# Journal of Consciousness Studies

controversies in science & the humanities

Vol. 26, No. 3-4, March/April 2019

# Representing Ourselves: Reflexive Approaches to the Function of Consciousness

Special Issue edited by Max Jones, Takuya Niikawa and Roly Perera

4	About Authors	
8	Editorial Introduction	Max Jones, Takuya
	1	viikawa & Koly Perera
17	Attributing Awareness to Others: The Attention Schema Theory and Its Relationship to Behavioural Prediction	Michael S 4 Graziano
	Relationship to Denavioural Trediction	menuel 5.11. Graziano
38	Bayesian Frugality and the Representation of Attention	Krzysztof Dołęga & Joe Dewhurst
64	Explaining the Enduring Intuition of Substantiality: The Phenomenal Self as an Abstract 'Salience Object'	Wania Wiese
88	Consciousness (Unconsciously) Designs Itse	elf Axel Cleeremans
112	Import Theory: The Social Making of Consciousness	Wolfgang Prinz
131	Cartesian Creatures: Watching Ourselves Watching the World	d Roly Perera
155	Representing Our Options: The Perception of Affordances for Bodily and Mental Action	of on <i>Tom McClelland</i>
181	Getting the World Right: Perceptual Accura and the Role of the Perceiver in Predictive Processing Models	cy Tobias Schlicht & Elmarie Venter

#### CONTINUED

207	Minimal Self-Awareness from Within:	Anna Ciaunica &
	A Developmental Perspective	Laura Crucianelli
227	Consciousness, Language, and the Possibility of Non-human Personhood:	
	Reflections on Elephants	Don Ross
252	Are There Degrees of Self-Consciousness?	Raphaël Millière

Published in the UK and USA by Imprint Academic Ltd., PO Box 200, Exeter EX5 5YX, UK. World Copyright © Imprint Academic Ltd., 2019. No part of any contribution may be reproduced in any form without permission, except for the quotation of brief passages in criticism and discussion. The opinions expressed in the articles and book reviews are not necessarily those of the editors or publishers. The publishers allow authors to post their own versions of their papers from the Journal of Consciousness Studies on their personal websites and their own institution's archive, but only when two full years have elapsed after publication. Please note that all institutions subscribing to the Journal of Consciousness Studies get immediate free electronic access via Ingenta®. JCS is indexed and abstracted in: Social Sciences Research Index<sup>®</sup>, ISI Alerting Services (includes Research Alert<sup>®</sup>), Current Contents<sup>®</sup>: Social and Behavioral Sciences, Arts and Humanities Citation Index<sup>®</sup>, Current Contents<sup>®</sup>: Arts & Humanities Citation Index<sup>®</sup>, Social Scisearch<sup>®</sup>, PsycINFO<sup>®</sup> and The Philosopher's Index. ISSN 1355 8250 (print) ISSN 2051 2201 (online)

2

# Tobias Schlicht and Elmarie Venter

# Getting the World Right

# Perceptual Accuracy and the Role of the Perceiver in Predictive Processing Models<sup>1</sup>

Abstract: Predictive processing is often presented as a unifying framework for perception, action, and cognition, being able to explain most (if not all) mental phenomena (Hohwy, 2013; Clark, 2016): with regard to perception, the brain harbours a generative model issuing top-down expectations that are matched against bottom-up sensory feedback. Mismatches lead to error messages and model updates until the brain is 'getting it right'. The core notion of prediction error minimization commits the framework to a specification of accuracy conditions. We therefore turn to issues related to the determination of satisfaction (or accuracy) conditions discussed in the philosophy of perception. In particular, we rely on work by Recanati (2007) who shows that the accuracy conditions of perceptual content are partly determined by the intentional mode and the perceiver (or self). It is important to notice that the self can enter the specification of accuracy conditions in two ways, namely as subject or as object. Even if

Correspondence:

Tobias Schlicht and Elmarie Venter, Ruhr-University Bochum, Universitätsstraße 150, 44801 Bochum, Germany. Email: tobias.schlicht@rub.de; elmarie.venter@ruhr-uni-bochum.de

Journal of Consciousness Studies, 26, No. 3-4, 2019, pp. 181-206

<sup>&</sup>lt;sup>1</sup> Work on this paper was supported by funding from the Volkswagen Foundation for the research project 'Situated cognition: Perceiving the world and understanding other minds' and by the German Science Foundation for the Research Training Group on Situated Cognition (GRK 2185/1).

we do not perceive ourselves explicitly as an object, we always implicitly represent ourselves as subject. We discuss whether and how these two ways of self-representation can be respected in the predictive processing framework.

#### 1. Introduction

Proponents of *predictive processing* (PP) approaches claim to have found 'the first truly unifying account of perception, cognition and action', combining important insights from earlier frameworks in cognitive science, e.g. connectionist ideas, work in robotics, and work on dynamic self-organization and on the embodied mind (Clark, 2016, pp. 2, 10). But although the framework has already been successfully applied in the explanation of a range of cognitive phenomena, it also puts a number of constraints on the kind of explanation it provides. This forces us to conceive of phenomena like consciousness, selfconsciousness, and self-representation in a particular way. In this paper, we investigate some of the implications for the integration of self-representation into the framework. Central to our discussion is a differentiation between two distinct aspects or dimensions of self, namely, self-as-subject and self-as-object. We illustrate this distinction via the ways in which the self can enter the determination of the accuracy conditions of a perceptual state. Given how these perspectives on the self have been framed in philosophical debates, their adequate integration into the framework seems to be hindered by certain constraints inherent to the framework. Thus, our claim is that if we take the distinction to be important then it is a substantive challenge for the PP framework to account for both dimensions of the self. But if we take the PP framework at face value then we might have to give up some characteristics of self-representation that have been defended in philosophical debates.

We proceed as follows. In Section 2, we provide a sketch of central theoretical elements of the PP framework and highlight some constraints for the explanation of cognitive phenomena. The general outline of the framework forces its proponents to conceive of mental phenomena in a certain way to the exclusion of other ways. This is particularly interesting for phenomena like consciousness and self-awareness. Apart from some general commitments, there is room for dissent, especially concerning the alleged commitment to mental representations (Clark, 2015; Gładziejewski, 2016; Kiefer and Hohwy, 2018; Hutto and Myin, 2017; Bruineberg, 2017). Given the focus of

this Special Issue on Representing Ourselves, we focus on versions of the PP framework which provide explanations in terms of representations. We are aware of alternative variations, but given that enactivists like Hutto and Myin (2017), for example, eschew representations tout court they would most certainly approach phenomena related to the self very differently. For reasons of space we thus simply assume a representationalist reading of the framework in this paper. Section 3 then introduces the distinction between self-as-subject and self-asobject via a discussion of the ways in which the self can enter the determination of accuracy conditions of perception. Many proponents of the PP framework introduce it as an account of perception on the level of neurocognitive architecture, i.e. implicating that it can explain how the brain can 'get the world right' (Hohwy, 2013, p. 2). We take this claim at face value and confront it with an analysis of the conditions of veridical perception in a representational theory of perception. Here, we draw upon seminal work by Searle (1983; 2015), Siegel (2010), and Recanati (2007), resulting in the claim that, properly understood, the accuracy conditions for a given perceptual experience involve implicit reference to features of the spatiotemporal aspects of the perceiver. In Section 4, we return to the PP framework by considering ways in which it can integrate the two dimensions or aspects of self in its account of veridical perception. This involves an extensive critical look at some recent PP approaches to the self, in particular the account developed by Hohwy and Michael (2017) and by Newen (2018).

### 2. Predictive Processing (PP)

#### 2.1. The basics

The core idea of the PP framework is that perception is a process constituted by a delicate balance and interplay of top-down knowledge and bottom-up sensory signals (Hohwy, 2010, p. 135). As Frith puts it: 'When we perceive something, we actually start on the inside: a prior belief, which is a model of the world in which there are objects in certain positions in space' (Frith, 2007, p. 126). The brain maintains a generative probabilistic model of the world, which yields (top-down) expectations or predictions about (the causes of) sensory input, which are in turn compared to the actual (bottom-up) sensory input. To see a structured world is to use existing knowledge about the state and structure of the world to generate models of sensory perturbations from the top down. In this sense, the brain is burdened with unavoidable uncertainty because there is no one-to-one mapping of worldly causes and sensory signals (Gładziejewski, 2016, p. 561). As Clark (2013, p. 183) puts it, 'the task of the brain, when viewed from a certain distance, can seem impossible: it must discover information about the likely causes of impinging signals without any form of direct access to their source'. The top-down predictions are selected based on statistical estimation. The statistical estimates function as neural representations of the current state of the world (Wiese and Metzinger, 2017, p. 3). The predictions meet the incoming signals and can be either accurate or inaccurate. Once the prediction is accurate, nothing more needs to be done, and the generative model of the world that issued the respective top-down predictions can count as accurate with respect to the state of the world. The model can be said to be 'getting the world right'. If the predictions are inaccurate, prediction error (defined as the mismatch between top-down and bottom-up information) is transmitted to optimize the estimates (models) of hidden causes. A more formal picture of PP involves 'random variables and other mathematical functions which suggests that the brain computes the functions described by those models, and that neural states represent values of variables (or numbers, vectors, and matrices)' (Wiese, 2017, p. 720). Relying on Egan's (2014) account of mathematical and cognitive content, we agree with Wiese's (2017) claim that the essential components mentioned above are mathematical contents of representations in a PP system.

These components work together to fulfil the basic function and primary objective of the whole system, which is to minimize the deviation of top-down predictions from bottom-up sensory information. Indeed, in this framework, *all* cognitive phenomena are conceived as serving the same function:

Perception, action, and attention are but three different ways of doing the very same thing. All three ways 'must be balanced carefully with each other in order to get the world right. The unity of conscious perception, the nature of the self, and our knowledge of our private mental world is at heart grounded in our attempts to optimize predictions about our ongoing sensory input. (Hohwy, 2013, p. 2)

Similarly, Friston introduces his formulation of the PP framework in terms of the free-energy principle as a 'unified brain theory' (2010), and Clark praises the 'fundamental attraction of these accounts' which lies in 'their ability to offer a deeply unified account of perception, cognition, and action' (2013, p. 186). Consequently, Hohwy and

Michael (2017, p. 366) ask: 'If all we ever do is minimize prediction error, what could it mean to have a *self*?'

The PP framework posits some essential components, i.e. hierarchical generative models, estimates, predictions and prediction errors, precision estimates, and corrections or updates of the generative models (Wiese, 2017, p. 716). Moreover, the brain makes use of all possible channels to serve its goal to 'make better models' (Frith, 2007, p. 127). Given its outlook, the framework constrains the range of possible theories of mental phenomena and it prescribes the shape of such theories. For example, if proponents of the PP framework intend to provide a theory of consciousness, then the aspect of consciousness that it can illuminate will have to be explainable in terms of function (contra Chalmers' notion of phenomenal consciousness), namely, the function of prediction error minimization, it will also be constrained by attention — in line with Dehaene's (2014) global workspace theory but in contrast to Block's (2007) overflow theory and it will be based on a hierarchical generative model which presumes a 'cognitive' and mechanistic theory of consciousness (e.g. Dennett, 1991). Similarly, one might expect that the framework's contribution to an explanation of self-consciousness will proceed naturally via the integration of self-representation into the hierarchical generative model (see Newen, 2018). It is not our aim here to formulate and defend a complete list of such constraints. Rather, we want to focus on aspects of self-representation and evaluate the prospects to account for them in the terms provided by the framework.

Representation implies the possibility of misrepresentation. This also applies to self-representation. In the next section, we highlight two aspects of this phenomenon in the context of discussing the accuracy conditions of perceptual states.

# 3. Perceptual Content and the Determination of Accuracy Conditions

Representational approaches to perception conceive of perception as a form of intentionality (e.g. Crane, 2001; Siegel, 2010; Searle, 2015). If the representational content of the perceptual state presents the world in a way that coincides with reality, then the perceptual representation is accurate. If not, it is inaccurate. On such accounts, illusory perception and hallucination are conceived as cases of misrepresentation. If I see the two lines in the Müller-Lyer illusion as different in length, while they are actually equal in length, my perceptual state

misrepresents the length of the lines. Naturalistic approaches to intentionality capture this by positing some physical mechanisms or vehicles which function as representations. That is, naturalists assume that intentionality can be naturalized, contra Brentano's dictum that intentionality 'is characteristic exclusively of mental phenomena. No physical phenomenon exhibits anything like it' (Brentano, 1874/1995, p. 88). And such naturalization involves the tasks of demonstrating how some physical entity can represent something, i.e. stand in for something else. That is, it involves formulating 'the set of conditions that make it the case that something is functioning as a representational state' and 'the set of conditions that make it the case that something functioning as a representation has the specific content that it does' (Ramsey, 2016, p. 4). Ramsey (2007) calls this the 'job description challenge' faced by defenders of mental representations. The PP framework is intended as providing neurocomputational principles that can illuminate how the brain enables perception of the world, i.e. by matching inputs with prior expectations (Clark, 2013; Hohwy, 2013). The link between representationalism (or intentionalism) about perception and a representationalist version of the PP framework is thus provided by the additional constraint of naturalism. Intentionalists about perception without any shares in naturalism or physicalism (e.g. Crane, 2001) need not worry about this link. But all proponents of the PP framework are naturalists of sorts, aiming at a mechanistic explanation of cognitive phenomena. With respect to the specification of the mental representations playing a role in the PP explanation of perception, there is again room for dissent. While Kiefer and Hohwy (2018, p. 3) defend a 'mature internalist semantics centered on the notion of structural representation', Clark states that the predictive processing framework 'involves representations that are action-oriented through and through', aiming 'to engage the world' rather than mirroring it in a completely action neutral way (Clark, 2015, p. 4).

At this point, it was only our aim at forging the link between the philosophical debate about the intentional character of perception and the representationalist version of the PP framework. To the extent that proponents of the PP framework aim at providing an account of how the brain achieves 'getting the world right', i.e. yields a match between the content of its predictions and expectations and the actual (causes of the) sensory inputs, such accounts presuppose not only accuracy conditions for perception but also a way how such accuracy conditions are determined. That is, we must have a look at the factors that enter the determination of accuracy conditions of perception.

According to the intentional theory of perception, perception has the same intentional structure as any other intentional state:<sup>2</sup> a *subject* is directed at an intentional *content* via some intentional *mode*. Modes are perceiving, imagining, remembering, believing, etc. Mental phenomena with the same mode can be directed at different contents and mental phenomena with different modes can be directed at the same content. I can see the coffee mug in front of me, I can remember and believe it to be standing there. I can also perceive the computer standing next to the coffee mug. In a perceptual state, an intentional object, the coffee mug say, is presented in some way or other in the content of the mental state (Crane, 2001). But given that it is one key feature of intentionality that the object presented in the content need not exist in reality, the intentional theory of perception can adequately account for hallucinatory and illusory experiences as cases of misrepresentation. In such cases, the mental state presents the world to be some way that deviates from the actual state of the world. A perceptual experience is veridical or accurate in case the world is as the content presents it to be (Searle, 1983; Siegel, 2010).

One dispute among proponents of intentional or representational approaches to perception concerns the determination of what Searle (1983) calls the 'satisfaction conditions' of the perceptual experience, or what we may call its accuracy conditions.<sup>3</sup> Are these conditions determined by the representational content or by the phenomenal character of the experience? Are there further factors that partly determine the state's accuracy conditions or are they fully determined by the content? Presupposing familiarity with the general outline of the intentional approach to perception (*cf.* Crane, 2001; Siegel, 2010), we would like to focus on these questions in order to connect this debate to the claims of PP accounts of perception.

<sup>&</sup>lt;sup>2</sup> When we speak of intentional *states*, we use this as shorthand for states, events, and processes. We want to remain noncommittal with respect to which of these ontological descriptions best captures perception.

<sup>&</sup>lt;sup>3</sup> In order not to presuppose that perceptual content is propositional, but to allow for perceptual content to be non-conceptual, we prefer the weaker notion of accuracy condition to the notion of a truth condition. For the purposes of this paper, it is only crucial that perceptual states, whether consciously experienced or not, are representational states with content. We also remain neutral with respect to the question as to whether only low-level or also higher-level properties can be perceived (*cf.* Siegel, 2010).

#### 3.1. Conscious and unconscious perception

In his recent discussions of perception, Searle argues that the 'phenomenological features' of what he calls the 'subjective visual field determine the conditions of satisfaction of the visual experience' (Searle, 2015). The 'raw phenomenology', i.e. how it seems to the perceiver, 'just is the presentation of those conditions of satisfaction' (ibid., p. 114). His intuition is that 'for something to be red in the ontologically objective world is for it to be capable of *causing* ontologically subjective visual experiences like this' (ibid., p. 122). And likewise, for something to be the object of a perceptual experience is for it to be experienced as the cause of the experience, Searle argues. Thus, a perceptual experience as of something red 'necessarily carries the existence of a red object as its condition of satisfaction' (ibid., p. 123). This determination of accuracy conditions by the phenomenology of experience seems to be in tension with Searle's assumption of direct realism, i.e. the claim that we can directly perceive the world as it is. Why is there a tension? If the phenomenal character of my perceptual experience determines its accuracy conditions, and if we make the existence and being-thus-and-so of a perceptual object (being red, say) dependent on its capacity to cause experiences of a certain kind in a perceiver, then the allegedly objective features of the object depend on the nature, peculiarities, and limitations of the perceiver's perceptual apparatus. Creatures with this kind of apparatus (humans, say) will experience the object in one way, while creatures with a different perceptual apparatus (bats, say) will experience it quite differently. Which phenomenal character determines which features the object in the world really possesses? This seems to be a problem for Searle's account.

Another problem concerns the possibility of unconscious or subliminal perception. If phenomenal features determine accuracy conditions, then it seems impossible for such representations to have accuracy conditions at all. To use a familiar example, patients suffering from blindsight (Weiskrantz, 1986) claim to be blind with respect to a certain area of their visual field. Yet, when presented with an object in this area and forced to guess what is there, a high percentage of their judgments are accurate, at least far above chance. Yet, they still claim not to be phenomenally conscious of the objects presented. On Searle's account, their unconscious perceptual states cannot be considered either as accurate or inaccurate. But in order for researchers to estimate that the blindsight patients are getting it right

in most cases, the unconscious information processing underlying their judgments must be subject to accuracy conditions. Searle acknowledges that these patients are 'getting something very much like the same information' as healthy subjects, but he does not discuss the consequences for the accuracy conditions of such information processing (Searle, 1997, p. 200). Since in such cases there is no phenomenology, the phenomenal features of perception cannot determine the accuracy conditions of such states. This observation carries over to non-pathological cases of subliminal priming. Brief exposure to certain stimuli modulates performance in subsequent cognitive tasks like recognition of subsequently presented stimuli (Bargh and Chartrand, 2000). Accurate representation of the subliminal stimuli is presupposed in our reasoning about the effect of such information processing. Thus, since cognitive neuroscientists rely on subliminal information processing to be more or less accurate in different situations, and since cases such as blindsight provide evidence for accuracy and inaccuracy on the subliminal level, Searle is simply wrong that the accuracy conditions must be specified in terms of the phenomenal features of perceptual states. Why is this important? If Searle were right, then this paper would be futile. Since we want to connect the claim that the brain is in the business of matching topdown expectations to incoming sensory input in the service of prediction error minimization to the issue of how accuracy conditions are determined according to the representational theory of perception, it is important that accuracy conditions do not only apply on the level of phenomenology.

## 3.2. Factors in the determination of accuracy conditions

Often, the intentional or representational account of perception is presented in a way suggesting that the representational content of a given mental state determines its satisfaction conditions. The belief 'John believes that snow is white' is taken to be true if snow is indeed white, and Susanna's visual perception of the coffee mug is accurate if indeed there is a coffee mug. Siegel (2010) calls this the Content View: 'The notion of representation is tied to the idea that experiences have contents, where contents are a kind of condition under which experiences are accurate, similar in many ways to the truth-conditions of beliefs' (*ibid.*, p. 4). But Searle (1983; 2015) emphasizes that the accuracy condition of my perceptual experience of the coffee mug in front of me is not simply given by the proposition that there is a coffee mug in front of me. On Searle's view, perception is self-referential in the sense that the perceptual state is only satisfied if exactly this coffee mug in front of me right now, not any other mug, caused the perceptual state. Searle highlights the importance of the specific position of the perceiver and their relation to the coffee mug with respect to the accuracy conditions of the resulting perceptual state.

Recanati (2007) offers a critical appraisal of Searle's account. He disagrees with Searle's proposal that content alone determines the conditions of satisfaction of an intentional state and proposes that we have to take into account further factors, reminding us of the other constituents of intentional states, namely subject and mode. In contrast to Searle, Recanati argues that the self-referentiality of perception is not part of the content but rather given by the intentional mode, in our case, perceiving:

That the state of affairs represented [e.g. there being a coffee mug there] causes the representation of that state of affairs is a condition that has to be met for the representation in question to count as a *perception* (rather than, say, an expectation). It follows that the self-referential condition is determined by the perceptual mode of the state, not by its content. (*ibid.*, pp. 131–2)

This is intuitively plausible, since my memory of there being a coffee mug in front of me (yesterday), my perception of the coffee mug right now, and my expectation that there will be a coffee mug in front of me tomorrow share the content *that there is a coffee mug in front of me*. The content of such different experiences can be simply stated, and yet, the various modes require different spatio-temporal features regarding the relation of perceiver and coffee mug to be in place. Only the perceptual mode requires a direct causal relation between perceiver and mug *right now* at the time of having the experience. The proposition *that there is a coffee mug 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. (*ibid.*, p. 135)

On Recanati's view, the necessary reference to the time and place of perception and to the causal relation obtaining between perceiver and intentional object are 'extra pieces of information' which are not explicitly represented in the content like the coffee mug, say, is represented. Rather, they enter the accuracy conditions only implicitly, via the '*functional role* of that representation, determined by its mode. Qua perception, the representation is caused by the situation around the subject, and it serves to guide the subject's action in that situation' (*ibid.*, p. 145). As this functional description indicates, in addition to the *mode* a further factor enters the determination of accuracy conditions, namely the perceiver or subject of experience. The next section is thus devoted to an analysis of the sense in which the subject partly determines the accuracy conditions of experience.

#### 3.3. Self-as-subject vs. Self-as-object

We could add more detail to the example of seeing the coffee mug in front of me by considering that, once I have seen it, I reach out my hand and grasp the mug in order to drink from it. I am then taking advantage of the affordances (Gibson, 1979) provided by my immediate environment, namely, the coffee mug having a particular design which allows me to grasp it with my hand. Put in terms of the PP framework, this is active inference: by issuing an action, the brain holds a given hypothesis constant and interferes with the state of the world in order for it to match the hypothesis rather than updating the hypothesis in order to match the state of the world.

We can envision an analogous contrasting situation in which I am acting like this while at the same time looking at myself in a mirror. Only in the latter case am I myself explicitly represented in the visual content, via seeing myself in the mirror, whereas in the former case only the coffee mug is explicitly represented in the content. What remains constant across both situations is that I am provided with proprioceptive and kinaesthetic information about my grasping from the inside. Since I am explicitly represented in the content of my visual experience in the mirror case, I can misrepresent myself as the intentional object of that very visual experience. I could be mistaken with respect to whom I am seeing in the mirror holding the coffee mug. It could be someone standing next to me who was also drinking coffee and reflected in the mirror. In such a case, the information from the outside would be at odds with the information about grasping I gain from the inside via proprioception. One might argue that since I myself am represented as an intentional object, the accuracy conditions of my perceptual state in the mirror case are partly determined by how I am represented while this is not the case in the other situation where I only have proprioceptive information about my grasping. No misidentification is possible in the case in which I gain information about grasping the mug only from the inside. Recanati, focusing on the subsequent perceptual judgment, explains this as follows:

The explanation of this phenomenon is straightforward, and it involves the distinction between mode and content once again. Information gained on the *proprioceptive/kinesthetic mode...* can only be about one's own body. This is a contingent fact, but it is a fact nonetheless. So the (explicit) content of the bodily experience is not a complete subjectpredicate proposition involving a certain person (myself) and a property... The content is a relativized proposition, true at a person, and the internal mode determines the person relative to which that relativized content is evaluated: myself. (Recanati, 2007, pp. 147–8)

But the subject of perception also enters the determination of the accuracy conditions in the case where she is not explicitly represented. Since the *mode* of proprioceptively perceiving myself holding the coffee mug already determines who is grasping the mug, the perceiving subject need not be explicitly represented in the perception and enters the accuracy conditions only implicitly. What matters is how the perceptual scene presents itself to the perceiver, although the subject and her situation are not themselves elements of the fact represented, namely that there is a coffee mug or that she is grasping the mug. Appealing to Perry's (1986) work, Recanati considers the subject as an 'unarticulated constituent'. To clarify, the subject of the perceptual state need not be explicitly mentioned in the content because it is already taken care of by the perceptual *mode*, indeed by its being in an intentional state. In the case under consideration, perceiving and grasping the coffee mug, I am only involved in the perceptual experience as a perceiver or subject if the source of information is proprioceptive, i.e. from the inside. This is not possible when we perceive others. We can see someone else holding a coffee mug but in such cases we always only gain information from the outside

Since information from the inside is also present in the case where I see myself in the mirror, I am part of the scene twice over, so to speak. I am present in two senses or along two dimensions, namely as perceiving subject and as perceived object. Consequently, I enter the content of this perceptual experience and the determination of its accuracy conditions in more than one way. Rather than merely being the perceiving subject, I am also part of the scene: the perceiver perceived. I am at the same time subject and object of the perceptual state. These two ways in which I can enter the scene differ importantly and have

consequences for a variety of phenomena, ranging from peculiarities of self-reference and self-awareness to features pertaining to the use of the first-person pronoun 'I' and epistemic issues to do with selfidentification.<sup>4</sup> For the purposes of this paper, we are bracketing the phenomenological dimension of perceptual experiences and focus on its representational dimension, thus we leave aside any discussion of consciousness and self-awareness. What is relevant, though, is the way Wittgenstein (1958), Shoemaker (1968), Evans (1982), Perry (1986), and others used the distinction of self-as-subject and self-asobject in discussions of the apparent impossibility to error through misidentification in the context of self-reference. When I have a bodily sensation, for example, as when I am feeling hungry or feel my hand grasping the cup, I might be mistaken as to what exactly I am feeling or holding. It might not be hunger but some related feeling associated with my gut, and it might not be a coffee mug but a bottle of beer. But since the relevant bodily sensation presents itself to me immediately, in a first-person subjective kind of way, I cannot be mistaken that it is me who is having the sensation (Zahavi, 2014; Schlicht, 2017). The question simply does not arise, or so many have argued. This seems intuitively true. Once I feel hungry, I am immediately aware of myself as the one who is hungry. I do not have to wonder and ask myself whose hunger that might be. The 'I' in such cases is implicitly given with the feeling of hunger, and thus the reference of the first-person pronoun in the ensuing self-ascription of the feeling is guaranteed simply because the self is an unarticulated constituent, as Perry called it.

Shoemaker (1968) argues that since no act of identification is needed, I cannot be mistaken about myself in such cases. He argues that it is a necessary aspect of situations like this one. But whether this immunity to error through misidentification is a necessary or contingent feature of bodily sensations is up for debate. Langland-Hassan (2015) discusses an empirical case of craniopagus twins, Krista and Tatiana Hogan, whose brains are connected in such a way that they share a portion of the thalamus. The thalamus has been shown to be an important relay station for sensory information to become conscious with every sensory system having a thalamic nucleus that sends sensory signals to the associated cortical areas (see *ibid*.). This

<sup>&</sup>lt;sup>4</sup> Not all of these effects of the two dimensions of self are important for the present discussion. See a comprehensive review in Longuenesse (2017).

connection enables the twins to get access to — and become aware of — each other's experiences such that the question of whose sensation one of them might be introspecting can reasonably arise. Langland-Hassan concludes that Shoemaker's claim was too strong, and that the impossibility to error can thus only be a contingent feature of such experiences. In the context of this paper, we do not want to engage in this discussion about necessity and contingency but are content with the claim that, in *typical* cases, the dimension of self-as-subject need not be questioned but is taken care of by the *mode* of the experience. But as we will see below, this may become problematic in the face of the PP framework such that we may have to give up on the claim that this is a special feature of self-representation.

To sum up: so far, we have presented the PP framework as a unified account of the neurocognitive architecture underlying perception, action, and other cognitive phenomena. The core notion of prediction error minimization commits the framework to the specification of accuracy conditions for perception (Section 2). In Section 3, we argued that accuracy conditions are determined independently of the phenomenal features of perception and that for such determination not only the content but also the intentional mode and the subject are crucial. As we have seen, the subject of experience can either only function as an implicit (unarticulated) constituent of the content or it can also be explicitly represented *as* an object like any other object. Thus, in the following concluding section we refer to this distinction using the terminology of 'self-as-subject' and 'self-as-object' and discuss how the versions of the PP framework which are committed to accuracy conditions for perception could deal with it.

# 4. The Subject in Predictive Processing Accounts of Perception

The role of the subject for intentional states is often underdeveloped in theories of intentionality. For example, even in Crane's (2001, p. <sup>/</sup>31) extensive discussion of the intentionality of perception and thought, 'the nature of the subject is not something which is within the scope of' his book. Searle (2001, p. 75) admits that he did not see any need for an account of the subject of perceptual experience but accepts 'with the greatest reluctance' that it is at least a necessary element in a theory of practical reason and rational action. Within the context of

the PP framework, only a few recent contributions address the issue of the subject or self,<sup>5</sup> but they do not yet connect with debates on intentionality and the representational structure of perception (Moutoussis *et al.*, 2014; Limanowski and Blankenburg, 2013; Frith, 2007; Hohwy and Michael, 2017). An adequate account is still a desideratum. Thus, it is not surprising that those contributions, which address the subject or self-representation in the context of predictive processing, appeal to existing work on the self from other contexts, in particular to theories involving self-models. Limanowski and Blankenburg (2013), Hohwy and Michael (2017), and, to some extent, Seth, Suzuki and Critchley (2012) rely on Metzinger's (2003; 2009) notion of the phenomenal self and his explanatory tool of a selfmodel.

### 4.1. Body- and agent-models

Metzinger (2003; 2009) claims that there are no selves and that no one ever had or was a self. What we need to account for though, he argues, is the phenomenal experience of being a self, i.e. our becoming and being self-aware. The aim of his theory is to show how this feeling can arise once we treat it as the result of complex information processing in the central nervous system. Thus, all we need for an explanation of this feeling is a 'self-model': a coherent inner model of the agent (or more generally, system) as a unit and totality. A self-model is a representation generated by the brain, spanning information from all available sources including — in our case — all information about the body and its inner milieu. Here, Metzinger relies heavily on Damasio's work on how the brain monitors and regulates the overall (homeostatic) state of the body in order to assure that it remains within the bounds needed for survival (Damasio, 1999; 2010).<sup>6</sup> The Minimal

<sup>&</sup>lt;sup>5</sup> For the purposes of this paper, we use the notions of self and subject interchangeably. Some authors prefer to talk of the 'self', others take the notion of 'subject' to be less committal in metaphysical respects. We are aware of the conceptual differences and problems in the context of using notions like 'subject' and 'self'. Although we prefer the weaker notion of a subject of experience because it has — traditionally — weaker metaphysical implications, we follow other contributors to the debate, for example Hohwy and Michael (2017), and adopt their use of the notion of self. We would like to note though that neither notion is supposed to imply any stance on the metaphysical nature of the self (see Schlicht, 2016; 2017, on this issue).

<sup>&</sup>lt;sup>6</sup> On the connection between Metzinger's self-model theory and Damasio's work on the self see also Schlicht (2018).

Phenomenal Self (MPS) is identical to the conscious experience of being a self, i.e. of '...being a distinct, holistic entity capable of global self-control and attention, possessing a body and a location in space and time' (Blanke and Metzinger, 2009, p. 7).

Since we have only restricted space here, we cannot outline Metzinger's account in its fullest detail, but instead we focus on how it can be utilized to account for the self within the PP framework. Limanowski (2017), for example, notes the striking commonalities of the self-model theory (SMT) with the PP framework:

Most notably, SMT suggests a hierarchy of phenomenal self-modeling, ranging from pre-reflective, 'minimal' self-representations like a firstperson perspective, body self-identification, or spatio-temporal selflocation... to complex cognitive self-representations. Such selfmodeling can be well-described in terms of active inference, whereby the 'self' (in all its cognitive-to-minimal dimensions) is a sophisticated hypothesis about the organism's environment which is generated by the brain's hierarchical generative model, and which tries to maximize evidence for its own existence. (Limanowski, 2017, p. 9)

Newen (2018) also appeals to self-models within his pattern account of the self, arguing that sensory input simultaneously produces representations of the external world and of the cognitive system itself. While grasping the coffee mug, he argues, neural signalling carries not only information about the mug, but also egocentric information about my hand, its grip, position, and movement. Self-related and objectrelated information are but two sides of the same coin. But he emphasizes that neither of them has priority since they are carried by the same neural signal.

Hohwy and Michael (2017) also connect their discussion to Metzinger's self-model theory. They claim (a) that 'the self is an inferred model of endogenous, deeply hidden causes of behavior', and (b) that 'accurate self-models arise and are sustained as a natural consequence of humans' skill in modeling and interacting with each other' (*ibid.*, p. 363). That is, the use of agent-models shapes children's self-development because through active inference infants 'approximate the intentional agents they take others to be' (*ibid.*, p. 381).<sup>7</sup> To illustrate how the distinction between aspects of self-

<sup>&</sup>lt;sup>7</sup> To be fair, in personal communication John Michael distanced himself from the strong claim that the self-model is a 'consequence' or 'result' of the application of agentmodels to others. It is possible, he argues, that infants come equipped with a very

representation as elaborated in the last section affects the PP framework, we will focus on their first claim but will also make some critical remarks on the second.

Hohwy and Michael (2017, p. 369) appeal to Metzinger's notion of a self-model which is integrated in a whole model of the world. This world-model, according to the PP framework, consists of a repertoire of beliefs about the hidden (exogenous) causes in the world and their interactions. Causes within the body are called endogenous, and include 'an integrated net of character traits, biases, reaction patterns, affections, standing beliefs, desires, intentions, base level internal states, and so on' (ibid.). Given that some of the causes of sensory changes pertain to the agent's body, it is quite natural that the internal model represents bodily causes, including 'the more deeply hidden internal causes of the agent, which interact with each other (e.g., fear plus hunger gives one trajectory of sensory input, fear plus pain gives another) and in turn with worldly influences (e.g. fear and presence of tigers vs. fear and no presence of tigers)' (ibid., p. 368). An agent needs to track motor commands in order to draw a distinction between changes in the environment that result from her own action and those changes that are due to other factors in the environment (Hohwy, 2007, p. 2). Thus, modelling such deeply hidden causes is necessary for the successful minimization of prediction error. Hohwy and Michael propose to 'conceive of this internal model of endogenous causes as a representation of the self, suggesting that 'agents model the self as a hierarchy of hidden, endogenous causes and... that the self is identical to these causes' (2017, p. 369). Consequently, 'the self is just one set of causes' and can be modelled like anything else out there in the world can be modelled as a cause (*ibid.*, p. 374). This model of causes is characterized by a hierarchy of layers that mirrors the timescales of action, ranging from (milli-)seconds in the case of reaching for a coffee mug, say, to hours, days, months, etc. in cases of travelling, studying, etc.

In light of their claim that 'the brain is fundamentally an organ for prediction error minimization' (*ibid.*, p. 363) and 'prediction error is the one quantity the brain can compute' (*ibid.*, p. 376), Hohwy and Michael develop their account of the self only in the context of generative models and updating in response to prediction errors. In

shallow self-model (perhaps a minimal body-model), and that social cognition only shapes and enriches that bodily self-model.

accordance with the explanatory constraints set by the PP framework that we mentioned in the second section, this is consequential. And if this is right, then any aspect of the self that is not subject to prediction error and updating is bound to be left out of this conception. That is, if the dimension of self-as-subject is grounding the (typical) immunity to error through misidentification, escaping the need for correction or updating altogether, then a focus on self-models which are prone to prediction errors cannot deliver the whole picture. In the PP framework of self-models, as developed by Hohwy and Michael, only the dimension of self-as-object can be respected and considered, and only what is explicitly represented in the self-model can be matched against sensory feedback. Thus, prediction error will only concern the self as it is represented as intentional object in a given perceptual state. Similarly, Moutoussis et al. (2014) and Limanowski and Blankenburg (2013) consider the self only as represented in a self-model and, given the prediction error minimization framework, updating a model is all that ever happens in terms of perception, cognition, and action. According to Hohwy and Michael, 'the body is nothing special' and 'representation of the body is nothing special either; it is just one among many causes that get represented in the internal model used for prediction error minimization' (2017, p. 368). And yet, specifying the exact relation between self and body is apparently not so easy, especially since they attempt to defend a 'more metaphysically robust' account of self-representation than Metzinger has been willing to grant.

According to Hohwy and Michael, 'the self is modeled in perceptual inference, as the system learns what its own self is' (ibid., p. 375). And it does that by learning patterns of active inference, i.e. action. This fits well with their second claim that we develop accurate selfmodels only as a consequence of applying agent-models to others. This is a variant of the old debate between subjectivity-first accounts and intersubjectivity-first accounts in the philosophical and psychological literature. The latter position holds that we only acquire a concept of self and consecutive self-knowledge after being able to represent and interact with others. But what we have identified above as the subject dimension of the self is necessarily systematically prior to any intersubjective generation of a self-model. For, once I apply an agentmodel to someone else in order to make sense of his or her behaviour, the self-as-subject is already presupposed as author of this cognitive act. Hohwy (2007, p. 4) seems to acknowledge this in previous work when he states: '[i]t seems that there is an immediate, pre-reflective experience of minimal self or "mineness" as one voluntarily performs bodily movements (at least in non-pathological cases). A central aspect of this experience of mineness is the sense of agency such that I experience the movement as intended, initiated and controlled by me.' If we assume that pre-reflective experience presupposes (implicit) self-representation, then this implies the primacy of self-representation in the execution of voluntary action, also for Hohwy. This comes close to an acknowledgment of the priority of self-as-subject as it has been applied in discussions of pre-reflective self-awareness (e.g. Zahavi, 2014).

Finally, Hohwy and Michael's restricted view of the self (as object) is also evident in their claim that the self-model develops '...as a natural consequence of humans' skill in modeling and interacting with each other' (Hohwy and Michael, 2017, p. 377). When we observe others and apply an agent-model to them in order to make sense of their actions, we represent them explicitly, as peculiar intentional objects, namely embodied agents in the world; we cannot represent them in the same self-as-subject sense as we implicitly represent ourselves in all our cognitive acts. This is an important asymmetry between representing ourselves and representing others even though we can (perceptually) represent them as agents (Kinzler and Spelke, 2007, p. 258). Holding that the self-model only arises as a result of the application of agent-models to others is misleading and implausible in a twofold sense: first of all, the asymmetry between representing self and other is not taken seriously. Indeed, a neutral agent-model that can be applied to any entity that exhibits self-propelled biological motion does not capture the special function and significance that the selfmodel plays for the embodied agent herself in contrast to the application of an agent-model to others. Secondly, the claim is simply implausible with regard to the intimate role that the body-model plays as the 'fundamental' element of the self-model (see footnote 8). This body-model is arguably present from the very start. It is constantly updated and enriched and is of course also central for the execution of action. Active inference is integral to this picture as it is suggested that children update their agent-models through interaction with the world and other agents; this process continues into adulthood and ultimately determines our folk psychology (Friston, 2012). With this outline and criticism of Hohwy's and Michael's position in mind let us return to the dimension of self-as-subject.

# 4.2. Agency, affordances, and the self-as-subject

If the reasoning regarding the dimension of self-as-subject in Section 3 is on the right track, then for all perceptual states that do not explicitly represent the self as simply one among other intentional objects, no prediction error will affect the self (as implicitly represented as perceiver or agent). But as we noted above, the self is always implicitly represented as providing the anchor point from which the perceptual scene is perceived. Frith refers to the subject as 'the invisible actor at the center of the world' (2007, p. 109) and argues that the 'brain embeds us in the world and then hides us' (ibid., p. 100). But he does not elaborate on the philosophical implications of the prediction error minimization framework for the conception of the subject of experience. This echoes Recanati's observation: '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' (Recanati, 2007, pp. 147-8). We can interpret the absence of any act of identification as the absence of inference regarding the selfas-subject. But if Recanati is correct, the PP framework should account for the dimension of self-as-subject as well, which seems difficult given the constraints set by the framework and the fact that neither of the authors mentioned above take it into account.

This seems to lead us into a dilemma. Either we defend the special dimension of self-as-subject and may not be able to account for it in terms of the PP framework. Or we opt for the strength of the PP framework but then have to give up on the special status of the dimension of self-as-subject. But the dilemma is merely apparent; we can think of it more softly as two options: accepting the second horn constitutes the option of downplaying the dimension of self-as-subject since it does not make self-representation significantly different to representation of the external world. All aspects of self are then subject to prediction error minimization. Consequently, all self-related information can also be misrepresented, such that the alleged immunity to error through misidentification does not find any grounding in (implicit or explicit) self-representation. While Hohwy and Michael (2017) do not consider the self-as-subject at all, Newen (2018) is aware of the notion and sums it up as follows: it involves 'the self as (1a) bearer of one's sensations and perceptions, (1b) as the agent of one's action, (1c) as the owner of one's body parts, (1d) as the center of one's visual perspective and (1e) as the center of one's cognitive perspective (including experiencing oneself as the author of one's thoughts). Furthermore, *the self as subject* is characterized by its immunity to error through misidentification' (Newen, 2018, p. 5). But Newen allows for this information to be incorrect in the sense that 'in rare cases the proprioceptive information is evaluated incorrectly, e.g. one may think one's legs are crossed when they actually are not' *(ibid.*, p. 5). Of course, one could argue that the case of craniopagus twins mentioned above supports this option: *all* aspects of the self are subject to prediction error, although in typical cases the dimension of self-as-subject is not affected. In this case, Recanati's analysis would simply be rebutted by holding that all information regarding the subject can be inaccurate and the reference of the subject of perception is not always guaranteed by the framework. Newen seems to want to allow for self-related information to be wrong while the reference to the self (conceived as embodied organism) remains constant.

We would like to end this paper with some considerations in favour of a second option that attempts to have the cake and eat it too, since it can show how the dimension of self-as-subject could be accommodated by the PP framework. How can we respect Recanati's point and capture both dimensions of the self adequately in the PP framework without postulating any duality of entities or a further self-model? Let's consider Clark's action-oriented and embodied interpretation of the PP framework. Clark argues that paying full respect to the coupling of the subject, conceived as embodied agent, in its immediate surroundings amounts to replacing the picture in which perceptual representations 'mirror' the state of the world with a picture in which representations are action-oriented (Clark, 2015). His reading of the PP framework differs markedly from Hohwy's internalist conception. Coming close to Newen's emphasis on the common origin of selfrelated and world-related information, Clark (1997, pp. 153-7) develops the notion of an action-oriented representation in order to capture the embodied and situated character of action-guiding perception. Such representations are not symbol-like but (a) actionspecific in that they present the world in terms of actions possible for this particular agent, (b) egocentric in that they are coded from the perspective and standpoint of the present agent, and (c) contextsensitive in that they are to be exploited in the here and now but may not be useful in another situation (see also Wheeler, 2005, p. 199). The action-oriented character of my perceptual representation of the coffee mug in front of me guarantees that it is always already geared to me as the agent possessing this specific set of sensorimotor capacities in order to take advantage of the affordances (Gibson,

1979) provided by the situation around me. Even in the presence of others, this specific layout of my immediate surroundings presents itself to me in such a distinctive way that it can be represented for no other agent. Affordances arise from the coupling of a particular agent with their surroundings.8 Sitting at my desk right now, a number of objects afford being grasped by me while others are out of reach but are within reach to my co-author. In social contexts, we always visually perceive an affordance space where the environment affords some actions for me, other actions for someone else, and yet other actions for both of us. The important point here is that while I may share some affordances in this situation with another agent, and I may even represent some of another agent's set of affordances (Bruineberg, Chemero and Rietveld, 2018), there will always be a difference with respect to which action possibilities arise for me rather than anyone else. That is, we can see how action-oriented representations always implicitly carry information about the self-as-subject, or rather, selfas-agent. Once we represent an object as affording some specific action for me then this is coupled to a 'motor map' and 'motor instructions' needed for the execution of such actions, e.g. grasping by reaching out my arm and forming my hand using the right kind of grip aperture (Wheeler, 2005, pp. 85, 196; cf. Rizzolatti and Sinigaglia, 2008). The important point about action-oriented representations is that they convey information about and to the subject or agent in question without explicitly representing the agent (as object). The information conveyed in such action-oriented representations concerns the subject in their capacity to act on the world. While things can go awry also in such cases of grasping and exploiting affordances, e.g. in alien hand syndrome or apraxia, the possible error does not so much concern the identity of the agent in question but only more specific aspects of the bodily movement. That is, even in such cases there is no misidentification. We think that this reasoning underscores the importance of action (active inference) and the embodied and enactive character of perception for the integration of the self-as-subject in the PP framework.

Given this reasoning and our criticism of Hohwy's and Michael's treatment above, we think that Clark's version of the framework is better suited to capture both dimensions of the self. As we stressed

<sup>&</sup>lt;sup>8</sup> See Siegel (2010) for an argument to the effect that affordances are represented in perception and thus have accuracy conditions.

earlier, distinguishing the self-as-subject from the self-as-object is not intended as introducing two different selves. The distinction picks out two important dimensions that only pertain to the special notion of self or subject. Focusing on agency and all factors involved in the specification and consumption of affordances as relational properties between a particular agent and her environment, it is possible to adequately address the implicit articulation of the subject in the accuracy conditions of perception, construed as action-oriented through and through.

#### 5. Conclusion

The predictive processing account of cognition is committed to the view that prediction error minimization is one of the primary functions of the brain. The representational nature of prediction error, whether constructive or non-constructive in nature, commits the framework to the specification of accuracy conditions for perception. We argued that accuracy conditions can be determined independently of the phenomenal features of perception and showed how the subject can enter the determination of these accuracy conditions in two ways, as (explicitly represented) object and as (only implicitly represented) subject. We proposed in this paper that the PP framework should be able to account for both dimensions of self, and argued that Clark's (2015) 'radical' PP account is best suited to achieve this. One feature of the PP account is that perception presents the environment always from the viewpoint of the perceiver, representing a range of possible actions within this environment. Action-oriented representations are not independent, action-neutral, and allocentric representations of objects, but always respect the agent's abilities. By comparison, the recent account offered by Hohwy and Michael (2017) seems importantly incomplete.

#### References

- Bargh, J.A. & Chartrand, T.L. (2000) The mind in the middle: A practical guide to priming and automaticity research, in Reis, H.T. & Judd, C.M. (eds.) *Handbook* of Research Methods in Social and Personality Psychology, pp. 253–285, Cambridge: Cambridge University Press.
- Blanke, O. & Metzinger, T. (2009) Full-body illusions and minimal phenomenal selfhood, *Trends in Cognitive Sciences*, **13** (1), pp. 7–13.
- Block, N. (2007) Consciousness, accessibility, and the mesh between psychology and neuroscience, *Behavioral and Brain Sciences*, 30, pp. 481–548.
- Brentano, F. (1874/1995) Psychology from an Empirical Standpoint, London: Routledge.

Bruineberg, J. (2017) Active inference and the primacy of the 'I can', in Metzinger, T. & Wiese, W. (eds.) *Philosophy and Predictive Processing*, Frankfurt am Main: MIND Group.

Bruineberg, J., Chemero, A. & Rietveld, E. (2018) General ecological information supports engagement with affordances for 'higher' cognition, *Synthese*, [Online], https://doi.org/10.1007/s11229-018-1716-9.

Clark, A. (1997) Being There, Cambridge, MA: MIT Press.

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

Clark, A. (2015) Embodied prediction, *Open MIND*, Frankfurt am Main: MIND Group, [Online], https://open-mind.net/papers/embodied-prediction.

Clark, A. (2016) Surfing Uncertainty: Prediction, Action, and the Embodied Mind, Oxford: Oxford University Press.

Crane, T. (2001) Elements of Mind, Oxford: Oxford University Press.

Damasio, A. (1999) The Feeling of What Happens, New York: Harcourt.

Damasio, A. (2010) Self Comes to Mind, New York: Basic Books.

Dehaene, S. (2014) Consciousness and the Brain: Deciperhing How the Brain Codes Our Thoughts, New York: Viking Penguin.

Dennett, D.C. (1991) Consciousness Explained, New York: Penguin.

Egan, F. (2014) How to think about mental content, *Philosophical Studies*, **170**, pp. 115–135.

Evans, J. (1982) The Varieties of Reference, Oxford: Clarendon Press.

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

Friston, K. (2012) Prediction, perception and agency, *International Journal of Psychophysiology*, 83 (2), pp. 248–252.

Frith, C. (2007) *Making Up the Mind: How the Brain Creates Our Mental World*, Hoboken, NJ: John Wiley & Sons.

Gibson, J.J. (1979) The Ecological Approach to Visual Perception, London: Routledge.

Gładziejewski, P. (2016) Predictive coding and representationalism, *Synthese*, **193** (2), pp. 559–582.

Hohwy, J. (2007) The sense of self in the phenomenology of agency and perception, *Psyche*, **13** (1), pp. 1–20.

Hohwy, J. (2010) The hypothesis testing brain: Some philosophical applications, in Christensen, W., Schier, E. & Sutton, J. (eds.) ASCS09: The Proceedings of the 9th Conference of the Australasian Society for Cognitive Science, pp. 135– 144, North Ryde: Macquaurie Centre for Cognitive Science.

Hohwy, J. (2013) The Predictive Mind, Oxford: Oxford University Press.

Hohwy, J. & Michael, J. (2017) Why should any body have a self?, in de Vignemont, F. & Alsmith, A. (eds.) *The Subject's Matter: Self-Consciousness* and the Body, Cambridge, MA: MIT Press,

Hutto, D.D. & Myin, E. (2017) Evolving Enactivism: Basic Minds Meet Content, Cambridge, MA: MIT Press.

Kiefer, A. & Hohwy, J. (2018) Content and misrepresentation in hierarchical generative models, *Synthese*, **195** (6), pp. 2387–2415.

Kinzler, K.D. & Spelke, E.S. (2007) Core systems in human cognition, *Progress in Brain Research*, 164, pp. 257–264.

Langland-Hassan, P. (2015) Introspective misidentification, *Philosophical Studies*, 172 (7), pp. 1737–1758.

- Limanowski, J. (2017) (Dis-)attending to the body, in Metzinger, T. & Wiese, W. (eds.) *Philosophy and Predictive Processing*, Frankfurt am Main: MIND Group.
- Limanowski, J. & Blankenburg, F. (2013) Minimal self-models and the free energy principle, *Frontiers in Human Neuroscience*, 7, art. 547.
- Longuenesse, B. (2017) *I, Me, Mine: Back to Kant, and Back Again*, New York: Oxford University Press.
- Metzinger, T. (2003) Being No One, Cambridge, MA: MIT Press.
- Metzinger, T. (2009) *The Ego Tunnel: The Science of the Mind ad the Myth of the Self*, New York: Basic Books.
- Moutoussis, M., Fearon, P., El-Deredy, W., Dolan, R.J. & Friston, K.J. (2014) Bayesian inferences about the self (and others): A review, *Consciousness and Cognition*, 25, pp. 67–76.
- Newen, A. (2018) The embodied self, the pattern theory of self, and the predictive mind, *Frontiers in Psychology*, **9**, art. 2270.
- Perry, J. (1986) Perception, action, and the structure of believing, in, Perry, J. (1993) *The Problem of the Essential Indexical: And Other Essays*, Oxford: Oxford University Press.
- Ramsey, W. (2007) *Representation Reconsidered*, Cambridge: Cambridge University Press.
- Ramsey, W. (2016) Untangling two questions about mental representation, *New Ideas in Psychology*, **40**, pp. 3–12.
- Recanati, F. (2007) Perspectival Thought: A Plea For (Moderate) Relativism, Oxford: Clarendon Press.
- Rizzolatti, G. & Sinigaglia, C. (2008) Mirrors in the Brain, Oxford: Oxford University Press.
- Schlicht, T. (2016) Selves, or something near enough, in von Guhem E. & Graetzel, S. (eds.) *Life, Body, Person and Self*, Freiburg: Alber.
- Schlicht, T. (2017) Experiencing organisms: From Mineness to subject of experience, *Philosophical Studies*, **175** (10), pp. 2447–2474.
- Schlicht, T. (2018) Does separating intentionality from representation imply radical enactivism?, *Frontiers in Psychology*, **9**, art. 1497.
- Searle, J.R. (1983) *Intentionality: An Essay in the Philosophy of Mind*, Cambridge: Cambridge University Press.
- Searle, J.R. (1997) The Mystery of Consciousness, New York: NY Review of Books.

Searle, J.R. (2001) Rationality in Action, Cambridge, MA: MIT Press.

- Searle, J.R. (2015) *Seeing Things as They Are: A Theory of Perception*, New York: Oxford University Press.
- Seth, A.K., Suzuki, K. & Critchley, H.D. (2012) An interoceptive predictive coding model of conscious presence, *Frontiers in Psychology*, **2**, art. 395.
- Shoemaker, S.S. (1968) Self-reference and self-awareness, The Journal of Philosophy, 65 (19), pp. 555–567.
- Siegel, S. (2010) *The Contents of Visual Experience*, New York: Oxford University Press.
- Weiskrantz, L. (1986) *Blindsight: A Case Study and Implications*, Oxford: Oxford University Press.
- Wheeler, M. (2005) *Reconstructing the Cognitive World: The Next Step*, Cambridge, MA: MIT Press.
- Wiese, W. (2017) What are the contents of representations in predictive processing?, *Phenomenology and the Cognitive Sciences*, **16** (4), pp. 715–736.

Wiese, W. & Metzinger, T. (2017) Vanilla PP for philosophers: A primer on predictive processing, in Metzinger, T. & Wiese, W. (eds.) *Philosophy and Predictive Processing*, Frankfurt am Main: MIND Group.

Wittgenstein, L. (1958) The Blue and Brown Books, Oxford: Blackwell.

Zahavi, D. (2014) Self and Other: Exploring Subjectivity, Empathy, and Shame, Oxford: Oxford University Press.