Philosophical Psychology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/cphp20

"Minimal self" locked into a model: exploring the prospect of formalizing intentionality in schizophrenia

Marianne D. Broeker & Matthew R. Broome

To cite this article: Marianne D. Broeker & Matthew R. Broome (2025) "Minimal self" locked into a model: exploring the prospect of formalizing intentionality in schizophrenia, Philosophical Psychology, 38:1, 331-352, DOI: 10.1080/09515089.2024.2365935

To link to this article: https://doi.org/10.1080/09515089.2024.2365935

Published online: 13 Jun 2024.

PHILOSOPHICAL PSYCHOLOGY
2025, VOL. 38, NO. 1, 331–352
https://doi.org/10.1080/09515089.2024.2365935

Routledge
Taylor & Francis Group

"Minimal self" locked into a model: exploring the prospect of formalizing intentionality in schizophrenia

Marianne D. Broeker[a] and Matthew R. Broome[b]

[a]Department of Experimental Psychology, University of Oxford, Oxford, UK; [b]School of Psychology, University of Birmingham, Birmingham, UK

ABSTRACT

Computational psychiatry is a quickly evolving discipline that aims to understand psychopathology in terms of computational, hence algorithmic processes. While cognitive phenomena, especially beliefs or ways of "reasoning", can more easily be formalized, meaning re-described in mathematical terms and then entered computational models, there is speculation as to whether phenomenology might be formalizable too. In other words, there are speculations in terms of what aspects of the human experience, rather than specific cognitive processes alone, can enter computational models. Here, we explore the possibility of formalizing and modeling a phenomenological account of schizophrenia, using the concepts of "minimal self" and "intentionality". To test the applicability of these concepts for formalization and modeling, we first aim to clarify some misunderstandings around the very nature of minimal self and intentionality, namely: whether these concepts entail a "minimal" sense of self, or might be better described in "transparent" sensory-integration and information processing terms. We then try to apply the concepts to a computational logic based on Marr's levels of description, a fundamental account for understanding the rationale of computational psychiatry. Overall, we are asking via what conditions phenomenology can enter a computational logic.

ARTICLE HISTORY
Received 10 March 2023
Accepted 31 May 2024

KEYWORDS
Phenomenology; Bayesian brain hypothesis; predictive processing; delusions in psychosis; minimal self; intentionality

Introduction

The question whether phenomenology can be naturalized has been addressed before, (Mishara, 2007b; Zahavi, 2004), where the focus has been on the possibility of a biological naturalization. In this paper, we are asking if the phenomenology of schizophrenia, which will be referred to as psychosis interchangeably, can be mathematically described, hence

CONTACT Marianne D. Broeker Marianne.broeker@psy.ox.ac.uk Department of Experimental Psychology, University of Oxford, 56 Woodstock Road, Oxford OX2 6HS, UK

formalized and then computationally modeled, as described by Marr's (2010) levels.

We take a phenomenological approach to schizophrenia to determine whether "pure experience", hence, "what it is like" to be schizophrenic, can be broken down into components or units which can enter computer models.

A phenomenological perspective not only allows for an understanding of the very experience of schizophrenia, but the experience itself can be informative and lead to a potential "explanation" of schizophrenia. This has been emphasized by Zahavi (2003), leaning on Husserl's division of phenomenology into phenomenological psychology, which is more descriptive, and transcendental phenomenology, which is more constitutive of a phenomenon. Here, the phenomenon itself is taken to be the condition of possibility and meaning. The transcendental is a field of evidence embedded within mundanity rather than a formal construction of principles deduced to explain or justify mundanity. Zahavi (2003, 2004) suggests that descriptive phenomenology, which is often described as the "empirical dimension", can lead to transcendental phenomenology. However, he also remarks that transcendental phenomenology is often ignored, and phenomenological psychology is understood as encompassing "phenomenology" as a whole. This leads to an impoverished, non-technical, and superficial understanding of phenomenology; it is either understood as pure description or as some kind of "introspective account" rather than as something that can broaden our understanding of the phenomena.

To determine whether a phenomenological account of schizophrenia can enter a computational model, we need a clear outline of it. Within the phenomenological discourse around schizophrenia, there are two main camps, which have been labeled "Apollonian" and "Dionysian" Mishara (2007b).[1] The Apollonian account focuses on changes in self-consciousness or ipseity (Sass & Parnas, 2003), whereas the Dionysian account focuses on changes in sensory and integrative processes, as constitutive of schizophrenia (Mishara, 2007b). Both use the term "minimal self", though the term is more and more rejected in the Dionysian camp. When it is used, it is tied to the "sensory self", while in the Apollonian camp, it is connected to the subjectively experiential self. Thus, points of contention and division between both approaches are definitions of "subjectivity" and "minimal self", especially the question of how "reflective" and "actively aware" a minimal self is, vs. how transparent and devoid of an actual self it is.

To enter computational models, phenomena must be clearly circumscribed. In this article, we first attempt to settle the controversy regarding a phenomenological understanding of self-disorder in schizophrenia, using the concepts minimal-self, intentionality as well as sensory processing. Based on a unified understanding of self-disorder in schizophrenia, we

then ask if this understanding can be operationalized and computationally modeled. Thus, we evaluate the suitability of the concept's "intentionality" and "minimal self" to be broken down into operationalizable and formalizable units, to then enter computational models. We do not claim to ascertain if and how other phenomenological accounts can or cannot be modeled. However, we provide steps which can be applied to other conditions to help determine whether they are formalizable and can be modeled computationally.

We show that only specific conceptualizations of phenomena, namely when they are tied to very circumscribed processes that can be isolated into single units, have the potential to enter into computational models. Furthermore, computational models always follow a deterministic logic, where a specific input is constitutional of a specific output, and thus, modeling a phenomenon requires knowing which inputs universally lead to which outputs. Certain conceptions of the minimal self constrain these requirements. Thus, depending on the very understanding of a phenomenon, formalization and modeling may or may not be a possibility.

An Apollonian understanding of phenomenology and schizophrenia

Within a phenomenological perspective on schizophrenia, the discrepancies between the Apollonian vs. Dionysian understanding make it difficult to find the circumscribed processes that are needed in order to formalize and model the phenomenon. We will first look at the Apollonian vs. Dionysian understanding of self and how it is changed in schizophrenia by understanding minimal self and intentionality. Intentionality relates to minimal self as a means of connecting it to the intersubjective world, thus, both concepts can be understood together.

Husserl, Merleau-Ponty, and Watsuji outline intentionality as a tacit, implicit process that describes the directedness and connection of the lived-body (self) to the intersubjective world. This understanding of intentionality goes beyond and has little to do with sensory processing alone, or circumscribed mental states such as reflection, knowledge, or direct awareness.

The philosophical understanding of intentionality differs from the everyday use of the term, where it describes purposefulness, commitment, and deliberate actions, and is psychologically related to "higher order", reflective and deliberate cognition.

In a phenomenological understanding, intentionality binds the self, the body, to the world, as bodies are directly embedded into it (Krueger, 2020; Zahavi, 2003). Intentionality thereby determines how the world shows up to us. Intentionality has also been described as a central feature of consciousness, grasping the meaning of the world for us.

Through the connection between self, hence body and world, intentionality brings about that we always experience things around us in a particular way, implicitly giving them subjective meaning and character (Krueger, 2020). This is the "how" of an object's appearance to consciousness and it reflects the constituting, yet implicit and automatic, activity of the subject's perception of and engagement with the world. Implicit and automatic refer to the fact that these processes take place without deliberate or volitional direction, or higher-order thinking.

Intentionality has been described by multiple philosophers: Husserl coined the term "operative intentionality" (Husserl, 1962, 2014), to describe the process that gives meaning prior to conscious reflection. Husserl also emphasized "intentional qualities", which refer to the fact that the same object can be given in different modes or qualities: e.g., an apple can be desired, judged (for it going bad), disliked etc., differences that do not amount to sensory differences when perceiving the apple.

Merleau-Ponty's (1962) "pre-reflective intentionality" describes a pre-reflective directness toward the lived world, through which the body is part of the subject-world system or body-subject concrete, rather than intentionality being about objects that are reflected upon. The "pre-reflective" moving body is intentional in itself, "reaching out towards the world" – where bodily space and external space form a practical system. As opposed to explicit reflection and knowledge, the intentional arc is described as comprehension, vision, and motion. Intentionality is more clearly apparent in our desires, evaluations, and the landscape we see, than in objective knowledge. In other words, intentionality is phenomenological "comprehension" rather than active reflection or a deliberate bringing to awareness. Merleau-Ponty gives the example of love, an act that is directed without needing a reflective understanding of it: the impulse that carries one toward someone does not need aboutness or a distinguishable intentional object (Merleau-Ponty, 1962; Reuter, 1999).

Both Husserl and Merleau-Ponty point to the concept of "horizon" that describes the anticipatory structure of intentionality. Perception consists of two elements: seeing what is presented (from my vantage point) and what is possible (from other vantage points). As pointed out by Krueger (2020), Watsuji Tetsuro (1996) extended the concept of intentionality by highlighting its inherently social and relational character, described by the word "betweenness" = "aidagara". Intentionality is relational rather than confined to the head; it is neither a feature of a person's mental state, nor a collection of sensations (Watsuji, 1996, p. 73). It is precisely the between-ness that regulates how things show up in consciousness as meaningful objects of experience.The horizon is modulated not only by the body-subject concrete, but also by the social milieu.

In sum, intentionality can be described as an embedded process that comes about through a connection between world and self, rather than being a circumscribed mental state. Minimal self has been defined as pre-reflective, tacit level of selfhood (Nelson et al., 2014).

It refers to the implicit first-person quality of consciousness, as all experiences articulate themselves in the first-person perspective as "my" experience. The concept of minimal self refers to a structural dimension of consciousness that is considered a necessary feature of all experiences given to the subject. It provides the foundation and prerequisite for more complex forms of selfhood: personal, narrative, and social identity. Zahavi (2003) famously stated that all our experiences have a subjective or first-person character ("for-me-ness" or ipseity). According to Zahavi (2003), every conscious state, perception, emotion, recollection, abstract belief, etc., has a certain subjective character, a phenomenal quality of "what it is like" to live through or undergo that state. This is in fact what makes the mental state in question conscious in the first place. Hence, the reason we are "aware" – in a minimal sense – of our experiences, and can distinguish them from one another, is precisely because there is something 'it is like' to be in those states.

The phenomenal state itself does not need to be an object of (higher-order) reflection to be intentional, but it still influences "how" we perceive the world. Further, subjectivity, the phenomenal side, is not a self-enclosed mental realm; rather subjectivity and world, as it shows up, are codependent and inseparable (Merleau-Ponty, 1962, p. 430). Thus, (minimal) self and world are connected through intentional engagement; intentionality can only come about through a minimal self that drives the directness toward the world.

Within the phenomenal or subjective state, Zahavi (2003) points to an important distinction: 1.) What the object is like for the subject, i.e., the properties the object is experienced to have. 2.) What the experience of the object is like for the subject (Carruthers, 1998; McIntyre, 1999), i.e., the mode in which something is experienced, which is reminiscent of Husserl's "intentional qualities". The mode of access can take different forms of givenness: perceived, imagined, desired, contested, etc. These experiential properties are not properties like spicy or gray; they are properties pertaining to these different types of access, different qualities of mine-ness. This has also been called "First-person givenness". In other words, experience is given non-inferentially, non-critically and non-reflectively, but with a primitive built-in self-reference. "Mine-ness" is not a worldly property, nor the property of the stimulus, nor the sensory processing directly mapping onto that, but an experiential property. Every experienced mental state involves a "givenness".

Importantly, both the experienced qualities of the object and the givenness of it, do not need to be actively reflected on or brought to awareness, in order to be experienced and to bring about the intentional directedness toward the intersubjective world. All conscious acts are intrinsically self-conscious. Thus, both intentionality and minimal self come about through subjective engagements with the world, rather than being concrete mental states.

In schizophrenia, changes in ipseity therefore mean an instability of the self-world, tapping onto changes in minimal-self and intentionality, thus, the self's directedness toward the world.

Dionysian objections to the Apollonian minimal self approach -I

The terms "awareness" and "reflection" have led to the misunderstanding that something deliberately reflective is built into our way of experiencing. Furthermore, the idea that minimal self alludes to a reduced "sense of self" rather than being sensory or transparent, has generated confusion.

Within the Dionysian camp, highlighting changes in sensory-integrative processes while omitting the "reflective" sphere, Mishara (2007a) proposes that a reduced minimal self in schizophrenia involves a dysfunctional pre-attentive binding between sub-components of the self (I, me, and mine), prior to the emergence of the self as unitary experience in awareness. Alterations in the minimal self emerge through impaired sensory-motor integration. Mishara (2007a) defines the concept of minimal-self as automatic and implicit, and states that it therefore stands in opposition to the supposedly "reflective" or "actively self-aware" minimal self of the Apollonian approach.

His conception rests on Husserl's "'passive synthesis'" the assumption that automatic processes underlie the organization of our experience. Within this view, minimal self is described as a stream of consciousness which involves conscious and automatic processes. Mishara stresses that we do not have reflective access to this process, which is "the self", as the stream is always ahead of us. Mishara highlights that for Husserl, the stream of transcendent consciousness, the self, is passively constituted. Thus, the "self" is not a mental representation but an ongoing process of integration, where access is only retrospective. In psychosis, the self is captivated in the present moment; the ability to transcend current experience is compromised. According to Mishara, symptoms of psychosis, like delusions, are a way of preserving the self minimally, when the process of self-transcendence is impaired.

Mishara's "subcomponent view of the self", based on sensory-motor integration, differs from Zahavi's (2003) concept of minimal self, where the first-person perspective, the pre-reflective self-awareness is always

present in conscious experience as all conscious experiences are infused with a "quality of mine-ness" or "ownership", which stands in the Apollonian tradition. According to Mishara, the claim that we have direct reflective access to a pre-reflective self equates "'I'" with "'me'": it does not acknowledge an ongoing but mutually exclusive shifting of naively experiencing and reflecting attitudes within the subject (see Husserl, 1962; Mishara, 2007a; von Weizsaecker, 1968). Thus, Mishara draws on a supposed dichotomy between experiencing and reflecting, relating the former to Husserl's "passivity synthesis" and the latter to active self-reflection within the ipseity-disorder model of the Apollonian camp. Husserl's "'passive synthesis,'" i.e., the automatic processes underlying the organization of our experience, has also been acknowledged by Binswanger, Blankenburg, Schwartz and Wiggins, who view schizophrenia as arising from the breakdown of Gestalt-perception and resulting anomalous perceptual experiences. These approaches are conceptualized in opposition to the IDM: he disruption of experience in schizophrenia is thought to lay on the level of pre-intentional passive synthesis, not at the level of higher-order "'intentionality'", which Mishara sees at the core of the Apollonian; hence, ipseity-disorder approach. According to Mishara (2007a), by remaining at the level of intentionality, the ipseity model does not reach "'deeply'" enough into the automatic processes shaping our experience. In contrast, the Dionysian approach holds that the disruption of perceptual Gestalt, and sensory-motor integration in schizophrenia rests on passive, automatic processes outside of our direct awareness that are prior to, and the precondition for intentionality and reflectivity.

Responses to objections -I

Some problems with this critique of minimal self and the supposed dichotomy between the Dionysian and Apollonian approach can be pointed out:

(a) The Dionysian view is defined in opposition to the ipseity-disorder view, hence the Apollonian approach. The opposition rests on the assumption that "minimal self" and intentionality, within the Apollonian view, imply a form of "active" or "reflective" self-awareness. This is illustrated by Mishara citing Sass & Parnas: "We are self-aware through our practical absorption in the world of objects (.) pre-reflective self-awareness... entails a built-in self-reference' (Sass & Parnas, 2003, p. 430). Mishara (2007a) then goes on to state that in the Dionysian approach, absorption and self-awareness are mutually exclusive. However, this characterization misrepresents pre-reflective self-awareness, which is mistaken for a "higher order", deliberately reflective process, rather than for

a phenomenal state of being directly embedded into a bodily self-world structure. Within the Apollonian approach, self-awareness is a process that cannot come about through conscious, deliberate reflection or and everyday understanding of intentionality. The process described by Husserl, Merleau-Ponty and Watsuji is very likely to emerge in a rather passive, not reflected upon, fashion, and therefore, Mishara's proposal is not a valid point of distinction. Minimal self, within the IDM or Apollonian approach, does not refer to a form of self-awareness that exists apart from and alongside experience. Basic "self-awareness" is not reflectiveness and therefore, experiencing and reflecting come down to the same processes within both camps. That an experiential givenness is distinct from an active reflection has also been stressed by consciousness science (e.g., Brown et al., 2019; Rosenthal, 1993).

(b) When talking about the "hard problem" of consciousness (what makes an experience "what it is like"), David Chalmers (1997) distinguished between the psychological (functional and causal) and the phenomenological side, advocating to ignore the latter and focus on the former. However, by making this distinction, he points to the existence and fundamental difference of the two entities.

When referring to sensory-integration, Mishara addresses the "psychological", rather than the "phenomenological" level. Thus, in Chalmers words, this type of "minimal self" could be described as a "pseudo-minimal self" or "pseudo-intentionality". Sensory processes probably underpin the "minimal self" but aren't reducible to such. In other words, the "minimal self" cannot be equated to sensory-integrative information processing, which is necessary but not sufficient.

Dionysian objections to the Apollonian minimal self approach -II

Feyaerts et al. (2024) remark, when suggesting a revision of the IDM, that as minimal-self entails "self-reflection", it points to its "essentialist" nature and thus goes beyond a "transparent" experiential state. The authors question whether there truly is an aspect of "self" that separates "self-experience" from "ordinary experience", and how this aspect might then be disrupted in psychosis. They argue that the very definition of "mine-ness" makes it difficult to comprehend how something that basic and "essential" could be disturbed if it is a necessary condition for consciousness itself. "mine-ness", in the way in which they criticize it, has been described as having an "all-or-nothing" quality. They further argue that the minimal self view confuses the idea that every experience is had by someone with the idea that every experience is necessarily experienced as "mine".

In short, they criticize the "essentialist" nature of minimal-self in self-disorder, whether "mine-ness" or subjectivity ought to be found in experience generally and could sequentially be disturbed in psychosis. In addition, they argue that alterations in schizophrenia involve a heightened or exaggerated sense of self, given the experienced solipsism and ontological paranoia. As they cannot answer what the "minimal self' is or how it could potentially vary or be disturbed, they propose that instead of a disrupted "ipseity" the very "subjectivity" component itself may be removed and the "transparency view"(Dretske, 1995; Howell, 2023; Prinz, 2006) might be adapted instead. The transparency view questions that a sense of self is present in immediate, unreflective experience. Rather, they propose that experience is transparent and impersonal; it is directly focused on the world, without self-awareness. As experience is transparent, the world is directly presented to us as an autonomous, subject-independent domain of reality (Dretske, 1995). The transparency view states that the experience of worldly objects does not contain a sign of "self-as-subject", or any phenomenal ownership of experience, but merely presents the world in its phenomenal qualities: "There is no more to the quality of one's experiences in experiencing blue than there is to the color blue since the color blue is the color one experiences" (Dretske, 1995, p. 85). In sum, Feyaerts et al. take the stance that "ordinary unreflective consciousness" does not have a minimal-self or anything subjective in it.

On the contrary, in psychosis, the loss of impersonal "transparency" leads to a pervasive sense of subjectification, hence, an increase in subjectivity that is normally absent (Feyaerts et al., 2024).

Responding to objections-II

Some problems with this critique of minimal self, in particular the denial of minimal self, can be pointed out:

(a) As minimal self and intentionality are intrinsically tied together, we can use Zahavi's (2003) two elements of experience: (1) the subjective experience of the object; and the (2) experiential givenness, to directly argue against the "transparency view". (1) For the experience of objects, empirically, we know that the perception of the same object can differ greatly on a subjective level: The example of color perception, e.g., when looking at individual differences or differences across ethnic groups, strongly indicates that the world is not experienced in a "subject-independent" way (Emery & Webster, 2019; Schloss et al., 2015) as proposed by Dretske: Emery and Webster (2019) showed that judgments of color between individuals can vary widely, and in ways that are not readily explained by differences in sensitivity or the

environment. Furthermore, Phillips (2019) has shown cultural and educational differences in the perception of color, illusions, and depth, which could all be described as rooted in object perception. Thus, there is ample evidence against the transparency view of object perception. Furthermore, there is empirical evidence for the givenness of experience, the mode, which changes how the same object is experienced. e.g., Barrett and Moshe (2009) have shown that the brain's ability to see in the present incorporates a representation of the affective impact of those visual sensations, which varies inter-individually. Thus, affective responses, or the mode in which an object is given, support vision from the very moment that visual stimulation begins. This givenness does not come down to differences in sensory processing.

(b) Lastly, based on Husserl, Merleau-Ponty, and Watsuji's description of intentionality, the very feeling of "mine-ness" could in itself be transparent, precisely because it isn't experienced as a separate object of reflection. In other words, the minimal self view does not state that a sense of self is present in the immediate experience, which again points to a misunderstanding of minimal self and its ascriptions in terms of "self-awareness". One of these misunderstandings is highlighted by the statement of Mishara (2007b) where Dreyfus (2007) argues that a smooth and successful interaction with the world can, and, perhaps, often must dispense with self-awareness. Dreyfus is clearly alluding to a higher-order form of awareness here, which is fundamentally different from pre-reflective self-awareness. Thus, the transparency view isn't opposed to the "minimal self", as the "minimal self" is itself a background and implicit phenomenon.

Taken together, the main misunderstandings about minimal self take place within the debate of how "reflective" and "self-aware" the it is. Intentionality and "minimal-self" have been misunderstood as a reflective (higher-order) phenomenon, as explicit and separate objects of awareness.

Following Husserl's and Merleau-Ponty's (etc.) understandings of it, this is not accurate: operative intentionality, intentional qualities (desires, moods. . .), pre-reflective intentionality, and the body-subject concrete all indicate that intentionality and minimal self do not entail direct "reflective self-awareness", active reflection or any other higher-order process. Watsuji provides a further argument against the "actively reflective" character of the minimal self, by pointing to its relational character. Lastly, Zahavi's two elements of experience, and the provided empirical evidence, clearly show that experience does entail a subjective, minimal self, which might very well be transparent in its own givenness and have a sensory underpinning.

When modeling psychopathology, the definition and understanding of concepts from phenomenology greatly matter when these concepts are applied to psychosis: do we understand self-disorder only in terms of disrupted sensory processes (Gestalt, etc), which would be easier to isolate into circumscribed units, or do we understand it in terms of the lived-body, intersubjectivity, which might be underpinned by sensory processes, rather than coming about through those?

Now that we have an understanding of the elements of psychosis when it is defined from a phenomenological angle, we can explore the question of whether these elements – intentionality and minimal-self – can be broken down into formalizable units and can enter a computational model.

Situating implicit or "transparent" intentionality and minimal self at the "phenomenological level" does not exclude nor rule out the possibility that they are underpinned by biological processes. However, they are not constituted through these processes, nor can they be reduced to them, since these processes represent the psychological or functional, rather than a phenomenological level. In short, the self is grounded in, but not reducible to sensory processes.

Computational levels

Computational psychiatry builds on the assumption that mental states, such as cognition and behaviors, can be explained computationally with algorithmic models that map input to output. In this paper, we are asking whether phenomenology, too or more specifically, particular concepts from phenomenology, can enter computational models. While computational psychiatry aims to reform treatment approaches, more fundamental motivations target the formalization of human cognition in general, and psychopathology specifically, which largely intersects with motivations of predicting and controlling human behavior.

A good presentation of computational models is provided in Marr's levels of analysis (Marr, 2010) in which the computational level serves as a theoretical foundation, defining the mapping between an input, or initial state, and an output, or resulting state, constituting a mental capacity. This highlights the dependency of the mental capacity or form of cognition on the input, e.g., incoming sensory information or other types of information. The algorithmic level describes the mechanism of how the output capacity is realized and optimized, i.e., how the input is transferred into the output. Lastly, the implementation level states where the capacity is realized within the (physiological) system.

While defining and mapping inputs to outputs provide the theoretical framework, the algorithmic level provides the explanation of the mapping that constitute the phenomenon or capacity. Thus, computational models of

cognition aim to uncover the mechanistic algorithms behind emotions, thoughts, behaviors, or symptoms of psychopathology. Likewise, aberrations in algorithms are thought of as being the underlying causes of specific psychopathological phenomena.

Explaining mental states computationally means describing phenomena mechanically in terms of inputs that lead to specific outputs and finding the algorithms of that transformation. In computational psychiatry, the realizing algorithms are mainly, though not exclusively, thought of as inferential. In terms of theory-building, it remains to be seen to what extent computational psychiatry provides new insights, rather than it just being a formalized re-description of cognitive and often inferential, processes (e.g., inferential reasoning in psychosis). The computational approach has been applied to various psychiatric and neurological disorders such as positive, often isolated symptoms of schizophrenia (Adams et al., 2013; Corlett, 2019; Sterzer et al., 2018); autism (Lawson et al., 2014); Parkinsons disease (O'Callaghan et al., 2017); anorexia (Gadsby & Hohwy, 2021); addiction (Smith et al., 2020) and depression (Barrett et al., 2016). Inference is generally understood as a method of (rational) reasoning based on (input) evidence, whereas Bayesian inference, that is abductive inferences under uncertainty (Coltheart, 2010; Mathys et al., 2011), is mostly used in computational psychiatry and is a specific type of inferential computation. The role of uncertainty (Hohwy, 2013), as well as the notion of evidence to reduce uncertainty, are central to Bayes, where uncertainty of the outside world is dealt with through Bayes decision rule (Fahlman et al., 1983; Huys et al., 2011). Within the Bayesian Brain hypothesis (BBH), Bayes' rule refers to the aim of the organism to predict what is happening next: a world state, a perception, an emotion, etc., where prior knowledge, and therefore internally generated predictions, are combined with new sensory evidence (input) from the environment or the body, leading to new estimates (posterior) of what is about to happen. Thereby, the goal is to minimize the errors resulting from predictions (Clark, 2013; Friston, 2005; Hohwy, 2013). This was first conceptualized with the notion of free energy as a measure of the discrepancy between actual features of the outside world and their internal representations (Hinton & Zemel, 1993), where action (active inference) and perception (perceptual inference) result from free energy minimization (Friston, 2010, 2005). If a prediction error occurs, it is typically transmitted back to an internal hierarchical knowledge system to update higher-level prior expectations, leading to dynamic interactions.

Most theories on schizophrenia, especially those on clinical delusions, focus on alterations in hierarchical inferences (Adams et al., 2013; Friston, 2005; Sterzer et al., 2018) across various interdependent levels of processing, starting from lower inferences at less abstract levels (i.e.,

perception of lower-level features) to higher-level inferences on abstract concepts (i.e., estimation of hidden world states or direct inferential reasoning).

Conceptualizing the brain as a hierarchical prediction and information processing machine, the BBH aims to establish a mechanistic link between biological processes that implement information processing (Friston, 2010), quantitative computational models, and symptoms of psychopathology. The assumption that cognition generally and psychopathology specifically correspond to information processing as well as to aberrations thereof, represents an ontological assumption, or background belief, about the nature of cognition and the causal structure of the world (Maatman, 2024). It also lies within the deficit approach of psychopathology, where inferential aberrations and psychopathology are generally thought of as deficits. However, what constitutes an aberration in information processing needs to be defined. Thus, we need a framework that determines how information is most optimally processed, based on given inputs and outputs, to then establish aberrations thereof. Based on Marr's levels, theories about the computational level are often regarded as normative, e.g., rational or optimal normative (Oaksford & Chater, 2009). Therefore, Bayesian modelers argue that computational theories are idealized optimization, serving rational agents.

Since Bayes' rule is a rational, logico-mathematical rule within the classical BBH information processing follows a strategy whereby minimizing the prediction error is considered optimal. However, this Bayesian rule is only optimal if no other contingent factors are present. If, in a given situation or task, contingent facts are known to influence the relationship between input and output, what is optimal may become contingent on these other, known, influential factors. The optimal processing strategy, which is defined a priori in order to establish aberrations thereof, therefore also depends upon the structure and contingent facts of a given situation or task, which can be called context-dependency of optimality (Gigerenzer, 2008), or embedded optimality. Contingent factors might be the structure of the environment or task, the limitations of the agent, or factors that have previously been considered "irrational" in comparison to processing in accordance with Bayes' rule, such as processing costs of motivational influences (Williams, 2021; Williams & Montagnese, 2020). Importantly, when we define a framework of optimal information processing to map input to output, we assume to have knowledge of all contingent factors that this mapping depends on, which then gives leverage to interpret what is most optimal. Thus, optimality is defined a priori, in a fixed, general, and universally applicable way, as well as in reference to a causal structure of the world. In computational psychiatry, contingent factors are often controlled for, so that underlying Bayesian forms of rational processing can be laid bare.

To summarize, computational psychiatry is based on specific ontological commitments, some of them being:

(1) Mental capacities consist of (unconscious) information processing algorithms that can be formalized (= psychological assumption vs. epistemological assumption).
(2) *Psychopathology* evolves around minimizing general uncertainty (free energy Karl Fristen) of the immediate external environment. This also implies that pathology evolves in the realm of a shared single, external, objective world that entails uncertainties and hidden states of the world. Thus, computational psychiatry entails the assumption that aberrations in free-energy minimization are driving *psychopathology*
(3) *Psychopathology* unfolds in rational or optimal decision-making processes, as these are mostly captured in the experimental paradigms and computational models.

Phenomenology and computational psychiatry

Based on the characterization of computational psychiatry outlined above, we will now consider whether a phenomenological understanding of psychosis, focusing on changes in minimal self and intentionality, can be formalized and enter an "input-algorithm-output" based logic. Thus, we ask if this specific kind of phenomenological understanding of psychosis, changes in minimal self, and intentionality, can be formalized or if it constrains formalization.

Using Marr's levels, we firstly must ask what needs to enter a model in terms of inputs, which then evolve into changed or aberrant "intentionality" and "minimal self" at the output level, as these are the phenomena whose (phenomenological) aberrations we are interested in. We would then try to find the algorithm that maps a specific and a priori fixed input to changes in minimal self at the output level. Some problems with this approach can immediately be pointed out:

a.) First, formalization requires adequate and concise definitions of the constructs, henceforth phenomena, that enter the model, and which result into the output states we are interested in. In other words, for a computational model, we need concrete and circumscribed definitions of the outcome state, thus, robust and measurable phenomena. This is a dubious assumption to make for DSM categories already, as the DSM is descriptive and does not assert etiological or ontological claims; computational research still takes DSM definitions as starting points for explanatory accounts of a given phenomenon, i.e., delusions (as an outcome state), and

wants to find an algorithmic "cause" of that outcome state, based on fixed inputs (e.g., inferential information about hidden states).

The debate around minimal self and intentionality demonstrates that these phenomena, especially in reference to their application to psychosis (alterations of "minimal self") aren't conclusively defined. Although we tried to settle the debate between the view that includes and the one which excludes the "self" (Apollonian vs. Dionysian), definitions of the terms still don't constitute robust, circumscribed, one-dimensional phenomena, which could easily be operationalized.

Rather, minimal self represents a multitude of cohesive and continuous phenomena, where singling out and operationalizing specific parts would destroy the phenomenon as a whole in its embodied "self-world concrete" so that we would no longer be modeling the phenomenon in question. In short, it seems almost impossible to break-down a phenomenon like minimal self into circumscribed parts to find an algorithmic explanation of it, which would need to go beyond looking at isolated aspects, like perception, alone. Further, psychological processes, which would be easier to operationalize since they constitute robust phenomena, e.g., sensory processing and integration, are often confused for a phenomenological explanation (Mishara, 2007a). Due to Husserl (1962) and Zahavi (2003, 2004), these are, however, different levels of description. While both are necessary, the psychological, i.e., functional, level (e.g., sensory processing), does not equate to the phenomenological one.

According to their understanding, if minimal self is mainly described in sensory-processing terms, we are talking about "pseudo minimal self", or "pseudo-intentionality", which might underpin minimal self and intentionality, but cannot be equated to these phenomena. In sum, input and output states need to represent concrete phenomena or robust and measurable entities, to enter a computational model. While it might be interesting to build a computational model of isolated processes, which might add up to the minimal self, one needs to be clear what these processes do and do not represent.

b.) The algorithmic level in Marr's terms is confined to the head, where inputs are received and are directly constitutive of the output. In other words, the output is directly derived from the input and the algorithm that transforms input into output, which can be called the "input-dependency" of a phenomenon, while all factors, inputs and outputs, must be known. Looking at the individual components of Marr's levels, inputs are determined by the task an agent is performing and directly constitute the data a computational model is built on. The input that is given or stipulated in the task (perceiving, reasoning, etc.) is assumed to be directly constitutive of the "output", which is the phenomenon. Therefore, the exact "input" which we believe leads to the resulting "output" (e.g., delusions or changes in

minimal self) needs to be isolated. According to phenomenological conceptions of minimal self and intentionality, both aren't confined to the head, but directly come about through the interaction with the world, through a subject directly acting on the world in non-inferential ways. Thus, if there would be an algorithm, it wouldn't be in the head, it would be in the embodied subject-world concrete. This has been described by the concepts of operative-intentionality (Husserl, 1962) and pre-reflective intentionality (Merleau-Ponty, 1962), with the "body-subject concrete" acting on the intersubjective world, as well as by the "intentional arc", where intentionality, as the output state, is apparent in our desires, our evaluations and the landscape we see, rather than in objective knowledge or concrete mental states that are confined to the head.

Furthermore, it is not clear which general and inter-individually applicable "inputs" would constitute "outputs" of that kind, since intentionality does not arise through specific sensory or informative inputs, but rather in subjectively meaningful actions toward the world, thus, through the "constituting activity of the subject" itself. This also includes the minimal self, which is not confined to the head either, but rests on the stability of the self-world structure (Nelson et al., 2014). Hence, what counts as input vs. output is far from straightforward because intentionality, minimal self and the world seem circular, without specific inputs deterministically leading to the phenomena as outputs. Furthermore, it isn't obvious how the concept of "between-ness" should enter a computational model, and if it constitutes the input or the output. Similarly, the concept of horizon, the anticipatory structure of intentionality varies intersubjectively, and therefore cannot be captured in a universal, subject-independent, and homogeneous input structure.

Further, when trying to model these concepts, the "intentional qualities" of experiences would need to be accessed – which poses an additional problem for any "a prior" defined input and outcome measures: Phenomenological states, the first-person givenness or form of access always occurs in a mode (Zahavi, 2003), and is composed of both a.) what the object is like for the subject, hence, the properties of the experienced objects, and b.) what the experience of the object is like, in which mode it is given. Again, both aspects can hardly be defined as stable and intersubjectively fixed output states, given their very subjective and variable character. Experienced properties (output) of objects (inputs) vary intersubjectively, rather than humans experiencing the world in a "transparent" way, as has been argued above. Furthermore, the mode in which an object is given poses an additional challenge for computational psychiatry. Modes like very clear beliefs (hence objectifiable knowledge) or circumscribed mental states might be operationalizable as outcome states. However, a model would have difficulties capturing desires or other "modes of access" a priori,

which has been pointed out by (Yon et al., 2020) stating that belief-like, but not desire-like states (e.g., drives, goals, etc.) can enter computational models. Even if we could model desires etc., we would need to predict which inputs would trigger the specific modes an object is given in, across individuals, which is an impossible task. Thus, the first-person givenness, including the mode of access and the "intentional qualities", are antithetical to inferential models of mental capacities, where the output is directly predicted from and contingent on the input. In phenomenology, as demonstrated by the modes of access, the same objects (as inputs), are precisely not taken as leading to intersubjectively unified outputs: the same apple being desired vs. being disliked makes the deterministic logic of computational models obsolete for flexible outcome states.

c.) Lastly, there is a problem with the given uncertainty as well as the approach of dealing with that uncertainty. While a task (giving input data to collect an output state) requires disambiguating "uncertainty", uncertainty is embedded into the input structure and given in a universally homogenous way. Furthermore, the transformative algorithm, that stipulates how the ambiguous information is being dealt with, relies on deeply normative assumptions, as it stipulates a universally applicable "optimal" transformation of the input to output. However, what is considered optimal is arbitrary: sometimes it is more optimal to give more weight to prior knowledge and sometimes to incoming sensory evidence. This is called "Complete class theorem", whereby studies establish different "optimals" even if the same phenomena are being tested. In sum, uncertainty of the world and the optimal way of disambiguating that uncertainty are assumed to apply to all individuals homogeneously.

Within the transparency view, if there were no subjectively experienced qualities nor a mode of givenness, the assumption of predictability from inputs to outputs would hold, as the world and worldly objects would be given in direct and uniform ways.

However, the transparency view (the view without a "minimal self") is highly improbable from a philosophical (Rosanthal, 1980), neuroscientific, psychological (Azzopardi & Cowey, 2001) or empirical angle (Emery & Webster, 2019; Schloss et al., 2015), as many empirical results, e.g., individual differences in color perception, indicate the contrary, that e.g., our perception of the objects in the world and the mode of their givenness is always shaped by a "minimal-self", or something minimally subjective. Thus, neither perception itself, nor the mode of access (desire, goals) could universally enter computational models. In other words, the very aspects of experience, how an object is given and the mode of its givenness, vary intersubjectively and are never directly predictable from the sensory inputs, henceforth the stimuli alone. In other words, the experiential properties and the mode of experience cannot be predicted from the properties

of the sensory input stimulus alone, even under unambiguous conditions. Thus, a transformative algorithm could only make sense once we know an outcoming state that does not vary inter-individually, since different outcome states clearly warrant different transformative algorithms. Thus, building a model that tries to find inputs that universally predict subject-independent outputs seems highly implausible, for a phenomenological understanding of minimal self generally, but also in psychosis.

Reducing uncertainty that is supposed to affect individuals in unified ways rests upon a naïve form of "realism". This can also be called an "ontological misunderstanding" of science as naïve, where the subject matter is taken for granted, reality is assumed to be out there to be discovered in an observer-independent fashion. Furthermore, the need to reduce uncertainty and the role of this reduction in psychopathology can be contested as well. As pointed out by Dreyfus (2007) humans deal with disambiguation in a way that uncertainty is reduced due to subjective goals and concerns in a given situation. Lastly, humans have an ambiguity tolerance and don't need to reduce uncertainty at all costs.

In terms of the notion of optimality, the general assumption that an optimal vs. aberrant way of reducing uncertainty ultimately leads to psychopathology can be challenged. Computational psychiatry depends on the notion of rationality or optimality to define and gauge psychopathology, where violations to rationality are of an inferential kind. Inferential reasoning impairments are therefore central to computational psychiatry, and therefore, computational psychiatry entails a fundamental cognitive assumption about humans as rational agents and is an extension of the cognitive tradition. This is in sharp contrast to any phenomenological understanding of psychosis with the help of concepts like "minimal self" and "intentionality".

Conclusion

We illustrated the difficulty in understanding the phenomenology of schizophrenia, how it is applied to descriptive vs. transcendental understandings, and how that has helped to clarify that a purely descriptive phenomenology only represents one side of the coin which can, nonetheless, potentially lead to a more transcendental understanding, as long as it starts with the suspension of rule-based hypotheses for that phenomenon and uses the pure descriptive account of the first-person as starting point. We then showed that self-disorders cannot be fully understood in terms of the sensory level, e.g., the sensory vacuum (Postmes et al., 2014) or, the subcomponent view alone (Mishara, 2007a), as these represent the "psychological or functional level", which is, according to

Chalmers, distinct from the phenomenological one. Thus, minimal self and intentionality might be underpinned by sensory and integrative processes, but cannot be reduced to these, as they only constitute "pseudo-minimal self" or "pseudo-intentionality". Husserlian phenomenology has often been described as the most likely candidate for the project of naturalization (Mishara, 2007b; Zahavi, 2003), an idea which has often resulted from a biased understanding of his concepts, e.g., focusing on passive synthesis or his rejection of mathematical equations at the time. A detailed outlining of concepts such as "intentionality" and "minimal-self" helps to highlight specific aspects of phenomenology that go beyond "pure description" (Watsuji, Merleau-Ponty, etc.), or sensory-processing, while asking to what extent the "passivity synthesis" can be reduced to sensory processing in the first place.

In this article, we have shown that the question of whether phenomenological concepts can be formalized or not depends on the very definition of these concepts. If phenomenology is mainly understood as Gestalt or perpetual processes, as these present somewhat more circumscribed phenomena, it might be possible to build a computational model of such a phenomenology. However, that comes at the risk of employing an impoverished understanding of phenomenology, by focusing on the psychological (functional) side alone. We do emphasize that the psychological or functional side is important too, but one might then not be able to claim the "formalization of phenomenology".

Thus, the process of operationalizing certain phenomena might destroy the coherence that the phenomenon depends on. The broad definition of intentionality and minimal-self and their location in the embodied subject-(intersubjective) world-concrete, as well as its relational character contradicts its confinement to the head and the concretization of it into operationalizable units.

In sum, a rich, rather than impoverished, understanding of phenomenology makes it difficult to see how a phenomenological understanding of "minimal self" disorder changes in schizophrenia could enter computational model based on Marr's levels.

Note

1. The two approaches, pre-reflective and disruption to pre-intentional perceptual-automatic processing (existential phenomenological approach), have been labeled Apollonian and Dionysian approaches, respectively.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Marianne D. Broeker http://orcid.org/0000-0002-2181-8915

References

Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., & Friston, K. J. (2013). The computational anatomy of psychosis. *Frontiers in Psychiatry*, *4*. https://doi.org/10.3389/fpsyt.2013.00047

Azzopardi, P., & Cowey, A. (2001). Why is blindsight blind? In B. De Gelder (Ed.), *Out of mind: Varieties of unconscious processes* (pp. 3–19). Oxford University Press.

Barrett, L. F., & Moshe, B. (2009). See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364* (1521), 1325–1334. https://doi.org/10.1098/rstb.2008.0312

Barrett, L. F., Quigley, K. S., & Hamilton, P. (2016). An active inference theory of allostasis and interoception in depression. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160011. https://doi.org/10.1098/rstb.2016.0011

Brown, R., Lau, H., & LeDoux, J. E. (2019). Understanding the higher-order approach to consciousness. *Trends in Cognitive Sciences*, *23*(9), 754–768. https://doi.org/10.1016/j.tics.2019.06.009

Carruthers, P. (1998). Natural theories of consciousness. *European Journal of Philosophy*, *6* (2), 203–222. https://doi.org/10.1111/1468-0378.00058

Chalmers, D. J. (1997). *The conscious mind: In search of a fundamental theory*. OUP USA.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Coltheart, M. (2010). The neuropsychology of delusions. *Annals of the New York Academy of Sciences*, *1191*(1), 16–26. https://doi.org/10.1111/j.1749-6632.2010.05496.x

Corlett, P. R. (2019). Factor one, familiarity and frontal cortex: A challenge to the two-factor theory of delusions. *Cognitive Neuropsychiatry*, *24*(3), 165–177. https://doi.org/10.1080/13546805.2019.1606706

Dretske, F. (1995). *Naturalizing the mind*. MIT Press.

Dreyfus, H. L. (2007). The return of the myth of the mental*. *Inquiry*, *50*(4), 352–365. https://doi.org/10.1080/00201740701489245

Emery, K. J., & Webster, M. A. (2019). Individual differences and their implications for color perception. *Current Opinion in Behavioral Sciences*, *30*, 28–33. https://doi.org/10.1016/j.cobeha.2019.05.002

Fahlman, S. E., Hinton, G. E., & Sejnowski, T. J. (1983). Massively parallel architectures for AI: Netl, thistle, and boltzmann machines. *Proceedings of the Third AAAI Conference on Artificial Intelligence* (pp. 109–113).

Feyaerts, J., Nelson, B., & Sass, L. (2024). *Self-disorders in schizophrenia as disorders of transparency: An exploratory account* [Preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/2uak6

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. Article 2. https://doi.org/10.1038/nrn2787

Friston, K. J. (2005). Hallucinations and perceptual inference. *Behavioral and Brain Sciences*, *28*(6), 764–766. https://doi.org/10.1017/S0140525X05290131

Gadsby, S., & Hohwy, J. (2021). Predictive processing and body representation. In C. Colin (Ed.), *Routledge Handbook of Bodily Awareness*. Routledge. https://doi.org/10.31234/osf.io/zvfx2

Gigerenzer, G. (2008). *Rationality for mortals: How people cope with uncertainty*. Oxford University Press.

Hinton, G. E., & Zemel, R. (1993). Autoencoders, minimum description length and helmholtz free energy. *Advances in Neural Information Processing Systems*, *6*. https://proceedings.neurips.cc/paper/1993/hash/9e3cfc48eccf81a0d57663e129aef3cb-Abstract.html

Hohwy, J. (2013). *The predictive mind*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199682737.001.0001

Howell, R. J. (2023). *Self-awareness and the elusive subject*. Oxford University Press.

Husserl, E. (1962). *Husserliana: Band IX: Phänomenologische Psychologie: Vorlesungen Sommersemester 1925*. Martinus Nijhoff.

Husserl, E. (2014). *Ideas: General introduction to pure phenomenology*. Routledge.

Huys, Q. J. M., Moutoussis, M., & Williams, J. (2011). Are computational models of any use to psychiatry? *Neural Networks*, *24*(6), 544–551. https://doi.org/10.1016/j.neunet.2011.03.001

Krueger, J. (2020). Watsuji, intentionality, and psychopathology. *Philosophy East and West*, *70*(3), 757–780. https://doi.org/10.1353/pew.2020.0053

Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00302

Maatman, F. O. (2024). Psychology's theory crisis, and why formal modelling cannot solve it. Preprint. https://doi.org/10.31234/osf.io/puqvs

Marr, D. (2010). *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press.

Mathys, C., Daunizeau, J., Friston, K., & Stephan, K. (2011). A Bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00039

McIntyre, R. (1999). Naturalizing phenomenology? Dretske on Qualia. In J. Petitot, F. Varela, B. Pachoud, & J.-M. Roy (Eds.), *Naturalizing phenomenology: Contemporary phenomenology and cognitive science* (pp. 429–439). Stanford University Press.

Merleau-Ponty, M. (1962). *Phenomenology of perception*. Routledge. http://catdir.loc.gov/catdir/enhancements/fy0909/95165035-d.html

Mishara, A. L. (2007a). Is minimal self preserved in schizophrenia? A subcomponents view. *Consciousness and Cognition*, *16*(3), 715–721. https://doi.org/10.1016/j.concog.2007.07.009

Mishara, A. L. (2007b). Missing links in phenomenological clinical neuroscience: Why we still are not there yet. *Current Opinion in Psychiatry*, *20*(6), 559. https://doi.org/10.1097/YCO.0b013e3282f128b8

Nelson, B., Parnas, J., & Sass, L. A. (2014). Disturbance of minimal self (ipseity) in Schizophrenia: Clarification and current status. *Schizophrenia Bulletin*, *40*(3), 479–482. https://doi.org/10.1093/schbul/sbu034

Oaksford, M., & Chater, N. (2009). Précis of bayesian rationality: The probabilistic approach to human reasoning. *Behavioral and Brain Sciences*, *32*(1), 69–84. https://doi.org/10.1017/S0140525X09000284

O'Callaghan, C., Hall, J. M., Tomassini, A., Muller, A. J., Walpola, I. C., Moustafa, A. A., Shine, J. M., & Lewis, S. J. G. (2017). Visual hallucinations are characterized by impaired sensory evidence accumulation: Insights from hierarchical drift diffusion modeling in Parkinson's disease. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *2*(8), 680–688. https://doi.org/10.1016/j.bpsc.2017.04.007

Phillips, W. L. (2019). Cross-cultural differences in visual perception of color, illusions, depth, and pictures. In K. D. Keith (Ed.), *Cross-cultural psychology: Contemporary themes and perspectives* (2nd ed., pp. 287–308). Wiley Blackwell. https://doi.org/10.1002/9781119519348.ch13

Postmes, L., Sno, H. N., Goedhart, S., van der Stel, J., Heering, H. D., & de Haan, L. (2014). Schizophrenia as a self-disorder due to perceptual incoherence. *Schizophrenia Research*, *152*(1), 41–50. https://doi.org/10.1016/j.schres.2013.07.027

Prinz, J. (2006). The emotional basis of moral judgments. *Philosophical Explorations*, *9*(1), 29–43. https://doi.org/10.1080/13869790500492466

Reuter, M. (1999). Merleau-Ponty's Notion of Pre-Reflective Intentionality. *Synthese*, *188*. https://doi.org/10.1023/A:1005144911619

Rosenthal, D. M. (1993). Higher-order thoughts and the appendage theory of consciousness. *Philosophical Psychology*, *6*(2), 155–166. https://doi.org/10.1080/09515089308573085

Sass, L. A., & Parnas, J. (2003). Schizophrenia, consciousness, and the self. *Schizophrenia Bulletin*, *29*(3), 427–444. https://doi.org/10.1093/oxfordjournals.schbul.a007017

Schloss, K. B., Hawthorne-Madell, D., & Palmer, S. E. (2015). Ecological influences on individual differences in color preference. *Attention, Perception, & Psychophysics*, *77*(8), 2803–2816. https://doi.org/10.3758/s13414-015-0954-x

Smith, R., Schwartenbeck, P., Stewart, J. L., Kuplicki, R., Ekhtiari, H., & Paulus, M. P. (2020). Imprecise action selection in substance use disorder: Evidence for active learning impairments when solving the explore-exploit dilemma. *Drug and Alcohol Dependence*, *215*, 108208. https://doi.org/10.1016/j.drugalcdep.2020.108208

Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., & Corlett, P. R. (2018). The predictive coding account of psychosis. *Biological Psychiatry*, *84*(9), 634–643. https://doi.org/10.1016/j.biopsych.2018.05.015

Tetsuro, W. (1996). *Watsujiō Tetsur's Rinrigaku: Ethics in Japan*. State University of New York Press.

Von Weizsaecker, V. (1968). *Der Gestaltkreis, Theorie der Einheit von Wahrnnehmung und Bewegung* (4th ed.). Thieme [1948].

Williams, D. (2021). Epistemic irrationality in the Bayesian Brain. *The British Journal for the Philosophy of Science*, *72*(4), 913–938. https://doi.org/10.1093/bjps/axz044

Williams, D., & Montagnese, M. (2020). Bayesian psychiatry and the social focus of delusions. preprint. https://doi.org/10.13140/RG.2.2.27852.23683

Yon, D., Heyes, C., & Press, C. (2020). Beliefs and desires in the predictive brain. *Nature Communications*, *11*(1), 4404. https://doi.org/10.1038/s41467-020-18332-9

Zahavi, D. (2003). Intentionality and phenomenality: A phenomenological take on the hard problem. *Canadian Journal of Philosophy*, *29*, 63–92. https://doi.org/10.1080/00455091.2003.10717595

Zahavi, D. (2004). Phenomenology and the project of naturalization. *Phenomenology and the Cognitive Sciences*, *3*(4), 331–347. https://doi.org/10.1023/B:PHEN.0000048935.94012.4e