Self unbound: ego dissolution in psychedelic experience

Chris Letheby[*,‡] and Philip Gerrans[‡]

Department of Philosophy, University of Adelaide, North Terrace, Adelaide, South Australia 5005, Australia

*Correspondence address. Department of Philosophy, University of Adelaide, North Terrace, Adelaide SA 5005, Australia. Tel: +61 08 8313 4249, Fax: +61 08 8313 4341. E-mail: cerletheby@gmail.com

‡Both authors contributed equally to the development of the arguments and the writing of the manuscript. C.L. proofread, formatted, and prepared the manuscript for submission.

Abstract

Users of psychedelic drugs often report that their sense of being a self or 'I' distinct from the rest of the world has diminished or altogether dissolved. Neuroscientific study of such 'ego dissolution' experiences offers a window onto the nature of self-awareness. We argue that ego dissolution is best explained by an account that explains self-awareness as resulting from the integrated functioning of hierarchical predictive models which posit the existence of a stable and unchanging entity to which representations are bound. Combining recent work on the 'integrative self' and the phenomenon of self-binding with predictive processing principles yields an explanation of ego dissolution according to which self-representation is a useful Cartesian fiction: an ultimately false representation of a simple and enduring substance to which attributes are bound which serves to integrate and unify cognitive processing across levels and domains. The self-model is not a mere narrative posit, as some have suggested; it has a more robust and ubiquitous cognitive function than that. But this does not mean, as others have claimed, that the self-model has the right attributes to qualify as a self. It performs some of the right kinds of functions, but it is not the right kind of entity. Ego dissolution experiences reveal that the self-model plays an important binding function in cognitive processing, but the self does not exist.

Key words: psychedelic; self; psilocybin; LSD; binding and multisensory integration; hallucinogen

Introduction

In this paper, we argue that 'ego dissolution' in psychedelic experience is a window on the nature of self-awareness. By 'self' here we mean an entity, substance, or bare particular that instantiates properties. Just as the mind infers that properties are possessed by objects, it infers that the self is an object whose continued existence explains the co-occurrence of physical and psychological attributes. We argue that the mind models the self as standing in the same relation to its properties as objects in general stand to theirs. Under this strategy, higher-level models predict that features represented at lower levels are attributes of objects. Thus, for example, in perceptual and sensorimotor feature binding the higher-level modelling of objects is used to integrate representations of features into coherent wholes. We see and feel objects, not concatenations of features, as a result of this top-down process. Similarly, we argue, the self is postulated by higher-level processes as an entity to facilitate the binding or integration of information. We argue that this explains a Cartesian intuition that the self is a simple indivisible entity. In psychedelic experience these integrative processes are disrupted, leading to the phenomenology of ego dissolution.

Representing the self as an entity which sustains interoception, affection, cognition, and perception facilitates the integration of information and allows preferential allocation of

resources to self-relevant information through the coordination of the salience and emotional processing systems. The former allocates processing resources adaptively by biasing cognition and the latter coordinates the processing of information relevant to the goals and interests of the organism (e.g. detecting and responding to danger is the interest served by the emotion of fear).

We support this idea with evidence from two related sources. The first is research by Sui and Humphreys (2015) on cognitive binding and the role of the self-model in enhancing it. The postulation of a self to which information is relevant makes a difference to the integrity of representations. However, the mechanisms involved model that self as a heuristic, a way of making information 'sticky', rather than as a way of tracking the fluctuating cognitive fortunes of an actual entity. Self-awareness is the experience of cognitive processes in which these binding processes are intact. Sui and Humphreys, however, are not very precise about the nature of the self. Their work suggests that the self is the entity to which things matter. We suggest that emotional processes play a crucial role here since they tell the organism how and why things matter.

The second is evidence provided by the phenomenon of 'ego dissolution' reported in psychedelic experience. We argue that evidence suggests that psychedelics target mechanisms on which self-binding depends. Psychedelic experience degrades these binding processes, enabling subjects to experience cognition not bound by self-models. We emphasize that the 'self' which dissolves in psychedelic experience is not an actual entity or an object of perception, interoception, or introspection but an entity inferred by the mind to predict the flow of experience in and across cognitive modalities. As Thomas Metzinger put it: 'There is just no entity there, no individual substance, and scientifically we can predict and explain everything we want to predict and explain in a much more parsimonious way' (quoted in Marshall 2016).

We can contrast Metzinger's view with Hohwy and Michael's (2017) 'causal realist interpretation of the reference of intentional terms and of the self' which argues that the causal power exerted by the object(s) of self-representation implies a causally efficacious entity underlying the hierarchy of integrative processes. Sui and Humphreys similarly argue that the binding power of self-representation cannot be explained by a self that is essentially a theoretical or narrative abstraction. Ultimately we will argue, on the basis of the psychedelic evidence, that Sui and Humphreys, and Hohwy and Michael, are right to reject a purely narrative view of the underpinnings of self-awareness: the self-model is deeper, more robust, and more cognitively ubiquitous than that. But we will side with Metzinger and against Hohwy and Michael in denying that anything, including the self-model, possesses the right attributes to qualify as a self.

The issues are subtle since all sides agree on the integrative role of the self-model and the predictive hierarchical nature of self-modelling. Our contention is that the phenomenology of ego dissolution is easier to explain if the subject is experiencing the disintegration of a system whose integration she normally experiences in terms of an indivisible mental substance. On this basis, we propose that the right attributes to qualify as a self are those represented by the self-model, but not instantiated by anything actual: indivisibility, substantiality, and strict identity over time, as well as distinctness from and ownership/authorship of experiences, thoughts, and feelings. The concept of a self does not just delineate a functional role (cause of behaviour) but posits a realizer of that role with specific attributes.

From the non-instantiation of those attributes we conclude that the self-model fails to refer (compare the elimination of caloric in the mechanical theory of heat).

We offer a diagnosis of this phenomenon in terms of self-binding (i.e., the preferential enhancement of cognitive binding for self-relevant information). We discuss the nature of binding and self-binding, arguing that self-binding does not entail the existence of an object to which attributes are bound—though it does require the representation of an object, to which representations of attributes are bound. We then turn to the main sections of the paper that explain the phenomenon of ego dissolution in psychedelic experience. We describe the phenomenology and neural correlates and show that ego dissolution is best understood as a phenomenon of unbinding. In a final section, we situate our account in the conceptual framework of contemporary debates and address some puzzles and remaining issues. In brief, our conclusion is that even in florid psychedelic experience the self-model is never entirely destroyed. Rather, as the coherence it normally imposes degrades, we become aware that our normal experience of unity depends on a modelling process. Just as disorders of feature binding help disclose the nature of object representation, ego dissolution discloses the nature of self-awareness.

Furthermore, our normal experience of unity compels the inference not just that we are a self, but that that self is a Cartesian substance. The self-model is a hypothesis of a unitary and persisting entity—a bare particular—which owns and inhabits the body, which thinks the thoughts, which feels the feelings, and which was present in past experience and will be present in future experience. It turns out that controversy over the neural basis of self-awareness turns on whether the mechanisms which produce the experience of self-awareness perform their integrative tasks by modelling the self as a simple indivisible substance.

Binding

The notion of binding solves a problem first discussed for visual representation. It arises from evidence that different visual features of objects, such as colour, shape, and motion, are processed in separate areas of visual cortex. These representational elements are combined into coherent, unified percepts by feature binding processes. Evidence for the existence of feature binding mechanisms comes from conditions in which they fail, such as 'illusory conjunctions' of features induced experimentally or resulting from pathology (Burwick 2014). An example might be a subject presented with a green circle and a blue square who misperceives a blue circle and a green square.

Binding can be described and explained at the phenomenal (unity of experience), cognitive (coherence and integration), and neural (mechanism) levels (Revonsuo 1999). There is no single theory of the nature of binding. Some theorists treat it as a global phenomenon involving the integration of information across large-scale networks. Others focus on local or modular integration of percepts or even perceptual elements, proposing that such modular integration is necessary before candidate representations can be globally integrated. However, in the absence of a final theory it remains true that cognition and experience require the coherent functioning of spatially and temporally distributed neuronal populations to produce integrated representations.

The resurgence of interest in the phenomenon by neuroscientists in the 1990s produced different theories of binding. It has been explained in terms (i) of phase synchrony (peaks

and valleys of waveforms occurring simultaneously) across distributed circuitry (Singer 1999); (ii) convergence zones: specialized circuitry that integrates signals fed forward from lower level systems (Damasio 1989). Such convergence zones could be hierarchically organized, elements of a percept being individually unified then fed forward to higher levels to be integrated into the overall representation of a perceived scene; and (iii) integrative properties of layer V pyramidal cells (Bachmann 2015). Refined versions of this idea suggest that information thus bound passes through a processing bottleneck enabling it to become the object of metacognitive processes. (i)–(iii) fit well with attentional and global broadcast theories of consciousness, which make availability to executive processing the essence of consciousness. On these views unbound information cannot be the object of conscious awareness. (iv) is a Hebbian theory that proposes that binding consists in the frontally regulated construction and maintenance of transient activity patterns in frontal–posterior (especially parietal) circuitry, stabilizing and integrating perceptual representations, which then become available to executive functioning (Ballard *et al.* 1983; van Essen *et al.* 1992).

Here, we want to note several aspects of binding theory. It is proposed to explain the unity and stability of conscious experience and it implies a neural mechanism.[1] It explains the unity of conscious experience in terms of functional coupling between cognitive subsystems and there is a strong implication that such coupling requires frontal involvement. Relatedly, the disruption, alteration, or disappearance of consciousness is produced by the disruption of frontal–posterior coupling (Mashour 2013). This last point reminds us that consciousness comes in different forms and that any theory of the role of binding in consciousness should help explain (ultimately) sleep, dreaming, coma, alert waking, and psychosis, as well as the state we are interested in here: psychedelic experience.

We think that the phenomenon of ego dissolution is a case where understanding the nature of phenomenal unbinding can illuminate the nature of binding: in this case, the binding that underlies the phenomenon of self-awareness.

Binding and Predictive Processing

When the binding problem and tentative solutions were first proposed the predictive processing theory had not undergone its recent revival[2] but we note that it provides an elegant framework for understanding the binding of information. Predictive processing theories of cognition treat the brain as a neurally instantiated processing hierarchy, with generative models at higher levels predicting bottom-up inputs from lower levels. The discrepancy between predicted and actual input takes the form of an error signal. Error signals are cancelled either by updating the model or by taking action to alter the discrepant input (Hohwy 2013).

This framework provides an elegant solution to the binding problem. Regularities in sensory input over time constrain the combinations of features into objects that will be predictively successful. So the binding of features together into coherent

percepts is determined by the probabilistic model which best matches present and past inputs, and the choice of probabilistic model is determined by regularities in the inputs. Binding, on this view, is essentially a kind of top-down abductive inference. If I am fielding in a game of cricket, the patterns in the inputs to my visual system are efficiently and reliably predicted by a model that posits various entities—a batter, a bat, a ball—each possessing various attributes such as colour, shape, location, speed, and direction of motion. Thus, object perception is a matter of hypothesizing an ontology that makes the flow of sensory input intelligible and predictable.

Our central idea here is that self-representation can be understood using the predictive framework in the same way as object representation. The self-model represents a persisting object in order to make sense of, unify, and predict ongoing patterns of egocentric, salient, autobiographical experience.

The notion of egocentric experience reflects the fact that experience is spatially and temporally organized from a perspective or location such as point of view or origin of movement. For example, sensorimotor control depends on integration of sensory input at a point or points internal to the organism.

Salience refers to that fact that the organism is bombarded by information only a fraction of which is relevant to its goals and interests. Each of those goals, which range from phylogenetically ancient (maintaining homeostasis) to recent (maintaining social reputation) creates a dimension of salience which requires the mind to allocate processing resources adaptively. The ability to represent these multiple dimensions of salience depends on hierarchically organized neural circuitry, which focuses perception and cognition, biasing cognitive systems to the processing of relevant information. This salience network (SLN) coordinates the interaction of neural systems that detect and evaluate relevance, bias processing accordingly and, importantly, allow us to feel the consequences as affective states.

Thus, we are organisms who experience and act in the world from a unique physical (spatiotemporal) and affective orientation. In the process, the mind makes models to explain the way the world makes us feel as we move through it. When we deliberate on and communicate the resultant experience we represent the integrated functioning of these egocentric and autobiographical models as unified through time in a single entity.

In effect the salience system links the egocentric self-models that manage the physical/bodily interface with autobiographical models which situate the organism historically. The salience system tells the organism which information matters and how it matters, moment to moment and through time. In effect the salience system is constantly creating the illusion of substantial selfhood by binding information into a representation not of the world in itself, but of the world as it matters to the organism. As Seth put it:

> emotion and embodied selfhood are grounded in active inference of those signals most likely to be 'me' across interoceptive and exteroceptive domains. In humans, self-related predictive coding simultaneously engages multiple levels of self-representation, including physiological homeostasis, physical bodily integrity, morphology and position, and—more speculatively—the metacognitive and narrative 'I'.
> (2013, 570).

The narrative 'I' is Dennett's (1991) narrative self or Damasio's (2010) autobiographical self, which allows us to experience

---

1 Note that cognitive binding and the unity of consciousness are not synonymous, but the former is a serious contender for an explanation, at least partial, of the latter; cf. Revonsuo (1999).

2 The central insights of the predictive processing framework go back at least to Helmholtz (1925) and arguably to Kant (Swanson 2016).

high-level (attempted) control of the integrated functioning of the egocentric and salience systems.

We think that Dennett is correct that the narrative self is a model which abstracts from the integrated functioning of a hierarchy of egocentric and salience systems[3] to posit a simple unified entity of which those integrated features are attributes. This view is consistent with the idea that as functioning disintegrates we will no longer feel like unified entities.

The Architecture of Subjectivity

That the SLN is activated in such a variety of conditions, ranging from basic perceptual processing to high-level metacognition, is a consequence of the way minds evolved. The mind did not evolve to range impartially over all the information available in the world and represent its causal and conceptual structure. It allocates its resources to information which is rewarding or valuable. We define salient information as that which is represented as 'potentially rewarding for the organism'. This allows for the context sensitivity and multidimensionality of salience. It also fits with the neurocomputational interpretation of the role of the mesolimbic dopamine system as a reward prediction system whose activity makes information salient by increasing gain in target circuitry.

The self-model allows us to feel that the experiences resulting from the interplay of the salience system with other (cognitive, emotional, and affective) systems whose activity it coordinates belong to a consistent unified entity: the self. One interesting feature of ego dissolution is that it produces highly salient experiences (they have the 'dopamine halo') unmoored from normal affective responses and the sense that they are happening to the self. Our hypothesis is that this is because the self-model is no longer smoothly integrating these experiences by attributing them to the same entity.

A consistent finding in the case of self-referential processing is activity in a network of midline structures whose hubs are involved in emotion, motivation, salience, and the switching of attention between neutral/external and self-relevant information (Qin and Northoff 2011). These structures are less active when attention is focused 'outward' so to speak on the perceptual environment or on problem solving with an impersonal aspect. It seems clear that stable, context-sensitive activity in this midline-centred network is an essential aspect of cognitive function. Organisms need to be able to allocate cognitive resources appropriately according to whether the context requires accurate representation of the world (concrete or abstract) or the significance of the world to their goals and interests.

The cortical midline structures repeatedly implicated in self-reference divide into two distinct, large-scale functional networks. The SLN, which we have been discussing, is centred on key hub regions such as the anterior cingulate cortex (ACC), implicated in error detection and task switching, and the anterior insular cortex (AIC), implicated in interoceptive processing and conscious emotional feelings (Seth 2013). Meanwhile, the much-discussed default mode network (DMN) is centred on key hub regions such as the medial prefrontal cortex (MPFC),

---

3 By egocentric systems here we merely mean those cognitive systems that use an egocentric reference point (bodily, perceptual, narrative) to produce their representations. Contrast these with allocentric or decontextualized representational systems which for example represent relations that hold between entities irrespective of subjective perspective.

posterior cingulate cortex (PCC), and inferior parietal lobule (IPL; Davey *et al.* 2016).

The DMN was originally identified in neuroimaging studies of the resting brain (Raichle *et al.* 2001). It has since been found to be activated by many self-referential tasks, including 'prospection' or 'mental time travel'—the autobiographical simulation of past and future experience—and theory of mind (Spreng and Grady 2010). Of course, mind-wandering in task-free conditions often concerns self-referential themes. The pioneering studies of Damasio, and much subsequent research, support the view that the MPFC is centrally involved in the attribution of personal relevance to (that is to say, the egocentric evaluation of) actual or simulated autobiographical episodes (D'Argembeau 2013). In relation to the different levels of self-modelling, the DMN is implicated in higher-level, narrative self-representation, as opposed to the more minimal embodied form of self-awareness supported by the SLN.

The PCC is a functionally enigmatic region with considerably higher metabolic activity and structural connectivity than most brain areas, leading some to describe it as the 'core node of the DMN' (Davey *et al.* 2016, 390). In their analysis aimed at delineating the substrates of self-reference within the DMN, Davey *et al.* found that the optimal model was one in which 'self-related processes were driven by PCC activity and moderated by the regulatory influences of MPFC.' Activity in the left IPL, which is known to play a role in the retrieval of semantic information, was also part of the core self-reference network identified by this analysis.

Many intriguing findings support a key role for the PCC in self-reference. Neurofeedback studies of meditators suggest that PCC activity co-varies not with mind wandering as such, but with the experience of becoming 'caught up' in a train of thought. Brewer proposes that the PCC is involved in getting 'caught up in experience', whether it be a particularly compelling rumination or a drug craving—and numerous results find meditation downregulating PCC activity (Brewer *et al.* 2013; Brewer and Garrison 2014). Experimental results implicate the PCC in processes such as integrating information about spatial self-location and body ownership (Guterstam *et al.* 2015). According to one recent model, the PCC is a key hub involved in 'tuning' the connectivity of many other brain networks, regulating the balance between internal and external attention, as well as the breadth of attention (Leech and Sharp 2014). These observations about PCC function cohere with findings from psychedelic neuroscience which we will discuss below. First, however, we turn to the generic function of 'self-binding' which seems to be implemented by these self-processing networks.

Self-binding

The concept of self-binding comes from the work of Sui and Humphreys (2015) who proposed it to explain why cognition across domains and levels is enhanced for self-relevant information. Subjects asked to classify words on the basis of self-relatedness or meaning remember more self-related words, and moreover remember more episodic details surrounding the learning of those words, suggesting enhanced mnemonic binding. Meanwhile, in face processing studies, subjects are quicker to recognize their own face than friends' or strangers' faces, whether upright or inverted. Friends' faces also have an advantage over strangers', but only in the upright condition, suggesting that self-reference confers a unique Gestalt or integrative advantage. (Other results militate against an explanation of this effect in terms of mere familiarity.)

Neurocognitive networks implicated in self-representation

Network	Key nodes	Aspects of self-representation	Result of down-regulation	Relevant studies
Default Mode Network (DMN)	Posterior cingulate cortex (PCC), medial prefrontal cortex (MPFC), inferior parietal lobule (IPL)	Narrative self; mental time-travel; judgement, planning, and goals	Dissolution of narrative/cognitive self (personality, history, goals, ownership of thoughts); compromised binding of embodied to narrative representations	Carhart-Harris et al. (2012); Palhano-Fontes et al. (2015); Bouso et al. (2015); Tagliazucchi et al. (2016); Speth et al. (2016)
Salience Network (SLN)	Anterior cingulate cortex (ACC), anterior insular cortex (AIC)	Embodied self; emotional feeling; salience	Dissolution of embodied self (changes to body boundaries, spatial self-location, personal relevance of emotional feelings)	Lebedev et al. (2015); Tagliazucchi et al. (2016)

Sui and Humphreys also cite evidence that self-reference enhances coupling between different temporal stages of information processing, possibly related to temporal binding, and enhances functional coupling between specific brain regions such as the aforementioned DMN, implicated as a neural substrate for autobiographical thought.

Sui and Humphreys also show that there is an attentional and decision-making enhancement effect for self-related information which speeds up sequential processing by enhancing 'binding between different states of processing' (Sui and Humphreys 2015, 724). They noted that these effects could best be explained in terms of an integrative role for self-related processing and hence they employed the concept of self-binding. Given the necessity for binding in establishing coherence they then examined a role for functional coupling of distributed circuitry consistently implicated in self-referential effects. Dynamic causal modelling of fMRI data showed 'self-reference enhances the neural coupling between regions concerned with a core self-representation (vmPFC) and with distinct domain-specific regions associated with different components of the self, including self-related attention (LpSTS)' (Sui and Humphreys 2015, 724). Meta-analyses of fMRI data reinforced the finding that 'there is enhanced neural coupling for self-processing (vs. other-related processing) between the vmPFC/ pregenual anterior cingulate (pACC) and several other regions including the bilateral anterior insula, left striatum, right thalamus, and amygdala' (Sui and Humphreys 2015, 724).

Further evidence that self-relevance enhances processing comes from the role of affect in perceptual binding. It is well-established that perceived objects evoke immediate affective responses. Lebrecht *et al.* (2012) found that even paradigmatically neutral objects such as coffee cups and clocks evoke measurable 'micro-valences', concluding that 'valence should be construed as a property of object representations'—that is, the affective response is bound to the percept, forming a unified representation, rather than being merely associated with it. The view that affect is bound to perception is supported by studies showing the role of the amygdala in driving eye saccades in face perception, directing the gaze to emotionally salient stimuli.

At higher levels, affect is bound to memories and imaginative states invoked in reverie or deliberation, a process that seems to depend particularly on the vmPFC. This suggests that the vmPFC is recapitulating at a higher level the role played by

the amygdala in perceptual processing; namely, coordinating the construction of self-relevant representations. One nice demonstration of this phenomenon arises in a distinction drawn between episodic (the episodic representation of previously acquired information) and autobiographical memory (in which the subject feels herself present as part of the recollected experience) by Asaf Gilboa:

> the Ventromedial PFC … does so by establishing a *self-related* retrieval template that sets up the parameters against which retrieved memories are evaluated based on (or which give rise to) a general intuitive 'feeling of rightness'. Thus, the monitoring of the veracity and perhaps cohesiveness of autobiographical memories is primarily mediated by the ventromedial PFC.

(2004, 1345–6; our italics).

For Sui and Humphreys, the self plays far too important a role to be dismissed as a narrative fiction. They frame their argument in terms of the function of self-reference and argue that the integrative function they identify could not be performed by a fiction. They describe different elements of a self-referential processing system as constituting an 'integrative hub for information processing' (2015, 719). The key nodes of the network they identify are those implicated in emotional binding and affective experience, the enhancement of attention for emotionally salient and motivating information and the switching of attention according to context.

These systems together perform the function identified by Sui and Humphreys. At perceptual and sensory levels, they ensure that salient information is preferentially attended to and integrated into coherent percepts. At higher conceptual or metacognitive levels, they ensure that the same is true of the representations that figure in memories, stories, and causal explanations. In a review of the neural correlates of self-binding, which are the substrates of the networks we outline above, Lou *et al.* (2017) concluded that 'self-awareness is an integral function of all conscious experiences, binding conscious experiences together into a single construct with a sense of unity of consciousness'.

This is a very strong claim. It implies that conscious experience would not arise in the absence of self-binding. The work of Sui and Humphreys suggests that this is not the case. Their subjects did not lose consciousness when perceiving self-irrelevant information. The same is true for patients with lesions to

components of this network. However, such patients do seem to have compromised abilities in tasks (such as autobiographical simulation, personal decision making, and social interaction) that we pretheoretically describe as requiring self-awareness. As we will show, the neuroscience and phenomenology of psychedelic ego dissolution put further pressure on the view that self-binding is necessary for consciousness. Indeed, psychedelic consciousness, in which the mechanisms of self-binding are compromised, is usually described as a state of increased or intensified consciousness, leading Carhart-Harris *et al.* (2014) to suggest that self-referential processes constrain consciousness rather than enabling it.

Recognizing the role of self-binding and self-models leaves questions about the mechanisms of binding unanswered. As we noted earlier many candidates have been proposed, including various forms of temporal correlation (synchronization of cells or functional coupling of brain regions; Singer 1999), convergence of modular input streams in high-level integrative brain regions (Damasio 1989), and even intracellular processes implemented in cortical pyramidal cells (Bachmann 2015).

At this point, evidence about the actions of psychedelic drugs and the experience of ego dissolution becomes relevant. Under the influence of psychedelics, we continue to represent the world and our bodies, but attention is captured and allocated in different ways, experience seems both more intense and less personal, and salience, affective feeling, and motivation become detached from personal goals and history.

Psychedelic Ego Dissolution

'Classic' (serotonin 2a receptor agonist) psychedelics such as mescaline, psilocybin, and dimethyltryptamine (DMT, a key constituent in the South American beverage ayahuasca) have a long history of religious, spiritual, and medicinal use in various cultures (Sessa 2012). Serious scientific interest in these substances was sparked by the accidental discovery in the 1940s of the extremely potent psychoactive effects of lysergic acid diethylamide (LSD). During the 1950s and 1960s, psychedelics were studied intensively as models of psychotic and mystical states, as psychotherapeutic agents, and as instruments for mapping the varieties of human consciousness (Osmond 1957). However, uncontrolled use of the drugs in the context of the hippie counterculture led to socio-political controversy, culminating in the virtual cessation of human psychedelic research for some decades.

Since the early 1990s this research has been slowly but steadily resuming, with greater methodological rigour and aided by the many new technical and theoretical innovations in the mind and brain sciences. Several small studies have revealed preliminary evidence for the safety and therapeutic efficacy of carefully controlled psychedelic administration (Griffiths *et al.* 2016; Ross *et al.* 2016; dos Santos *et al.* 2016a). Meanwhile, neuroimaging investigations have begun to identify neural correlates of the changes in cognition and consciousness caused by psychedelics (dos Santos *et al.* 2016b).

The remarkable variability and alleged ineffability of psychedelic experiences pose barriers to describing these states. But some progress has been made on charting the phenomenology of the 'antipodes of the mind' (Huxley 1954; Masters and Houston 1966; Shanon 2002). Psychedelics alter many aspects of experience: sense perception, emotion, cognition, and the apprehension of time and space. But of all their effects, perhaps none is more provocative than the profound alteration to the ordinary sense of self or ego known as 'ego dissolution'. The counter-cultural icon Alan Watts, responsible for popularizing both Eastern religion and psychedelics in the 1960s, wrote: 'there are certain types of change which are usual enough to be considered characteristic of psychedelics: the sense of slowed or arrested time, and the alteration of "ego boundary"—that is, of the sensation of one's own identity' (1964). Psychedelic subjects often report that their sense of being a self, or 'I', distinct from the rest of the world 'out there', is weakened, altered, or abolished during the intoxication.

Ego dissolution experiences often occur in the context of mystical states in which the ordinary sense of self is replaced by a sense of union with an ultimate reality underlying all of manifest existence—the famous 'cosmic consciousness' experience. Shanon describes the content of this apparent metaphysical realization as 'idealistic monism with pantheistic overtones' (2002, 163). The propensity of psychedelics to occasion such experiences goes some way to explaining their history of religious use. Indeed, intellectuals such as Watts and Aldous Huxley were initially drawn to psychedelics due to a prior interest in mysticism. With respect to ego dissolution, it is worth noting that apprehending the non-existence of the individual self is a central goal of Buddhist meditation (Albahari 2014). There is evidence that mystical states are important for the therapeutic effects of psychedelics (Garcia-Romeu *et al.* 2016) so explaining the ego dissolution experience is a crucial step in theorizing the mechanisms of psychedelic treatment.

Early theoretical speculations about psychedelic ego dissolution strikingly anticipated the key ideas we will develop in the rest of the paper. Savage (1955) and Klee (1963) both suggest explanations of ego dissolution in terms of failures of mechanisms of cognitive integration, and view these compromised perceptual (particularly interoceptive) integrative mechanisms as the basis of the self-model: 'Lacking a reliable inflow (or integration) of stimuli, particularly from his body, the subject has lost much of the basis for his self-percept' (ibid., 465). Both authors also liberally refer to the breakdown of cognitive processes of expectation (cf. prediction) in explaining psychedelic ego dissolution. These early speculations sit very comfortably with currently influential theoretical frameworks in cognitive neuroscience, and suggest a bidirectional, mutually reinforcing relationship between self-representation and cognitive integration.

Another important point is that ego dissolution is not an all-or-nothing affair. Different aspects of self-awareness may be more or less disrupted in different ways on psychedelics. Various authors have suggested that some vestiges of self-awareness are preserved in most, if not all, psychedelic experiences (Pahnke 1969; Shanon 2002), which helps explain the puzzling fact that autobiographical memories can apparently be formed of these putatively selfless episodes (Metzinger 2005). Moreover, there are aspects of psychedelic phenomenology which may not be explicitly described as ego dissolution but are still readily explained by our account. Subjects often find their attention drawn to stimuli which they normally would not notice; as Watts puts it, psychedelics 'make the spotlight of consciousness a floodlight which… brings to light unsuspected details—details normally ignored because of their lack of significance' (2002). Attention is no longer guided exclusively by adaptive and egocentric goals and agendas; salience attribution is no longer bound to personal concern.

Subjects often become less defensive and better able to view their own thoughts and feelings dispassionately—one main rationale for early therapeutic use of these drugs (Eisner and Cohen 1958). It is worth pointing out that subjects are not confused about their bodily or cognitive boundaries here; they

know they are having the experience. Rather it seems as if the experiences, although intense and fascinating, are no longer automatically attributed to an entity. A very common experience is to see one's own dysfunctional emotional or behavioural patterns, and the possibility of alternatives, with striking clarity (Shanon 2002, 162–3). There is a tendency towards decentring and the objectification of self-related phenomena which ordinarily are taken very personally and evoke strong emotional reactivity (Soler *et al.* 2016). All of this suggests that, even when the ego is not felt to dissolve altogether, the contents of consciousness are less filtered through considerations of self-relevance than is usual. Various cognitive processes which are usually tightly coupled to self-representation become decoupled from it.

Psychedelic experiences which lead subjects to strongly endorse descriptions of 'ego dissolution' tend to feature dramatic changes, such as the weakening of bodily boundaries or the loss of the sense of ownership for thoughts and actions. The account we will propose suggests that such paradigmatic ego dissolution experiences are on a continuum with, and susceptible to the same kind of explanation as, less dramatic experiences which may not be described as 'ego dissolution' (cf. Letheby 2017). These phenomena all result from disruption to predictive self-binding processes; the differences depend on which specific aspects of the self-model are compromised, and to what extent.

Ego Dissolution as Unbinding

When Carhart-Harris *et al.* (2012) published the first human fMRI studies of serotonergic psychedelics their results caused quite a stir. Not only did they find that intravenous psilocybin caused an overall reduction in neural activity (measured by BOLD signal), apparently contradicting earlier PET studies showing metabolic hyperactivity on psilocybin (Vollenweider *et al.* 1997). The decreases were concentrated mainly in the DMN. Psilocybin decreased activity in DMN regions but also the integrity of the network as measured by functional connectivity analyses. The normal functional coupling between key DMN hubs diminished—in particular, between the MPFC and the PCC. Despite apparent inconsistency with PET results, the fMRI results cohered with MEG investigations in the same study. These electrophysiological measures revealed another striking result: decreased alpha power (which has been linked to inhibitory processes) in the PCC correlated strongly with psychometric ratings of ego dissolution.

Not all of these findings have been replicated in all subsequent psychedelic fMRI studies. But various observations support a role for downregulation of the DMN, and the PCC in particular, in psychedelic ego dissolution. Palhano-Fontes *et al.* (2015) reported decreased BOLD signal in DMN nodes including the MPFC and PCC during ayahuasca intoxication. Meanwhile, thinning of the PCC has been observed in long-term ayahuasca users, with the degree of thinning correlated with both psychometric scores for trait 'self-transcendence' and the extent of prior ayahuasca use (Bouso *et al.* 2015). Interestingly, these ayahuasca users equalled or outperformed matched controls on tests of neuropsychological function, suggesting that this PCC thinning is not linked to cognitive impairment. The connection with findings about PCC deactivation in 'effortless awareness' meditation is obvious, and bolstered by the finding that acute ayahuasca intoxication increases mindfulness-related capacities (Soler *et al.* 2016).

Matters have been complicated by recent apparently conflicting findings concerning the neural correlates of ego dissolution. Lebedev *et al.* (2015) reported that disintegration not of the DMN but of the SLN correlated with psilocybin-induced ego dissolution. Referring to the apparent conflict with prior results, Lebedev *et al.* note that while the DMN is frequently linked to the narrative self, the SLN has been linked with a more minimal or embodied sense of self. The distinction between narrative and minimal selfhood is commonplace in recent literature (e.g. Blanke and Metzinger 2009) and will form part of our explanation of this pattern of results too.

Lebedev *et al.* also found ego dissolution significantly correlated with reductions in interhemispheric connectivity and in functional coupling between the medial temporal lobes (MTL) and cortical regions. This connects with earlier findings of MTL–DMN decoupling under psilocybin (Carhart-Harris *et al.* 2014)—though in the earlier studies, this decoupling did not significantly predict ego dissolution, while in the Lebedev *et al.* study, the MTL decoupled from broader cortical regions than just the DMN.

Most recently, Tagliazucchi *et al.* (2016) found that ego dissolution induced by LSD correlated with increased global functional connectivity, with these increases resulting from a breakdown in the integrity of resting-state networks. Here, the DMN and SLN were both affected. The strongest specific correlation with ego dissolution was decoupling of the parahippocampal cortex (in the MTL) from the precuneus—a key DMN node anatomically adjacent to the PCC. In this study, disintegration of the DMN also tracked reduced mental time travel to the past, consistent with this network's hypothesized role in this process (Speth *et al.* 2016).

What are we to make of these varied results? The major networks whose breakdown is linked to psychedelic ego dissolution have all independently been linked to one aspect or another of self-representation. Lebedev *et al.* suggest a need for fine-grained psychometric instruments capable of distinguishing between alterations to different aspects of self-awareness. Unfortunately, the ego dissolution inventory recently introduced and validated by Nour *et al.* (2016), while an important step, does not make such discriminations, so for now we are stuck with theoretical speculation.[4] However, our account will predict that Lebedev *et al.* are correct: different aspects or levels of the self-model are disrupted by psychedelics in different conditions, compromising the integrity of distinct self-binding processes.

We submit that an elegant resolution of these issues is as follows. Sui and Humphreys are correct that the self is more than a mere narrative posit. There is a robust self-representation implemented in cortical midline nodes of the DMN and SLN which plays a causal role in processing at all levels, both enabling and emerging from cognitive binding processes. This model includes both the representation of the self as a persisting entity as well as the various perceptual, embodied, and narrative representations of attributes, goals, and events which are bound to that entity. The plastering of these representations with 'sticky "I, me, mine" labels that... prove painful to remove' (Kingsland 2016, 285) explains the anxiety and dysphoria which sometimes attend ego dissolution experiences (albeit relatively rarely in controlled settings).

The core idea is that the self-binding function identified by Sui and Humphreys—the enhancement of binding for self-relevant information—is explained by the supposition that

---

4   It is possible that this limitation might be overcome by the use of additional techniques such as 'microphenomenological' interviewing (Millière 2017).

there is a self-model functioning according to predictive coding principles. This model is a complex, hierarchical representation of an object or entity having various attributes: spatiotemporal location, a history, personality traits, ownership of a body, and ownership and authorship of thoughts, feelings, and actions. The binding of representations of stimuli which are incorporated into the self-model occurs via top-down predictions generated by that model. And there is a congenital bias in favour of predictions generated by the self-model specifically because this is an adaptive feature of cognitive architecture for any entity that needs to parse experiences into internally and externally caused, to allocate resources preferentially to self-relevant experiences, and to communicate and interact consistently with others.

Meanwhile, the multi-layered nature of the self-model explains the apparently variable neural correlates of the same coarse-grained subjective report of 'ego dissolution'. Savage suggests that on LSD '[changes] in body ego feeling usually precede changes in mental ego feeling and sometimes are the only changes' (1955, 11), evincing an early clinical recognition that different aspects of the self-model might be affected to varying extents by psychedelics. This common temporal sequence, from blurring of body boundaries and loss of sense of ownership for body parts through to later loss of sense of ownership for thoughts, speaks further to the hierarchical architecture of the self-model. (Note that this account entails the testable prediction that psychedelic-induced disruption to SLN and DMN integrity should be selectively associated with disruption to embodied and narrative aspects of self-consciousness, respectively.)

However, it does not follow from the existence of a robust, causally efficacious, and multi-layered self-model that the entity represented by this model exists. We agree with Metzinger that it does not, and reject Dennett's identification of the self with narrative self-models as well as Hohwy and Michael's identification of the self with hierarchical predictive self-models. This is because the self-model does not posit the existence of a narrative or a model; it posits the existence of a substance, a prime mover, a persistent entity that underlies, owns, and initiates thoughts, experiences, and actions. In accordance with predictive coding principles of cognitive efficiency, the postulation of a bare particular to which events and attributes are bound is an economical means to prediction and explanation of the flow of information.

Hohwy and Michael (2017) argue that the self-model qualifies as a self because it performs many of the functions attributed to the self, such as being a 'hidden cause' of thought, emotion, and behaviour. It is true that the self-model does these things. But on our view it is more like an existential placebo than a successfully self-referential model. The expectation that a drug treatment will work is a mental representation which is capable of performing some of the functional role which it attributes to its posited object—viz., the pharmacological activity of a chemical. But this does not mean such an expectation is a pharmacological activity of a chemical. It does some of the right stuff, but it is the wrong kind of thing.

One way of putting the point, then, is that the concept SELF is not purely functionally defined but has crucial 'substantial' metaphysical commitments; playing part of the functional role of a self is an insufficient qualification for selfhood. Certainly in some cases, the correct inference to make from empirical discoveries is that a pretheoretically posited entity does in fact exist, but simply has a very different nature than we assumed. But the placebo case just presented, and the case of caloric

mentioned earlier, show that sometimes the correct inference clearly is an eliminative one; the explanatory work previously delegated to the putative entity is being done by something so different that identification is not on the cards. The pretheoretic concept of SELF has sufficiently central and defining metaphysical commitments that it too merits such eliminative treatment.

Or consider the case of objects understood as substances or particulars. The mind cannot help but attribute stably co-occurring properties to objects; this is a successful predictive strategy installed by evolution. But it is not necessarily the case that stabilization of a collection of properties depends on unification in an underlying entity. In the case of the self the stability and regularity we experience very plausibly depends on a network of hidden causes. The network plays the integrative role attributed to the self, but unlike the self it is not a simple, single thing. It has complex internal structure. Given this fact, we think that the unity and simplicity postulated by the self-model are erroneous. Ego dissolution helps expose this error by decomposing the self.

An example is the relationship between syndromes and their causes. Sometimes the co-occurrence of symptoms in a syndrome is because of an underlying disease, in which case the inference to a single hidden cause is correct. Often, however, a syndrome is the product of a network of interacting causes not themselves causally unified at a deeper level.

Hohwy and Michael point out that we make and revise judgements about ourselves which imply the existence of a unified entity (e.g. 'I' am generous, 'I' am stingy). They reasonably ask: what grounds such judgements if there is no entity to which they refer? Instrumentalists in the philosophy of science have long pointed out that the predictive adequacy of theoretical posits does not entail the existence of such posits. We are advocating an instrumentalist interpretation of the role of the self-model.[5] Of course this is consistent with the idea that in many cases predictive adequacy is served by accuracy. There are, for instance, good reasons for the success of a model that locates the causes of my voluntary actions within the boundaries of my body. But the existence of a substantial entity initiating the actions is not among those reasons.

5   It is common ground between Hohwy and Michael and ourselves that hidden endogenous causes which produce regular patterns of integrated experience are tracked by the multilevel self-model. That both the model and the (relatively) invariant regularities exist is not in dispute, which is why Hohwy and Michael are happy to regard the model as the 'real' (i.e. causally efficacious) self. On our view, the representation of causes of invariant patterns of experience is insufficient to account for the feeling we have of being a substantial self, in the same way as the existence of stability in patterns of perceptual experience is insufficient to account for the perceptual experience of substantial objects. The reason is that perceptual binding depends on the postulation of objects as part of the process. Our contention is that self-binding exploits an analogous cognitive strategy at all levels of the hierarchy. Tracking endogenous causes of integrated experience postulates a substance around which properties cohere. So our claim is that we experience not only the emergent result (endogenously integrated patterns of experience) but the locus of integration itself. The experience of ego dissolution allows us to experience 'unbinding' and understand the binding role played by the postulation of a substantial self.

Our argument, then, turns crucially on the claim that the content of the self-model is that of an enduring substance which is distinct from all the mental, emotional, and physical activity of which it is supposed to be the owner. As Metzinger (2005) points out, our 'phenomenal avatar' (conscious self-model) is 'transparent': one does not feel like an avatar encoded in a 'biological data format', one feels like a unitary, persisting substance, or entity. The prevalence of this Cartesian intuition in philosophical discussions of selfhood (as exhibited by standard responses to Parfit's famous teleportation thought experiments) supports this psychosemantic claim. Some may find this claim introspectively dubitable. But it is precisely the ubiquity of this sense of 'I' that makes it difficult to isolate phenomenologically—and this is the reason why phenomena such as ego dissolution are theoretically valuable (Millière 2017). As Savage says in his discussion of LSD phenomenology, 'By and large the individual is not aware of the ego boundaries of his mind and body and becomes aware of them only when a change has occurred in them' (1955, 6).

Psychedelics, by deconstructing the avatar, render it opaque and acquaint subjects directly with its representational nature (cf. Letheby 2015, 2016). Alterations in feelings of 'mineness' or ownership, the sense of bodily boundaries, and so forth put pressure on the predictive hypothesis of a unitary entity underlying and persisting throughout experiences. The subsequent diminution in the sense of solid selfhood shows subjects that this sense is ultimately just one more conscious experience, rather than a transcendental precondition of all such experiences. The transformative existential shock which often attends this discovery testifies to the fact that a mere avatar is not what we, in the ordinary and sober course of things, deeply feel and take ourselves to be.

Conclusion

Dennett (1992) famously argued that the self is a 'centre of narrative gravity' akin to the centre of gravity of an object: an abstraction, but no less real or predictively useful for that. On the predictive coding construal of the integrative self, the self-model functions as a centre of representational gravity—not just in narrative processes but also in lower-level affective, bodily, and spatial processes. It is a representation of the 'point' around which everything else revolves. But crucially, it is not a representation as of a mere point, but as of a particular. This is why we reject Dennett's metaphysics of selfhood, as well as Hohwy and Michael's claim that the self-model has enough of the right attributes to qualify as a self.

If our conjectures are right, the content of the self-model is Cartesian: it is of a substance or an entity which has the properties and experiences. And here we agree with Metzinger that there 'is just no entity there, no individual substance, and scientifically we can predict and explain everything we want to predict and explain in a much more parsimonious way' (quoted in Marshall 2016). Thus, our view is structurally analogous to Mackie's (1977) error theory of morality, conjoining the psychosemantic claim that the self-representation is as of a Cartesian substance with the ontological claim that no such substance exists. It is also substantially similar to Metzinger's views about the self-model, developed in the context of a specific case and a mechanistic proposal (self-binding on predictive processing principles) about how the model performs its adaptive functions for the organism.

All of this fits smoothly with the idea that the DMN and SLN implement narrative and embodied aspects of self-representation, respectively. The self-representation as we conceive it serves as an organizing principle at different levels and in different domains of processing. Perceptual representations are organized in an egocentric space; interoceptive representations are interpreted as signals of adaptively relevant events or 'core relational themes' (Prinz 2004); and narratives are structured around the fortunes and prospects of a protagonist. At all levels, salience is attributed, attention directed, and information integrated in accordance with the relevance of information to the organism's goals. As James Kingsland (2016, 209) puts it, 'we have evolved into an ape that takes things personally'.

Funding

References

Albahari M. Insight knowledge of no-self in Buddhism: an epistemic analysis. *Philos Imprint* 2014;**14**:1–30.

Bachmann T. How a (sub)cellular coincidence detection mechanism featuring layer-5 pyramidal cells may help produce various visual phenomena. *Front Psychol* 2015;**6**:1947. doi:10.3389/fpsyg.2015.01947.

Ballard DH, Hinton GE, Senjowski TJ. Parallel visual computation. *Nature* 1983;**306**:21–6.

Blanke O, Metzinger T. Full-body illusions and minimal phenomenal selfhood. *Trends Cogn Sci* 2009;**13**:7.

Bouso JC, Palhano-Fontes F, Rodríguez-Fornells A *et al.* Long-term use of psychedelic drugs is associated with differences in brain structure and personality in humans. *Eur Neuropsychopharmacol* 2015;**25**:483–92.

Brewer JA, Garrison KA, Whitfield-Gabrieli S. What about the "self" is processed in the posterior cingulate cortex?. *Front Hum Neurosci* 2013;**7**:647. doi:10.3389/fnhum.2013.00647.

Brewer JA, Garrison KA. The posterior cingulate cortex as a plausible mechanistic target of meditation. *Ann N Y Acad Sci* 2014;**1307**:19–27.

Burwick T. The binding problem. *WIREs Cogn Sci* 2014;**5**:305–15.

Carhart-Harris RL, Erritzoe D, Williams T *et al.* Neural correlates of the psychedelic state as determined by fMRI studies with psilocybin. *Proc Natl Acad Sci U S A* 2012;**109**:2138–43.

Carhart-Harris RL, Leech R, Hellyer PJ *et al.* The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. *Front Hum Neurosci* 2014;**8**:20. doi:10.3389/fnhum.2014.00020.

D'Argembeau A. On the role of the ventromedial prefrontal cortex in self-processing: the valuation hypothesis. *Front Hum Neurosci* 2013;**7**:372. doi:10.3389/fnhum.2013.00372.

Damasio AR. The brain binds entities and events by multiregional activation from convergence zones. *Neural Comput* 1989;**1**:123–32.

Damasio AR. *Self Comes to Mind: Constructing the Conscious Brain.* New York: Pantheon, 2010.

Davey CJ, Pujol J, Harrison BJ. Mapping the self in the brain's default mode network. *Neuroimage* 2016;**132**:390–7.

Dennett D. *Consciousness Explained.* Boston: Little Brown & Co, 1991.

Dennett D. The self as a center of narrative gravity. In: Kessel F, Cole P, Johnson D (eds.), *Self and Consciousness: Multiple Perspectives*. Hillsdale: Erlbaum, 1992.

Dos Santos RG, Osório FL, Crippa JAS et al. Antidepressive, anxiolytic, and antiaddictive effects of ayahuasca, psilocybin, and lysergic acid diethylamide (LSD): a systematic review of clinical trials published in the last 25 years. *Ther Adv Psychopharmacol* 2016a;**6**:193–213.

Dos Santos RG, Osório FL, Crippa JAS et al. Classical hallucinogens and neuroimaging: a systematic review of human studies: hallucinogens and neuroimaging. *Neurosci Biobehav Rev* 2016b;**71**:715–28.

Eisner BG, Cohen S. Psychotherapy with lysergic acid diethylamide. *J Nerv Ment Dis* 1958;**127**:528–39.

Garcia-Romeu A, Kersgaard B, Addy PH. Clinical applications of hallucinogens: a review. *Exp Clin Psychopharmacol* 2016;**24**:229–68.

Gilboa A. Autobiographical and episodic memory—one and the same? Evidence from prefrontal activation in neuroimaging studies. *Neuropsychologia* 2004;**42**:1336–49.

Griffiths RR, Johnson MW, Carducci MA et al. Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: a randomized double-blind trial. *J Psychopharmacol* 2016;**30**:1181–97.

Guterstam A, Björnsdotter M, Gentile G et al. Posterior cingulate cortex integrates the senses of self-location and body ownership. *Curr Biol* 2015;**25**:1416–25.

Helmholtz H. *Treatise on Physiological Optics* [Translated from the 3rd German ed. Edited by Powell J.]. Rochester: Optical Society of America, 1925.

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

Hohwy J, Michael J. Why should any body have a self? In: de Vignemont F, Alsmith AJT (eds.), *The Subject's Matter*. Cambridge, MA: MIT Press, PsyArXiv, 2017. http://osf.io/fm4cr.

Huxley A. *The Doors of Perception*. London: Chatto and Windus, 1954.

Kingsland J. *Siddhartha's Brain: The Science of Meditation, Mindfulness and Enlightenment*. London: Robinson, 2016.

Klee GD. Lysergic acid diethylamide (LSD-25) and ego functions. *Arch Gen Psychiatry* 1963;**8**:461–74.

Lebedev AV, Lövdén M, Rosenthal G et al. Finding the self by losing the self: neural correlates of ego-dissolution under psilocybin. *Hum Brain Mapp* 2015;**36**:3137–53.

Lebrecht S, Bar M, Barrett LF et al. Micro-valences: perceiving affective valence in everyday objects. *Front Psychol* 2012;**3**:107. doi:10.3389/fpsyg.2012.00107.

Leech R, Sharp DJ. The role of the posterior cingulate cortex in cognition and disease. *Brain* 2014;**137**:12–32.

Letheby C. The philosophy of psychedelic transformation. *J Conscious Stud* 2015;**22**:170–93.

Letheby C. The epistemic innocence of psychedelic states. *Conscious Cogn* 2016;**39**:28–37.

Letheby C. Naturalizing psychedelic spirituality. *Zygon* 2017;**52**, in press.

Lou HC, Changeux JP, Rosenstand A. Towards a cognitive neuroscience of self-awareness. *Neurosci Biobehav Rev*, 2017. doi:10.1016/j.neubiorev.2016.04.004.

Marshall R. All about the ego tunnel. *3AM Magazine*, 2016. http://www.3ammagazine.com/3am/all-about-the-ego-tunnel/ (8 December 2016, date last accessed).

Mackie JL. *Ethics: Inventing Right and Wrong*. New York: Penguin, 1977.

Mashour GA. Cognitive unbinding: a neuroscientific paradigm of general anesthesia and related states of unconsciousness. *Neurosci Biobehav Rev* 2013;**37**:2751–9.

Masters REL, Houston J. *The Varieties of Psychedelic Experience*. New York: Holt, Rinehart and Winston 1966.

Metzinger T. Précis: being no one. *Psyche* 2005;**11**:1–35.

Millière R. Looking for the self. *Front Hum Neurosci*, 2017;**11**:245. doi: 10.3389/fnhum.2017.00245.

Nour MM, Evans L, Nutt D et al. Ego-dissolution and psychedelics: validation of the ego-dissolution inventory (EDI). *Front Hum Neurosci* 2016;**10**:269. doi:10.3389/fnhum.2016.00269.

Osmond H. A review of the clinical effects of psychotomimetic agents. *Ann N Y Acad Sci* 1957;**66**:418–34.

Pahnke WN. The psychedelic mystical experience in the human encounter with death. *Harvard Theol Rev* 1969;**62**:1–21.

Palhano-Fontes F, Andrade KC, Tofoli LF et al. The psychedelic state induced by ayahuasca modulates the activity and connectivity of the default mode network. *PLoS One* 2015;**10**:e0118143.

Prinz JJ. *Gut Reactions: A Perceptual Theory of Emotion*. Oxford University Press, 2004.

Qin P, Northoff G. How is our self related to midline regions and the default mode network? *Neuroimage* 2011;**57**:1221–33.

Raichle ME, MacLeod AM, Snyder AZ et al. A default mode of brain function. *Proc Natl Acad Sci U S A* 2001;**98**:676–82.

Revonsuo A. Binding and the phenomenal unity of consciousness. *Conscious Cogn* 1999;**8**:173–85.

Ross S, Bossis A, Guss J et al. Rapid and sustained symptom reduction following psilocybin treatment for anxiety and depression in patients with life-threatening cancer: a randomized controlled trial. *J Psychopharmacol* 2016;**30**:1165–80.

Savage C. Variations in ego feeling induced by d-lysergic acid diethylamide (LSD-25). *Psychoanal Rev* 1955;**42**:1–16.

Sessa B. *The Psychedelic Renaissance*. London: Muswell Hill Press, 2012.

Seth AK. Interoceptive inference, emotion, and the embodied self. *Trends Cogn Sci* 2013;**17**:565–73.

Shanon B. *The Antipodes of the Mind*. Oxford: Oxford University Press, 2002.

Singer W. Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 1999;**24**:49–65.

Soler J, Elices M, Franquesa A et al. Exploring the therapeutic potential of ayahuasca: acute intake increases mindfulness-related capacities. *Psychopharmacology* 2016;**233**:823–9.

Speth J, Speth C, Kaelen M et al. Decreased mental time travel to the past correlated with default-mode network disintegration under lysergic acid diethylamide. *J Psychopharmacol* 2016;**30**:344–53.

Spreng RN, Grady CL. Patterns of brain activity supporting autobiographical memory, prospection, and theory of mind, and their relationship to the default mode network. *J Cogn Neurosci* 2010;**22**:1112–23.

Sui J, Humphreys GW. The integrative self: how self-reference integrates perception and memory. *Trends Cogn Sci* 2015;**19**:719–28.

Swanson LR. The predictive processing paradigm has roots in Kant. *Front Syst Neurosci* 2016;**10**:79. doi:10.3389/fnsys.2016.00079.

Tagliazucchi E, Roseman L, Kaelen M et al. Increased global functional connectivity correlates with LSD-induced ego dissolution. *Curr Biol* 2016;**26**:1043–50.

van Essen DC, Anderson CH, Felleman DJ. Information processing in the primate visual system: an integrated systems perspective. *Science* 1992;**255**:419–23.

Vollenweider FX, Leenders KL, Scharfetter C *et al.* Positron emission tomography and fluorodeoxyglucose studies of metabolic hyperfrontality and psychopathology in the psilocybin model of psychosis. *Neuropsychopharmacology* 1997;**16**:357–72.

Watts A. A psychedelic experience: fact or fantasy? In: Solomon D (ed.), *LSD: The Consciousness-Expanding Drug*. New York: Putnam, 1964.