Creating a World in the Head: The Conscious Apprehension of Neural Content Originating from Internal Sources

Stan B. Klein and Judith Loftus

Department of Psychological and Brain Sciences, University of California, Santa Barbara

Stan B. Klein https://orcid.org/0000-0002-7754-014X

The authors thank Robert Kunzendorf for his expertise, insights and inspiration.

Correspondence concerning this article should be addressed to Stan B. Klein, Department of

Psychological and Brain Sciences, University of California, Santa Barbara, 551 UCEN Road,

Santa Barbara, CA 93106, United States. Email: Klein@psych.ucsb.edu.

Abstract

Klein, Nguyen, & Zhang (in press) argued that the evolutionary transition from respondent to agent during the Cambrian Explosion would be a promising vantage point from which to gain insight into the evolution of organic sentience. They focused on how increased competition for resources -- in consequence of the proliferation of new, neurally sophisticated life-forms -- made awareness of the external world (in the service of agentic acts) an adaptive priority. The explanatory scope of Klein et al (in press) was limited to consideration of the conscious apprehension of externally sourced content – i.e., content delivered from the sensory registration of objects occupying phenomenal space. But consciousness – at least for humans -- takes its objects from internal as well as external sources. In the present article we extend their analysis to the question of how internally sourced content (i.e., mental states) became the object of conscious apprehension.

Creating a World in the Head: The Conscious Apprehension of Neural Content Originating from Internal Sources

In a recent paper, Klein, Nguyen, and Zhang (in press) argued that sentience1,2 was a critical component of the adaptive solution to recurrent problems organisms faced in their evolutionary past. Specifically, evolution endowed organisms with the capacity to adopt an agentic stance toward increasingly complex and unpredictable environmental demands placed on them during the so-called Cambrian Explosion (henceforth CE). One consequence of taking an agentic stance was that it required the organism to project its internally situated representation3 of reality into a multidimensional phenomenal space4 existing outside its brain. Projection into an external space, in turn, required the evolution of sentience. Sentience was thus both the consequence of agentic behavior and its modus operandi.

Klein et al (in press) offered an analysis of the evolution of sentience based on evidence and argument. In the next section, we present a summary of their deliberations. Those interested in the conceptual foundation on which our summation is based are referred to Klein et al (in press).

A Brief Review of the Evolutionary Origins of Sentience

*1a. The Beginning*

 During the early part of the Paleozoic Era (of which the Cambrian was the first geological period), most lifeforms were unicellular and simple. It was not until the CE (which spanned approximately 25 million years beginning around 545 million years ago) that complex, multicellular organisms within the subphylum Vertebrata -- including mammals, birds, reptiles and fish – first appeared.

For early members of the vertebrate lineage, behavior likely consisted in movement occasioned by genetically transmitted action schemata in concert with ontogenetic adjustments (i.e., modifications acquired in the organism’s lifetime) executed in response to sensory detection of environmental stimuli. Activity for such organisms was not an intentional effort to act on one’s surroundings. It simply was movement in response to stimuli. Prior to the CE, organisms almost exclusively were respondents (Klein et al, in press).

*1b. Respondent versus Agent*

 For a respondent, behavior is caused, not chosen. Most typically, the cause is non-sentient detection of an environmental contingency. Once commenced, behavior continues to completion along a predetermined path, unaccompanied by awareness of having been issued from a self or directed toward objects existing beyond the respondent.

An agent, in contrast, behaves intentionally. Its behavior is deliberately chosen and intentionally directed toward effecting change in a world external to the agent. Prior to completion, agentic acts are subject to modification and correction based on the agent’s goals and interpretation of the situation (for discussion see Ferrero, 2022; MacMurray 1957/1969; Moreno, 2018; Pickering, 2024; D.M. Walsh, 2015).

*1c. A Necessary First Step toward Agentic Behavior is to Appreciate there is a World in which to Behave*

 For pre-CE organisms, behavior originated within, and operated on, neurally-housed representations. Environmental stimuli were nothing over and above the brain states5 enabled by neural systems designed to gather information about the world (e.g., electromagnetic radiation) and translate them into electrochemical spike trains (e.g., Aljadeff et al., 2016; Gerstner & Kistler, 2002).

If neural representation benefits survival, natural selection has no adaptation-driven imperative to extend reality beyond its cranial confines. For a system so designed, the physical world has no need of observer-independent realization: It exists for the organism as it exists in the organism. Accordingly, no meaningful distinction can be drawn between the physical world and its neural instantiation. The world was that which was in the organism’s head.

The transition from respondent to agent required the organism to transform predetermined, inwardly conceived and directed acts into intentional behaviors targeting objects positioned in a multidimensional space outside its body. (e.g., James, 1904; Pereira, 2018; Pribram, 2004; Rudrauf et al., 2017; Velmans, 2007, 2009). To fashion a world external to the organism, neural activities must be phenomenologically projected onto the space external to the brain in which they originate (e.g., Pribram, 2004; Velmans, 2007). This process – called “phenomenal projection” (for discussion, see Pereira, 2018; Pribram, 2004; Velmans, 2007, 2009) -- served both as the product of and occasion for observation of the physical world. To experience a world consisting of objects and their relations requires those objects be fitted with properties in virtue of which they can be individuated. Sensory registration had expanded to include conscious perception.

*1d. Why Evolution Favored Agency*

 There are clear adaptive advantages accompanying acts performed in the service of reasoned deliberation (e.g., Klein et al, in press; MacMurray 1957/1969; Pickering, 2024; D.M. Walsh, 2015). An agent acts on its environment in virtue of being in its environment. In consequence, agentic acts can be tailored to the contingencies as they present and altered in accord with perceived changes of circumstance.

Prior to the CE, all taxa possessed of motility acted from response. However, as competition for resources intensified in consequence of the expansion of behavioral competencies enabled by the CE (e.g., (e.g., Feinberg & Mallatt, 2016; Ginsburg & Jablonka, 2007, 2019), responses that could be deliberately fitted to the demands of an increasingly unpredictable world would be favored by natural selection. The CE provided a context in which acts issuing from agency would have adaptive advantages over acts based on response when navigating the hazards posed by Cambrian existence.

*1e. From Agency to Sentience*

There is nothing inherent in the sensorial registration of external reality to suggest its phenomenal composition should be assembled into objects (e.g., Kant, 1998; Locke, 1689-1700/1975). This, of course, is reflected in James’s (1890) well-known posit that an infant’s initial experience of the world consists in a “blooming, buzzing confusion” (p. 488) of sensorial content. A sentient being provides structure to its world by imposing forms on the energetic patterns apprehended by its sensory organs.

There is longstanding debate regarding how an organism learns to individuate objects of experience (e.g., Aristotle, Hume, Kant, Locke, Mill, Plato). While proposals differ on specifics (for discussion see Hummel, 2013; Liter, & Bülthoff, 1998; Varzi, 2019), there seems to be consensus that object discrimination involves identification of properties that cohere repeatedly in space and time (e.g., James, 1890; Kant, 1998; Locke, 1690/1731; MacMurray, 1957/1969; Piaget, 1954; Spelke, 1990). Such properties are used to fashion single, undivided objects from the “blooming, buzzing confusion” of the whole. In short, properties that recurrently coalesce in the sensorium constitute and identify the object that possesses them.

We cannot navigate the world independent from our ways of partitioning it into objects. But, to behave agentically toward those objects the organism must be aware that (a) it is behaving, and (b) its behavior targets objects situated in the space external to the organism. Without such awareness, its behavior would that of a respondent rather than an agent.

But there can be no object-oriented agentic acts absent sentient registration of object defining properties (e.g., Klein et al, in press; MacMurray, 1957/1969; Orilia & Paoletti, 2022). The evolution of sentience -- the feeling of “what it is like for organism X to experience property Y” (e.g., the color of the sky or the pain of a bee sting; e.g., Chalmers, 1996; Hacker, 2002; T. Nagel, 1974) – is thus necessary for (a) populating phenomenal space with phenomenal objects, (b) behaving toward those objects in an agentic manner and (c) the passage from sensorial detection to conscious perception.

 In summary, natural selection’s answer to problems posed by the CE was to change respondents to agents. To act as an agent, an organism must differentiate the target of its behavior from non-targeted objects occupying phenomenal space. This is accomplished by breaking the organism free of its neural mooring and positioning it within a phenomenal space outside its brain. To enable this new way of “being in the world,” external space was populated with phenomenal objects whose presence could be detected by sentient registration (i.e., “the feeling of what it is like to experience X”) of the properties of which those objects were composed.

2. The Origin of Internally Sourced Conscious Reality

 Klein et al (in press) took the position that the transition from respondent to agent would be a promising vantage point from which to gain insight into the evolution of organic sentience. They focused on how increased competition for resources -- in consequence of the proliferation of new, neurally sophisticated life-forms -- made awareness of the external world (in the service of agentic acts) an adaptive priority. Consequently, the explanatory scope of their article was limited to consideration of the conscious apprehension of externally sourced content – i.e., content delivered from the sensory registration of objects occupying phenomenal space.

But consciousness – at least for humans -- takes its objects from internal as well as external sources (e.g., Chalmers, 1996; James, 1890; Johnson & Raye, 1981; Klein, in press; Levine, 2003; Kunzendorf, 2015; McGinn, 1991 Robinson, 2008; Tallis, 1991). The goal of the present article is to extend the analysis in Klein et al. (in press) to the question of how internally sourced content became an object of conscious apprehension.

*2a. Externally and Internally Sourced Mental Content*

In what follows, we use the terms “external” and “internal” to designate whether content provided to consciousness was culled from sensory experience or self-generated cerebration.6 Content originating from the physical world is assembled from sensory innervation: Sense organs target externally located stimuli which they transform into electrochemical discharges. The resulting activity is used to compile neural representations that capture (to varying degrees of fidelity; e.g., Bartlett, 1938; Munsterberg, 1909) the manner in which the stimuli were encoded at the time at which they first were experienced.

Internally sourced content, in contrast, is generated within the organism. Though often derived from external sources, once housed in the central nervous system (CNS) such content undergoes computational transformations (for discussion see Bruner, 1973; Bartlett, 1932; Guillery, 2017; Klein, in press; Klein et al., 2002) prior to recruitment for production of mental states such as imagery, inner monologue, decisions, judgments and memory.7 In contrast to externally generated content, internally sourced content is made to available to consciousness via activity originating within the CNS.

*2b. Sidebar: Our use of the Term “Mental State”*

In the preceding section we used the term “mental state”. This construct has a long, contentious history among psychologists and philosophers (for modern treatments, see Apperly, & Butterfill, 2009; Berger, 2014; Carruthers, 2015; Crane, 2015; Goldstein, 1994; J. Nagel, 2013; Perner, 1991; Searle, 1991). Therefore, explicit specification of our use may help avoid confusion and misunderstanding. While not everyone will agree with our construal, there should be little question of our intended meaning.

A mental state consists in both contentual (i.e., the intentional objects of consciousness)8 and qualitative (i.e., the subjective feel of those objects) features (e.g., Pernu, 2017). That is, for a state of the brain to qualify as “mental” there must be “something it is like” (e.g., T. Nagel, 1974) for the organism to be in that state.

It is worth mention that mental states have non-experiential neural events supporting their conscious realization (e.g. e.g., Searle, 1991; Klein, 2015a; Strawson, 2009). While these non-experiential preconditions are necessary for enabling a mental state, they are non-mental in the sense that they are mechanisms that help make the mental state possible, but are not the mental state per se. They conceivably could (and often do) go on without there being any conscious awareness of their operation (for discussion see Klein, 2015a, Strawson, 2009).9,10

In the remainder of this article, we address one type of internally sourced mental state –imagery. The multitude of intentional objects that populate human consciousness – e.g., belief, desire, memory, inner monologue, thought, fantasy, knowledge, judgment, hope -- came long (in evolutionary time) after imagery became an intentional object. They will not be addressed herein. Our reasons for focusing on mental imagery are dictated by the principles of evolutionary biology in conjunction with arguments made by Klein et al (in press). They are presented in the following section.

3. The Evolution of Internally Generated Conscious Imagery

Prior to the emergence of vertebrate life, internal monologues, judgments, beliefs and a host of other mental states had yet to make an appearance. This is not to say pre-Cambrian organic life was incapable of judgments, decisions, problem resolution, etc. But, to the extent these abilities were part of the organism’s repertoire, they were enabled by nonconscious response rather than conscious deliberation -- a claim whose justification is mandated by acceptance of the thesis that consciousness emerged during the CE.

*3a. Flux and Stability in Mental Representation*

Sentient beings -- in consequence of compositional and perspectival changes objects undergo over time and the multiplicity of contexts in and from which they are encountered -- are bathed in flux (e.g., Cornford, 1941, 1957; Klein, 2019a; Noonan, 1989).11 This poses a problem for creatures whose evolutionary viability depends, in large part, on their capacity to engage agentically with their surroundings. To serve as the focus of agentic behavior, an object must appear sufficiently consistent to permit its perceptual identification and re-identification (e.g., Brennan, 1988; Klein et al., in press; Mead, 2002; Noonan, 1989; Sider, 2001). But representational stability is difficult to attain when assignment of an object’s individuating properties (e.g., size, shape, mass, color, orientation, etc.) are compromised by ambiguities resulting from componential, contextual, and perspectival variation.

The capacity to transform the flux of external reality into relatively stable mental representations was evolution’s answer to this challenge (e.g., Klein, 2019a). In the early stages of vertebral evolution, creatures had recourse only to rudimentary mechanisms of object stabilization (e.g., sensitization, generalization) by which to navigate the chaotic world of experienced variation (e.g., Eccles, 1989; Kaufman, 1974; Mostofsky, 1965; V. Walsh & Kulikowski, 1998; Young, 1976). Such mechanisms, being largely reflexive, were unable to support sustained and flexible agentic engagement with the environment (e.g., Klein, 2019a; Klein et al, 2002). What was needed was a means of stabilizing object representations in the service of agentic behavior.

To lessen the experienced variability of objects in the physical world, natural selection fitted the neural architecture with mechanisms capable of supporting the stabilization of ontogenetically acquired content (e.g., consolidation; e.g., Dudai, 2004; McGaugh, 2000; Nadel & Moscovitch, 1997). Although not themselves the objects of consciousness, these representational structures served as the formative basis for the conscious identification and categorization of objects occupying phenomenal space.

But a representational structure, no matter how stable, is little more than a pointless appendage unless accompanied by mechanisms capable of making it available to the right systems at the right times. And this is exactly what the act of perception accomplished. During perception, stabilized representations were selectively recruited from the CNS, providing the interpretive framework within which sensory innervations arriving from external world were organized, identified and made available as objects of consciousness (i.e., perceived). In this way, the phenomenal world acquired a stability and consistency in which agentic acts could be meaningfully enacted.

*3b. Why Mental Imagery? The Principle of Evolutionary Conservatism*

By hypothesis, consciousness, in its initial incarnation, consisted in the perceptual registration of external objects whose neural stabilization made them amenable to agentic treatment. This expanded the range and manner in which the organism could address adaptive challenges encountered during the highly competitive conditions characterizing the CE. Additional modifications to this architecture subsequently were incorporated to the extent they enhanced the rate that the organism successfully solved new and/or residual challenges.

Natural selection does not respond to recurrent problems organisms faced in their evolutionary past with de novo production of complex, metabolically costly phenotypic systems. Rather, it modifies the design of existing structures in ways that enhance the organism’s ability to survive and reproduce (e.g., Barkow, Cosmides. & Tooby, 1992; Dawkins, 1976; Klein et al., 2002; Mayr, 1983; Sherry & Schacter, 1987; Williams, 1966). We call this the principle of “evolutionary conservatism”.11

Given the logic of evolutionary conservatism, it is reasonable to assume that the inner sourced content targeted for conscious apprehension initially consisted in nonconscious representational structures employed in the production of perceptual imagery. By extending the purview of consciousness to include these previously insentient representations, agentic acts could be directed toward objects located in an “internal or inner space” (a notoriously vague – both descriptively and phenomenologically – “site” of mental states and happenings: see *Section 3d. What is Inner Space?*). The organism now could consciously experience internally as well as externally sourced imagery.

There is evidence consistent with the proposition that internally and externally sourced imagery drew on many of the same neural structures (i.e., evolutionary conservatism). For example, studies have shown that many areas of the sensory cortex recruited during perception overlap with those active during visual, auditory, tactile and olfactory imagery (e.g., Dijkstra, Bosch and van Gerven (2019; Djordjevic, Zatorre, Petrides, Boyle, & Jones-Gotman, 2005; Freeman, James McCarthy 3rd, & Jolesz, 1990; Ganisa, Thompson, & Kosslyn, 2006; Lee, Han, Park, 2016; Stevenson, & Case, 2005; Zatorre, & Halpern, 2005). In addition, it long has been known that visual mental imagery selectively interferes with visual perception and that auditory mental imagery interferes with auditory perception (e.g., Craver-Lemley, & Reeves, 1990; Perky, 1910; Segal, & Fusella, 1970; but see Hopkins, 2012). Though far from conclusive, these findings are consistent with the hypothesis that many of the same neural structures participated in the evolution of perceptual and mental imagery.

*3c. The Adaptive Benefits of Mental Imagery*

 A case can be made that internally generated imagery served a number of adaptively beneficial functions. For instance, it allowed the organism to anticipate, imagine and in other ways experiment with content that once existed exclusively as nonconscious representations. Forming and maintaining a mental image of an object permits the organism to manipulate the object agentically from the safety of mental space -- enabling assessment of anticipatory possibilities free from external consequence.

Such capabilities also would be highly valued when searching for prey and avoiding predation. When the target of perception -- be it tracking elusive prey, avoiding a stealthy predator or fetching a bouncing ball – has a non-constant sensory presence, the agent can remain “on target” and “on task” even when the object of interest no longer is present to perception. In short, conscious awareness of internally sourced imagery helped bridge the epistemological gap created by “out of sight, out of mind”.

*3d. What is Inner Space?*

Consciousness must, of logical necessity, be directed toward some “other” that serves as its object (e.g., Brentano, 1995; Earle, 1955, 1972; Husserl, 1964; Klein, 2012; Neuhouser, 1990; Rossman, 1991; Zahavi, 2005).For humans, the distinction between self and other is as basic as that between life and death. We certainly appear to reside in a physical world outside our body. This space is home to the objects and events toward which we behave. While, as we have argued, our experience of (not necessarily the reality of) the physical world is a projection of a phenomenal space fashioned by the CNS, our daily involvement with external reality affords experiential assurance we are not solipsists.

 At some point, either during or following the CE, conscious registration broadened to include imagery positioned within the self (in philosophical parlance this often is called the “subject”; e.g., Earle, 1955, Kant, 1998; Klein, 2012; Zahavi, 2005). This inner space (i.e., the mind) is the experiential reality of hominin life (e.g., Klein, 2015b, 2016). But, in stark contrast to the experienced location of external reality, the placement of our inner mental topography is a matter of considerable debate.

 Although well beyond the scope of this article to attempt a scholarly treatment of deliberations on the location of experiential reality, it briefly should be noted that human intuition is not a reliable guide to its placement. For example, the ancient Greeks often located mental states in the torso rather than in the brain (e.g., Sullivan, 1999). And, as advocates of Cartesian philosophy see it the mind12, being an independent, immaterial substance, has contact with, but not location within, the material brain (e.g., Descartes, 1984; Almog, 2002).

 Despite debate, virtually all contemporary psychologists, philosophers and neuroscientists accept the mind as the product of the CNS (for discussion and review, see Clark, 2009; Dehaene, 2014; Joshua, Babu, & Jayaraj, 2020; Koch, Massimini, Boly, & Tononi, 2016; León, & Zahavi, 2023; Prinz, 2012). However, there is a significant difference between point of origination and experienced location. Sadly, there is little empirical evidence pointing to a specific experiential placement for inner space.

Some of its properties, however, have been explored. Most of these pertain to the modality of visual imagery. For example, psychological investigations have demonstrated that the experienced terrain of visual inner space shares a number of properties with perceived external space – such as extension, object location and dimensionality (e.g., Kosslyn, 1980, 1994; Kosslyn, & Alper, 1977; Kunzendorf, & Reynolds, 2004-2005; Pinker, 1980; Shepard, & Metzler, 1971).

 Suffice it to say that although inner mental space is an experiential certainty (it widely is held to be the aspect of reality of which we can be most certain; e.g., Gallagher & Zahavi, 2008; Midgley, 2014; Strawson, 2009; Wittgenstein, 1958), little is known about the placement or psychological topography of our internally sourced phenomenology.

*3e. Situating Objects in External and Internal Space: Personal Ownership*

 Since, consciousness of inner and external space share many underlying mechanisms (i.e., the principle of evolutionary conservatism), how are the internal and external objects of consciousness assigned to their corresponding locations?

One possibility is suggested by the concept of personal ownership (e.g., Albahari, 2006; Klein, 2015a; Lane, 2012). By personal ownership, we mean that self-generated (as opposed to sensorially-derived) intentional objects are felt as belonging both to and within oneself -- that my mental states take place in my head. What individuates a mental state as distinctly and exclusively taking place in one’s head (i.e., the inner space of the mind) is that I intuitively sense -- without need for intuition or inference -- that the content of that state is uniquely and infallibly authored from within (for comprehensive treatment see Albahari, 2006: Klein, 2012, 2015a: Klein & Nichols, 2012; Lane, 2012; Shoemaker, 1968; Stephens, & Graham, 2000; Zahavi, 2011).

Cases involving disruption of personal ownership lend support by showing that intentional objects can be sourced internally, yet be felt as unowned. When this happens, the content of consciousness still is apprehended, but the feeling that the content “belongs to me” no longer is secured. In consequence, the intentional object is treated as an external happening presenting to one’s senses.

For example, in certain clinical conditions, intentional objects can be present in awareness, yet lack the feeling that they are personally owned. When this occurs, the content that serves as the intentional object is treated as alien to the self (i.e., as external in origin). This can be seen in symptomology accompanying pathologies including, but not limited to, schizophrenic hallucinations (e.g., Bentall, 1990; Frith, 1992; Freeman & Garety, 2003) and somatoparaphrenic denial of body part ownership (e.g., Nightingale, 1982; Vallar & Ronchi, 2009). Misplacement also makes a non-clinical, nightly appearance when personally-authored dream narratives are experienced as external reality (e.g., Klein, 2019b).

The finding that individuals can experience mental content absent a feeling that this content belongs to “me,” shows that the relation between mental content and personal ownership can come undone. And this, in turn, endorses the conclusion that consciousness and its intentional objects are ontologically real and functionally independent aspects of mind (for discussion see Albahari, 2006; Klein, 2012, 2014, 2015a; Lane, 2012).

Returning to the question of localization, something like personal ownership – most likely an evolutionary precursor (personal ownership requires a sense of self, and this likely was missing from early vertebral life; e.g., Gallup, & Anderson, 2020; Povinelli, & Cant, 1996; Reiss, & Morrison, 2017; Sedikides, Skowronski, & Dunbar, 2006) – provided the grounding from which externally and internally sourced objects could be positioned in phenomenal and mental space. The occasional misappropriation of intentional objects in modern humans suggests the possibility that locational slippage may have been far more common in our ancestors (a possibility raised by Jaynes, 1976).

4. Conclusions

 One of the most basic experiential distinctions for sentient beings (at least human sentient beings) is that between subject and object. This dichotomy has been the target of centuries of scholarly discourse. In philosophy and psychology, the issue often has been framed as the relation between self and non-self. While various proposals to make sense of the apparent bipartite division of reality have long been on display – e.g., materialism (radical, non-reductive), eliminativism, panpsychism, parallelism, idealism (solipsistic, absolute), phenomenalism, dualism (substance, property), pluralism – a definitive treatment is not in sight. In this article we offered a speculative meditation on the topic.

Importantly, we made no claim to understanding the nature of consciousness per se: We have absolutely no idea how insentient matter gives rise to an organism’s ability to experience life. Rather, we drew on considerations from psychology, philosophy and evolutionary biology in the hope we might acquire an understanding of (a) how life transitioned from insentient matter to organic consciousness, and (b) how the conscious registration of physical (objective) reality broadened to include experiential (subjective) reality.

We cannot vouch that our deliberations adhere to the reality of the evolutionary journey leading to consciousness and it localizations. We have tried to limit our reflections to things we believe logically warranted and consistent with what can be said to be known about the evolutionary progression of organic life. Some will feel our journey has led us astray (e.g., the ontological deflationism of metaphysical claims; e.g. Chalmers, Manley, & Wasserman, 2009; Lowe, 2001). Perhaps. We do, however, think informed speculation on this topic is absolutely necessary if we hope to make progress toward understanding who we are and where we are.

References

Aljadeff, J., Lansdell, B.J., Fairhall, A.L., & Kleinfeld, D. (2016). Analysis of neuronal

spike trains, deconstructed. Neuron, 91, 221–259.

Albahari, M. (2006). Analytical Buddhism: The two-tiered illusion of self. Houndsmills, UK: Palgrave Macmillan.

Almog, J. (2002). What am I? Descartes and the mind-body problem. Oxford, UK:

 Oxford University Press.

Apperly, I., & Butterfill, S. (2009). Do humans have two systems to track beliefs and

 belief‐like states? Psychological Review, 116, 953-970.

Armstrong, D.M. (1981). (1981). The nature of mind and other essays. Ithaca, NY:

 Cornell University Press.

Barkow, J. H., Cosmides, L., & Tooby, J. (1992). The adapted mind: Evolutionary

psychology and the generation of culture. New York, NY: Oxford University

Press.

Bartlett, F. C. (1932). Remembering. London, UK: Cambridge University Press.

Bentall, R. P. (1990). The illusion of reality: A review and integration of psychological

 research on hallucinations. Psychological Bulletin, 107, 82–95.

Berger, J. (2014). Mental states, conscious and nonconscious. Philosophy Compass, 9,

 392-401.

Block, N. (1995). On a confusion about a function of consciousness. Brain and

 Behavioral Sciences, 18, 227—247.

Brennan, A. (1988). Conditions of identity: A study in identity and survival. Oxford,

 UK: Oxford University Press.

Brentano, F. (1995). Descriptive psychology. London, UK: Routledge.

Broad, C.D. (1925). The mind and its place in nature. London, UK: Kegan Paul,

 Trench, Trubner & Co Ltd.

Brown, R. (2006). What is a brain state? Philosophical Psychology, 19, 729–742.

Bruner, J.S. (1973). Beyond the information given: Studies in the psychology of

 knowing. New York, NY: W.W. Norton & Company.

Carruthers, P. (2015). Perceiving mental states. Consciousness and Cognition 36,

498-507.

Chalmers, D.J. (1996). The conscious mind: In search of a fundamental theory.

 Oxford, UK: Oxford University Press.

Chalmers, D.J. (2018). The meta-problem of consciousness. Journal of Consciousness

Studies, 25, 6–61.

 Chalmers, D., Manley, D. & Wasserman, R. (2009). Metametaphysics: New essays on

 the foundations of ontology. Oxford, UK: Oxford University Press.

Clark, A. (2009). Spreading the joy? Why the machinery of consciousness is (probably)

 still in the head. Mind, 118, 963-993.

Cornford, F. M. (1941). The Republic of Plato. London, UK: Oxford University Press.

Cornford, F.M. (1957). Plato and Parmenides. New York, NY: The Liberal Arts Press.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J.

 Barkow, L. Cosmides, & J. Tooby (Eds.), The adapted mind: Evolutionary

 psychology and the generation of culture (pp. 163–228). New York, NY: Oxford

 University Press.

 Crane, T. (2015). The mental states of persons and their brains. Royal Institute of

 Philosophy Supplement, 76, 253-270.

 Craver-Lemley, C., & Reeves, A. (1992). How visual imagery interferes with vision.

 Psychological Review, 99, 633– 649.

 Dawkins, R. (1976). The selfish gene. New York, NY: Oxford University Press.

Dehaene, S. (2014). Consciousness and the brain: Deciphering how the brain codes our

 thoughts. New York, NY: Penguin Books.

Descartes, R. (1984). The philosophical writings of Descartes, 2 volumes. Cambridge,

 UK: Cambridge University Press (edited by J. Cottingham).

Djordjevic, J., Zatorre, R. J., Petrides, M., Boyle, J.A, & Jones-Gotman, M. (2005).

 Functional neuroimaging of odor imagery. Neuroimage, 24, 791–801.

Djordjevic, J., Zatorre, R.J., Petrides, M., & Jones-Gotman, M. (2004). The mind’s nose.

 Psychological Science. 15, 143–148.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram?

 Annual Review of Psychology, 55, 51– 86.

Earle, W. (1955). Objectivity: An essay on phenomenological ontology. New York,

 NY: The Noonday Press.

Earle, W. E. (1972). The autobiographical consciousness. Chicago, IL: Quadrangle

Books.

 Eccles, J.C. (1989). Evolution of the brain: The creation of consciousness. London, UK:

 Routledge.

 Feinberg, T., & Mallatt, J. (2016). The ancient origins of consciousness. Cambridge,

 MA: The MIT Press.

 Ferrero, L. (2022). The Routledge handbook pf philosophy of agency. New York, NY:

 Routledge.

 Freeman, D., & Garety, P. A. (2003). Connecting neurosis and psychosis: The direct

 influence of emotion on delusions and hallucinations. Behaviour Research and

 Therapy, 41, 923–947.

 Frith, C.D. (1992). The cognitive neuropsychology of schizophrenia. East Sussex,

 UK: Psychology Press.

 Gallagher, S., & Zahavi, D. (2008). The phenomenological mind. New York, NY:

 Routledge.

 Gallup, G. G., Jr., & Anderson, J. R. (2020). Self-recognition in animals: Where do we

 stand 50 years later? Lessons from cleaner wrasse and other species. Psychology

 of Consciousness: Theory, Research, and Practice, 7, 46–58.

 Gerstner, W. & Kistler, W. (2002). Spiking neuron models: Single neurons,

 populations, plasticity. Cambridge, UK: Cambridge University press.

 Ginsburg, S., & Jablonka, E. (2007). The translation to experience: Limited learning and

 limited experiencing. Biological Theory, 2, 218–230.

Ginsburg, S., & Jablonka, E. (2019). The evolution of the sensitive soul: Learning and

 the origins of consciousness. Cambridge, MA: The MIT Press.

 Goldstein, I. (1994). Identifying mental states: A celebrated hypothesis refuted.

 Australasian Journal of Philosophy 72, 46-62.

 Guillery, R. (2017). The brain as a tool: A neuroscientist’s account. Oxford, UK:

 Oxford University Press.

 Hacker, P.M.S. (2002). Is there anything it is like to be a bat? Philosophy, 77, 157–174.

 Hopkins, R. (2012). What Perky did not show. Analysis, 72, 431-439.

 Hummel, J.E. (2013). Object recognition. In D. Reisburg (Ed.), Oxford Handbook of

 Cognitive Psychology (pp. 32-46). Oxford, UK: Oxford University Press.

 Husserl, E. (1964). The phenomenology of internal time-consciousness. Bloomington,

 IN: Indiana University Press.

 James, W. (1890). Principles of psychology (Vol 1). New York, NY: Henry Holt and

 Company.

 James, W. (1904). Does ‘consciousness’ exist? The Journal of Philosophy, Psychology,

and Scientific Methods, 1, 477–491.

 Joshua, A., Babu, Y., & Jayaraj, G. (2020). Engine of consciousness in brain - a review.

 International Journal of Research in Pharmaceutical Sciences, 11, 659-667.

 Jaynes, J. (1976). The origin of consciousness in the breakdown of the bicameral mind.

 Boston, MA: Houghton Mifflin.

 Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. Psychological Review, 8,

 67–85.

 Kant, I. (1998). The Cambridge edition of the works of Immanuel Kant: Critique of pure

 reason. New York, NY: Cambridge University Press. (Translated by P. Guyer &

A.W. Wood).

 Kaufman, L. (1974). Sight and mind. New York, NY: Oxford University Press.

 Kim, J. (1998). Philosophy of Mind. Oxford, UK: Westview Press, Inc.

 Kirk, G.S., Raven, J.E., & Schofield, M. (1983). The presocratic philosophers (2nd ed.).

 Cambridge, UK: Cambridge University Press.

 Klein, S. B. (2014). The two selves: Their metaphysical commitments and functional

 independence. Oxford, UK: Oxford University Press.

 Klein, S.B. (2015a). The feeling of personal ownership of one’s mental states: A

 conceptual argument and empirical evidence for an essential, but

 underappreciated, mechanism of mind. Psychology of Consciousness: Theory,

 Research, and Practice, 2, 355–376.

 Klein, S.B. (2015b). A defense of experiential realism: The need to take

 phenomenological reality on its own terms in the study of the mind. Psychology

 of Consciousness: Theory, Research, and Practice, 2, 41–56.

 Klein, S.B. (2015c). What memory is. Wiley Interdisciplinary Reviews: Cognitive

 Science, 6, 1–38.

 Klein, S. B. (2016). The unplanned obsolescence of psychological science and an

 argument for its revival. Psychology of Consciousness: Theory, Research, and

 Practice, 3, 357–379.

 Klein, S.B. (2018). Remembering with and without memory: A theory of memory and

 aspects of mind that enable its experience. Psychology of Consciousness: Theory,

 Research, and Practice, 5, 117–130.

Klein, S.B. (2019a). An essay on the ontological foundations and psychological

 realization of forgetting. Psychology of Consciousness: Theory, Research, and

 Practice, 6, 292-305.

Klein, S.B. (2019b). The phenomenology of REM-sleep dreaming: The contributions of

 personal and perspectival ownership, subjective temporality, and episodic

 memory. . Psychology of Consciousness: Theory, Research, and Practice, 6,

 55-66.

Klein, S.B. (in press). Consider the source: An examination of the effects of externally

 and internally generated content on memory. Psychology of Consciousness:

 Theory, Research, and Practice.

 Klein, S. B., Cosmides, L., Tooby, J., & Chance, S. (2002). Decisions and the evolution

 of memory: Multiple systems, multiple functions. Psychological Review, 109,

 306–329.

 Klein, S.B., & Nichols, S. (2012). Memory and the sense of personal identity. Mind,

 121, 677–702.

 Klein, S.B., Nguyen, B.N., & Zhang, B.M. (in press). Going out of my head: An

 evolutionary proposal concerning the “Why” of sentience. Psychology of

 Consciousness: Theory, Research, and Practice.

 Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of

 consciousness: Progress and problems. Nature Reviews Neuroscience, 17,

 307–321.

 Kosslyn, S.M. (1980). Image and mind. Cambridge, MA: Harvard University Press.

 Kosslyn, S.M. (1994). Image and brain. Cambridge, MA: The MIT Press.

 Kosslyn, S.M., & Alper, S.N. (1977). On the pictorial properties of visual images:

 effects of image size on memory for words. Canadian Journal of Psychology, 31,

 32-40.

Kunzendorf , R.G. (2015). On the evolution of conscious sensation, conscious

 Imagination, and consciousness of self. Amityville, New York: Baywood

 Publishing Company.

 Kunzendorf, R.G., & Reynolds, K. (2004-2005). On the cognitive function of visual

 images and the development of individual differences. Imagination, Cognition

 and Personality, 24, 245-257.

Lane, T. (2012). Toward an explanatory framework for mental ownership.

 Phenomenology and the Cognitive Sciences, 11, 251–286.

Lee, S.E., Han, Y., &, Park, H.W. (2016). Neural activations of guided imagery and

music in negative emotional processing: A functional MRI Study. Journal of Music Therapy, 53, 257-278.

 León, F., & Zahavi, D. (2023). Consciousness, philosophy, and neuroscience. Acta

 Neurochirurgica, 165, 833–839.

 Levine, J. (2003). Knowing what it is like. In B. Gertler (Ed.), Privileged access:

 Philosophical accounts of self-knowledge (pp. 45–53). Hampshire, United

 Kingdom: Ashgate.

 Liter, J.C., & Bülthoff, H.H. (1998). An introduction to object recognition. Zeitschrift

 fur Naturforschung, 53c, 610-621.

 Locke, J. (1690/1731). An essay concerning human understanding. London, UK:

 Edmund Parker.

 Losos, J. B. (2008). Phylogenetic niche conservatism, phylogenetic signal and the

 relationship between phylogenetic relatedness and ecological similarity among

 species. Ecology Letters, 11, 995–1007.

 Lowe, E.J. (2001). The possibility of metaphysics: Substance, identity, and time.

 Oxford, UK: Oxford University Press.

 MacMurray, J. (1957/1969). The self as agent. London, UK: Farber and Farber.

 Mayr, E. (1983). How to carry out the adaptationist program. American Naturalist, 121,

 324–334.

 McGaugh, J. L. (2000). Memory -- a century of consolidation. Science, 287, 248–251.

McGinn, C. (1991). The problem of consciousness: Essays toward a resolution. Oxford,

 UK: Blackwell Publishers.

Mead, G. H. (2002). The philosophy of the present. Amherst, NY: Prometheus Books.

Midgley, M. (2014). Are you an illusion? New York,NY: Routledge.

Moreno, A. (2018). On minimal autonomous agency: natural and artificial. Complex

 Systems 27, 289-313.

Mostofsky, D.J. (1965). Stimulus generalization. Stanford, CA: Stanford University

 Press.

Münsterberg, H. (1909). On the witness stand: Essays on psychology and crime. New

York, NY: Doubleday.

Nagel, J. (2013). Knowledge as a mental state. Oxford Studies in Epistemology, 4,

275-310.

Nagel, T. (1974). What is it like to be a bat? Philosophical Review, 83, 435–450.

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the

 Hippocampal complex. Current Opinion in Neurobiology, 7, 217–227.

Neuhouser, F. (1990). Fichte’s theory of subjectivity. New York, NY: Cambridge

 University Press.

Nightingale, S. (1982). Somatoparaphrenia: A case report. Cortex, 18, 463– 467.

Noonan, H. (1989). Personal identity (2nd ed.). New York, NY: Routledge.

Orilia, F., & Paoletti, M. P. (2022). Properties. In E.N. Zalta (Ed.), The Stanford

 encyclopedia of philosophy. Metaphysics Research Lab, Stanford University.

Pereira, A., Jr. (2018). The projective theory of consciousness: From neuroscience to

 Philosophical Ppsychology. Trans/Form/Ação, 41(spe), 199–232.

Perky, C. (1910). An experimental study of imagination. American Journal of

 Psychology, 21, 422–452.

Perner, J. (1991). Understanding the representational mind. Cambridge, MA: The MIT

 Press.

Pernu, T. K. (2017). The five Mmrks of the mental. Frontiers in Psychology. 8: 1084.

 doi:10.3389/fpsyg.2017.0108.

Piaget, J. (1954). The construction of reality in the child. New York, NY: Basic Books.

Pickering, A. (2024). What is agency? A view from science studies and cybernetics.

 Biological Theory, 19, 16-21.

Pinker, S. (1980). Mental imagery and the third dimension. Journal of Experimental

 Psychology: General, 109, 354–371.

Pinker, S. (1997/2009). How the Mind Works. New York, NY: W.W. Norton &

 Company.

Povinelli, D.J., & Cant, G.H. (1996). Arboreal clambering and the evolution of self-

 Conception. The Quarterly Review of Biology, 70, 393-421.

Pribram, K.H. (2004). Consciousness reassessed. Mind and Matter, 2, 7–35.

Prinz, J. (2012). The conscious brain: How attention engenders experience.

 New York, NY: Oxford University Press.

Reiss, D., & Morrison, R. (2017). Reflecting on mirror self-recognition: A comparative

 view. In J. Call, G. M. Burghardt, I.M. Pepperberg, C.T. Snowdon, & T. Zentall

 (Eds.), APA handbook of comparative psychology: Perception, learning, and

 cognition (pp. 745–763). American Psychological Association.

Robinson, D. N. (2008). Consciousness and mental life. New York, NY: Columbia

 University Press.

Rosenthal, D.M. (1991). The nature of mind. New York, NY: Oxford University Press.

Rossman, N. (1991). Consciousness: Separation and integration. Albany, NY: State

 University of New York Press.

Rudrauf, D., Bennequin, D., Granic, I., Landini, G., Friston, K., & Williford, K. (2017).

 A mathematical model of embodied consciousness. Journal of Theoretical

 Biology, 428, 106–131.

Russell, B. (1921). The analysis of mind. London, UK: George Allen and Unwin, Ltd.

Ryle, G. (1949). The concept of mind. New York, NY: Barnes & Noble, Inc.

Searle, J. R. (1991). Consciousness, unconsciousness and intentionality. Philosophical

Issues, 1, 45–66.

 Sedikides, C., Skowronski, J.J., & Dunbar, R.I.M. (2006). When and why did the human

 self evolve? In M. Schaller, J.A. Simpson, & D.T. Kenrick (Eds.), Evolution and

 social psychology (pp. 55–80). Philadelphia, PA: Psychology Press.

Segal, S.J., & Fusella, V. (1970). Influences of imagined pictures and sounds on

detection of visual and auditory signals. Journal of Experimental Psychology, 83, 458– 464.

 Shepard, R.N., & Metzler, J. (1971). Mental rotation of three-dimensional objects.

 Science, 171 (3972), 701–703.

 Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems.

 Psychological Review, 94, 439–454.

Shoemaker, S. (1968). Self-reference and self-awareness. The Journal of Philosophy,

 65, 555–567.

Sider, T. (2001). Four-dimensionalism: An ontology of persistence and time. Oxford,

 UK: Oxford University Press.

Spelke, E.S. (1990). Principles of object perception. Cognitive Science, 14, 29-56.

Stephens, G. L., & Graham, G. (2000). When self-consciousness breaks: Alien voices

and inserted thoughts. Cambridge, MA: The MIT Press.

Stevenson, R.J., 7 Case, T.I. (2005). Olfactory imagery: a review. Psychonomic Bulletin

 and Review, 12, 244-264.

Strawson, G. (2009). Mental reality (2nd ed.). Cambridge, MA: The MIT Press.

Sullivan, S. D. (1999). Sophocles, use of psychological terminology: Old and new.

 McGill-Queen's University Press, Montreal, Canada.

Tallis, R. (1991). The explicit animal: A defense of human consciousness. London,

 UK: MacMillan Academic and Professional LTD.

Vallar, G., & Ronchi, R. (2009). Somatoparaphrenia: A body delusion. A review of the

 neuropsychological literature. Experimental Brain Research, 192, 533–551.

Varela, F.J., Thompson, E., & Rosch, E. (1993). The embodied mind: Cognitive science

 and human experience. Cambridge, MA: The MIT Press.

Varzi, A. (2019). Mereology. In E.N. Zalta (Ed.), The Stanford encyclopedia of

 philosophy. Metaphysics Research Lab, Stanford University.

Velmans, M. (2007). Reflexive monism. Journal of Consciousness Studies, 15, 5–50.

Velmans, M. (2009). Understanding consciousness (2nd ed.). New York, NY:

 Routledge.

Walsh, D.M. (2015). Organization, agency, and evolution. Cambridge, UK: Cambridge

 University Press.

Walsh, V., & Kulikowski, J. (1998). Perceptual constancy: Why things look as they do.

 Cambridge, UK: Cambridge University Press.

Williams, G. C. (1966). Adaptation and natural selection: A critique of some current

evolutionary thought. Princeton, NJ: Princeton University Press.

 Wittgenstein, L. (1958). The blue and brown books (Rush Rees, Ed.). New York, NY:

 Harper and Row.

Yoo, S-S., Freeman, D.K., McCarthy 3rd, J.J., & Jolesz, F.A. (1990). Neural substrates

 of tactile imagery: a functional MRI study. Neuroreport, 14, 581–585.

Young, J.Z. (1976). The evolution of memory. Burlington, NC: Scientific Publications

 Division, Carolina Biological Supply Company.

Zahavi, D. (2005). Subjectivity of selfhood: Investigating the first-person perspective.

 Cambridge, MA: The MIT Press.

Zahavi, D. (2011). The experiential self: Objections and clarifications. In M.

 Siderits , E. Thompson , & D. Zahavi (Eds.), Self, no self: Perspectives from

 analytical phenomenological and Indian traditions (pp. 56–78). Oxford, UK:

Oxford University Press.

 Zatorre, R.J., & Halpern, A.R. (2005) Mental concerts: Musical imagery and auditory

 cortex. Neuron, 47, 9–12.

Footnotes

1. Sentience is the capacity to experience life (Klein et al., in press). More formally, an organism is sentient if and only if there is “something it is like” for “organism X to be in mental state Y” (e.g., Hacker, 2002; T. Nagel, 1974).

Sentience is what most philosophers have in mind when discussing phenomenal consciousness (e.g., Block, 1995; Klein, 2015a; Strawson, 2009). Phenomenal consciousness is a mental state characterized by the manner in which it presents itself subjectively – that is, the way it feels to its possessor (e.g., Block, 1995; Chalmers, 1996; Hacker, 2002; T. Nagel, 1974).

2. In this article, we use the words sentience, consciousness, subjectivity, awareness and experience interchangeably. While this treatment may seem overly-inclusive (for review see Van Gulick, 2022), our identification of sentience with phenomenal consciousness (see Footnote 1) helps explain our usage (a similar synonymy is adopted by Chalmers, 2018).

3. In our usage, the term “representation” refers to an information-bearing neural structure (i.e., a pattern of neural activity possessing semantic properties) whose meaning may be subject to conscious apprehension (i.e., a mental representation).

4. In Klein et al (in press), the phenomenal space outside the organism was described as three-dimensional. Unfortunately, there are no a priori reasons to attribute a capacity for three-dimensional representation to Cambrian vertebrates. Accordingly, we believe it more prudent to describe the phenomenal space experienced by animals during the CE as “multidimensional”.

5. A useful examination of the properties of and criteria for qualifying as a “brain state” is provided by Brown (2006).

6. Klein et al., (2002) used the terms inceptive and derived to label what we call externally and internally sourced mental content. In this paper, we adopt the latter terminology since it seems better attuned to the distinction we are trying to capture.

7. It is well known that numerous changes (e.g., addition, subtraction, recombination) can, and typically do, take place between and encoding and retrieval (e.g., Bartlett, 1932; Nadel & Moscovitch, 1997).

8. All conscious states have content -- that is, they are about something. Strictly speaking, these “somethings” are called the “intentional objects of consciousness.” (e.g., Brentano, 1995) For ease of exposition, we often will omit the predicate “intentional” when referring to the objects of conscious apprehension.

9. An analogy may help. A theatrical play consists in a great deal of behind the scenes activity (e.g., financing, venue selection, auditions), but, strictly speaking, none of this activity is the play per se.

10. The thesis that “all things are in flux” can be traced to Pre-Socratic Greek antiquity (cf. late 6th or early 5th century BCE; e.g., Cornford, 1941, 1957; Kirk, Raven, & Schofield, 1983). The protagonists divide into two camps distinguished primarily by their metaphysical commitments -- that is, those who posit change as the nature of reality, and those who regard change as the appearance of an unchanging reality that lays behind it. It is important to recognize that despite differences in assignment of ontological status, both camps accord change a central role in the physical world (the appearance of change, after all, is an experience, and experiences are happenings realized in a subcategory of physical reality -- i.e., the brains of sentient creatures).

11. The principle of evolutionary conservation should not be confused with phylogenetic (or phylogenetic niche) conservatism – a term which refers to the tendency for species to retain their ancestral traits (e.g. Losos, 2008).

12. The concept of “mind” has proven notoriously difficult to fit with a set of propositional “truths” that realize consensual agreement (e.g., Armstrong, 1981; Broad, 1925; Kim, 1998; Pinker, 1997/2009; Rosenthal, 1991; Russell, 1921; Ryle, 1949; Varela, Thompson, & Rosch, 1993). As an article of grammar, “mind” takes, as its adjectival form, the word “mental”. Accordingly, we use the word “mind” to mean the collection of sub-experiential processes required for having a mental state, in addition to the mental states they enable (e.g., Klein, 2015c, 2018).