Consciousness and Cognition 85 (2020) 103029

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

The minimal self hypothesis

Timothy Joseph Lane *

Graduate Institute of Mind, Brain and Consciousness, Taipei Medical University, Taipei, Taiwan
Brain and Consciousness Research Centre, TMU Shuang-Ho Hospital, New Taipei City, Taiwan
Institute of European and American Studies, Academia Sinica, Taipei, Taiwan

ARTICLE INFO

Keywords:
Minimal self
Thought insertion
Unresponsive wakefulness syndrome
NREM sleep
General anesthesia
Levels of consciousness
Neuroimaging
Default mode network

ABSTRACT

For millennia self has been conjectured to be necessary for consciousness. But scant empirical evidence has been adduced to support this hypothesis. Inconsistent explications of "self" and failure to design apt experiments have impeded progress. Advocates of phenomenological psychiatry, however, have helped explicate "self," and employed it to explain some psychopathological symptoms. In those studies, "self" is understood in a minimalist sense, sheer "for-me-ness." Unfortunately, explication of the "minimal self" (MS) has relied on conceptual analysis, and applications to psychopathology have been hermeneutic, allowing for many degrees of interpretive latitude. The result is that MS's current scientific status is analogous to that of the "atom," at the time when "atom" was just beginning to undergo transformation from a philosophical to a scientific concept. Fortunately, there is now an opportunity to promote a similar transformation for "MS." Discovery of the brain's Default Mode Network (DMN) opened the door to neuroimaging investigations of self. Taking the DMN and other forms of intrinsic activity as a starting point, an empirical foothold can be established, one that spurs experimental research and that enables extension of research into multiple phenomena. New experimental protocols that posit "MS" can help explain phenomena hitherto not thought to be related to self, thereby hastening development of a mature science of self. In particular, targeting phenomena wherein consciousness is lost and recovered, as in some cases of Unresponsive Wakefulness Syndrome (UWS), allow for design of neuroimaging probes that enable detection of MS during non-conscious states. These probes, as well as other experimental protocols applied to NREM Sleep, General Anesthesia (GA), and the waking state, provide some evidence to suggest that not only can self and consciousness dissociate, MS might be a necessary precondition for conscious experience. Finally, these findings have implications for the science of consciousness: it has been suggested that "levels of consciousness" (LoC) is not a legitimate concept for the science of consciousness. But because we have the conceptual and methodological tools with which to refine investigations of

Abbreviations: ACC, anterior cingulate cortex; ALFF, Amplitude of Low Frequency Fluctuations; BA9, Brodmann Area 9; BOLD, blood-oxygen-level-dependent; CRS-R, Coma Recovery Scale-Revised; DES, Descriptive Experience Sampling methodology; DMN, Default Mode Network; DoC, Disorders of Consciousness; EASE, Examination of Anomalous Self-Experience; EMCS, Emerging from Minimally Conscious State; ERP, Event-related Potentials; GA, general anesthesia; GABA, *gamma*-Aminobutyric acid; GNW, global neuronal workspace; HG, Heschl gyrus; IFT, isolated forearm technique; Levels of consciousness, LoC; mPFC, medial prefrontal cortex; MCS, Minimally Conscious State; MCS-, Minimally Conscious State MINUS; MCS+, Minimally Conscious State PLUS; MS, Minimal Self; MSH, Minimal Self Hypothesis; NREM, non-rapid eye movement sleep; PACC, Perigenual Anterior Cingulate Cortex; PCC, Posterior Cingulate Cortex; PCI, Perturbational Complexity Index; PFC, prefrontal cortex; RDoC, Research Domain Criteria; SFG, Superior Frontal Gyrus; SON-FV, subject's own name presented by a familiar voice; SSM, Sleep State Misperception; UWS, Unresponsive Wakefulness Syndrome; tACS, transcranial alternating current stimulation; TBI, Traumatic Brain Injury; TPJ, Temporoparietal Junction.

* Address: Graduate Institute of Mind, Brain and Consciousness, Taipei Medical University, 250 Wuxing Street, Taipei, Taiwan.
E-mail address: timlane@tmu.edu.tw.

https://doi.org/10.1016/j.concog.2020.103029
Received 7 September 2019; Received in revised form 7 September 2020; Accepted 26 September 2020
Available online 19 October 2020

MS, we have the means to identify a possible foundation—a bifurcation point—for consciousness, as well as the means by which to measure degrees of distance from that foundation. These neuroimaging investigations of MS position us to better assess whether LoC has a role to play in a mature science of consciousness.

"The universal conscious fact is not 'feelings and thoughts exist,' but 'I think' and 'I feel.'" (James, 1981, p. 221)

1. Introduction

The epigraph describes the relationship between self and consciousness as integral. The intuitive ground for this characterization has been widely shared among philosophers for millennia (Barresi & Martin, 2011). The intuition strikes many as so natural that self is felt to be related to experience, just as branch is related to branch-bending (Shoemaker, 1996, p. 10). Plainly stated: "Experience is impossible without an experiencer" (Strawson, 2011, p. 253; Frege, 1956, p. 299). Let us dub this the Minimal Self Hypothesis (MSH): it is necessarily the case that whenever there is a conscious experience, there is a self.

Philosophy is not the only discipline to investigate this relationship. Psychologists have been no less interested, though progress has been elusive (Klein, 2014). Although the psychological sciences opened the door to empirical investigations of how self and consciousness are related, many conceptual problems remain unresolved. As for problems dogging the scientific investigation of "consciousness" (e.g. Chalmers, 1996)[1], sufficient progress has been made in development of experimental protocols (Mashour, Roelfsema, Changeux, & Dehaene, 2020), methods for measurement (Massimini & Tononi, 2018), and mathematical modeling (Kleiner, 2020), that signs of progress are evident. The same cannot be said for "self."

Neither is there a consensus as regards the characteristic features of "self," nor is there a consensus concerning what can rightly be regarded as its paradigmatic cases (Olson, 2002). Fortunately, over the last two decades, new opportunities for investigation of self have emerged, due in part to the discovery of the Default Mode Network (DMN). The DMN comprises a set of brain regions that evince spontaneous increase in activity when subjects are at rest, viz. when not performing a task (Buckner, 2012; Raichle et al., 2001). This discovery is relevant because the DMN's intrinsic activity contributes to neuronal instantiations of self (Gusnard, Akbudak, Shulman, & Raichle, 2001; Sheline et al., 2009; Smallwood & Schooler, 2015).

Naturally, the DMN's functions are not confined to the self, nor are self-related functions confined to the DMN (Davey, Pujol, & Harrison, 2016). Nevertheless, discovery of the DMN's functions has helped pave the way for neuroimaging investigations of self. What is more, these investigations are not limited to probes of regional activity: neuroimaging is also used to probe temporal (Knyazev, 2013; Knyazev, Slobodskoy-Plusnin, Bocharov, & Pylkova, 2011) and biochemical (Lou, Thomsen, & Changeux, 2020) substrates of self.

Nevertheless, to date neuroimaging investigations of self have tended to be diffuse and inchoate (e.g. Frewen et al., 2020), insufficiently focused to enable testing of the MSH. The intent here is to narrow the focus and explicate concepts in such a way that the MSH becomes experimentally tractable. That is to say, I identify characteristic features of "self," along with paradigmatic phenomena wherein those features can be fruitfully investigated. At the same time, I describe operations that can facilitate investigation.

Investigations of the atom are an exemplar of how to transition from philosophical to scientific work (Chalmers, 2009).[2] As with self, atoms engaged the attention of philosophers for millennia. Just as self has been taken to be fundamental to or the "basement" of consciousness, so too atoms were taken to be fundamental to matter. The modern concept of the "atom" has helped explain the great variety of different forms exhibited by matter; one goal here is to begin assessing whether "self" can play a similar role in helping to explain the great variety of different forms exhibited by consciousness.

In the case of the atom, investigations stalled for centuries. Those investigations only began to gain traction in the 19th century when a succession of models were operationalized in ways that made the atom amenable to scientific probes, probes that were able to incorporate applications of new technologies, e.g. the Crookes tube. It was by virtue of adding experimental constraints to conceptual innovation—notably, abandonment of the idea that atoms are indivisible—that the relevant physics was able to make progress. In part taking a cue from early understanding of the atom, and in part following the lead of recent investigators who have engaged empirical science by employing "self" to explain pathological conscious experiences, I focus on the Minimal Self (MS). The ambition is to articulate the relationship between neuroimaging investigations of MS and consciousness in such a way as to delineate a comprehensive approach to testing the MSH—self is the necessary foundation for conscious experience.

The manuscript comprises the following sections. First, I describe how "MS" has been used as a posit to help explain thought insertion. Second, adhering to work on psychopathology, in order to ensure consistency of referent, I explicate "MS." Third, I assess phenomenological psychiatry's methodology for interpreting the explanatory role of MS. Fourth, I introduce the possibility that MS and consciousness dissociate, and suggest how this dissociation might assist with assessing MSH. Fifth, I describe how neuroimaging

[1] For purposes of this manuscript I bracket concerns about conceptual foundations for the scientific study of consciousness—e.g. those raised by Chalmers (1996), Levine (2001), and Nagel (1974). I believe that, at least in part, these concerns should not be regarded as unique to studies of consciousness and that they harbor unreasonable expectations of scientific explanation (Lane, 2020).

[2] Many similar examples can be found in the history of science, e.g. the concept "rationality" (Lane, 2017). But the "atom's" evolution from a purely philosophical to a solidly scientific concept, because of its emphasis on what is minimal, is best suited to serve as an analogy for the relationship between MS and consciousness.

probes of the DMN and other aspects of the brain's spontaneous, intrinsic activity can be used to detect the neural substrate of MS. Sixth, I argue that neuroimaging probes of Disorders of Consciousness (DoC) can be used to assess MSH. Seventh, I limn the pattern of inference that suggests neuroimaging probes of DoC lend support to MSH. Eighth, I show how this approach can be extended into other areas of empirical investigation—sleep, anesthesia, and quotidian states. Finally, in a concluding section, I argue that methodological progress which follows from rendering MS empirically tractable stands to benefit the science of consciousness, in particular our understanding of DoC.

1.1. Self and thought insertion

Investigations of MS by those working at the intersection of phenomenology and psychiatry have been ongoing for more than two decades (e.g. Zahavi, 2011; Parnas & Sass, 2011; Gallagher & Zahavi, 2008). Consistent with the MSH, these investigators also treat MS as a "necessary, universal" component of consciousness. More formally, with "MS" understood as "Minimal Self" and "C" as "consciousness," the thesis can be expressed thus: C⇒MS.

Recently, Henriksen, Parnas, and Zahavi (2019) have extended these analyses, defending their version of the MSH. They argue that attempts to adduce evidence from cases of thought insertion to show that the MS is neither necessary nor universal, fail. This defense can serve as a starting point for refined investigation of how self and consciousness are related. Among other things, the strategy is to interpret the necessity-relationship as empirical and take steps toward identifying a causal nexus into which MS can be situated.

Critics of phenomenology's characterization of MS—and, by implication, the MSH—have argued that thought insertion is a counter-example, because patients claim some thoughts do not belong to self. These subjective reports from psychiatric patients seem to imply that MS is not necessary for consciousness. This empirically motivated challenge to MS is similar in spirit to Dalton's first steps at integrating theoretical and experimental analyses of the atom (Kedov, 1949): similar, both in that theoretical concerns are compelled to engage a new constraint, and in that the challenge precipitates a skein of conceptual and empirical responses. In the case of MS, Henriksen et al. point out that critics evince an inadequate clinical grasp of thought insertion, a contentious tendency to take patient claims at face value, and a penchant for cherry-picking examples (cf. Lane, 2015).

Not unlike Dalton's model of the atom (Kedov, 1949), phenomenological-psychiatry's methodology can serve as an impetus for development of more rigorous methods. Phenomenological psychiatry's approach comprises qualitative data collection that takes MS as a crucial posit. By so doing Henriksen et al. have been able to offer a "unifying account" of outwardly heterogeneous symptoms. In fact, there is even indication that inclusion of MS in this research program has helped usher in a "renaissance" of phenomenological psychopathology.

In short, Henriksen et al.'s strategy—details of which are addressed below—goes beyond mere refutation of attempts at using thought insertion to challenge this version of the MSH. They also provide a "new account" of how MS is affected in schizophrenia. The new account aspires to provide an adequate explanation of multiple schizophrenic symptoms.

Although MS can play a role in explaining some conscious experiences, the approach they adopt persists in leaning heavily upon conceptual analysis, analysis that allows for too many degrees of interpretive latitude, thereby limiting its empirical plausibility. Although these interpretations can be paired with psychometrics (e.g. Martin, Franck, Cermolacce, Coull, & Giersch, 2018), nevertheless, MS remains a construct for which there is no consensus as to how it can be measured or referenced in an experimental setting. Therefore, determination of whether it can indeed can help explain conscious experiences stands to benefit from extending its reach to include more phenomena and more methods for testing (Feynman, Leighton, & Sands, 2010, pp. 2-16). Extending the reach of the MSH carries with it an additional explanatory virtue—it helps minimize the possibility that hypotheses of narrow scope become locked into closed systems that shield them from empirical challenges (Wolpert, 1993, p. 134).

For these reasons, and motivated in part by research on the DMN, I adumbrate an alternative means of investigating MS, with regard to phenomena that have yet to be carefully investigated. Success in expanding the scope of phenomena covered and the methods for measuring MS, when combined, can promote greater clarity about MS's explanatory role. In other words, the purpose is to assess potentially converging lines of evidence for the role that MS is conjectured to play in the science of consciousness.

From the start, however, as was the case with the evolution of models of the atom, it is to be expected that when transitioning from purely conceptual to robustly experimental investigations, models of the MS are also likely to evolve. It is in this spirit that "MS" is explicated below. If the atom analogy is cogent, these considerations should be regarded as a point of departure, not an end point.

1.2. The concept of minimal self

Henriksen et al.'s concern with MS is not unique. Cognates are plentiful, including: "for-me-ness," "minimal selfhood," "minimal experiential self," "ipseity," along with various other terms, each referring to the same phenomenon (Lane, 2012). Emphasis on "minimal" is intended to mark a contrast with alternative understandings of "self." Among these less atomistic understandings are the narrative or autobiographical self (Damasio, 2010; Hardcastle, 2008), use of efference copies to distinguish between my movement and movement that emanates from the world (Churchland, 2013), the morally responsible self (Doris, 2015), or the bodily self (Lane, Yeh, Tseng, & Chang, 2017; Park & Blanke, 2019). The concern is not with how we create narratives, how we navigate, how we achieve moral agency, or how we experience our bodies. Instead, the fundamental claim is that "phenomenal consciousness necessarily entails reflexive (pre-reflective) self-consciousness in the sense that there is something it-is-like *for me* to have or live through experiences" (Henriksen et al., 2019). To a first approximation, and as a starting point for investigation, this is the sense of "self" referred to by the MSH.

For-me-ness is taken to follow from the 1st person, subjective character of experience (cf. Zahavi, 2014, p. 22). It contrasts with

what is inter-subjectively accessible: e.g., I experience my pain in a way that differs from how others experience my pain.[3] Tables and chairs are inter-subjectively available, but experiences, like thoughts, are not. Even when a thought is described as having been inserted, it is still "my" thought: the denial that a thought belongs to me is a denial of "personal ownership." But there is a self, more minimal—the individual who has the 1st person experience, and that self retains a "perspectival ownership" over alien, inserted thoughts. In other words, "epistemic asymmetry is preserved," for how I know these thoughts differs from how others know them (cf. Zahavi & Kriegel, 2016, pp. 44-45). In this respect, for-me-ness is preserved in passivity or in other psychopathological experiences, like thought insertion.[4]

The authors' point is not that the MS is unaffected in psychopathological experiences; on the contrary, for-me-ness is affected. Here they explicitly reference James's characterization of how our own thoughts and feelings are related to MS: in quotidian contexts, thoughts and feelings exude "warmth and intimacy." It is this warmth and intimacy that Henriksen et al. claim is absent during thought insertion and some other psychopathological experiences.

Henriksen et al. maintain that patient descriptions suggest a felt distance between experience and experiencer. What is more, they argue that these descriptions are not the result of self-observation or introspection; indeed, these descriptions are not the result of any *reflective* mental activity. Instead, these descriptions result from "involuntary self-witnessing." That this self-witnessing is involuntary is taken to imply that that MS, for-me-ness, is already affected at the pre-reflective level. Specifically, the pattern of inference is that because these experiences are forced upon the patient—the experiences are "involuntary"—therefore, it is MS that is disrupted.

1.3. The empirical, clinical support

But what empirical, clinical support is there for the claim that the pre-reflective MS is affected? The principal source is a scale—EASE: Examination of Anomalous Self-Experience—and its applications, the development of which has helped facilitate systematic investigations of anomalous experiences.[5] The scale originated in clinical phenomenological interviews with incipient schizophrenia spectrum patients; findings from these interviews, e.g. patient reports of "persisting identity void", in turn motivated further systematic data collection (Parnas, 2005, p. 237). Henriksen et al. aver that EASE-guided investigations of schizophrenia spectrum patients provide evidence of disturbance to the MS, "to some extent."

Because the claim is that the MS is disturbed in some psychopathologies, it behooves us to consider how the EASE scale's application during interviews. Parnas (2005, p. 238) emphasize that the interview should be semi-structured, open-ended and that, ideally, the interview should comprise "patient-doctor mutually interactive reflection." The investigator poses a question to which the patient responds, and the investigator might then reformulate the answer. The patient can then correct the reformulation. Moreover, because the goal is to identify essential features of the experience, the investigator's probes involve "imaginative variation," in an effort to identify and eliminate descriptions of the patient's experiences that are merely accidental or contingent. In short, the process is hermeneutic, involving copious reflection.

There is a problem with data collected in this way. They are adduced as support for the hypothesis that the pre-reflective, MS is disturbed, albeit not dissolved. But the interviews require "mutually interactive reflection" that includes potential reformulation of patient descriptions by the investigator, even potential elimination of some patient descriptions. This hermeneutic maneuver allows for too many junctures at which the data can be contaminated. If one of the main reasons for regarding these experiences as pre-reflective is that they are driven by "involuntary self-witnessing," a minimalist interview technique would be more likely to provide credible data. "Interactive reflection" that includes editing of patient reports might inform about MS, but data thus collected would be hard-pressed to confirm the hypothesis. In addition, falsification would be out of the question.

The point is not to dismiss the value of these investigations, especially in virtue of the importance and the difficulty of adequately characterizing the content of phenomenal consciousness. Even in the era of the Research Domain Criteria (RDoC), self-report matters (Lilienfeld & Treadway, 2016). Qualitative methodologies of the type cited by Henriksen et al. are important aids to understanding consciousness, at least to getting clear about the phenomenon in need of explanation. But having already discovered DMN, alternative

[3] Cases such as mirror-touch synesthesia (Banissy & Ward 2008; Medina & DePasquale, 2017; Salinas, 2017) and craniopagus twins who are connected at the thalamus (Lane, 2014; Langland-Hassan, 2015) suggest that this assumption should not be taken to be self-evident. But I will not develop this point here. For a detailed investigation of the possibilities of "mindmelding" see Hirstein, 2012.

[4] As is the case for epistemic asymmetry, just what "perspective" amounts to here is not self-evident: for example, in a case of somatoparaphrenia one patient was only able to have tactile experiences when told that someone else, other than self, would be touched (Lane, 2011). Importantly, under carefully controlled conditions, repeated multiple times over two days, it was a *necessary* condition for feeling a touch on the hand that the person was primed to expect someone other than self would be touched. Arguably, the necessary condition here was not for-me-ness; instead, it was for-*other*-ness.

[5] Naturally this is but one of many similar instruments that aid with investigations into such phenomena. Others include the *Cambridge Depersonalization Scale* (Sleigh et al., 2010), the Sense of Agency Scale (Tapal, Oren, Dar, & Eitam, 2017), and the Sense of Self Scale (Flury & Ickes, 2007). Moreover, MS, as characterized by phenomenology, has also been developed in the context of neuroscience investigations of bodily illusions (Blanke & Metzinger, 2009).

methods for investigating whether or how MS is related to consciousness are available: multi-modal neuroimaging methods can be applied with greater rigor,[6] can extend investigations of MS into other aspects of consciousness, and are less dependent on the presuppositions of phenomenology.

1.4. Where else to look for evidence of MS?

One problem with the version of MS described above is that it is "weak or thin" (Gallagher & Zahavi, 2008, p. 58). If it is so minimal that the most straightforward way of identifying it is to contrast it with how it is not knowable—viz. inter-subjectively—then self can begin to seem trivial, or insubstantial, not suited to the explanatory role it is intended to play.[7] Writing of this and other issues in this vicinity, Seager (2016, p. 153) notes: "as we weaken the concept of self that figures in experiences of for-me-ness we approach the limiting case that for-me-ness is nothing but phenomenal consciousness." When we add the concern that some empirical cases, e.g. mirror-touch synesthesia (see Fn. 3), suggest that even the proposed epistemic distinction may not be universally true, it can begin to seem that the phenomenon is not being characterized accurately.

It should also be noted that, arguably, the most scientifically prominent theory of consciousness—global neuronal workspace (GNW)—gives short shrift to self, treating it as nothing more than a mental representation of some aspect of "me"—e.g. my body, behavior, feelings or thoughts (Dehaene, 2014, pp. 23-24; Graziano, Guterstram, Branden, & Wilterson, 2020). Advocates of the GNW do not shy away from admitting that their concern is with access consciousness, not phenomenal. But arguments adduced to the effect that distinguishing access from phenomenal consciousness is nothing more than a chimera (Dennett, 2018), lend support to the idea that GNW is making progress on the project of understanding consciousness, despite disregarding the access-phenomenal distinction along with the entity that allegedly shadows or imbues phenomenal consciousness, the self.

Clearly, however, there is no consensus view, and there is insufficient warrant for proclamations that whatever phenomenal consciousness is, it will be explained, eventually, just so long as we continue to focus on understanding access. Moreover, I agree with Kriegel (2009), another advocate of the view that phenomenal consciousness involves for-me-ness, that there are facts of the matter about phenomenology and gaining clarity about these facts is necessary to developing an adequate explanation of consciousness (Lane, 2015). So where does this leave us?

There are means of investigating MSH that do not require "patient-doctor mutually interactive reflection." There are alternative methods that enable us to disentangle MS from consciousness, to determine whether it is pre-reflective, or in some other respect prior to consciousness. What is more, the scientific protocols in question adhere to a concept of "self" that is vanishingly thin. In many instances it is only necessary to gauge spontaneous reactions to stimuli like the sound of one's own name, singular 1st person pronouns, or similar proxies for MS. In this regard not only is the stimulus thin, reflection is not necessary.

Teasing self and consciousness apart in such a way that one avoids arbitrary dismissal of the possibility that the latter requires the former is not easily accomplished. But it can be approached by examining a particular type of phenomena. The relevant phenomena are conditions in which consciousness is lost or diminished, as in NREM sleep, general anesthesia (GA), or Disorders of Consciousness (DoC).

DoC in particular affords a unique opportunity, because recovery of consciousness can take time, even decades (e.g. Voss et al., 2006). The case of Martin Pistorius is illustrative in this regard (Pistorius, 2011): when he was 12 he fell ill, and within a few months had lapsed into a "vegetative state," now referred to as "Unresponsive Wakefulness Syndrome" (UWS). While suffering from UWS, for more than three years, he lacked consciousness of the environment as well as of self. But, beginning at about 16 years of age, he slowly re-emerged into consciousness: even though his re-emergence was not detected until he was 25, the recovery of consciousness seems to have been completed by the time he was 19. From onset of his injury until complete, albeit undetected, recovery, there was a seven-year period.

The protraction of time between loss and recovery of consciousness allows for specialized, periodic probes that can inform as to what is concealed by apparent inability to respond, latent mental processes, including MS. If evidence of MS can be detected under these special circumstances, we might be able to determine that in at least one important respect, MS *precedes* consciousness (cf. Sterzer, Mishara, Voss, & Heinz, 2016, p. 9). Obviously, this finding alone would not confirm that MS is necessary for each instantiation of phenomenal consciousness, but on a causal model of necessary and sufficient conditions, showing that MS precedes consciousness can help to shed light on how the two are related—whether in fact there is a MS that is necessary, universal, and pre-reflective.

The strategy then is to investigate circumstances in which MS and consciousness dissociate. Investigations should target cases wherein MS might be realized, even when participants are in a non-conscious state. When trained on these phenomena, the multi-modal tools of neuroimaging can help assess the MSH. What is more, this line of investigation might help clarify the explanatory

[6] The point here is not to give neuroimaging a free pass, at the expense of hermeneutics. It would be naïve to suggest that neuroimaging methodologies are unproblematic: to cite just one problem that is now widely understood, neuro-images are inferentially distant from neural activity (Roskies, 2010; Botvinik-Nezer et al., 2020), even when complemented by other methodologies, like machine learning (Carlson, 2018). But neuroimaging methods are evolving rapidly, in large part because they are situated within a thoroughly interdisciplinary milieu (Bowman, 2014), that helps reveal and constrain the chain of inferences, even for investigations of LoC (Schulman, 2013, pp. 128–131).

[7] Advocates of this version of the "MS" sometimes seem more intent on treating the concept as part of the explanation for phenomena like thought insertion; at other times, their concern is to emphasize that, despite appearances (e.g. "this thought does not belong to me"), MS is necessarily a part of all instantiations of phenomenal consciousness. Because the authors fail to keep the phenomena-to-be explained distinct from the explanatory framework, their explanation risks circularity (Hempel, 1965, p. 373).

role that MS can play in the developing science of consciousness. Focusing on instances wherein consciousness is lost throws into relief how MS is related to "levels of consciousness" (LoC). This focus also helps reveal how MS can serve as a legitimate theoretical concept for the maturing science of consciousness.

1.5. MS in the brain's spontaneous, intrinsic activity

Self, operationalized as how we react to self-related stimuli when performing a task, can be investigated by assorted neuroimaging techniques (e.g., Gusnard et al., 2001; Northoff & Bermpohl, 2004; Qin & Northoff, 2011). But phenomenologists have expressed concern that paradigms such as those that involve facial self-recognition or adjectival self-attribution only probe neural correlates of "objectifications" of self; these paradigms should not be regarded as providing insight into "experiencing oneself as a subject" (Zahavi & Roepstorff, 2011, p. 147). In other words, reclining in a scanner, when participants see photographs of themselves, and then render a judgment "that is me," the relevant sense of "self" is self-as-object (Lane & Liang, 2011, p. 80), an externalized view of self, the self that is inter-subjectively accessible, not the private self of 1st person experience. Vogeley and Gallagher (2011, p. 128), who are also concerned with the pre-reflective self, emphasize that because tasks like facial self-recognition require reflective evaluation of a visual stimulus, it follows that the self being investigated is not the pre-reflective self.[8] In sum, the objections are that these operations capture neither the subjective experience of self nor the pre-reflective self.

But not all neuroimaging experiments focus on task performance. Many of these experiments focus on neuronal instantiations of the self within the DMN or other aspects of resting state activity: the concern is with neuronal activity that *precedes* perception of a stimulus or performance of a task. Based upon what is observed of intrinsic, spontaneous activity—spatial, temporal, and biochemical—it is possible to predict whether subjects will judge stimuli to be self-related (Bai et al., 2016; Qin et al., 2016). For example, elevated levels of glutamate in the Perigenual Anterior Cingulate Cortex (PACC) predispose higher, spontaneous α power (8–12 Hz), and the higher the spontaneous α power the more likely a participant is to judge a stimulus (e.g. a picture or white noise) to be self-related (See Fig. 1). In other words, neuronal preconditions for self—spatial and biochemical (elevated PACC glutamatergic activity) as well as temporal (spontaneous α power)—are integral to the brain's ongoing, spontaneous activity (Lane, Duncan, Cheng, & Northoff, 2016).

With respect to any given task, this spontaneous, intrinsic activity is pre-reflective, by definition—because it precedes the stimulus. Naturally, after being presented with a visual or an auditory stimulus, participants in these experiments are required to render judgments—e.g. whether an emotional picture is self-referring or whether white noise contains their name or the name of someone else. Although the rendering of judgments involves reflection, *what drives those reflections is the prior neuronal activity*. The pre-stimulus neuronal activity comprises the substrate of dispositions for experiencing stimuli in a particular way—here those dispositions are self-intimating.

The intent of investigating DMN or other resting state activity is to identify an empirical foothold for investigating the types of neuronal activity that encode for-me-ness. Before looking closely at cases of self and consciousness dissociating, it is necessary to know what type of neuronal activity to look for when consciousness is either lost or diminished. Elevated PACC glutamatergic activity or α power are not instantiations of MS, as such. Nevertheless, they can serve as intimations of when the for-me-ness disposition is manifest.

Allowing that no operation can perfectly capture the concept of "MS", is it empirically possible to move beyond neuronal activity that is self-*related*, to activity that is *specific* to self? Meta-analytic investigations of self, familiarity, other, and the DMN show that while self, familiarity, and the DMN overlap much of the time, the PACC seems to be more engaged in neuronal processing that is specific for self (Qin & Northoff, 2011) (See Fig. 2). Moreover, fMRI studies that compared regional activations in response to the sound of one's own name and the sound of "familiar others" yielded similar results. The subject's own name is associated with anterior cortical regions—in particular, the Anterior Cingulate Cortex (ACC)—whereas merely familiar stimuli are associated with the Posterior Cingulate Cortex (PCC) and the Temporoparietal Junction (TPJ), that is to say, posterior cortical regions (Qin et al., 2012).[9]

To be clear—there is no center-of-self in the brain. Neither is there a temporal nor a spatial center. Nevertheless, as was made clear from the early descriptions of the DMN, it is possible to identify clues as to when self-referential activity is ongoing. At such times some features of neuronal activity are more prominent than others. Not only does this concentration on spontaneous, intrinsic activity provide us with clues as to how self is instantiated, it also helps to blunt the force of certain criticisms: that is, not all self-referential paradigms require an objectification of self nor do they inform only about a post-reflective version of self. Experimental findings based upon intrinsic activity provide an empirical foothold, wherefrom we can begin to assess and, crucially, evolve the MSH.

1.6. Minimal self in disorders of consciousness

DoC are severe neurological conditions caused by varied etiologies: traumatic brain injury, hypoxic-ischemic neuronal damage, and sometimes by neurodegenerative or congenital disorders. Damage results in disruption to functional brain networks and the

[8] It should be noted that recent attempts to relate EASE to neurocognitive and neurophysiological measures, with a special focus on source monitoring, also depend upon collecting data that requires judgment and reflection (Nelson et al., 2019).

[9] It goes without saying that investigations of neuronal self-specificity remain at a fledgling state of development. Even for the studies reported here, although they are consistent in highlighting the role of the ACC, some suggest the caudal ACC is more relevant to self-specificity, while others suggest the perigenual ACC is more relevant. Of even greater importance, identifying regional activity is only a part of the problem; temporal and biochemical activity are equally important. The focus on the ACC here should be regarded as a means of gaining a foothold, and not as a definitive statement of how self-specificity or MS is realized in the brain.

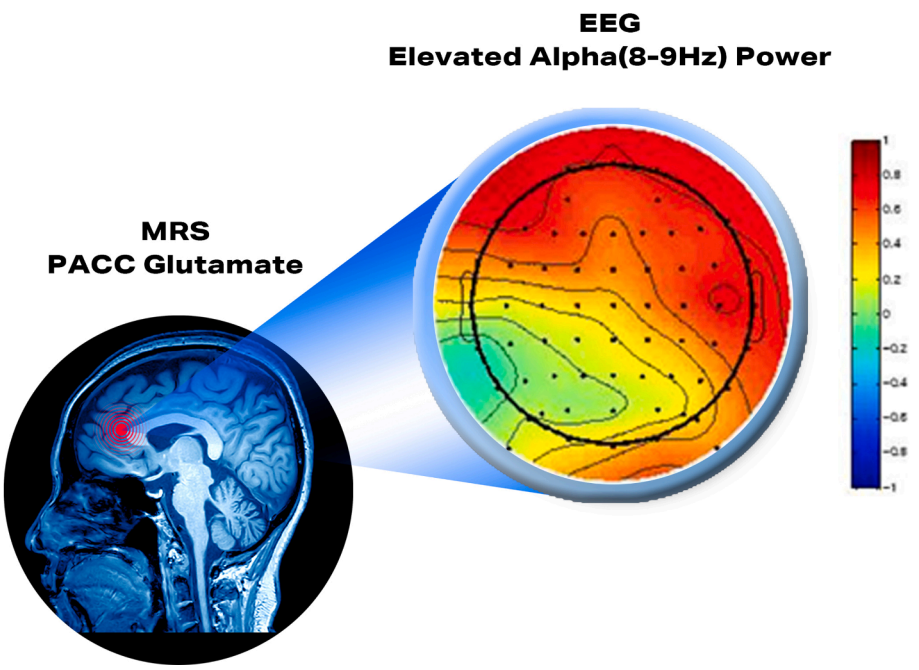

Fig. 1. Elevated levels of glutamatergic activity in the PACC—detected by magnetic resonance spectroscopy— predispose participant brains to evince higher, spontaneous alpha power; pre-stimulus alpha power inclines participants to experience stimuli—e.g. emotional pictures—as self-related. For EEG, the deeper the red hue, as seen in the frontal and central regions (roughly corresponding to the PACC's location), the more it seems that stimuli are experienced as self-related. Adapted from Figu. 4A, Bai et al. (2016).

neurotransmitter systems that underlie them (Bomalaski, Claflin, Townsend, & Peterson, 2017; Ciurleo, Bramanti, & Calabro, 2013). A feature shared by cases of DoC in the UWS stage is wakefulness without conscious awareness, for protracted periods of time.[10] UWS patients exhibit sleep-wake cycles and cyclic eye opening, but they evince no behavioral indications of conscious awareness, either of the external world or of themselves (Owen, Schiff, & Laureys, 2009): pointedly, they fail to exhibit sustained, reproducible, purposeful, or voluntary motor responses.

Diagnosis is typically based upon clinical history and behavioral observation, including use of the Coma Recovery Scale-Revised, CRS-R (Bodien, Carlowicz, Chatelle, & Giacino, 2016), or similar scales. Among other items of behavior, close attention is paid to whether patients are able to track objects with their eyes, or track the movement of their own eyes, when they are looking into a mirror (Binder, 2016). But these indicators are limited in how much they are able to reveal about consciousness (Overgaard, 2009): for example, even just the probability of detecting visual pursuit varies, depending upon the time of day, being most likely at 10:30 am and 3:00 pm, and least likely at postprandial times (Riganello, 2016).[11] If visual fixation or pursuit is detected in the clinical context, that behavior is taken to be an indication of consciousness, albeit "minimal" (Heine, Laureys, & Schnakers, 2016, pp. 26-29).[12] In virtue of the need to monitor behavior carefully, regularly, and at distinct time intervals, it is easy to see how a Minimally Conscious State (MCS) might go undetected, and a patient might be diagnosed as UWS, lacking even periodic episodes of conscious awareness. It is, therefore, not surprising that about 40% of DoC patients are misdiagnosed (Schnakers et al., 2009; van Erp et al., 2015). Stated plainly, absence of motor output does not entail absence of consciousness (Monti, Coleman, & Owen, 2009). It is important to keep this fact in mind when assessing the relationship between self and consciousness: at the current stage in our understanding of these two phenomena, arguments grounded in clinical or experimental research can only be regarded as provisional, since we cannot be empirically certain, in any given case, whether consciousness is absent. The relevant neuroscience remains at a nascent stage. The way forward is a strategy that includes ongoing conceptual refinement which informs and which, in turn, is informed by ongoing empirical work. Transformation from philosophy to science requires that experimental and conceptual work proceed in tandem.

One strategy for penetrating the façade of patient unresponsiveness, is by using auditory stimuli of various types, including the subject's own name presented by a familiar voice (SON-FV), in order to stimulate neuronal activity, while scanning with fMRI (Qin

[10] For a brief survey, written for a popular audience, that covers much of the research on DoC, beginning with work carried out in the first years of this millennium, see Owen, 2017.

[11] In as yet unpublished data, the author has observed that visual pursuit is sometimes manifest during the night, shortly after patients awaken spontaneously and when, unprompted, they open their eyes.

[12] MS and MCS, both contain the term "minimal." The MSH implies that MS is prior to consciousness. An alternative view is that MS and consciousness are not dissociable, such that neural correlates of one are also correlates of the other. Even if the alternative view is correct, however, "minimal" self and "minimal" consciousness are distinct, both conceptually and empirically.

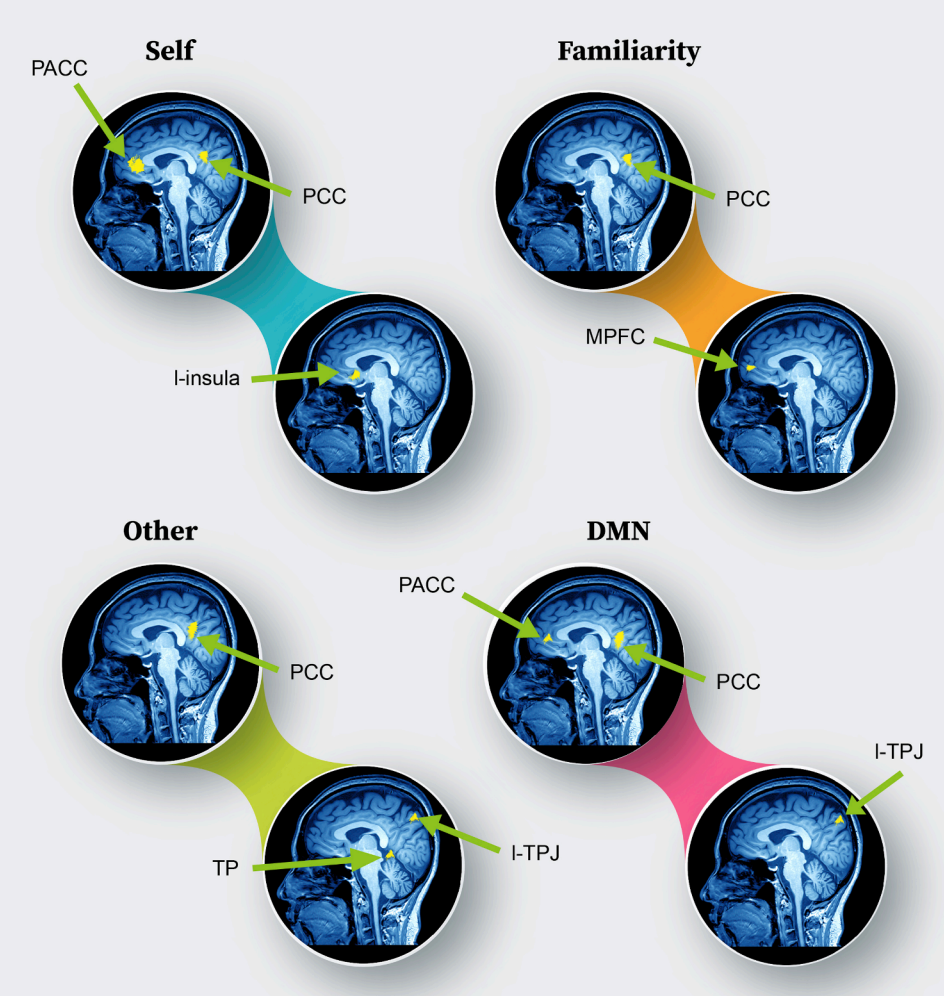

Fig. 2. Pictured are the results of a meta-analysis that compares self-specific stimuli (e.g. the sound of one's own name or a photo of self), to familiar stimuli (e.g. the face of someone familiar), to "others" (e.g. the face of a stranger), and these to the Default Mode Network (viz. the pattern of activity associated with the resting state or mind-wandering). Self-specific stimuli appear to be strongly associated with activity in the PACC. Adapted from Qin & Northoff, 2011, figure one.

et al., 2010). In this study, 11 DoC patients participated: seven had been diagnosed as UWS, while four were MCS. Scores from the CRS-R evinced a positive correlation with signal changes in the ACC; recall that the ACC was implicated previously in self-relatedness and self-specific investigations that focused on the brain's intrinsic, spontaneous activity (e.g., Bai et al., 2016). Findings from the DoC study revealed that the higher the LoC, the higher the degree of neuronal activity when processing these SON-FV stimuli. One biochemical factor that might underlie this difference in neuronal response is $GABA_A$ receptor binding in the ACC, as it has also been shown to predict the potential for improvement among DoC patients (Qin et al., 2015).

Most germane to my concerns here—the relationship between MS and consciousness—two of the seven UWS patients who evinced significant ACC neuronal response to the SON-FV stimuli during the scanning session, improved to MCS three months later. One plausible interpretation of these findings is that *even when patients are unconscious*, the neural substrate of MS responds to appropriate stimuli. If this interpretation approximates the truth, there is then at least this one respect in which *self precedes consciousness*. The capability of exhibiting neuronal response to this ever-so-thin stimulus can be a precursor of consciousness.

Keep in mind that the relevant definition of MS under consideration here is that it is a reflexive—pre-reflective—sense of for-me-ness. SON-FV is not only a minimal representation of self it also aptly captures the concept of "for-me-ness," at least as well as any alternative operation. What is more, to the best of our current knowledge—viz. CRS-R scores—some unconscious patients reacted to stimuli pre-reflectively. And some among those who reacted to the stimulus, later recovered, thereby suggesting the empirical possibility that MS can precede consciousness.

Obviously this is but one, short, empirical step. Can additional evidence be adduced to suggest that self is related to consciousness, in the manner described by the MSH? Huang et al. (2014) also focused on self in DoC patients; however, they used a distinct protocol, one that provides further support for the conjectured relationship. In this study eleven patients (six UWS, five MCS) were given a "task"

in the form of two questionnaires—one contained self-referential questions; the other, questions of simple fact. Self-referential questions were of the type, "Have *you* ever been to Missouri?"; questions of simple fact were of the type, "Are there 60 s in one minute?" As with the case of hearing SON-FV, self-referential processing in the ACC correlated significantly with patients' degree of consciousness. Consistent with the conjecture that self might be foundational for or prior to consciousness, when comparing self to factual questions, MCS patients evinced greater signal differentiation than did UWS patients. What is more, the two DoC patients with the highest ACC signal changes recovered after two months, suggesting that self-referential activity in this region might be "driving" the degree of consciousness (Huang et al., 2014, p. 2005). Here too self seems to be a precursor.

Both of these protocols target the relationship between self and consciousness. Both suggest that self can precede recovery of consciousness. But the stimuli differ. The simpler of the two is SON-FV; the more complex, paired questionnaires. For the former it might be argued that ACC signal changes are not indicative of a sense of self; perhaps the brain is only reacting to an over-learned stimulus. It would be less credible, however, to argue that sense of self is lacking when "responding" to personalized questions. The latter are not paradigmatic for over-learned stimuli. On the contrary, a natural interpretation of the difference in neuronal response to factual and self-referential questions is that personalized questions are evoking a sense of self.

Results from Qin et al. and Huang et al. are consistent. But they share specific methodological defects. Two shared problems are that both include small sample sizes for the DoC patients and modest follow-up assessments. Wang et al. (2015) attempted to correct for these problems by collecting usable data from 66 patients and by using CRS-R to follow up with each patient, 3, 6, and 12 months after fMRI data acquisition. Each patient was monitored for blood-oxygen-level-dependent (BOLD) signals elicited while hearing SON-FV. Of the 66 patients, 39 were diagnosed as UWS, and of these only seven evinced no activation whereas 16 evinced significant activation in the primary auditory cortex, and an additional 16 evinced yet more extensive activation that included the Heschl gyrus (HG). For the final group, within 12 months, those who had exhibited a higher level BOLD signal that extended into the HG, 12 of 16 (75%) recovered to either MCS or to "emerging from" MCS (EMCS). Of these 16, for 13 the etiology of their UWS was traumatic brain injury (TBI) and of these 13, 12 (92%) recovered to MCS or to EMCS. In short, for those who suffered from traumatic brain injury, if hearing SON-FV evinces BOLD signals that extend into the HG, their likelihood of recovering consciousness, at least to the MCS level, exceeds 90%. In short, this methodologically more thorough investigation, in that it suggests self is a precursor for consciousness, lends support the MSH.

Wang et al.'s investigation differs in at least one obvious way from the previous studies: whereas the other studies correlate self-referencing with neuronal activity within the ACC, here it is the HG that seems to be playing a crucial role. It is when hearing SON-FV that BOLD signals extending into the HG are evinced, and it is this extension of neuronal activity that correlates with the likelihood for achieving an elevated level of consciousness. In what respect might the HG be related to self?

An answer is suggested by the observation of increased HG activation *during spontaneous inner speech* (Hurlburt, Alderson-Day, Kühn, & Fernyhough, 2016). Hurlburt et al. employ a Descriptive Experience Sampling methodology (DES) that requires participants to record their inner experiences when signaled by a beeper, at semi-random intervals. The beep signals the participant to answer one question—"What was occurring in your inner experience at the moment of the beep?" (Hurlburt & Heavey, 2001). To invoke Hulburt's favored analogy, the goal is to "catch experience in flight." As is the case with investigations of the DMN and other resting state activity, the focus is on spontaneous, intrinsic activity.

DES is not without critics (Weisberg, 2011), but advocates argue that with just a short period of training, participants are able to describe inner experience with a fidelity that is not inferior to descriptions of external experience. Indeed, participants seem able to capture experience in "pristine" form; investigators emphasize that participants need merely describe what is directly apprehended, while avoiding a preference for any particular experiences and while putting aside presuppositions (Hurlburt & Heavey, 2015). In other words, to the extent that DES is successful, it enables participants to characterize thoughts or sensations that are directly apprehended at a given moment in time, "pristine inner experiences" that are pre-reflective, undisturbed by acts of observation, and that are "intimately personal, produced of, by, and for the individual in the individual's own manner" (Hurlburt, Alderson-Day, Fernyhough, & Kühn, 2017). Given the "intimately personal," pre-reflective nature of these experiences (Fernyhough, 2016), it may well be that the HG plays a role analogous to the ACC, contributing to neuronal instantiations of self.

1.7. Neuroimaging evidence and inferential support for the MSH

Several empirical premises conjoin to support the MSH. The pattern of inference is thus: First, neuroimaging evidence suggests that multiple types of spontaneous, intrinsic activity modulate whether we experience stimuli as self-related. Second, intrinsic activity that precedes stimulus presentation modulates how we experience stimuli—e.g. whether self-related or not—implying that judgements about the stimuli are driven by *pre-reflective*, neuronal activity. Third, further evidence suggests that we can move beyond self-*relatedness*, to identification of neuronal processes that are self-*specific*. Fourth, presenting UWS patients with the SON-FV stimulus (as a proxy for MS), evinces this self-specific neuronal response in a subset of patients. It is this subset of patients who later show signs of elevated LoC. Fifth, a distinct protocol also applied to UWS patients, one that employs the 1st person personal pronoun, a protocol not easily dismissed as eliciting mere overlearned neuronal activity, converges on the same result—MS precedes elevation of LoC. Sixth, an investigation that corrected for methodological shortcomings in the previous UWS studies replicated those findings. Conjoined premises 1–6 support phenomenological conjectures about MS—here dubbed the MSH: MS is empirically necessary for consciousness. In this context, empirical necessity implies that evoked responses to MS stimuli are prior to elevation of LoC.

Needless to say, the evidence does not entail the conclusion. The intent is not to deduce a conclusion from empirical propositions. Instead, the intent is to demonstrate the empirical possibility that the conjectured relationship between MS and consciousness used to explain psychopathological cases can be further tested, by examining distinct phenomena, and while employing distinct

methodologies.

When James asserted, "the universal conscious fact is not 'feelings and thoughts exist,' but 'I think' and 'I feel," he could only derive this conclusion through conceptual analysis. Many similar conjectures, likewise dependent upon conceptual analysis, can be traced at least to Aristotle; in some cases, self is prioritized, while in others consciousness is prioritized (Humphrey, 2011, p. 91). Indeed, the phenomenological interpretation of psychopathology described above is continuous with this tradition of conceptual analysis.

Here, however, the strategy is to take a cue from how the concept of the "atom" evolved from the purely philosophical to the robustly empirical. By building on findings about the DMN, we have a chance to increase the degree to which conceptual refinement is motivated by experimental work. This can be achieved by targeting a set of phenomena wherein self and consciousness might dissociate.

As was the case with the atom, the transition from conceptual to empirical is strewn with potential pitfalls. Several are addressed here. First, available evidence does allow for alternative interpretations. Second, the evidence might be confounded by difficulties pertaining to the diagnosis of DoC. And, third, proxy stimuli like SON-FV might fail to capture aspects of MS, such as "the 1st person perspective" or "subject-of-experience."

One worry pertinent to interpretation of evidence described above is the respect in which "self' might not be minimal. Schiff's (2010) Mesocircuit Hypothesis suggests an alternative: it calls attention to covariation between activity in the central thalamus—-damage to which is known to produce DoCs—and activity in the ACC. But Schiff emphasizes the ACC's role in the performance of a variety of cognitive functions, noting that "frontal executive systems" can drive or reciprocally increase activity in concert with the central thalamus. In other words, on Schiff's interpretation, it might be an "executive" or a higher-order self that is playing the crucial role in driving conscious systems.

An interpretation similar in spirit, albeit not in detail, is Dehaene, Lau, and Kouider (2017): they identify a role for self, understood as "self-monitoring," the capacity for a system to monitor its own information processing, a capacity that can dissociate from consciousness—where "consciousness" is operationalized as reportability. Here too the ACC plays an important role, especially as regards error detection. But here too the emphasis is on meta-cognitive executive functions, not "self" understood as pre-reflective. Going forward, one empirical possibility that warrants investigation is whether the relevant concept of "self" and its relationship to consciousness might be better understood as an executive or metacognitive function. Making progress on distinguishing MS from higher-order conceptions of "self" will require a pincers maneuver—further conceptual refinement of "MS" vis-à-vis "executive systems" in parallel with neuroimaging experiments designed to tease these two apart (Lane, 2020). Probes designed to distinguish between minimal and higher-order conceptions of "self" will require paradigms that control for the confounding influences of report, reflection, and judgment neural substrates (e.g. Tsuchiya, Wilke, Frassle, & Lamme, 2015; Block, 2019).

A second problem concerns the possibility that patients thought to be suffering from UWS might actually be conscious, at least to some degree, which would not be surprising, given the high rate of misdiagnosis for DoC (Schnakers et al., 2009; Vanhaudenhuyse et al., 2018).[13] If this were found to be the case for the investigations described above, it would suggest at least two alternative interpretations of the data: first, it could be that MS is not prior to consciousness, but it is a reliable indicator of the capacity for consciousness, possibly because in fact the two are not dissociable. If this were the case, it would still be consistent with the view of MS under consideration here—C⇒MS.

Another alternative is the converse, MS⇒C: perhaps conscious experience is necessary for self (Humphrey, 2011). If this were the case, it would still be consistent with the view that the two are dissociable, albeit in the sense that conscious experience can occur independently of MS (Lane, 2012, 2015). Both hypotheses can be investigated, in tandem, by methods that target MS, conjoined with methods that target global consciousness (e.g., Casarotto et al., 2016). Given the current state of knowledge, both C⇒MS and its converse, MS⇒C, are viable.

Although both hypotheses are viable, the studies described above suggest that MS paradigms do help to distinguish among LoC, in such a way that lends support to the MSH. DoC are classified into distinct stages (Bruno, Vanhaudenhuyse, Thibaut, Moonen, & Laureys, 2011): after having been diagnosed as UWS, as patients begin to recover, if they exhibit the ability to track salient objects with their eyes or the ability to smile in response to emotional content, they are classified as MCS MINUS (MCS−). If they exhibit the ability to follow commands or to answer yes/no queries, they are classified as MCS PLUS (MCS+). And, if they exhibit the ability to communicate more fully, they are classified as having emerged from MCS (EMCS).

Both Qin et al. (2010) and Huang et al. (2014) discovered that degree of self-specific processing in the ACC correlates with LoC—whether UWS, MCS-, MCS+, or EMCS. In like manner Wang et al. (2015) discovered that higher level BOLD signals extending into the HG correlate with recovery into MCS or EMCS. These findings suggest that signal changes in response to MS-stimuli are tracking something indicative of LoC: the relevant neuronal changes are in step with LoC elevation. Therefore, even if some of the patients in these studies have been misdiagnosed, the findings still support the MSH. MS is necessary for consciousness.

Finally, a third problem is that the experimental operations might fail to capture the concept "MS". As touched upon above, if SON-FV simply primes overlearned, stored representations, then the evoked neuronal responses do not betoken for-me-ness. This worry, however, would only be compelling if self-specific neuronal responses were detectable only when responding to the SON-FV protocol. But Huang et al. (2014) identified a similar pattern when contrasting responses to two distinct questionnaires—self-specific and matters-of-fact. Moreover, results of Qin et al.'s (2012) meta-analysis cohere with those findings.

[13] To ensure accuracy when diagnosing DoC patients, it is necessary to consider many factors, including but not limited to posture (Wilson, Dhamapurkar, Tunnard, Watson, & Florschutz, 2014) and body temperature, as a proxy for circadian rhythms (Blume et al., 2017).

There is no consensus as regards what experimental operations are optimal at probing for-me-ness. But in quotidian circumstances responding to questions about oneself are paradigmatic, albeit in a non-technical sense. In a technical, experimental context, since results for SON-FV and singular 1st person pronouns converge on the same results, provisionally, both can be adopted as reliable probes.

And there is yet another reason to suggest that the coupling of questionnaires is an effective means of probing for-me-ness. The very fact that the different questionnaires evoke contrasting neuronal responses requires an explanation (Lipton, 1993). That this explanatory problem is relevant to the MSH is underlined by the convergence of results: SON-FV and contrasting questionnaires yield similar neuronal effects. In the case of contrasting questionnaires, because the principal difference between causal stimuli is that one set involves self while the other does not, it follows that the neuronal differences are best explained by the difference between self-specificity and matters-of-fact.

Despite the convergence of results, one might yet harbor doubts as to whether the relevant neuronal activations are indicative of a for-me-ness accompanied by self-experience. But *the hypothesis under consideration here is not that UWS patients necessarily undergo a self-experience* whenever they exhibit, say, specific signal changes in the ACC upon hearing the sound of SON-FV, or when comparing two sets of questions. As the history of empirical investigations into atomic models shows, the concepts will evolve in conjunction with novel investigations. In the case of the atom, what was once regarded as indivisible came to be seen as divisible. In like manner, although pre-theoretical intuitions might imply that for-me-ness is necessarily accompanied by self-experience, this may prove not to be the case. To invoke Levine's locution (Levine, 2001, p. 7), "what-it-is-like-for-me" may prove to comprise dissociable elements, such that for-me-ness is a component that can be instantiated independently of what-it-is-like. Indeed, the MSH implies that for-me-ness is the foundation upon which recovery of what-it-is-like is grounded—foundational in that it can be prior to.

Moreover, ambiguity in current usage of terms like the "1st person perspective" is already suggestive of how understanding of the relevant concepts might evolve. For example, locutions like the "1st person perspective" could be spatial metaphors that represent a "purely geometrical feature," an embodied perspective for seeing or hearing; they could, on the other hand, imply a system that is able to focus attention (Blanke & Metzinger, 2009, pp. 7–8). And it is the latter that is more suggestive of self-experience. But the main point here is to begin to move away from mere conceptual analysis and adumbrate a path forward for more robust empirical investigation of the MSH. By analogy to the "atom", what was once thought to be indivisible—what-it-is-like-for-me—might prove to have distinct components.

1.8. Tracking MS in other dissociations of consciousness and self

In keeping with the Feynman et al. (2010) vision of science, might there be additional areas into which this type of exploration can be extended? Although the caveats discussed in the previous section apply here as well, the received view is that other states are also paradigmatic for the loss or diminution of global consciousness. Among these are NREM sleep and General Anesthesia, GA (e.g., Brown, Lydic, & Schiff, 2010; Massimini & Tononi, 2018). To be sure, as we learn more empirical detail about each state, it becomes ever more difficult to trace a discrete line between conscious and non-conscious states (e.g. Windt, Nielsen, & Thompson, 2016; Bademosi et al., 2018; Banks et al., 2020). Nevertheless, at this stage in development of the science of consciousness, NREM and GA can still be assumed to differ from waking states in LoC.

The examples presented in this section—NREM, GA, and quotidian wakeful states—are intended as proposals for how to extend explanatory reach of the MSH. The goal is to determine whether investigations of MS can precipitate development of a robust research program, one that like investigations of the atom leads to new discoveries—e.g. electrons and their wave-like motion. Successful investigation of MS would not be predicated upon the explanatory adequacy of any individual line of research; rather, it should be predicated upon the degree to which findings cohere "with a broad range of evidence from independent sources" (Petersen, 2016, p. 2).

It has long been known that objective measures of sleep onset dissociate from the subjective perception or experience of sleep onset, both in insomniacs and in healthy subjects (Yang, Han, Yang, Su, & Lane, 2010). This discrepancy between subjective and objective measures is often referred to as Sleep State Misperception (SSM). Using simultaneous EEG-MRI (for high resolution data on temporal and spatial dimensions), Hsiao et al. (2018) discovered that for healthy subjects, in the temporal dimension, SSM is associated with higher α EEG activity during NREM. That is to say: relative α activity was found to be higher when subjects "misperceived" the state that they were in, thinking themselves to be awake when objective measures suggest otherwise. Consistent with this finding, Krystal, Edinger, Wohlgemuth, and Marsh (2002) discovered that paradoxical insomnia—subjective report of sleep difficulty despite normal polysomnographic recordings—is also associated with elevated α activity during NREM. Hsiao et al. (2018) opine that these findings suggest "sustained α rhythms might be a mediating factor in subjective awareness." Now recall Bai et al. (2016): They discovered that elevated levels of PACC glutamatergic activity predispose subjects to have higher, spontaneous α power, and this in turn disposes subjects to experience stimuli as self-related. An interpretation of SSM, one consistent with Bai et al., is that subjects feel they are awake, because of this sustained, self-related neuronal activity.

As for the spatial dimension, the total power of the BOLD fMRI signal was measured for the low frequency range (0.01–0.1 Hz), by examining the Amplitude of Low Frequency Fluctuations (ALFF). Here it was discovered that SSM is characterized by higher ALFF values in the left Superior Frontal Gyrus (SFG). This finding might be relevant to the MSH, because Goldberg, Harel, and Malach (2006) found significant activation in the left SFG, that extended medially to the ACC, when participants were engaged in a self-reflection task. They also found that this self-related processing dissociates from sensory-motor processing; indeed, the two processes can be mutually antagonistic. When combined, these findings for temporal and spatial dimensions suggest that as we descend into light sleep, consciousness of much sensory-motor processing is being lost, but self-related processing remains active. Interpreting these SSM findings in the context of the DoC discussion above, there might be a respect in which self precedes consciousness, during recovery from UWS,

and that it can also linger, as conscious content is being lost during the descent into sleep.

Might something similar obtain in GA? GA relies upon a complex cocktail of drugs that prevent movement, block memory formation, reduce anxiety, and relieve pain (Crowder, Palanca, & Evers, 2017), but not necessarily diminish consciousness (Eger & Sonner, 2006). Properly understood, however, GA is the attempt to induce a reversible "coma" (Brown et al., 2010). Nevertheless, expected and actual effects of GA remain points of contention (Mashour & Avidan, 2015), and there are numerous ongoing attempts to integrate investigations of GA with investigations of consciousness (e.g., Alkire, Hudetz, & Tononi, 2008). Moreover, there are many measures whereby loss of consciousness in GA allows for drawing comparisons to DoC and NREM (Casarotto et al., 2016; Liu et al., 2017). In principle then, here too possible dissociations of consciousness and self can be investigated.

One method for investigating self-related processes during GA is the "isolated forearm technique" (IFT). Russell (1993, 2006), for example, administers an anesthetic cocktail, but applies a tourniquet on the patient's arm, in order to prevent the muscle relaxant from entering the patient's hands and wrist. In this way, paralysis ("neuromuscular blockade") of the hand can be prevented, thereby enabling the patient to communicate by using hand movements. While the patient is under GA, prerecorded messages are played through headphones, and Russell *addresses each patient by name* (Russell, 2006, p. 347): e.g. "Name, name, this is Dr. Russell speaking: if you can hear me, open and close the fingers of your right hand." Russell demonstrated that, at least according to certain conventional monitoring techniques, response to commands could be achieved even in "general anesthesia with deep hypnosis." In the more recent study (2006), Russell included only 12 patients; each of the patients responded to commands, despite being at levels that satisfied conditions for "surgical anesthesia."

In the earlier study Russell (1993) included 32 patients. Not only were 23 able to squeeze his hand upon "hearing" their names, followed by his command, what is more, 20 were able to squeeze his hand, yet again, when queried as to whether they were experiencing pain. In sum, Russell demonstrated that when satisfying conditions for surgical anesthesia, patients could respond to commands and report contents of their sensory state, very much like what is observed in MCS.

The relevance of self is made explicit by Sanders, Tononi, Laureys, and Sleigh (2012, p. 951), who suggest that patient IFT responses are similar to what occurs during lucid dreams. That is to say, GA is similar to dreams in which patients remain physiologically asleep but, metaphorically, self "awakens" as the dream is ongoing (Windt, 2015, p. 111). IFT response resembles lucid dreaming, because both the "anesthetized" and "dreamers" become capable of intentionally performing diverse actions. In the case of lucid dreams, volitional eye movements are used to report on the dreams (Baird, Mota-Rolim, & Dresler, 2019, p. 305-306). This is analogous to the opening and closing of fingers during GA. Although characterization of these behaviors as intentional or volitional is defeasible (Fischer & Truog, 2015), the ability to answer a question—and then a follow-up question about what is being experienced—is paradigmatic for deliberate action (Marcel, 2004).

In keeping with the interpretation of DoC, NREM, and GA presented here, it is worth noting that those who frequently experience lucid dreams have relatively more grey matter volume in Brodmann Area 9 (BA9). These subjects also evince a stronger BOLD signal in BA9, when asked to evaluate their own thoughts, as compared to conditions when they are not monitoring their thoughts (Filevich, Dresler, Brick, & Kuhn, 2015). This is consistent with what was found for NREM, as BA9 occupies a portion of the SFG, the region that exhibits higher ALFF values in subjects who claim not to have been asleep, even though objective measures suggest they were asleep.

Previously investigations of GA did not attend to self per se. Therefore, the chain of inference that links these studies to the MSH is more oblique than it is for UWS. Recently, however, Sleigh, Warnaby, and Tracey (2018) have suggested that a framework for GA that includes self can increase "taxonomic precision," thereby making it possible to explain variability in GA experience. The taxonomy includes an analogue of MS, the "core self," regarded as minimal in that it involves only the "feeling of existence." Metaphorically, it serves as a "pilot light" for consciousness that is not extinguished until thoroughly degraded under GA. Moreover, this "pilot light" is underlain by activity in the Salience Network, of which the ACC is a crucial component (Menon, 2011). And, descriptions of emergence from GA suggest what the transition from bare for-me-ness to what-it-is-like-for-me might be like (Sleigh et al., 2018, p. 235): "the mind is conscious but conscious of nothing." To a first approximation, bare for-me-ness might usefully be regarded as consciousness without content.

Finally, even for quotidian, wakeful states it is possible to design neuroimaging tasks in such a way as to focus on MS. Tasks that are properly designed reveal how MS can both precede consciousness and reflective self. High time resolution investigations can capture "thin" distinctions between pre-reflective and reflective self. By hypothesis, time epochs associated with MS should emerge quickly—even prior to relevant conscious content—and should precede time epochs associated with reflective self. Indeed, Esslen, Metzler, Pascual-Marqui, and Jancke (2008; cf. Wolff et al., 2019) have demonstrated the feasibility of such an approach, using EEG data complemented by source-localization analysis.

The task requires participants to respond both to self-specific stimuli (e.g. "I am happy) and to other-referencing, (e.g. "She/He is happy"). To be sure, the main intent of the experiment is not to draw the contrast between self and other; drawing that contrast is merely a means for highlighting "I." The purpose of the experimental design is to distinguish among aspects of self. Toward this end, to enable teasing apart of MS and reflective self, each of the three words in a sentence were presented to participants for 200 ms; each word was then followed by a fixation cross of 500 ms duration, and the final fixation cross was followed by a question mark. In keeping with the main intent of the experiment, event-related potentials (ERP) of 700 ms were calculated after self—after the proxy "I"—and after the adjective (e.g. happy). Intriguingly, distinctive activation triggered by "I" alone is detectable as early as 134 ms post stimulus. Because it can take 250 ms or more, post stimulus, for an unexpected event or unexpected information to become conscious (Dehaene, 2014, pp. 121–131; Gray, p. 7, 2004; Pöppel, 2009), it seems that MS might indeed be able to precede consciousness even during ordinary wakefulness.

Moreover, results of the conjoined ERP analyses evince a clear regional distinction between pre-reflective and reflective selves: while for the reflective self only dorsal parts of the medial prefrontal cortex (mPFC) were activated, for MS there was strong

involvement of the ventral (mPFC). This finding corresponds to what Farb et al. (2007) discovered, using a different protocol. The finding is especially revealing, because dorsal and ventral seem to be functionally distinct; the dorsal mPFC is more involved in complex, reflective, cognitive operations, whereas the ventral mPFC is more involved in spontaneous, emotional and affective processes. The latter's involvement with affect laden processes seems to be a consequence of its dense interconnectedness with brain regions linked to emotion and motivation—the nucleus accumbens, the amygdalae, and the basal ganglia.

It would be foolhardy to claim that the operations of the Esslen et al. protocol succeed in overcoming all conceptual worries related to focusing on the MS as opposed to other aspects of self. But they do cohere with other investigations that regard the MS, as realized through activity in the ventral mPFC, to be the "entrance" to consciousness (Schore, 2003). It is in this respect that they help to confirm the MSH.

In order to bring consideration of MSH full circle, recall that when explicating the relationship of thought insertion to "MS" and "perspectival ownership," Henriksen et al. (2019) cite James's characterization—our own thoughts have "warmth and intimacy." If this is indeed a distinctive feature of MS's relationship to conscious thought, the characterization accords with the finding that the ventral mPFC mediates emotional processes. To extrapolate from these findings, it may well be the case that in thought insertion, the ventral mPFC is activated, but its connectedness to the nucleus accumbens, the amygdalae, and the basal ganglia is disrupted. That is to say, inserted thoughts exhibit for-me-ness, understood as perspectival ownership. But inserted thoughts lack "personal ownership", viz. the feeling of "warmth and intimacy**."** In the context of investigations of MS described here, it seems that MS is neuronally realized, but emotional content is absent.

1.9. Discussion and implications for the science of consciousness

First, baldly stated, for cases of UWS we have empirical evidence of a "thin" self—MS—that involves no reflection and that precedes consciousness. Neuronal response, while unconscious, to SON-FV, the singular 1st person pronoun, or other proxies for MS suggest that what-it-is-like-for-me might be divisible. Rather like discovery that atoms are divisible, what-it-is-like might prove to be dissociable from for-me-ness. Moreover, that proxies for MS can be used to predict whether an unconscious patient will recover consciousness, suggests that MS might be empirically necessary for phenomenal consciousness. The evidence from UWS lends defeasible support to the MSH.

Second, some evidence suggests that while the contents of phenomenal consciousness are being lost during NREM or GA, self can linger. To invoke a metaphor, after the fire of consciousness is lost, the ember of self continues to smolder. This is consistent with the view that MS is the "entrance" to consciousness. Or, in this instance, more to the point, the exit. The experience of SSM or of consciousness without content upon emerging from GA are attempts to characterize the liminal state between what-it-is-like and for-me-ness.

And, third, the sluggishness of consciousness makes it possible that even for quotidian states, what-it-is-like and for-me-ness can be teased apart. MS proxies like "I" can precipitate neuronal activity more swiftly than is typically achieved by neural correlates of conscious perception. Findings from these investigations suggest that sufficiently high temporal resolution provides a means for investigating MSH, without appeal to disorders of or altered states of consciousness.

In sum, not only might MS and consciousness dissociate, some evidence suggests that MS precedes consciousness and that it persists after consciousness is lost. If the relationship between self and consciousness is as described, an empirical argument can be made that self is a necessary, universal component of phenomenal consciousness. In keeping with the Feynman et al. (2010) vision of science, the theoretical foundation of MSH provides a platform wherefrom investigations of psychopathology, DoC, NREM, GA, and even quotidian states can be pursued experimentally, within a unified framework.

A problem for the conjectured relationship between MS and consciousness is that diagnosing UWS is shadowed by distrust of the clinical-behavioral diagnostic criteria. CRS-R assessments of LoC must be taken with a grain of salt; despite their enduring clinical utility, DoC clinical scales are just rough-and-ready approximations. But as Quine and Neurath observed (Quine, 1960, p. 3): science is like a boat which, "if we are to rebuild it, we must rebuild plank by plank while staying afloat in it." In other words, CRS-R, and other behavioral measures of LoC are planks on the boat that are indispensable at this point in scientific inquiry, but that will certainly be revised substantially, as we become better able to characterize and explain consciousness. Just as techniques for investigating the atom evolved, so too will they evolve for MS and consciousness. Keeping this proviso in mind, CRS-R and similar measures remain useful tools as we bootstrap ourselves into more sophisticated means of assessing the MSH.

Another problem might be that in developing the argument that DoC and other instances of lost or absent consciousness provide optimal conditions for testing MSH, I tack between neuronal and mental activity, potentially conflating the two. The intent is not to suggest a reductionist strategy; instead, the approach is agnostic as regards mind–body theories. Emphasis on neuronal activity is taken as one means of gaining empirical traction: when trying to determine whether an instantiation of self is pre-reflective, it is useful to identify intrinsic or evoked neuronal activity that has previously been shown to correlate with or predispose for self-specific processing. Taking empirical steps requires a foothold.

Finally, a third problem is that although I have identified specific regional, temporal, and biochemical activity as neural correlates of or predispositions for MS, mental states can be realized in distinct ways (Edelman & Gally, 2001; Figdor, 2010; Sajid, Parr, Hope, Price, & Friston, 2020). Given that investigations of MSH are nascent, it would be premature to claim that any particular neural correlate or predisposition universally and uniquely mediates realization of MS. There is also a methodological concern in this vicinity. Botvinik-Nezer et al. (2020) have demonstrated that results for neuroimaging studies can vary greatly, even when analyzing the same dataset and when testing the same hypotheses. Here too then the Quine-Neurath analogy is apt: what we have identified thus far are planks on a boat. While the boat can lead to deeper understanding, the boat itself must be rebuilt, plank-by-plank as we pursue a more

mature understanding of the MSH.

MSH, as rendered by James (1981, p. 221) is: "The universal conscious fact is…'I think' and 'I feel.'" For James this universal relationship is self-evident. This view is consistent with phenomenological psychiatry: the "I" is miscible with and necessary for phenomenal consciousness. If the interpretation of lost-and-recovered consciousness presented here approximates the truth, "I" is in fact a necessary precondition for conscious thoughts or feelings. For practical purposes, SON-FV is as thin and weak a representation of self as can be operationalized, and it is a stimulus that betokens for-me-ness. That this stimulus can evoke self-specific neuronal responses, which in turn predict recovery of consciousness, suggests that the hypothesized relationship between MS and consciousness is substantive. The empirically robust, experimentally driven approach described here is the first step in getting straight about the details of this relationship, and it points the way toward a viable research program. But where might this research program lead?

To cite one possibility, the cases of lost-and-recovered consciousness discussed above—UWS, NREM and GA—are all taken to be phenomena for which the global Level of Consciousness (LoC) is scalable along a single dimension. In a clinical context, for example, when measuring with instruments like the CRS-R, MCS is scored as higher than UWS. Recently, novel quantitative means for measuring LoC have been developed, e.g. the Perturbational Complexity Index (PCI). The PCI quantifies EEG responses to magnetic stimulation, in order to calculate complexity of causal interaction among thalamocortical circuits. An algorithm applied to the EEG signals compresses them, like a zip file, yielding a numerical score between 0 and 1, with 0.31 marking the cut-off between conscious and non-conscious (Casarotto et al., 2016). But Bayne et al. (2016, 2017) are dubious that any concept of LoC, including PCI, should be regarded as a legitimate construct in the science of consciousness. Perhaps the concept "MS" described here can help to show what explanatory role LoC can play in a mature science of consciousness.

Bayne et al., as a conceptual point, opine that a creature is conscious if and only if it has a subjective-point-of-view, which, intuitively, does not admit of degrees. We can vary in how many objects or properties we are conscious of, but this fact does not imply that one person is more conscious than another. Second, Bayne et al. argue that states cannot be given a determinate ordering relative to one another.

One implication of their critique is that if, say, we compare mild propofol sedation to REM sleep, the two cannot be accommodated on an absolute scale. Bayne et al. suggest that these states are only comparable with respect to distinct dimensions. They do not show how their speculation applies to any particular case, but they do propose two possible dimensions. One dimension involves "gating" of conscious contents; the other, "global availability," or a function for guiding cognitive and behavioral processes. Presumably, at least in principle, mild sedation might rank higher than REM on the gating dimension, while REM might rank higher than mild sedation on the global availability dimension. Or, vice versa.

Applying their speculation to DoC, it could turn out to be the case that on one dimension a given MCS patient might resemble UWS patients. On another dimension, that same patient might resemble other MCS patients. Their suggestion is that, in time, we will discover that dimensions like "gating" or "global availability" will enable us to better assess consciousness, because we are able to apply relative scales. There will be no need to pigeonhole everyone into an absolute scale.

As to whether a subjective-point-of-view admits of degrees, now that we have a foothold for tracking MS in conscious and non-

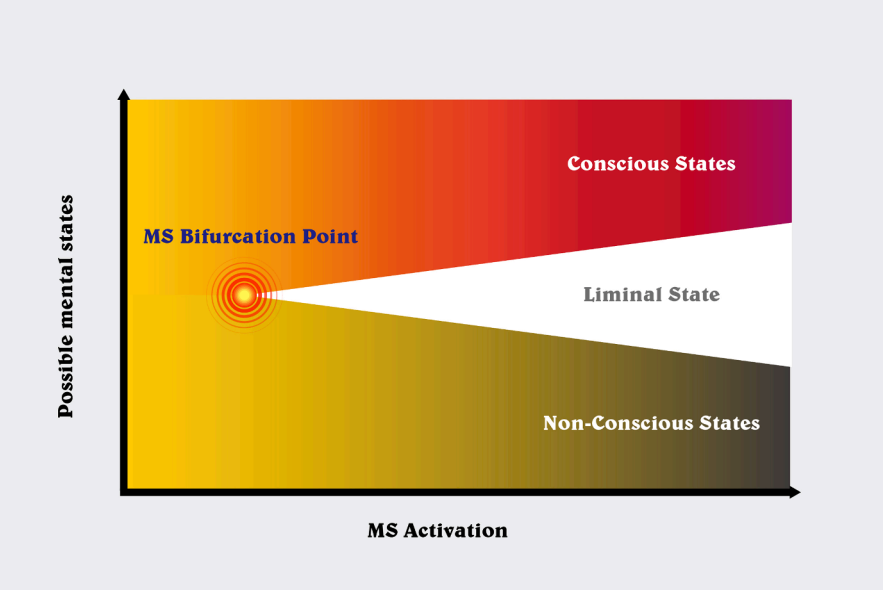

Fig. 3. This schematic figure represents MS activation on the horizontal axis, and variability of mental states on the vertical. Findings described here suggest that MS is spread across a mental landscape encompassing conscious and non-conscious states: that is, what-it-is-like-for-me might be dissociable. The schema depicts the possibility that neuronal increase in MS activation (for-me-ness) figures prominently in the rebooting of consciousness. When a bifurcation or tipping point is achieved, conscious states can be realized. Examples of liminal states include the feeling of being awake, although by most objective measures one is asleep, and emergence from GA when consciousness is recovered, albeit without content.

conscious states, we need no longer be constrained by the limits of conceptual analysis. This issue can be addressed experimentally. Because we are able to investigate the temporal, spatial and biochemical dimensions of MS, we can better assess whether subjective-point-of-view admits of degrees. Among other things, we have already adduced support for the claim that for-me-ness can obtain both with and without what-it-is-like. As to whether this counts against Bayne et al.—whether this can be taken to be a difference of degree—naturally much turns upon how "subjective-point-of-view" is unpacked. But when transitioning from philosophy to science, precisely the point is to reduce reliance on pre-theoretical intuitions and to entertain concepts that reflect the best science of the day, even if those concepts are counter-intuitive

On the view considered here, MS seems to be spread across a mental landscape encompassing conscious and non-conscious states. It remains to be seen, however, whether the rebooting of consciousness—emergence from UWS, GA, or NREM—has a distinct bifurcation or tipping point. Analogous to PCI, it could be that an algorithm which "zips" together multi-modal data from protocols that target MS will help identify a distinct point, and degrees of distance from that point. But it is an open question, one amenable to empirical probes, whether incremental accretions in MS or some other parameter triggers an abrupt qualitative change that marks emergence into consciousness or elevation of LoC (See Fig. 3).

As to the "gating" of content and "global availability" for performing specific functions, taking these to be the bases for the multidimensional state space presupposes that global states of consciousness can be adequately explained just so long as we thoroughly understand perceptual processes and cognitive functions. Although the final theory of consciousness might reveal that consciousness is reducible to an amalgam of cognitive and perceptual processes, since subjective experience is what we are trying to explain, it is striking that MS or other self-specific neuronal activity is neglected. If we only had the conventional *cognitive* neuroscience paradigms available for probing consciousness, this neglect would be less problematic. But because we now have paradigms for probing MS, or other aspects of self, and because self is the single feature most closely associated with consciousness, neglect of these paradigms becomes an obstacle to explaining consciousness (Lane, 2020). How might use of these paradigms matter, if the intent is to explain gating and global availability?

If James was correct that the universal conscious fact is "I feel." "I", understood as MS, might turn out to be the crucial bifurcation point. When MS is absent contents are gated; when gated, those contents are not globally available. Indeed, that MS might be the gate is suggested by the SON-FV probes of UWS, and that degrees of distance from that gate are measurable is suggested by the "smoldering embers" of self detected in those who have fallen into NREM sleep or who have been administered GA.

Bayne et al. are correct that clinical instruments like CRS-R are inadequate. But as planks that make bootstrapping possible, they do have a legitimate role to play in the emerging science of consciousness. More importantly, limitations of CRS-R do not imply that LoC should be discarded. One problem with the framework advocated by Bayne et al. is that they have omitted the feature of mind most closely associated with consciousness, the feature that might be most critical to explaining global consciousness—self.

Acknowledgments

My research is supported by funding from the Taiwan Ministry of Science and Technology (MOST106-2410-H-038-003-MY3, 104-2420-H-038-001-MY3), the Taiwan Ministry of Education (DP2-109-21121-01-N-07-01), as well as by a clinical trial grant from Taiwan's Ministry of Health and Welfare (IIT-1091-9). I am grateful to Patricia Huang for designing the figures. For discussions from decades past, I express gratitude to Carl G. Hempel, Julian Jaynes, Herbert Simon, and Tony O'Dea. I do as well express gratitude to the family members or guardians of more than 160 patients suffering from DoC whom I have been permitted to assess for inclusion in a clinical trial. And, I declare that I have no conflicts of interest to disclose.

References

Alkire, M. T., Hudetz, A. G., & Tononi, G. (2008). Consciousness and anesthesia. *Science, 322*(5903), 876–880.

Bademosi, A. T., Steeves, J., Karunanithi, S., Zalucki, O. H., Gormal, R. S., Liu, S., et al. (2018). Trapping of Syntaxin 1a in presynaptic nanoclusters by a clinically relevant general anesthetic. *Cell Reports, 22*(2), 427–440.

Bai, Y., Nakao, T., Qin, P., Chaves, P., Heinzel, A., Lane, T., et al. (2016). Resting state glutamate predicts pre-stimulus alpha increase during self-relatedness—A combined EEG-MRS study on rest-self overlap. *Social Neuroscience, 11*(3), 249–263.

Baird, B., Mota-Rolim, S. A., & Dresler, M. (2019). The cognitive science of lucid dreaming. *Neuroscience and Biobehavioral Reviews, 100*, 305–323.

Banks, M. I., Krause, B. M., Endemann, C. M., Campbell, D. I., Kovach, C. K., Dyken, M. E., et al. (2020). Cortical functional connectivity indexes arousal state during sleep and anesthesia. *NeuroImage, 211*(1), Article 116627.

Barresi, J., & Martin, R. (2011). History as prologue: Western theories of the self. In S. Gallagher (Ed.), *The oxford handbook of THE SELF* (pp. 33–56). Oxford: Oxford University Press.

Bayne, T., Hohwy, J., & Owen, A. M. (2016). Are there levels of consciousness? *Trends in Cognitive Sciences, 20*(6), 405–413.

Bayne, T., Hohwy, J., & Owen, A. M. (2017). Reforming the taxonomy in disorders of consciousness. *Annals of Neurology, 82*, 866–872.

Binder, H. (2016). Clinical evaluation of residual brain function and responsiveness in disorders of consciousness. In M. M. Monti, & W. G. Sannita (Eds.), *Brain function and responsiveness in disorders of consciousness* (pp. 37–49). New York: Springer Science+Business Media.

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

Block, N. (2019). What is wrong with the no-report paradigm and how to fix it. *Trends in Cognitive Sciences, 23*(12), 1003–1012.

Blume, C., Lechinger, J., Santhi, N., del Giudice, R., Gnjezda, M.-T., Pichler, G., et al. (2017). Significance of circadian rhythms in severely brain-injured patients. *Neurology, 88*, 1933–1941.

Bodien, Y. G., Carlowicz, C. W., Chatelle, C., & Giacino, J. T. (2016). Sensitivity and specificity of the coma recovery scale-revised total score in detection of consciousness. *Archives of Physical Medicine and Rehabilitation, 97*(3), 490–492.

Bomalaski, M. N., Claflin, E. S., Townsend, W., & Peterson, M. D. (2017). Zolpidem for the treatment of neurologic disorders: A systematic review. *JAMA Neurology, 74*(9), 1130–1139.

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., et al. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature, 582*, 84–88.

Bowman, F. D. (2014). Brain imaging analysis. *Annual Review of Statistics and Its Application, 1*(1), 61–85.
Brown, E. N., Lydic, R., & Schiff, N. D. (2010). General anesthesia, sleep, and coma. *New England Journal of Medicine, 363*(27), 2638–2650.
Bruno, M.-A., Vanhaudenhuyse, A., Thibaut, A., Moonen, G., & Laureys, S. (2011). From unresponsive wakefulness to minimally conscious PLUS and functional locked-in sydromes: Recent advances in our understanding of disorders of consciousness. *Journal of Neurology, 258*, 1373–1384.
Buckner, R. L. (2012). The serendipitous discovery of the brain's default network. *Neuroimage, 62*, 1137–1145.
Carlson, T., et al. (2018). Ghosts in machine learning for cognitive neuroscience: Moving from data to theory. *NeuroImage, 180*, 88–100.
Casarotto, S., Comanducci, A., Rosanova, M., Sarasso, S., Fecchio, M., Napolitani, M., et al. (2016). Stratification of unresponsive patients by an independently validated index of brain complexity. *Annals of Neurology, 80*(5), 718–729.
Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory*. New York: Oxford University Press.
Chalmers, A. (2009). *The scientist's atom and the philosopher's stone: How science succeeded and philosophy failed to gain knowledge of atoms*. Dordrecht: Springer.
Churchland, P. S. (2013). *Touching a nerve: The self as brain*. New York: W. W. Norton & Company.
Ciurleo, R., Bramanti, P., & Calabro, R. S. (2013). Pharmacotherapy for disorders of consciousness: Are 'awakening drugs really a possibility? *Drugs, 73*, 1849–1862.
Crowder, C. M., Palanca, B. J., & Evers, A. S. (2017). Mechanisms of anesthesia and consciousness. In P. G. Barash, B. F. Cullen, R. K. Stoelting, N. K. Cahalan, M. C. Stock, R. Ortega, S. R. Sharar, & N. F. Holt, (Eds.) Clinical Anesthesia 8th Edition (pp. 219–240).
Damasio, A. (2010). *Self comes to mind: Constructing the conscious mind*. New York: Pantheon.
Davey, C. G., Pujol, J., & Harrison, B. J. (2016). Mapping the self in the brain's default mode network. *NeuroImage, 132*, 390–397.
Dehaene, S. (2014). *Consciousness and the brain: Deciphering how the brain codes our thoughts*. New York: Viking.
Dehaene, S., Lau, H., & Kouider, S. (2017). What is consciousness and could machines have it? *Science, 358*, 486–492.
Dennett, D. C. (2018). Facing up to the hard question of consciousness. *Philosophical Transactions of the Royal Society B, 373*, 20170342.
Doris, J. M. (2015). *Talking to Our Selves: Reflection, ignorance and agency*. New York: Oxford University Press.
Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences, 98*(24), 13763–13768.
Eger, E. I., & Sonner, J. M. (2006). Anaesthesia defined (Gentleman, this is no humbug). *Best Practice & Research Clinical Anaesthesiology, 20*(1), 23–29.
Esslen, M., Metzler, S., Pascual-Marqui, R., & Jancke, L. (2008). Pre-reflective and reflective self-reference: A spatiotemporal EEG analysis. *NeuroImage, 42*, 437–449.
Farb, N. A., Segal, Z. V., Mayberg, H., Bean, J., McKeon, D., Fatima, Z., et al. (2007). Attending to the present: Mindfulness meditation reveals distinct neural modes of self-reference. *Social, Cognitive, and Affective Neuroscience, 2*(4), 313–322.
Fernyhough, C. (2016). *The voices within*. New York: Basic Books.
Feynman, R., Leighton, R. B., & Sands, M. (2010). *The Feynman Lectures on Physics Vol III, New Millennium Edition*. New York: Basic Books.
Figdor, C. (2010). Neuroscience and the multiple realization of cognitive functions. *Philosophy of Science, 77*, 419–456.
Filevich, E., Dresler, M., Brick, T. R., & Kuhn, S. (2015). Metacognitive mechanisms underlying lucid dreams. *The Journal of Neuroscience, 35*(3), 1082–1088.
Fischer, D. B., & Truog, R. D. (2015). What is a reflex? A guide for understanding disorders of consciousness. *Neurology, 85*(6), 543–548.
Flury, J. M., & Ickes, W. (2007). Having a strong versus a weak sense of self: The sense of self scale (SOSS). *Self and Identity, 6*, 281–303.
Frege, G. (1956). The thought: A logical inquiry. *Mind, 65*(259), 289–311.
Frewen, P., Schroeter, M. L., Riva, G., Cipresso, P., Fairfield, B., Padulo, C., et al. (2020). Neuroimaging and the consciousness of self: Review and conceptual-methodological framework. *Neuroscience and biobehavioral reviews, 112*, 164–212.
Gallagher, S., & Zahavi, D. (2008). *The phenomenological mind* (2nd ed.). London: Routledge.
Goldberg, I. I., Harel, M., & Malach, R. (2006). When the brain loses its self: Prefrontal inactivation during sensorimotor processing. *Neuron, 50*, 329–339.
Gray, J. (2004). *Consciousness: Creeping up on the hard problem*. New York: Oxford University Press.
Graziano, M., Guterstram, A., Branden, J., & Wilterson, A. I. (2020). Toward a standard model of consciousness: Reconciling the attention schema, global workspace, higher-order thought, and illusionist theories. *Cognitive Neuropsychology, 37*(3–4), 155–172.
Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences, 98*(7), 4259–4264.
Hardcastle, V. G. (2008). *Constructing the self*. Philadelphia: John Benjamins.
Heine, L., Laureys, S., & Schnakers, C. (2016). Behavioral responsiveness in patients with disorders of consciousness. In M. M. Monti, & W. G. Sannita (Eds.), *Brain function and responsiveness in disorders of consciousness* (pp. 25–36). London: Springer.
Hempel, C. G. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. New York: Free Press.
Henriksen, M. D., Parnas, J., & Zahavi, D. (2019). Thought insertion and disturbed for-me-ness (minimal selfhood) in schizophrenia. *Consciousness and Cognition*.
Hirstein, W. (2012). *Mindmelding: Consciousness, neuroscience, and the mind's privacy*. New York: Oxford University Press.
Hsiao, F. C., Tsai, P. J., Wu, C. W., Yang, C. M., Lane, T. J., Lee, H. C., et al. (2018). The neurophysiological basis of the discrepancy between objective and subjective sleep during the sleep onset period: An EEG-MRI study. *Sleep, 41*, 6. https://doi.org/10.1093/sleep/zsy056.
Huang, Z., Dai, R., Wu, X., Yang, Z., Liu, D., Hu, J., et al. (2014). The self and its resting state in consciousness: An investigation of the vegetative state. *Human Brain Mapping, 35*, 1997–2008.
Humphrey, N. (2011). *Soul dust: The magic of consciousness*. Princeton, NJ: Princeton University Press.
Hurlburt, R. T., Alderson-Day, B., Fernyhough, C., & Kühn, S. (2017). Can inner experience be apprehended in high fidelity? Examining brain activation and experience from multiple perspectives. *Frontiers in Psychology*, 103389/fpsyg.2017.00043.
Hurlburt, R. T., Alderson-Day, B., Kühn, S., & Fernyhough, C. (2016). Exploring the ecological validity of thinking on demand: Neural correlates of elicited vs. spontaneously occurring inner speech. *PLoS ONE, 11*(e0147932). https://doi.org/10.1371/journal.pone.0147932.
Hurlburt, R. T., & Heavey, C. L. (2001). Telling what we know: Describing inner experience. *Trends in Cognitive Sciences, 5*(9), 400–403.
Hurlburt, R. T., & Heavey, C. L. (2015). Investigating pristine inner experience: Implications for experience sampling and questionnaires. *Consciousness and Cognition, 31*, 148–159.
James, W. (1981). *The principles of psychology*. Cambridge, MA: Harvard University Press (Original work published 1890.).
Kedov, B. M. (1949). Dalton's atomic theory and its philosophical significance. *Philosophy and Phenomenological Research, 9*(4), 644–662.
Klein, S. (2014). *The two selves: Their metaphysical commitments and functional independence*. New York: Oxford University Press.
Kleiner, J. (2020). Mathematical models of consciousness. *Entropy, 22*(6), 609. https://doi.org/10.3390/e22060609.
Knyazev, G. G. (2013). EEG correlates of self-referential processing. *Frontiers in Human Neuroscience, 7*, 264.
Knyazev, G. G., Slobodskoy-Plusnin, J. Y., Bocharov, A. V., & Pylkova, L. V. (2011). The default mode network and EEG alpha oscillations: An independent component analysis. *Brain Research, 1402*, 67–79.
Kriegel, U. (2009). *Subjective consciousness: A self-representational theory*. New York: Oxford University Press.
Krystal, A. D., Edinger, J. D., Wohlgemuth, W. K., & Marsh, G. R. (2002). NREM sleep EEG frequency spectral correlates of sleep complaints in primary insomnia subtypes. *Sleep, 25*(6), 630–640.
Lane, T. (2012). Toward an explanatory framework for mental ownership. *Phenomenology and the Cognitive Sciences, 11*(2), 251–286.
Lane, T. (2014). When actions feel alien: an explanatory model. In T. W. Hung (Ed.), *Communicative Action* (pp. 53–74). Switzerland: Springer Science+Business Media.
Lane, T. (2015). Self, belonging, and conscious experience: A critique of subjectivity theories of consciousness. In R. Gennaro (Ed.), *Disturbed Consciousness: New Essays on Psychopathology and Theories of Consciousness* (pp. 103–140). Cambridge, MA: MIT Press.
Lane, T. (2017). Rationality and its contexts. In T.-. W. Hung, & T. J. Lane (Eds.), *Rationality: Constraints and contexts* (pp. 3–13). London: Academic Press.
Lane, T. (2020). Somebody is home. *Cognitive Neuropsychology, 37*(3–4), 193–196.
Lane, T., Duncan, N. W., Cheng, T., & Northoff, G. (2016). The trajectory of self. *Trends in Cognitive Sciences, 20*(7), 481–482.
Lane, T., & Liang, C. (2011). Self-consciousness and immunity. *The Journal of Philosophy, 108*(2), 78–99.

Lane, T., Yeh, S.-L., Tseng, P., & Chang, A.-Y. (2017). Timing disownership experiences in the rubber hand illusion. *Cognitive Research: Principles and Implications, 2*(1), 1–14.
Langland-Hassan, P. (2015). Introspective misidentification. *Philosophical Studies, 172*(7), 1737–1758.
Levine, J. (2001). *Purple haze: The puzzle of consciousness*. Oxford: Oxford University Press.
Lilienfeld, S. O., & Treadway, M. (2016). Clashing diagnostic approaches: DSM-ICD versus RDoC. *Annual Review of Clinical Psychology, 12*, 435–463.
Lipton, P. (1993). Contrastive explanation. In D. Ruben (Ed.), *Explanation* (pp. 207–227). New York: Oxford University Press.
Liu, X., Lauer, K. K., Ward, B. D., Roberts, C., Liu, S., Gollapudy, S., et al. (2017). Propofol attenunates low-frequency fluctuations of resting state fMRI BOLD in the anterior frontal cortex upon loss of consciousness. *Neuroimage, 147*, 295–301.
Lou, H. C., Thomsen, K. R., & Changeux, J.-P. (2020). The molecular organization of self-awareness. Paralimbic dopamine-GABA interaction. *Frontiers in Systems Neuroscience*. https://doi.org/10.3389/fnsys.2020.00003.
Marcel, A. J. (2004). The sense of agency: Awareness and ownership of action. In J. Roessler, & N. Eilan (Eds.), *Agency and self-awareness* (pp. 48–93). Oxford: Clarendon.
Martin, B., Franck, N., Cermolacce, M., Coull, J. T., & Giersch, A. (2018). Minimal self and timing disorders in schizophrenia: A case report. *Frontiers in Human Neuroscience, 12*, 132. https://doi.org/10.3389/fnhum.2018.00132.
Mashour, G. A., & Avidan, M. S. (2015). Intraoperative awareness: Controversies and non-controversies. *British Journal of Anesthesia, 115*, i20–i26.
Mashour, G. A., Roelfsema, P., Changeux, J.-P., & Dehaene, S. (2020). Conscious processing and the global neuronal workspace hypothesis. *Neuron, 105*(5), 776–798.
Massimini, M., & Tononi, G. (translation, F. Andersen) (2018). Sizing up consciousness: Towards an objective measure of the capacity for experience. Oxford, UK: Oxford University Press.
Medina, J., & DePasquale, C. (2017). Influence of the body schema on mirror-touch synesthesia. *Cortex, 88*, 53–65.
Menon, V. (2011). Large scale brain networks and psychopathology. *Trends in Cognitive Sciences, 15*(10), 483–506.
Monti, M. M., Coleman, M. R., & Owen, A. M. (2009). Neuroimaging and the vegetative state: Resolving the behavioral assessment dilemma? In N. D. Schiff, & S. Laureys (Eds). Disorders of Consciousness. Annals of the New York Academy of Sciences 1157, 81–89.
Nagel, T. (1974). What it is like to be a bat? *The Philosophical Review, 83*(4), 435–450.
Nelson, B., Lavoie, S., Gaweda, L., Li, E., Sass, L. A., Koren, D., et al. (2019). Testing a neurophenomenological model of basic self-disturbance in early psychosis. *World Psychiatry, 18*(1), 104–105.
Northoff, G., & Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Science, 8*, 102–107.
Olson, E. T. (2002). There is no problem of the self. In S. Gallagher, & J. Shear (Eds.) Models of the self (pp. 49–61).
Overgaard, M. (2009). How can we know if patients in coma, vegetative state, or minimally conscious state are conscious? *Progress in Brain Research, 177*, 11–19.
Owen, A. (2017). *Into the gray zone: A neuroscientist explores the border between life and death*. New York: Scribner.
Owen, A., Schiff, N. D., & Laureys, S. (2009). The assessment of conscious awareness in vegetative state. In S. Laureys, & G. Tononi (Eds.), *The neurology of consciousness: Cognitive neuroscience and neuropathology*. New York: Academic Press.
Park, H. D., & Blanke, O. (2019). Coupling inner and outer body for self-consciousness. *Trends in cognitive Sciences, 23*(5), 377–388.
Parnas, J., et al. (2005). EASE: Examination of anomalous self-experience. *Psychopathology, 38*, 236–258.
Parnas, J., & Sass, L. A. (2011). The structure of self-consciousness in schizophrenia. In S. Gallagher (Ed.), *The oxford handbook of the self* (pp. 521–546). Oxford: Oxford University Press.
Petersen, A. (2016). Consilience, clinical validation, and global disorders of consciousness. *Neuroscience of Consciousness*. https://doi.org/10.1093/nc/niw011.
Pistorius, M. (2011). *Ghost Boy*. New York: Simon & Schuster.
Pöppel, E. (2009). Pre-semantically defined windows for cognitive processing. *Philosophical transactions of the Royal Society of London B., 364*, 1887–1896.
Qin, P., Di, H., Liu, Y., Yu, S., Gong, Q., Duncan, N., et al. (2010). Anterior cingulate cortex activity and the self in disorders of consciousness. *Human Brain Mapping, 12*, 1993–2002.
Qin, P., Grimm, S., Duncan, N. W., Fan, Y., Huang, Z., Lane, T., et al. (2016). Spontaneous activity in default-mode network predicts ascription of self-relatedness stimuli. *Social Cognitive and Affective Neuroscience, 11*(4), 693–702.
Qin, P., Liu, Y., Shi, J., Wang, Y., Duncan, N., Gong, Q., et al. (2012). Dissociation between anterior and posterior cortical regions during self-specificity and familiarity: A combined fMRI-Meta-Analytic Study. *Human Brain Mapping, 33*, 154–164.
Qin, P., & Northoff, G. (2011). How is our self related to midline regions and the default-mode network? *NeuroImage, 57*, 1221–1233.
Qin, P., Wu, X., Duncan, N. W., Bao, W., Tang, W., Zhang, Z., et al. (2015). $GABA_A$ receptor deficits predict recovery in patients with disorders of consciousness: A preliminary multimodal [^{11}C]Flumazenil PET and fMRI study. *Human Brain Mapping, 36*, 3867–3877.
Quine, W. V. (1960). *Word and object*. Cambridge, MA: MIT Press.
Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences, 98*, 676–682.
Riganello, F. (2016). Responsiveness and the autonomic control—CNS Two-way interaction in disorders of consciousness. In M. M. Monti, & W. G. Sannita (Eds.), *Brain function and responsiveness in disorders of consciousness* (pp. 145–155). London: Springer.
Roskies, A. L. (2010). Neuroimaging and inferential distance: The perils of pictures. In S. J. Hanson, & M. Bunzl (Eds.), *Foundational issues in human brain mapping* (pp. 195–215). Cambridge, MA: The MIT Press.
Russell, I. F. (1993). Midazolam-alfentanil: An anesthetic? An investigation using the isolated forearm technique. *British Journal of Anasthesia, 70*, 42–46.
Russell, I. F. (2006). The narcotrend 'depth of anaesthesia' monitor cannot reliably detect consciousness during general anesthesia: An investigation using the isolated forearm technique. *British Journal of Anaesthesia, 96*(3), 345–352.
Sajid, N., Parr, T., Hope, T. M., Price, C. J., & Friston, K. J. (2020). Degeneracy and redundancy in active inference. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhaa148.
Salinas, J. (2017). *Mirror touch: Notes from a doctor who can feel your pain*. New York: HarperCollins.
Sanders, R. D., Tononi, G., Laureys, S., & Sleigh, J. (2012). Unresponsiveness ≠ Unconsciousness. *Anesthesiology, 116*(4), 946–959.
Schiff, N. (2010). Recovery of consciousness after brain injury: A mesocircuit hypothesis. *Trends in Neuroscience, 13*(1), 1–9.
Schnakers, C., Vanhaudenhuyes, A., Giacino, J., Ventura, M., Boly, M., Majerus, S., et al. (2009). Diagnostic accuracy of vegetative and minimally conscious state: Clinical consensus versus standardized neurobehavioral assessment. *BMC Neurology, 9*, 35.
Schore, A. N. (2003). *Affect regulation and the repair of the self*. New York: W. W. Norton.
Schulman, R. G. (2013). *Brain imaging: What it can (and cannot) tell us about consciousness*. New York: Oxford University Press.
Seager, W. (2016). *Theories of consciousness: An introduction and an assessment*. London: Routledge.
Sheline, Y. I., Barch, D. M., Price, J. L., Rundle, M. M., Vaishnavi, S., Snyder, A. Z., et al. (2009). The default mode network and self-referential processes in depression. *Proceedings of the National Academy of Sciences, 106*(6), 1942–1947.
Shoemaker, S. (1996). *The first-person perspective and other essays*. Cambridge: MIT Press.
Sleigh, J., Warnaby, C., & Tracey, I. (2018). General anesthesia as fragmentation of selfhood: Insights from electroencephalography. *British Journal of Anaesthesia, 121*, 233–240.
Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology, 66*, 487–518.
Sterzer, P., Mishara, A. L., Voss, M., & Heinz, A. (2016). Thought insertion as a self-disturbance: An integration of predictive coding and phenomenological approaches. *Frontiers in Human Neuroscience, 12*(10), 502.
Strawson, G. (2011). The minimal subject. In S. Gallagher (Ed.), *The Oxford Handbook of THE SELF* (pp. 253–278). New York: Oxford University Press.
Tapal, A., Oren, E., Dar, R., & Eitam, B. (2017). The sense of agency scale: A measure of consciously perceived control over one's mind, body, and the immediate environment. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2017.01552.

Tsuchiya, N., Wilke, M., Frassle, S., & Lamme, V. A. F. (2015). No report paradigms: Extracting the true neural correlates of consciousness. *Trends in Cognitive Sciences, 19*(12), 757–770.
van Erp, W. S., Lavrijsen, J. C. M., Vos, P. E., Bor, H., Laureys, S., & Koopmans, R. T. C. M. (2015). The vegetative state: Prevalence, misdiagnosis, and treatment limitations. *The Journal of Post-acute and Long-term Care Medicine, 16*(1), 85.e9–85.e14.
Vanhaudenhuyse, A., Charland-Verville, V., Thibaut, A., Chatelle, C., Tshibanda, J.-F. L., Maudoux, A., et al. (2018). Conscious while being considered in an Unresponsive Wakefulness Syndrome for 20 years. *Frontiers in Neurology, 9*, 691.
Vogeley, K., & Gallagher, S. (2011). Self in the brain. In S. Gallagher (Ed.), *The Oxford Handbook Self* (pp. 111–136). Oxford: Oxford University Press.
Voss, H. U., Uluç, A. M., Dyke, J. P., Watts, R., Kobylarz, E. J., McCandliss, B. D., et al. (2006). Possible axonal regrowth in late recovery from the minimally conscious state. *The Journal of Clinical Investigation, 116*(7), 2005–2011.
Wang, F., Di, H., Hu, X., Jing, S., Thibaut, A., Perri, C. D., et al. (2015). Cerebral response to subject's own name showed high prognostic value in traumatic vegetative state. *BMC Medicine, 13*, 83. https://doi.org/10.1186/s12916-015-0330-7.
Weisberg, J. (2011). Describing inner experience: A symposium debating Descriptive Experience Sampling (DES). *Journal of Consciousness Studies, 18*, 1.
Wilson, B., Dhamapurkar, S., Tunnard, C., Watson, P., & Florschutz, G. (2014). The effect of positioning on level of arousal and awareness in patients in the vegetative state or the minimally conscious state: A replication and extension of a previous finding. *Brain impairment, 14*(3), 475–479.
Windt, J. M. (2015). *Dreaming: A conceptual framework for philosophy of mind and empirical research*. Cambridge, MA: The MIT Press.
Windt, J. M., Nielsen, T., & Thompson, E. (2016). Does consciousness disappear in dreamless sleep? *Trends in Cognitive Sciences, 20*(12), 871–882.
Wolff, A., Di Giovanni, D. A., Gomez-Pilar, J., Nakao, T., Huang, Z., Longtin, A., et al. (2019). The temporal signature of self: Temporal measures of resting state EEG predict self-consciousness. *Human Brain Mapping, 40*(3), 789–803.
Wolpert, L. (1993). *The unnatural nature of science*. Cambridge, MA: Harvard University Press.
Yang, C. M., Han, H. Y., Yang, M. H., Su, W. C., & Lane, T. J. (2010). What subjective experiences determine the perception of falling asleep during sleep onset period? *Consciousness and Cognition, 19*, 1084–1092.
Zahavi, D. (2011). Unity of consciousness and the problem of self. In S. Gallagher (Ed.), *The oxford handbook of the self* (pp. 316–335). Oxford: Oxford University Press.
Zahavi, D. (2014). *Self and other: Exploring subjectivity, empathy, and shame*. Oxford: Oxford University Press.
Zahavi, D., & Kriegel, U. (2016). For-me-ness: What it is and what it is not. In D. O. Dahlstrom, A. Elpidorou, & W. Hopp (Eds.), *Philosophy of mind and phenomenology: Conceptual and empirical approaches* (pp. 36–53). New York: Routledge.
Zahavi, D., & Roepstorff, A. (2011). Faces and ascriptions: Mapping measures of the self. *Consciousness and Cognition, 20*, 141–148.