion of the background with a diachronic component so that a given experience not only co-occurs with synchronically occurring moods, thoughts, etc., but also "fits neatly into a sequence of perceptions" which constitutes my "embodied biography" (Slors and Jongepier 2014, 206). This is supposed to account for the mineness of *my* past experiences and the fact that *I* am remembering them now, for example. In effect, whole trains of thought, i.e. the whole stream of consciousness must be treated as *mine*. Of course, considering the stream of consciousness as a whole makes a lot of sense since sensations, thoughts, imaginations, feelings and so on are not only often meaningfully related *at a time* but also *across time*.

The resulting picture that Slors and Jongepier suggest is that "mineness… results from coherence with a diachronic background which is (1) subjectively embodied, (2) objectively embodied, and (3) biographical." (Slors and Jongepier 2014, 203) Subjective embodiment refers more or less to aspects of what Merleau-Ponty has called the phenomenal or lived body ('corps propre'), including the body schema, while objective embodiment refers more or less to aspects of what Merleau-Ponty (1964) has called the physical body or body as object ('corps objectif'). Finally, the biographical element forming the largest part of the background includes knowledge of one's feelings, memories,

intentions, whereabouts etc. So it is clear that the background is doing all the work such that "the coherence of an experience with aspects of the embodied biography it is a part of constitutes the mineness of that experience" (ibid., 206). The characteristic feature of mineness is supposed to consist in the *absence* of any further thought or attention which turns it into the unique familiarity with our experience that it is.

An important shortcoming of both Dainton's approach and the one developed by Slors and Jongepier is that merely relying on the co-occurrence of experiences (or representations) and (unattended) states like memories and thoughts in the background does not illuminate why a given experience of the blue sky should exhibit the feature of mineness. This is so because the question arises for *all* the elements that make up the background and for *all* other co-occurring states respectively. Simply saying, like Dainton, that because there are other co-occurring states, *this* experience presents itself as *mine*, is not sufficient if no explanation is given for why the co-occurring states should be treated as states that present themselves to *the same* subject.[5] It will be shown later in Sect. 4 how this problem can be solved by the proposal presented in this paper.

At first glance it seems that this absence of further thought is compatible with what Zahavi calls *pre-reflective* self-consciousness since this is also characterized by the lack or absence of a further or additional mental state or activity, namely reflection. This is also clear, for example, from Zahavi's criticism of Rosenthal's (2005) higher-order theory of consciousness which requires just such an additional thought in order to make a given perception (of the blue sky, say) phenomenally conscious and consequently something for *me*. Yet, while Zahavi draws a conclusion in favor of a minimal self because he grants it experiential reality, Slors and Jongepier want to resist this conclusion.[6] Likewise, their criterion for the absence of further thought, or cognition, is compatible with Blanke and Metzinger's notion of minimal phenomenal selfhood. Thus, it is not clear why Slors and Jongepier are so reluctant to accept minimal selfhood and thus subjects of experience.

As we saw above, the condition of co-occurrence and the notion of the phenomenal background lead us to the consideration of the unity of consciousness and the stream of consciousness and their respective relations to the subject of experience. Let us therefore have a closer look at the problematic of considering the whole stream of consciousness including its phenomenal (synchronic as well as diachronic) background as the necessary and sufficient condition for mineness.

[5] Note that this problem also arises for Rosenthal's (2005) higher-order theory, since he does not say why a given unconscious representation (of the blue sky, say) and a co-occurring non-inferential unconscious higher-order thought (representing the first-order representation) be states of *the same* subject. In conversation, Rosenthal pointed out that this should be treated as the "default" case, but it is surely not a sufficient explanation.

[6] It is possible that the disagreement between Zahavi and Slors and Jongepier is merely verbal, since at certain points Zahavi is satisfied to drop the notion of a minimal self in favor of the notion of the subjectivity of consciousness (2014).

2 The stream of consciousness and the subject of experience

In the debate about whether the structure of consciousness is egological or non-egological, critics of the former often rely on Hume's claims with respect to the stream of consciousness and his inability to find a subject in it. For example, Slors and Jongepier allude to Hume's famous observation that turning his attention inwards—'into what I call myself'—Hume was unable to "catch myself… without a perception" (Hume 1739/1958, 252). They also seem to be accepting Hume's conclusion that there is no self since they argue against theories that defend even a minimal notion of self. As has been argued above, Dainton also seems to tacitly presuppose Hume's bundle theory. In recent work, Kenneth Williford (2015a) has even used Hume's thesis as a *criterion* for a discussion of subjective character, i.e. he treats it as an intuition that a satisfying theory of mineness must meet. Therefore, it is worthwhile pausing here for a critical evaluation of Hume's intuition and its impact on the discussion of mineness. It will be helpful to discuss it in light of some interesting and important points that Kant made against Hume's claim since it highlights the problems of the accounts sketched so far and leads the way toward an alternative account.

Hume's observation is problematic if his description is supposed to form the only (or main) basis for rejecting the notion of self. First of all, it can be objected that Hume ignored the fact that he could only perform the task of turning his attention inwards because all these perceptions, thoughts etc. already presented themselves in a successive continuity which can be interpreted in a way that they already possessed the feature of mineness: they already presented themselves as *his* by constituting *his* stream of consciousness. And it was not just any one of Hume's perceptions that performed this task but it was rather what other philosophers call the subject of experience. It is simply impossible to perform this kind of turn inwards with respect to someone else's stream of consciousness. To presuppose that the commitment to the phenomenal fact of mineness and a subject of experience *implies* or *means* having to look for an *entity* of whatever kind may be misleading and fallacious. Hume could be interpreted more charitably by saying that he felt the subject was not *explicitly* represented in his stream of consciousness. That would leave open the possibility that it was represented implicitly, as accompanying all individual experiences (within the stream), which allowed attention to turn away from the experiences towards the experiencing subject. The latter is clearly possible phenomenally, so a theory of phenomenal experience should capture its possibility.

Relatedly, rejecting the notion of a subject of experience in favor of a mere *bundle* of perceptions seems to make it impossible to account for the possibility of the 'empirical consciousness of the identity of myself as subject', as Kant has pointed out. The fact that this seems to be phenomenally possible can be turned into an objection against Hume's account: I can become conscious of myself *as* the single, (synchronically as well as diachronically) identical common subject vis-à-vis my diverse and changing experiences (i.e. stream of consciousness). I do not identify myself with either one or many of my conscious representations. Neither do I identify myself with the whole stream of consciousness, as Williford (2015a)

would have it, recognizing that his view is counterintuitive (see below). Rather, I distinguish *myself* from my perceptions, thoughts etc. as their subject when I focus on or self-ascribe them. This surely does not require that I have to find some 'entity' in my stream of consciousness to which this feeling of identity refers.

Kant's devastating objection against Hume's bundle theory is that if the subject were nothing but the sum of different representations,[7] then the consequence should be that "I would have as multi-colored, diverse a self as I have representations of which I am conscious" (Kant 1998, B134). But, obviously, this flies in the face of our actual conscious experience. Kant even concedes to Hume that the self or subject of experience is nothing that we can intuit (perceive) in inner sense. Kant did not think that the "I" was always explicitly represented in inner sense. Kant also believes that, necessarily, Hume was going to stumble only across perceptions when turning inwards since, as Kant also holds, the 'I think' must only be able to *accompany* all my representations, to quote his famous slogan. This metaphor of 'accompaniment' means that the subject of experience is given merely" along with (not in)" conscious experiences (Kant 1998, B161). That is, neither does it *produce* representations nor does it *alter* them. This 'I' is also not an object, much like in the case of pre-reflective self-consciousness in Zahavi's theory.

It is noteworthy that on Kant's account, possible accompaniment by the 'I think' can be paraphrased in terms of the unity of consciousness: a manifold of representations can only be something *for me* if it belongs to one consciousness. What we regularly find is *combinations* of representations, i.e. co-occurring representations in the sense alluded to by Dainton and others. But what Dainton has not fully acknowledged is that among the regularities pertaining to these combinations (making up the phenomenal background) are also *logical* relations between two or more representations (as in deductive inference for example). That is, many of our beliefs, states of knowledge etc. are not only "coherent" (Slors and Jongepier 2014, 202) overall or "fit neatly into a sequence of perceptions" (ibid. 206) but stand to others in *logical* and *rational* relations. Kant has stressed that such logical relations cannot be presumed to pertain accidentally or result from the sensual information received. Therefore, he added a *spontaneous* mental activity that performs the relevant kinds of synthesis. This presumed activity, he believed, is part of an *epistemological* theory, it is not a property of an ego-entity.[8] In another context Kant explains that the 'I think' is a "simple and in content for itself wholly empty representation" (A345f/B404), *depending* on sensual input. That is, the 'I

[7] Here, the term 'representation' is the translation of the German word 'Vorstellung' which is a very general term in Kant's theory. In the *Critique of pure reason*, he draws the following distinction: "The genus is representation in general (*repraesentatio*). Under it stands the representation with consciousness (*perceptio*). A *perception* that refers to the subject as a modification of its state is a sensation (*sensatio*); an objective perception is a cognition (*cognitio*). The latter is either an intuition or a concept (*intuitus vel conceptus*)." (B376f) Consequently, the term 'representation' does not create any problems through the transfer from Kant's theory into the present intellectual context in which it is common to treat (most or all) mental states as representations. No commitment to Representationalism as a theory of phenomenal consciousness follows or is intended.

[8] For a recent acknowledgement and investigation of Kant's notion of "spontaneity" with respect to the contemporary cognitive neuroscience of consciousness see Hanna and Thompson (2014).

think' and its corresponding unifying spontaneity is a mere *function*—an act which, "without any empirical representation, which provides the material for thinking,… would not take place" (Kant 1998, B422n.). That is, the spontaneous mental function, understood as an act of integration, can only be apprehended in its relation to the various representations which must be given for the 'I think' to accompany them. This is equivalent to saying that it must be possible for the representations to be *integrated* in the same (total state of) consciousness. Consequently, the possible self-consciousness resulting from the performance of this function is not a consciousness of self as *object*, but as *subject*—a notion we also found in the context of Zahavi's theory of a subject of experience. In this very sketchy way, we can see how an epistemological theory of consciousness can accommodate mineness of experiences in the context of an account of the unity of consciousness without implying the need for the acceptance of a self as entity. Of course, these Kantian points need much more elaboration but that must be reserved for another occasion. The point was to show the shortcomings and counterintuitive implications of taking Hume's intuition too seriously by discarding the usefulness of the notion of a subject of experience (or minimal self) altogether.[9]

Finally, let's have a look at Williford's (2015a) proposal to account for mineness while meeting both Hume's intuition and the intuition that consciousness has a subject-object relational structure through the identification of the subject of experience with "episodes" or even the whole "stream of consciousness". If we follow Williford and identify the subject with complex conscious episodes (or even *the whole* stream of consciousness), then the question of mineness (or subjective character, the term Williford uses) only arises for episodes which, by definition, contain parts: "The episode is a unified whole, the differentiated qualities and objects appearing in/to it are like its parts…" (Williford 2015a, 10). Since he emphasizes that all episodes have parts, a single sensation of red, say, consequently does *not* count as an episode, because it can hardly be separated into parts; mineness would then only result from the sensation's being part of an episode —a "unified whole"—which contains more elements than the sensation of red. But since, trivially, everything always also is an improper part of itself, it could be argued that

[9] Searle (2001a) argues against Hume for the existence of a self in the context of his discussion of rational action and of freedom and neurobiology. He argues that while mental capacities like perception do not require to postulate a self, the fact that we have the phenomenal experience of freedom at various points in a decision-making process leads us to postulate such an entity, namely, an "irreducible non-Humean self". He admits such an entity "with the greatest reluctance" (2001a, 79–96). Actions do not simply happen to us, Searle argues, and we act for reasons which have to be consciously appreciated as such. Only agents can act in this way and "something is an agent […] if and only if it is a conscious entity that has the capacity to initiate and carry out actions under the presupposition of freedom" (2001a, 83). Neither a Humean bundle of experiences nor a notion of the self as 'center of narrative gravity' (Dennett 1991) are sufficient to capture this, Searle argues. Consequently, "'self' is simply the name for that entity which experiences its own activities as more than an inert bundle" (2001a, 93). Combined with Searle's view that consciousness is ontologically subjective and thereby irreducible, it seems that this irreducible ontological subjectivity must also hold for the self. But while this would seem to be incompatible with Searle's biological naturalism, according to which consciousness with *all* its features is caused by and realized in the brain (cf. Schlicht 2007), he also declares that there is "no metaphysical problem of the self" (2001b, 510f).

a single sensation of red *could* very well count as an episode. The central and principal problem for Williford's proposal is that he does not provide us with criteria for individuating conscious episodes. This leads to a dilemma:

1. If a single sensation of red *cannot* be an episode, then mineness for a sensation does not arise independently of this sensation's being an element of an episode. That is, Williford must provide a theory about how the *integration* of the sensation of red into a larger episode takes place and how mineness arises for it. A corollary of that horn of the dilemma is that this approach cannot provide an answer to the question that traditional varieties of representationalism (Tye 1995; Rosenthal 2005) and other theories have attempted to explain, namely what is responsible for an individual sensation of red's being something for *me*.
2. Alternatively, if a sensation of red *does* count as an episode, then this leads to the metaphysically (somewhat) extravagant view that the multiple individual sensations that may make up one's current state of consciousness are all identical to different selves. This seems to be Strawson's view, according to which a self lasts only as long as an individual state (or episode) of consciousness (Strawson 1997). But this view flies in the face of experience for the (Kantian) reasons given above. Accepting this horn of the dilemma therefore has the consequence that we would now need a story that helps us make sense of how the subject of the sensation of red is related to the subject that is identified with an auditory sensation of a loud sound etc. In effect, this would lead to a binding problem for the multitudinous "subjects" of experience, since in my view, we cannot be content with a multiplicity of conscious subjects. This calls again for an integration mechanism that produces such a unity. Though I can understand why one would now identify this resulting integrated single stream of consciousness with the subject of experience, I don't see any motivation to identify the episode "single sensation of red" with a subject of experience, if a more complex combination of episodes is needed anyway. Zeki and Bartels (1998) have proposed an alternative view according to which an individual sensation of red *can* be conscious without thereby being identified with a subject of experience. According to this view, every individual node of a perceptual system (visual, auditory etc.) can generate an "atom" of consciousness independently. Against this approach Bayne (2010) has objected that the phenomenal unity of consciousness seems to be a *deep* feature of consciousness just like subjective character, in the sense that it cannot break down and that phenomenal consciousness cannot occur without it. That is, prior to and independent of being integrated into the unity of consciousness an individual mental representation is not phenomenally conscious. Whether or not this is true, at any rate this is not the route that Williford is prepared to take. I conclude that the problem of individuating episodes leads either to the acceptance of problematic views like Zeki's or Strawson's theories of consciousness and self, or to the need for an integration account that explains how individual elements are combined into the one global conscious experience. My worry regarding Williford's proposal is that once we have such an integration-account, there

may be no need for his additional story in terms of the self-reflexivity of conscious episodes (which are identified with the subject of experience) in order to explain mineness.

As a result of this discussion of the attempt to reduce the subject to a bundle of perceptions or identify it with the stream of consciousness, including Hume's famous and popular version, it becomes clear that mineness must be discussed together with the unity of consciousness. There seem to be intimate relations between these aspects of conscious experience. Both Dainton as well as Slors and Jongepier take this seriously by alluding to the rich phenomenal background that any given sensation is always already embedded in. But although Slors and Jongepier have expanded Dainton's notion of a phenomenal background, it appears that they don't acknowledge the full implications of this route. It seems that the crucial conditions of fitting coherently together with other mental representations or co-occurring with them are insufficient to account for the mineness of conscious experience, since the question arises what makes *any* of these representations or the sum of them *mine*. Since these two coherentist and co-occurrence accounts must already appeal to the unity of consciousness and account for it, the question arises what kind of mechanism one could consider that yields such unity.

3 Consciousness, body, subject, and organism

At the outset, I emphasized two questions: With respect to experiences we may ask, what are the conditions that any given representation has to meet in order to be experienced as *mine*? And if such mineness commits us to the postulation of a phenomenal 'subject of experience', then how should we characterize this subject? The various considerations in Sect. 2 were intended to demonstrate that the acceptance of mineness does not imply the commitment to a self as a separate, metaphysically dubious entity. There is conceptual space for alternatives which can accommodate mineness and accept a subject of experience without such commitment, like in the accounts discussed by Zahavi and Metzinger. The aim of the rest of this paper is to sketch a theory of the (phenomenal) subject of experience that at the same time provides a condition that is responsible for the subjective character or mineness of conscious mental representations. First, I argue (1) that the structure of phenomenal consciousness implies the acceptance of a subject of experience and then I present an answer to the question regarding the source of mineness (2). In a third step (3), I argue that recent work in philosophy of mind and cognitive science supports this view of the subject, leading to the conclusion that the best way to conceive of it is by identifying it with the whole organism. Finally, I will show, in Sect. 4, how this fits into the predictive coding framework. An important step in the clarification of the sense in which the subject of experience is represented in experience (or in the generative model in such a framework), we have to respect the distinction between the subject- and the object-sense of subject mentioned earlier.

1. How does an adequate characterization of phenomenal consciousness support a case for the postulation of a subject? Phenomenally conscious experiences have

subjective character. A characteristic feature of a phenomenal state is, as Block puts it, that "one is in some way aware of having it", i.e. that its "content is in some sense 'presented' to the self". As is evident here, Block explicitly mentions the notion of self and draws the obvious connection to phenomenal consciousness and its subjectivity. But he does not provide a theory of the self or subject of experience, making his account of phenomenal consciousness importantly incomplete. Unlike Zahavi (2014), Block does not seem to allow for a close connection between phenomenal consciousness and either the self or self-consciousness, not even a pre-reflective self-consciousness, since he reserves the notion of self-consciousness for a sophisticated knowledge of one's own mental states, involving the possession of the concept of 'self' (Block 2002, 213). This is also evident from Block's criticism of Damasio's theory of consciousness and self. In line with the present emphasis on subjective character, Damasio argues that conscious experience involves a *sense* of self, and he argues that there *is* a self (his notion for the subject of experience), at the same time emphasizing that it is to be understood as "a process, not a thing" (Damasio 2011, 8) a very "elusive presence" (ibid., 9), grounded in a neural mechanism distributed over a cluster of connected brain structures. According to Damasio, the sense of self involved in conscious experiences—subjectivity—is ultimately grounded in a biological process and caused by specific brain structures which are responsible for the regulation of the whole organism's wellbeing. In order to maintain its identity and ensure survival, the organism's overall homeostatic state must remain within certain bounds. Because of their monitoring and regulating function with respect to this goal, Damasio calls these structures the unconscious biological 'proto-self'. Damasio (2011) takes to be the most important idea of his framework "the notion that the body is a foundation of the conscious mind" in the sense that these "proto-self structures are not merely *about* the body. They are literally and inextricably *attached* to the body". In Damasio's framework, these give rise to the phenomenally experienced sense of self, the 'core self', which in turn forms the basis of the 'autobiographical self'.[10] This comprises memories of the past and anticipations of the future, yielding a rich *conception* of oneself. Only this latter phenomenon presupposes the possession of a concept of self and is likely to be ultimately reserved for human beings (for philosophical details of Damasio's neurobiological theory see Schlicht 2008). In his review of Damasio's book, Block (2010) accuses Damasio of holding the strong claim that phenomenal experience is *dependent* upon self-consciousness in this richer sense which requires the possession of the concept of self. This is not Damasio's view. In fact, if Block's research program is to search for the neural correlates of consciousness, and if he acknowledges that phenomenally conscious experiences are "presented to the self", then Damasio's approach should be useful and welcome to clarify this issue. Yet, Block's surprising and

[10] It seems that Zahavi (2005, 138–140) associates Damasio's core self with his own notion of a minimal self (see also Gallagher 2000).

strict dismissal of Damasio's account helps highlight the fact that he lacks any alternative account despite making use of the notion of self in connection with phenomenally conscious experience.

2. This brief discussion was intended to highlight the contrast between Zahavi's (and Damasio's) position that phenomenal consciousness implies a pre-reflective sense of self and Block's account of phenomenal consciousness which remains incomplete. I take it as an important and worthwhile task to address this desideratum by developing a notion of the subject of experience that is both philosophically acceptable and empirically informed. In order to sketch such an account, let me start by picking up various elements of the discussion so far in order to sketch a condition for the mineness of experience first.[11] As we saw in Sect. 2, Dainton and Slors and Jongepier stand in the Kantian tradition by connecting mineness to the unity of consciousness.[12] The discussion of Williford's account showed the need for an *integration account*, illuminating the phenomenal fact that *different* representations share the same subjective character by presenting themselves as experiences of the same subject: *Me*.

As it turns out, Damasio's theory also helps in this regard by introducing the homeostatic state of the organism. The central idea which I want to put forward is that a mental representation (of the blue sky, say) becomes conscious and thus exhibits subjective character if its neural substrate is integrated—or "recruited" (Van Gulick 2004, 76)—into the neural substrate of the overall "biological" state of the organism. This claim needs unpacking and clarification.

Many theories of consciousness make use of the general idea of integration in different ways (e.g. Metzinger 1995; Dehaene 2014; Van Gulick 2004; Edelman and Tononi 2000; Damasio 2011). At any time, an organism will be in *one* overall experiential state with a total neural correlate. This state includes substrates of what's called creature consciousness—or state of vigilance (Dehaene et al. 2006), presumably involving the brainstem—and core substrates of the multitude of visual and other perceptions, thoughts, moods etc. that are part of this state. Thus, any given representation (of the blue sky, say) will only be a *modification* of the overall unified conscious state of the organism. The present proposal alludes to the need for some mechanism that *integrates* or *recruits* the individual representation into this unified total state in order to make the modification possible. This total experiential state of the organism contains variant and invariant aspects: Among the former are representations of objects and features of the world that come and go, depending on the focus of attention, as well as representations of mental states and of bodily states.[13] Among the latter are representations that are more stable across time. At this point, Damasio's account of the biological proto-self enters the story: the proto-self is construed as "an integrated collection of separate neural patterns that map,

[11] Schlicht (2015) contains a more detailed discussion of this issue.

[12] For a discussion of this issue in Kant's critical philosophy see Schlicht (2016).

[13] Edelman and Tononi (2000) have suggested that this state is supported by a neural pattern comprising large areas of the thalamo-cortical system.

moment by moment, the most stable aspects of the organism's physical structure" (Damasio 2011, 190).[14] The idea is that the whole organism, especially its homeostatic state, is monitored and regulated at all times by structures in the brain which represent the organism's overall bodily state. In case the necessary homeostatic balance needed for survival is disturbed,[15] the brain is forced to initiate measures to restore and maintain this balance. According to Damasio, in order for an object-representation to become conscious, its substrate must be integrated into, or connected with, the substrate of the representations of the body—the proto self: in this way, a connection between the organism (as represented in the brain) and its object-representations is established such that we can make sense of the important phenomenal fact that all conscious representations feel like something *for the organism*, that there is something it is like *for the organism* (in Nagel's words). The organism provides, as Damasio puts it, a "haven of stability and invariance" (Damasio 1999, 142, 153ff, see also Metzinger 2003, 161), i.e. just what we need in order to illuminate why conscious experiences should exhibit subjective character or mineness. Since it is always the organism's state that is represented, monitored and regulated and given that it must remain within a certain range, this approach has the means to make plausible why there should be something it is like for *the same organism* given the manifold of different individual representations. Metzinger agrees that in order to start constructing an account of mineness we have to look for the point of "*maximal invariance of content* in the conscious model of reality" (Metzinger 2003, 134). Metzinger argues that this invariance is most likely due to the organism and its bodily structures represented in the brain, since it is invariance (or maintenance of homeostatic balance) which is keeping the organism alive. Another advantage of this view is that it does not introduce a questionable additional entity, it avoids Williford's phenomenologically counterintuitive claim that the stream of consciousness should be identified with the subject of experience, and it avoids the various problems associated with versions of representationalism.[16]

3. This proposal draws a connection between mineness of experiences and the *pre-reflective self-consciousness* that is at play in phenomenally conscious experiences. Bermúdez (1998) as well as Blanke and Metzinger (2009) identify this most

[14] On the brain-stem level, these structures include the nuclues tractus solitarius, the parabrachial nucleus, the periaqueductal gray, the area postrema, hypothalamus and superior colliculus. In addition, some cerebral structures like the insula and anterior cingulate cortex constitute the proto-self (cf. Damasio 2011, 191). While the neuroanatomical details do not concern us here in this philosophical investigation, it is important to note that this perspective on the self is anthropocentric and (mostly) based on indirect evidence from lesion studies in relation to pathological conditions of patients that have to do with the loss of the sense of self.

[15] According to Damasio, this happens primarily via interaction with the world which results in object representations etc.

[16] Tye's (1995) PANIC-theory alludes to a specific way objects have to represented but has difficulties distinguishing conscious from unconscious representations. Rosenthal (2005) therefore introduces an additional higher-order thought that must represent the first-order representation. But this move introduces various problematic possibilities of misrepresentation. By refraining from introducing a representational relation, the present proposal avoids all these shortcomings.

primitive kind of self-consciousness with *bodily self-consciousness*. This is plausible since, as Metzinger argues, "bodily self-consciousness is, phenomenologically as well as functionally, the most important *source of invariance* human beings possess" (Metzinger 2003, 289). Part of this is what Bermudéz calls "somatic proprioception" which provides the organism with non-conceptual first-person contents and information about the embodied self that is immune to error through misidentification; it provides a way of registering the boundary between self and non-self and it is plausibly in place already at birth. Blanke and Metzinger (2009) associate this invariance with the global ownership of the organism's whole body. Many other authors have also stressed the importance and basic function of such bodily self-consciousness (e.g. de Vignemont 2011, 2014; Alsmith and de Vignemont 2012; de Preester 2007). It is noteworthy that Slors and Jongepier (2014, 203) include as an important part of the phenomenal background what they call "subjective embodiment": This includes "one's interoceptions, proprioceptions, one's sense of agency and ownership, and one's body image and body schema". The present proposal takes seriously the possibility of identifying the phenomenal with the metaphysical subject, namely the organism, to be understood as the claim that the phenomenal content representing the subject on the level of conscious experience inevitably reflects the organism. This is in line with Blanke's and Metzinger's account, since they identify the "representational content of MPS" as the "essence of selfhood" and characterize this content as "a representation of the entire body" (2009, 8). The aspects that typically become salient phenomenally when one turns attention away from a task or object towards oneself, are typically aspects to do with one's embodied subjectivity. Williford (2015b) and others may object to the present proposal that it is based on a confusion of the phenomenal subject with a metaphysical subject. While their view is about the former, my argument would only make sense with regard to the latter. But this is wrong. The proposal is exactly that it is (various bodily aspects of) the organism the representation of which grounds pre-reflective self-consciousness, understood as bodily self-consciousness. Take for example, my experience of trying to lift something which turns out to be too heavy. Being focused on the object while trying, I suddenly become aware of my body with regard to its (and consequently *my*) limitations. In cases like this, to which I also count pains, what's phenomenologically salient is my body. Importantly, the brain represents and regulates such aspects of the body. By way of integrating these representations with the object representations which result from the organism's interaction with the environment (pace Damasio), it makes sense to say that the former are presented to the organism such that there is something it is like *for the organism* to represent and experience these objects.[17]

[17] It is interesting to note that Nagel (1969), in criticizing Fodor's mechanistic theory of intentionality, argues that "perhaps we shall have to fall back on the idea of an organism or organismic system", but also advocates that we should see first whether the concept of a "person" is indeed a "dying notion". Thanks to an anonymous reviewer for pointing me to this passage.

A further line of support for the view advocated here comes from recent embodied cognitive science which stresses the importance of embodiment for cognition and consciousness. Some philosophers have argued that all these phenomena are essentially *embodied* in the sense that the body and various embodied ways of coping with the environment play a crucial role for these phenomena (Varela et al. 1991; O'Regan and Noë 2001; Clark 1997, 2008; Wheeler 2005; Gallagher 2005). Facial expressions and bodily postures are arguably partly constitutive of feelings and their expression; one's bodily position (e.g. walking upright), and thus one's bodily constitution (e.g. the position of the eyes) trivially determine what one can see. Moreover, perceiving an object from a certain point of view involves 'looking' by performing eye-, head- and whole body-movements. This supports the idea that perceiving is a skillful bodily activity, not merely a passive reception of information. The discovery and interpretation of *mirror neurons* (Rizzolatti and Sinigaglia 2006) can also be seen as providing evidence for the claim that perception does involve an important motor dimension, since mirror neurons become activated both when I perform a bodily action of some kind *and* when I merely observe someone else perform this or a similar action. Indeed, mirror neurons suggest that perception and action do not constitute two separate cognitive domains but rather form one dynamic unified system (Hurley 2001). If one is sympathetic to this broadly embodied view of cognition and perception as bodily *activities*, then the sense of agency associated with cognitive activities reveals the bodily aspects of the organism as important elements of the phenomenal background of the unified state of consciousness. Activities presuppose an agent; and sometimes the organism, achieving and performing, shows up phenomenally in a more or less salient way. Echoing Merleau-Ponty's (1964) philosophy of perception, it is the *whole organism* which *lives through* conscious experiences and is *intentionally directed* in various ways towards its environment (Hutto 2008; Thompson 2007).

On a side note, it is interesting that, by relying on prior work by Varela (1979) and Thompson (2007), MacKenzie (2010) identifies selves with organisms and characterizes them as autonomous, dynamic and self-organized systems. On this view, the defining trait of organisms is their self-organization and self-production or 'autopoiesis' (Maturana and Varela 1980): an organism consists of a biological network that self-produces a boundary setting it apart from its environment while at the same time allowing for the exchange of matter and energy with the environment (metabolism). In the case of humans, this semipermeable boundary is the body. This production and exchange can be said to result in the making of an 'identity' because the semi-permeable boundary marks a clear delineation of what the system consists of (self) and what is outside the system (non-self). The nature of the network also determines the biological needs of the system which in turn determine the system's intentional relations towards the environment. From the perspective of the system, its biological needs or "'concern'… turns an indifferent physicochemical world into an environment of biological significance" (Thompson 2007, 153). From the perspective of the organism some things then 'count' as food, for example, while others do not (Thompson 2007, 152ff). Thus, this continuous biological process of self-production and maintenance of identity constitutes a view of the self "as an

active, embodied, embedded, self-organizing process", and can be understood as "self-making" (MacKenzie 2010, 76)[18]

Thus, to conclude my argument so far, taking this motor and bodily dimension of consciousness and cognition into account against the background of Damasio's claims about the representation of the organism's body and overall biological state in the brain supports the notion that the subject of conscious experience should be identified with *the whole organism*, the animal, understood as an embodied agent, endowed with various cognitive, affective and sensorimotor capacities as well as basic non-conceptual and proprioceptive self-consciousness informing it about its body. Therefore, the missing aspect in Slors and Jongepier's account of the background is the recognition of the importance of the (bodily aspects of the) organism as a whole and of its representation in the brain. Given this background, it makes sense to say that a sensation of red becomes something *for the subject*, if its neural substrate is connected (via some integration mechanism) to the complex brain state supporting the representation of the organism (i.e. the proto-self-structures). Such a connection can yield a representation of the organism as representing (such-and-such) or, as Damasio (1999) calls it, "the self in the act of knowing".[19] It also solves Dainton's problem of having to provide a reason to believe his claim that the phenomenal background of a given experience somehow *constitutes* what it is like to be *me*. This is plausible if we identify the subject of experience with the organism that also enters the background as an invariant element in virtue of its representation in the brain (as a biological proto-self) which is part of the cluster into which an individual sensation must be integrated in order to exhibit mineness. By relying on the important function of integration, the answer to the question what makes a given sensation *mine* is given in connection with the answer to the question what is responsible for the unity of consciousness. The unified total state of the organism contains invariant elements like the representation of the organism's body (e.g. proto-self), and variable elements like representations of objects, sensations, thoughts etc.

In the final section, I would like to conclude with a brief discussion about how this proposal relates to the recently popular and explanatorily fruitful framework conceiving of such cognitive and representational brain activity in terms of predictive processing. Although its strength has been demonstrated with respect to the explanation of various perceptual and cognitive phenomena, such as attention,

[18] Kirchhoff and Froese (2017) discuss the connections between this kind of autopoietic enactivism and the predictive coding framework via Friston's (2010) "free energy principle". According to Lamb and Chemero (forthcoming), autopoietic systems form a subclass of dynamical systems. Both the free energy principle and the principles of dynamical systems theory apply to a greater class of systems than only to organisms.

[19] Given that Damasio supports his distinctions by extensive discussions of pathological cases, it is noteworthy that this sense of self (in the act of knowing) is what's absent in diseases like akinetic mutism. absence seizures and others, where core consciousness is momentarily suspended (although patients stay awake). Patients in this condition remain 'embodied selves', according to Damasio, but their first-person experience is gone. Since it depends on core consciousness, extended consciousness is also impaired here. In other less dramatic cases like amnesia or aphasia, core consciousness remains but only varieties of extended (or autobiographical) consciousness are impaired.

cognitive penetration etc., a clarification of the role of the perceiver, i.e. the subject of experience, is still a desideratum but obviously central to the topic of the present paper.

4 Predictive coding, perceptual contents, and the experiencing subject

It has already been stressed in Sect. 2 that Hume's non-egological account of consciousness ignored the fact all my perceptions, thoughts etc. already present themselves in a successive continuity as *mine* by constituting *my* stream of consciousness. Hume was looking for a self-entity by actively turning his attention inwards to *his* stream of consciousness which is impossible with respect to someone else's stream of consciousness. Such a shift of attention is only possible because all *conscious* experiences already exhibit the relevant feature of mineness. What Hume wanted to achieve is finding "himself" as represented in the contents of consciousness, independently of experiences. We already discussed his account from Kant's perspective. But Kant's ideas are not only useful for criticizing Hume's bundle theory, they also contain the roots for a very original and recently popular contribution to the debate about how the mind/brain achieves the complicated computational task of perceiving the world, namely by predicting the world.[20]

In a number of influential writings, Friston (2010), Hohwy (2013) and Clark (2013, 2016) defend the view that the brain is a prediction machine. This conception makes use of Kant's central insight that perception is to be explained by a delicate balance between top-down and bottom-up processing, which has been taken up by Hermann von Helmholtz in his account of perception as unconscious inference. The "delicious" (Dennett 2017, 167) basic idea of Bayesian predictive coding is that the brain constantly generates expectations and hypotheses (on many levels of complexity) about the possible causes of its sensory input. Such top-down expectations originate in the brain's model of the world; they are continuously compared to actual bottom-up input, yielding error messages of various magnitudes. Neural processing based on sensory input is reinterpreted as the processing of prediction error enforcing an update of the model or the hypothesis that issued the top-down expectations in the first place. This process continues until the model (more or less) matches the feedback and can thus count as accurate with respect to the state of the world. Furthermore, this process is "Bayesian" because it is supposed to obey Bayes' rule, according to which all hypotheses are expressed as probabilities (see Frith's (2007) exposition of this for detailed clarification). Given that the sensory input is most often ambiguous with respect to its cause, the brain is forced to formulate hypotheses about which source may cause the input. Several possibilities with different probabilities will then be evaluated. Viewed from this perspective, perception, action, attention, and other cognitive phenomena are all strategies of "doing the same thing" (Hohwy 2013, 2): The overarching goal (and

[20] Swanson (2016) discusses the roots of the predictive processing paradigm in Kant's philosophy of perception.

unifying principle) of the brain's perceptual and cognitive activities is to minimize prediction error, or "surprise".

Not all aspects of the model can be explained in terms of updating based on prediction errors. Although testing predictions against sensory feedback is a continuous process, in order to get started, the system (i.e. the brain) must harbor some innately specified "priors", i.e. fixed assumptions which yield the initial expectations and predictions. Since there must be a starting point, not all the prior knowledge needed for perception can be learned. Frith (2007, 128) provides some examples. The sun has been the only major source of light for millions of years, and given that the sunlight always comes from above, our brain is adapted to see concave objects as light on the top and dark at the bottom but convex objects the other way around. Most of the background knowledge needed for perception though will be learned by experience. For the purposes of this paper, this brief exposition of the key idea of the predictive coding hypothesis must suffice. The works by Clark (2013, 2016), Frith (2007), and Hohwy (2013) provide useful and detailed accounts of it.

According to Frith (2007, 117–119), the advantage of this framework is that it helps to evaluate the amount of information that a stimulus contains *with respect to the perceiver.* Shannon's (1948) information theory as such did not take the subject of experience into account. But since different perceivers have different background information, sensory input will be significant in different ways for different perceivers. Bayes' rule informs the brain to what extend its model of the world must be updated given sensory input (or feedback rather) and ensuing prediction errors. But this framework of "prediction error minimization" raises the issue of truth or accuracy conditions of perceptual experiences. On this view, the brain's model completely determines the content of my experience: Sensory input does not shape perception directly, it is better conceived of as "feedback to the queries issued by the brain" (Hohwy 2013, 2). Our top-down expectations "drive what we perceive", and consequently, "the hypothesis that is selected determines perceptual content (Hohwy 2013, 37). What we experience is not the world, from the first person point of view, "it *is* the predictions of the currently best hypothesis about the world" (Hohwy 2013, 48). Frith (2007, 111) captures this with his provocative slogan that "our perception of the world is a fantasy that coincides with reality". But how must the accuracy conditions be specified which decide about whether a given perception is veridical?

As Searle (1983, 2015) and Jackson (2004, 114f) have emphasized, the truth condition of my perceptual experience of the pencil in front of me is not simply that there is a pencil in front of me. Rather, the truth condition is that there is a pencil in front of me, causally interacting with *me* in a highly distinctive way right *now*. Only one of the many possible pencils on the desk is interacting with me in the sense of making this particular perceptual experience veridical. But the causal relation, Recanati (2007) has argued against Searle, is not part of the *intentional content*, but part of the *intentional mode*, such that because of this complication, the accuracy conditions of my perceptual experience are not solely determined by its content. The proposition that there is a pencil in front of me

> is meant to be evaluated with respect to a very specific situation, namely the subject's *perceptual* situation: a situation which the subject is causally affected by through his senses and which, in particular, causes the occurrence of the mental representation in question (Recanati 2007, 135).

As the example demonstrates, both the intentional mode (*perceiving*) and the subject of experience (*me*) enter the accuracy conditions of any particular perceptual state. But even though these aspects of the situation enter the accuracy conditions, they are not represented as such, because they enter in virtue of their functional role which is determined by the mode of perceiving. In this sense, perception is always situated. The accuracy conditions of my *imagining* or *remembering* a pencil in front of me differ to those for *perceiving* it precisely because of the particular way the intentional mode partly determines them, whereas the content remains constant. But if Recanati's analysis is right, then this raises the question in what way the subject enters the accuracy conditions and how this should be captured by the predictive coding framework. Consequently, any analysis of perception in terms of predictive coding will remain incomplete if this issue is not settled.

Not much has been written about this topic though. Limanowski and Blankenburg (2013) as well as Hohwy and Michael (forthcoming) discuss the framework by relying on Metzinger's (2003) self-model theory. Moutoussis et al. (2014) sketch an account that relates self and other within the framework and Friston and Frith (2015) emphasize the idea of alignment in social interaction between two agents. In my brief discussion here, I will restrict myself to the account presented by Hohwy and Michael. Their central idea is that through action, understood as active inference, we position ourselves in the realm of the hidden causes of sensory input. Given that the brain is in the business of inferring causes from input, it is (somehow) prompted to construct a self-model to capture the hidden causes that are generated through action. Hohwy and Michael rely on Metzinger's notion of a "self-model" but identify it primarily as a "full body model" which is a hierarchical structure of hidden, endogenous causes that will eventually be identified as the self. This hierarchy of layers mirrors the time scales of action, ranging from (milli-)seconds in the case of reaching for a bottle of wine, say, to hours, days, months etc. in cases of traveling, studying etc. By introducing a hierarchy of interacting layers of causes Hohwy and Michael attempt to capture what's distinctive about the subject. Hohwy defends the strong claim that "*all* we ever do" is minimizing prediction error, such that, consequently, the claim is that "*we* are like this… We manage to represent the world by being prediction error minimizers—neural mechanisms that realize Bayesian inference in the long run" (Hohwy & Michael, forthcoming). This idea seems natural given the framework but it suffers from various problems that I am going to point out briefly.

The first problem is that they remain unclear about what we should take the self to be. Their formulations vary dramatically between various options: If this quote above is to be taken literally, then they identify the self with a set of neural mechanisms instantiating prediction error minimization. This is in line with their claim that the self is constituted by a hierarchy of layers of hidden causes (exemplified in actions). But in other passages, they come close to identifying the

self with the body, which is also natural given that the body executes those actions. Thus, they claim that formulating hypotheses about hidden causes will automatically lead to a "full body model", since "the body is nothing special" just like "representation of the body is nothing special". They move back and forth between considering the agent's body and self, while remaining undecided about their exact relation. Body and self can neither be strictly identical—since the self is supposed to exhibit more abstract features like character traits which go beyond what is "represented in the body-model specifically"—, nor are they supposed to be separate entities. The authors also distance themselves from Dennett's (1991) narrative (fictional) approach to the self, since they are after "a more metaphysically robust account of self-representation". What they mean by this remains unclear, especially since they claim that being a self is "tied to existence or being". While it is straightforward that the body model represents the body, it is not so clear what the self-model represents, if there is no entity that exists beyond the organism, but the model is not supposed to represent the organism. If they do not intend to posit an additional entity but attempt to track the source of changes in the world as effected by actions, then it would be natural to identify the self with the organism since it is the organism, not a neural mechanism or model or mysterious inner cause, which is acting in the world. But they never seem to choose that option.

A second problem is that what Hohwy and Michael are trying (and able) to capture is only the self-as-object, in particular the self-as-bodily object. We introduced this distinction at the end of Sect. 2 and will elaborate it in this section. In order to see the impact of this objection, some background is needed. Although Frith does not elaborate on the philosophical implications of the prediction error minimization framework for our conception of the subject of experience, he refers to the subject as "the invisible actor at the center of the world" (2007, 109). It is natural to speak of the center of the world given that in perception, the scene in front of *me* is presented to *me* as I see it from *my* perspective. This invisibility is manifest in the fact that the subject need not be explicitly represented in the content of the perceptual experience in order to be an essential part of it. Recanati expresses this by saying that although *subject* and *mode* partly determine the accuracy conditions of a perceptual experience, "the subject and his situation are not themselves elements of the fact represented" (Recanati 2007, 145).

But to complicate matters, the subject can enter the perceptual experience in two different ways, which can be illustrated by an example also used by Recanati (2007, 146–7): Consider the situation where I am sitting on a chair and judge, based on proprioceptive information only, that my legs are crossed. Somehow I know this "from inside". Judgments based on proprioceptive information are immune to the error of misidentification (Shoemaker 1968), simply because no one is identified, no act of identification is needed. This is contingently so. Contrast this with the situation where I am sitting in front of a mirror and judge that my legs are crossed based on visual information. In this case, "the person to whom the property of having one's legs crossed is attributed is explicitly represented: it is 'articulated' or 'identified' in the content of the representation… In the former case that person is implicitly determined by the mode, but it is not represented, hence the statement is 'identification-free'" (Recanati 2007, 147). But in the mirror case, the perception is

about myself in two ways: I as subject enter the perceptual experience in both ways, as the object identified (based on visual information) and as the subject seeing the mirror and myself in it. In conclusion, Recanati observes with regard to the topic of mineness: "That the perception is mine and concerns my surroundings rather than someone else's, is something which is guaranteed by the architecture of the system, hence gives rise to immunity." (2007, 147–8)

How can we capture these two dimensions of the subject in the predictive coding framework? The subject dimension is nicely captured by Frith's (2007, 100) phrase that the "brain embeds us in the world and then hides us"—although he does not refer to this phenomenon explicitly and much less elaborates on it. Hohwy and Michael (forthcoming) have suggested a way how to account for the subject-as-object only, which seems straightforward since in the object-sense I can be treated just like any other object in the world, in terms of a self-model that falls under the scope of prediction error minimization. And given the prediction error minimization framework, updating a *model* is all that ever happens in terms of perception, cognition and action. Inference is central to prediction error minimization but absent in the case of an awareness of self-as-subject, since this is characterized by an absence of any act of identification. It is also evident from their heavy reliance on Metzinger's notion of a self-model, from their emphasis on the representation of the body, and from their emphasis of character and personality traits which go way beyond the notion of self-as-subject.

This is also evident from the fact that they claim the self-model to "arise… as a natural consequence of humans' skill in modeling and interacting with each other", emphasizing the development of character traits and personality features that individuate the person and distinguish them from others. They suggest that young children apply agent-models to understand others and in turn "shape their selves progressively to match the agent models they have been using" in social cognition. This claim is only plausible if it is taken to apply to the object-dimension of the self, not the subject-sense that is responsible for the impossibility to error through misidentification highlighted earlier.

The organism, conceived as object of experience is firmly embedded in the world and capable of interacting with it in various sensorimotor and cognitive ways. But the organism, conceived as subject of experience, is not represented in the contents of experience. As far as it is represented, it is always only represented as object. Therefore, if Recanati is right that the self-as-subject is not represented in the model, then it is not constantly updated based on sensory feedback. What can be updated is always the model of the self-as-object. Proponents of predictive coding theories need to capture this dimension in some way.

In this paper, I can only present this problem and formulate it as a task for future research to determine the place of the dimension of self-as-subject in the predictive coding framework. In any case, I think that the present discussion provides us with the constraint that its place must be outside the content of the self-model, i.e. outside the representational content of experience. A first option is to consider it as one of the very stable hyperpriors, highly abstract hypotheses that remain constant over long time scales (but can possibly break down in pathological cases). A second, more promising option is, like Hohwy and Michael could suggest, to associate the

self-as-subject with the dimension of agency (since agency may give rise to certain sensorimotor contingencies relative to the subject). Sensorimotor contingencies always have to be specified in relation to this specific organism and will be unique to this organism. Thus, this particular subject as organism will be implied in the contingencies. But Hohwy and Michael rely on Metzinger's self-model theory, and Blanke's and Metzinger's account of minimal phenomenal selfhood explicitly states that agency is not needed for it. So, such selfhood should already be associated with perceptual inference, not only with active inference. Thus, Hohwy and Michael would have to update their account. A third, related, option is to consider it as a default mode of the whole mechanism of the generative model since one of its central features is the absence of any act of identification. After all, the mode of perceiving provides information about the scene in front of *me* relative to *me* without saying anything about me. That is, from the perspective of the model, all information that is used in order to minimize prediction error, is information relative to the model itself. Since the model is generated by the perceiver's brain, it will be its default mode to take the perceiving organism—about which it has information—as the reference point of all expected and incoming information. But an evaluation and extensive discussion of these and possibly other options is a task for another occasion.

5 Conclusion

This paper has addressed an important feature of conscious experience, namely mineness, which is closely related to what Nagel has called a subjective point of view. It has been argued that the acceptance of mineness and the structure of phenomenal consciousness commit us to the acceptance of a subject of experience. Nevertheless, on the basis of Zahavi's phenomenological account it has been shown that this subject can be characterized independently of the notion of a self that is typically rejected, e.g. by cognitive scientists and philosophers. A discussion of recent attacks against even such a minimal conception of the subject from e.g. Dainton, Slors and Jongepier and Williford, demonstrated shortcomings and problems of their respective accounts. The typical reliance of these accounts on Hume's claims regarding the rejection of the self has been countered with considerations from Kant's philosophy of mind. It has finally been argued that the problems of these accounts can be solved by outlining a theory of the subject as organism. The subjective point of view is given by the constraints of the biological set-up of the organism *and* by the organism's needs in the homeostatic maintenance of itself. Damasio's theory of the biological proto-self, considerations about the embodied character of consciousness and cognition and an alignment of the pre-reflective self-consciousness associated with mineness with a basic bodily self-consciousness all support this notion. This brief look at recent evidence from the cognitive sciences suggests that the subject of conscious experience should be conceived of as *the whole animal*, i.e. as an embodied and embedded agent endowed with an arsenal of cognitive, affective and sensorimotor capacities. Thus, I take it to be a virtue of this view that it can integrate all this important work into a strong,

philosophically sound and parsimonious, as well as empirically informed account of the subject. Finally, the considerations of this paper have been applied to the recently popular predictive coding framework. Two important philosophical dimensions of the self have been emphasized, the self taken in its subject-sense and the self taken in its object-sense. Hohwy and Michael's recent account of the self has been discussed representatively and found wanting since it can only capture the subject of experience in the object-sense in terms of a self-model that is constantly updated and enriched during development. It remains a future task to demonstrate how the framework can handle the subject-dimension of the self.

References

Alsmith, A., & de Vignemont, F. (2012). Embodying the mind and representing the body. In A. Alsmith, & de F. de Vignemont (eds.), *The body represented/embodied representation. Review of philosophy and psychology*, Special issue, 3(1).

Bayne, T. (2007). Conscious states and conscious creatures: Explanation in the scientific study of consciousness. *Philosophical Perspectives (Philosophy of Mind), 21,* 1–22.

Bayne, T. (2010). *The unity of consciousness*. Oxford: Oxford University Press.

Blanke, O., Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends in Cognitive Sciences, 13*(1), 7–13.

Block, N. (1997). On a confusion about a function of consciousness. In N. Block, G. Güzeldere, & O. Flanagan (Eds.), *The nature of consciousness: Philosophical debates*. Cambridge: MIT Press.

Block, N. (2002). Concepts of consciousness. In D. J. Chalmers (Ed.), *Philosophy of mind. Classical and contemporary readings*. Oxford: Oxford University Press.

Block, N. (2010). What was I thinking? Review of Antonio Damasio, *Self comes to mind*: Constructing the Conscious Brain. *New York Times Book Review*, November 28, 2010.

Clark, A. (1997). *Being there. Putting brain, body and world together again*. Cambridge: MIT Press.

Clark, A. (2008). *Supersizing the mind*. Oxford: Oxford University Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204.

Clark, A. (2016). *Surfing uncertainty*. Oxford: Oxford University Press.

Crane, T. (2000). The origins of qualia. In T. Crane & S. Patterson (Eds.), *The history of the mind-body problem*. London: Routledge.

Dainton, B. (2008). *The phenomenal self*. Oxford: Oxford University Press.

Damasio, A. (2011). *Self comes to mind*. New York: Pantheon.

Damasio, A. R. (1999). *The feeling of what happens. Body and Emotion in the making of consciousness*. Mariner Books.

De Preester, H. (2007). The deep bodily origins of the subjective perspective: Models and their problems. *Consciousness and Cognition, 16*(3), 604–618.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences, 10*(5), 204–211.

Dennett, D. C. (2017). *From bacteria to Bach and back. The evolution of minds.* Allen Lane.

de Vignemont, F. (2011). A self for the body. *Metaphilosophy, 20*(3), 230–247.

de Vignemont, F. (2014). A multimodal conception of bodily awareness. *Mind, 123*(492), 989–1020.

Dehaene, S. (2014). *Consciousness and the brain. Deciphering how the brain codes our thoughts*. New York: Viking.

Dennett, D. C. (1991). *Consciousness explained*. New York: Basic Books.

Dolega, K., & Dewhurst, J. (2015). Curtain call at the cartesian theatre. *Journal of Consciousness Studies, 22,* 109–128.

Edelman, G., & Tononi, G. (2000). *A universe of consciousness*. New York: Basic Books.

Friston, K. (2010). The free-energy principle: A unified brain theory?. *Nature Reviews Neuroscience, 11*(2), 127–138.

Friston, K., & Frith, C. D. (2015). A duet for one. *Consciousness and Cognition, 36*, 390–405.
Frith, C. D. (2007). *Making up the mind: How the brain creates our mental world*. Oxford: Blackwell
Gallagher, S. (2000). Philosophical conceptions of the self. *Trends in Cognitive Sciences, 4*(1), 14–21.
Gallagher, S. (2005). *How the body shapes the mind*. Oxford: Oxford University Press.
Hanna, R., & Thompson, E. (2014). Neurophenomenology and the spontaneity of consciousness. *Canadian Journal of Philosophy, 33*(suppl. 1), 133–162.
Hobson, J. A., & Friston, K. J. (2014). Consciousness, dreams, and inference: The cartesian theatre revisited. *Journal of Consciousness Studies, 21*(1–2), 6–32.
Hohwy, J. (2013). The predictive mind. Oxford: Oxford University Press
Hume, D. (1739/1958). *A treatise of human nature*. In: L. A. Selby-Bigge (ed.). Oxford: Oxford University Press.
Hurley, S. (2001). Perception and action. Alternative views. *Synthese, 291,* 3–40.
Hutto, D. D. (2008). *Folk-psychological attitudes*. Cambridge: MIT Press.
Jackson, F. (2004). Representation and experience. In: H. Clapin, P. Staines, P. Slezak (Eds.), *Representation in mind*. Elsevier, pp. 107–124.
Kant, I. (1998). *Critique of pure reason*. (Transl. by P. Guyer, A. Wood). Cambridge: Cambridge University Press.
Kirchhoff, M. D., & Froese, T. (2017). Where there is life there is mind: In support of a strong life-mind continuity thesis. *Entropy, 19*(4), 169.
Kriegel, U. (2009). *Subjective consciousness*. Oxford: Oxford University Press.
Limanowski, J., & Blankenburg, F. (2013). Minimal self-models and the free energy principle. *Frontiers in Human Neuroscience, 7*.
MacKenzie, M. (2010). Enacting the self: Buddhist and enactivist approaches to the emergence of the self. *Phenomenology and the Cognitive Sciences, 9,* 75–99.
Maturana, H., & Varela, F. (1980). *Autopoiesis and cognition: The realization of the living*. Boston: Reidel.
Merleau-Ponty, M. (1964). *Phenomenology of perception*. London: Routledge.
Metzinger, T. (1995). Faster than thought. Holism, homogeneity and temporal coding. In T. Metzinger (Ed.), *Conscious experience*. Imprint Academic: Thorverton.
Metzinger, T. (2003). *Being no one*. Cambridge: MIT Press.
Metzinger, T. (2011). The no-self alternative. In S. Gallagher (Ed.), *The oxford handbook of the self*. Oxford: Oxford University Press.
Moutoussis, M., Fearon, P., El-Deredy, W., Dolan, R., & Friston, K. J. (2014). Bayesian inferences about the self: A review. *Consciousness and Cognition, 25*(1), 67–76.
Nagel, T. (1969). The boundaries of inner space. *The Journal of Philosophy, 66*(14), 452–458.
Nagel, T. (1974). What is it like to be a bat? *Philosophical Review, 83,* 435–450.
Noë, A. (2004). *Action in perception*. Cambridge: MIT Press.
O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 883–917.
Recanati, F. (2007). *Perspectival thought*. Oxford: Oxford University Press.
Rizzolatti, G., & Sinigaglia, C. (2006). *Mirrors in the brain*. Oxford: Oxford University Press.
Rosenthal, D. (2005). *Consciousness and mind*. Oxford: Oxford University Press.
Schlicht, T. (2007). *Erkenntnistheoretischer Dualismus*. Paderborn: Mentis
Schlicht, T. (2008). Selbstgefühl. Damasios Stufentheorie des Bewusstseins und der Emotion. In: E. Düsing (Ed.), *Geist und Psyche* (pp. 337–369). Würzburg: Königshausen Neumann.
Schlicht, T. (2015). Explaining subjective character. Representation, reflexivity or integration? Commentary on Kenneth Williford. In: T. Metzinger, J. Windt (Eds.), *OpenMind*.
Schlicht, T. (2016). Kant and the problem of consciousness. In: S. Leach & J. Tartaglia (Eds.), *Consciousness and the great philosophers*. London: Routledge
Searle, J. R. (1983). *Intentionality*. Cambridge, MA: MIT press.
Searle, J. R. (2001a). *Rationality in action*. Cambridge: MIT Press.
Searle, J. R. (2001b). Free will as a problem in neurobiology. *Philosophy, 76*(298), 491–514.
Searle, J. R. (2015). *Seeing things as they are*. Oxford: Oxford University Press.
Shoemaker, S. (1968). Self-reference and self-awareness. *The Journal of Philosophy, 65*(19), 555–567.
Slors, M. V. P., & Jongepier, F. (2014). Mineness without minimal selves. *Journal of Consciousness Studies, 21*(7–8), 193–219.
Strawson, G. (1997). The self. *Journal of Consciousness Studies, 4*(5/6), 405–428.

Swanson, L. R. (2016). The predictive processing paradigm has roots in Kant. *Frontiers in Systems Neuroscience, 10,* 79. doi:10.3389/fnsys.2016.00079.
Thompson, E. (2007). *Mind in life*. Cambridge: Harvard University Press.
Tye, M. (1995). *Ten problems of consciousness*. Cambridge: MIT Press.
Van Gulick, R. (2004). Higher-order global states. In R. Gennaro (Ed.), *Higher-order theories of consciousness*. Amsterdam: John Benjamin's.
Varela, F. (1979). *Principles of biological autonomy*. Boston: Kluwer Academic.
Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind. Cognitive science and human experience*. Mass: Cambridge.
Wheeler, M. (2005). *Reconstructing the cognitive world: The next step*. Cambridge: MIT Press.
Williford, K. (2015a). Representationalisms, subjective character and self-acquaintance. In T. Metzinger & J. Windt (Eds.), *Open MIND*. Frankfurt: Mind Group.
Williford, K. (2015b). Individuation, integration, and the phenomenological subject. In T. Metzinger & J. Windt (Eds.), *Open MIND*. Frankfurt: Mind Group.
Wittgenstein, L. (1958). *The blue and brown books*. Oxford: Blackwell.
Zahavi, D. (2005). *Subjectivity and selfhood*. Cambridge: MIT Press.
Zahavi, D. (2014). *Self and other*. Oxford: Oxford University Press.
Zeki, S., & Bartels, A. (1998). The autonomy of the visual systems and the modularity of conscious vision. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences, 353*(1377), 1911–1914.