

Review article

The *nature* of primary consciousness. A new synthesisTodd E. Feinberg^{a,*}, Jon Mallatt^b^a Icahn School of Medicine at Mount Sinai and Mount Sinai Beth Israel Medical Center, New York 10003, USA^b School of Biological Sciences, Washington State University, Pullman, WA 99164-4236, USA

ARTICLE INFO

Article history:

Received 7 April 2016

Accepted 20 May 2016

Keywords:

Primary consciousness
 Neurobiological naturalism
 Explanatory gaps
 Hard problem
 Subjectivity
 Evolution

ABSTRACT

While the philosophical puzzles about “life” that once confounded biology have all been solved by science, much of the “mystery of consciousness” remains unsolved due to multiple “explanatory gaps” between the brain and conscious experience. One reason for this impasse is that diverse brain architectures both within and across species can create consciousness, thus making any single neurobiological feature insufficient to explain it. We propose instead that an array of general biological features that are found in all living things, combined with a suite of special neurobiological features unique to animals with consciousness, evolved to create subjective experience. Combining philosophical, neurobiological and evolutionary approaches to consciousness, we review our theory of neurobiological naturalism that we argue closes the “explanatory gaps” between the brain and subjective experience and naturalizes the “experiential gaps” between subjectivity and third-person observation of the brain.

© 2016 Elsevier Inc. All rights reserved.

Contents

1. Introduction: What makes consciousness unique? The explanatory gap, the hard problem, and neurobiological naturalism.	114
2. Multiple explanatory gaps exist, not one: the neuroontologically subjective features of consciousness.	114
3. Large anatomical diversity means the gaps cannot have a single explanation	115
4. Consciousness is so diverse that many biological features must characterize it: the general and special features.	117
5. The general and special features explain the subjective features of consciousness.	119
5.1. The transition to mental images: the general and special features and the “explanatory gaps”.	119
5.1.1. Mental images and the transition to referral	120
5.1.2. Mental images and the transition to unity	120
5.1.3. Mental images and the transition to mental causation	121
5.1.4. Mental images and the transition to qualia	121
5.2. Affective consciousness and the transitions to the subjective features	122
6. Auto-ontological and allo-ontological irreducibilities: the “experiential gaps”	123
7. Neurobiological naturalism	123
Acknowledgements	124
References	124

* Corresponding author.

E-mail address: tfeinber@chpnet.org (T.E. Feinberg).

1. Introduction: What makes consciousness unique? The explanatory gap, the hard problem, and neurobiological naturalism

In his book *What Makes Biology Unique?* evolutionary biologist Ernst Mayr (2004) argues that while modern biology has disproved theories of *vitalism* – the idea that living organisms are animated by virtue of some fundamental force (*vis vitalis*) – nonetheless life is characterized by principles and functions that are in fact unique to biology among the sciences. Mayr cites among these unique features the great complexity of living systems, the refractoriness of biological systems to purely reductive approaches, the rampant emergence of novel features in biological systems, and the evolution of teleonomic (goal-directed) processes.

However, despite life's special and unique features, all its processes are explainable by natural principles: the heart is a pump; digestion is the physical and enzymatic breakdown of food; quadriplegia results from interrupted neural transmission through the upper spinal cord; ecosystems arise from the interactions of populations of living organisms with each other and the physical environment. The scientific basis life is not a conceptual mystery (Fry, 2000; Ginsburg & Jablonka, 2015). Mayr concludes that while a full understanding of biology does not require the positing of novel physical forces beyond those already known to physics, it does require an analysis of natural features and principles that are *unique* to biological systems.

Mayr's view is now the accepted dogma in biology. But when it comes to consciousness studies, many scholars from different disciplines have proposed that there is something fundamentally “different” about the biology of consciousness when compared to other biological phenomena (Chalmers, 2010; Koch, 2012; Schrödinger, 1967; Sperry, 1977). These scientists and philosophers agree with Mayr that while biology in general and the non-conscious brain functions can in fact be wholly explained by the known laws of physics and chemistry, consciousness presents a mysterious “explanatory gap” (Levine, 1983) between the physical properties of the brain and the subjective experiences that the brain thereby creates. They claim that whenever one attempts to explain subjective experience in terms of physics, chemistry or even neurobiology, there is always something “left out” of the equation, and something more is needed beyond the unique biological principles enumerated by Mayr. David Chalmers (1995) relates the explanatory gap to the “hard problem of consciousness”, which is the problem of how and why conscious experiences exist. John Searle traces the gap to the mutual irreducibility of subjective and objective points of view.

... consciousness has a first-person or subjective ontology and so cannot be reduced to anything that has third-person or objective ontology. If you try to reduce or eliminate one in favor of the other you leave something out ... biological brains have a remarkable biological capacity to produce experiences, and these experiences only exist when they are felt by some human or animal agent. You can't reduce these first-person subjective experiences to third-person phenomena for the same reason that you can't reduce third-person phenomena to subjective experiences. You can neither reduce the neuron firings to the feelings nor the feelings to the neuron firings, because in each case you would leave out the objectivity or subjectivity that is in question.

[Searle, 1997, p. 212]

Most problematically, the discontinuity is such that it seems unbridgeable when compared to the seamless unification of biology with the physical sciences. The challenge for a science of consciousness is to bridge or close the gap with a natural explanation.

To address this problem, we have formulated a theory called *neurobiological naturalism*. It was inspired by the earlier theory of *biological naturalism*, which Searle (1984, 2007) presented as a philosophical solution to the mind-body or mind-brain problem. *Biological naturalism* stated that mental phenomena are strictly biological, “caused by neurophysiological processes in the brain and are themselves features of the brain.” While we agree with Searle that consciousness arises exclusively through biological principles, we felt that more must be said about the special nature of consciousness in the natural world, so we extended the theory to include the neurobiological features that only consciousness has (Feinberg, 2012; Feinberg & Mallatt, 2016a).

Our theory of neurobiological naturalism is based on three tenets. First, in using natural science to solve the mind-brain problem one must begin with the biological features described by Mayr and others. Second, while consciousness is built upon the features shared by all life, it also depends on additional special neurobiological features. Third, because consciousness is so complex and multifaceted, a complete theory of consciousness that closes the explanatory gaps must use multiple approaches that integrate philosophical, neurobiological, and evolutionary principles. By using these principles, a natural elucidation of consciousness, the subjective mind, and the hard problem is possible.

2. Multiple explanatory gaps exist, not one: the neuroontologically subjective features of consciousness

We focus on the most basic kind of consciousness rather than on higher kinds of awareness. That is, we seek the neurological basis and evolutionary origins of *phenomenal consciousness* (Revonsuo, 2006, 2010), which is also called *primary consciousness* (Edelman, 1989) or *subjectivity* (Feinberg, 2012; Metzinger, 2003; Nagel, 1989; Searle, 1992, 1997; Tye, 2000; Velmans, 2000). As defined by Revonsuo:

Table 1

The neuroontologically subjective features of consciousness (NSFC). Adapted from Feinberg (2012).

1. <i>Referral</i> ^a	Conscious experiences are about (referred to) the outer world, the body, or affective states, but are not referred to the neurons that produce the experiences
2. <i>Mental unity</i> ^b	Consciousness is unified and bound into a relatively unified field of awareness in contrast to the divisible set of individual neurons that create it
3. <i>Mental causation</i> ^c	How the subjective mind can have causal influence on behavioral actions, the material body, and the outside world
4. <i>Qualia</i> ^d	Qualities, the subjectively experienced attributes such as colors, pains, sounds, etc.

^a *Referral*: Brain (1951), Feinberg (2012), Sherrington (1947), Velmans (2000), Velmans and Schneider (2007).^b *Mental unity*: Baars (1988), Baars et al. (2013), Bayne (2010), Bayne and Chalmers (2003), Dennett (1991), Edelman (2004), Feinberg (2012), Meehl (1966), Metzinger (2003), Sellars (1963), Teller (1992).^c *Mental causation*: Dardis (2008), Davidson (1980), Heil and Mele (1993), Kim (1998), Popper and Eccles (1997), Walter and Heckmann (2003).^d *Qualia*: Chalmers (1995, 1996), Churchland (1985), Churchland and Churchland (1981), Crick and Koch (2003), Dennett (1991), Edelman (1989), Jackson (1982), Kirk (1994), Levine (1983), Metzinger (2003), Revonsuo (2006, 2010), Searle (1992, 1997), Tye (2000). Edelman (2004) also considers the specific memories of such attributes to be qualia.

Phenomenal consciousness is the current *presence* of subjective experiences, or the *having* of subjective experiences. An organism possesses phenomenal consciousness if there is any type of subjective experience currently present for it. The mere occurrence or presence of any experience is the necessary and minimally sufficient condition for phenomenal consciousness.

[Revonsuo, 2006, p. 37]

It is the “gap” between the phenomenal or primary consciousness and the brain that we seek to close. But rather than there being a single gap, philosophers and scientists have actually identified multiple explanatory gaps between subjective experience and the material brain. We (Feinberg, 2012; Feinberg & Mallatt, 2016a, 2016b, chap. 1) have grouped these into four *neuroontologically subjective features of consciousness (NSFC)*: referral, mental unity, mental causation, and qualia. All these gaps must be closed or scientifically bridged if we are to understand the nature of consciousness.

The NSFC are summarized and referenced in Table 1. *Referral* means sensory experiences are perceived, never as if in the brain where they are constructed but as if in the outside world (from stimuli received by *exteroceptors* on the body surface), or inside the body (from stimuli received by *interoceptors*), or as an *affective state* of positive or negative feeling that involves the whole self. The experiential gap here is between the brain where sensation is actually created and to where the brain refers that experience. In *mental unity* the gap is between the divisible, discontinuous brain that consists of individual neurons and the unified, continuous field of awareness. This has been called the “grain problem” meaning that the divisible “grain” of the neurons of the brain is far coarser than the apparently seamless “grain” of subjective experience (Meehl, 1966; Sellars, 1963; Teller, 1992). *Mental causation* is the puzzle of how the subjective, seemingly immaterial mind can cause physical effects in the material world, including the physical body. *Qualia* are subjectively experienced qualities, both sensed traits (textures, smells, patterns of light) and affective states. Explaining qualia scientifically is the classical hard problem, the particular explanatory gap that has received the most attention.

3. Large anatomical diversity means the gaps cannot have a single explanation

If it were possible to identify a single or even a few physical features of conscious brains that could explain all four explanatory gaps, that could provide a biological solution to the puzzles of consciousness. However, to complicate matters greatly, our analysis of the neural basis of consciousness both within and across species reveals extensive diversity in the brain architectures that create subjective experience (Feinberg & Mallatt, 2016a).

The first illustration of this diversity is that the three different aspects of consciousness – exteroceptive, interoceptive, and affective – are associated with different brain regions and varied brain architectures (Fig. 1). For example, although some researchers who focus on humans and other mammals claim that all three aspects stem exclusively from the cerebral cortex and the thalamus (Barrett, Mesquita, Ochsner, & Gross, 2007; Berlin, 2013; Craig, 2010; Koch, Massimini, Boly, & Tononi, 2016), there is mounting evidence that affects arise subcortically in vertebrates (Aleman & Merker, 2014; Damasio, Damasio, & Tranel, 2012; Fabbro, Aglioti, Bergamasco, Clarici, & Panksepp, 2015; Feinberg & Mallatt, 2016a; Merker, 2007; O’Connell & Hofmann, 2011; Panksepp, 1998, 2011; Panksepp & Panksepp, 2013). Further, the circuits for the exteroceptive and affective aspects are organized differently. That is, the neural pathways of *exteroceptive* consciousness are physically organized into point-by-point representations of the outer world or of the receptor fields in the body and brain, in a topographic arrangement called “isomorphic mapping” (Feinberg & Mallatt, 2013; Kaas, 1997; Risi & Stanley, 2014), but the brain structures involved in *affective* consciousness do not require isomorphic organization (Fig. 7.1 in Feinberg & Mallatt, 2016a).

The second example of conscious diversity in mammalian (including human) brains is that the conscious pathway for smell differs from the pathways for the other senses. That is, only the smell pathway lacks a mandatory relay through the thalamus, while corticothalamic paths are held to be vital for consciously perceiving the other senses (vision, hearing, touch, etc.) (Gottfried, 2010; Shepherd, 2007). Smell has a minor thalamic path to the brain (Mori, Manabe, Narikiyo, & Onisawa, 2013), but this not essential for olfactory consciousness so the stated difference remains.

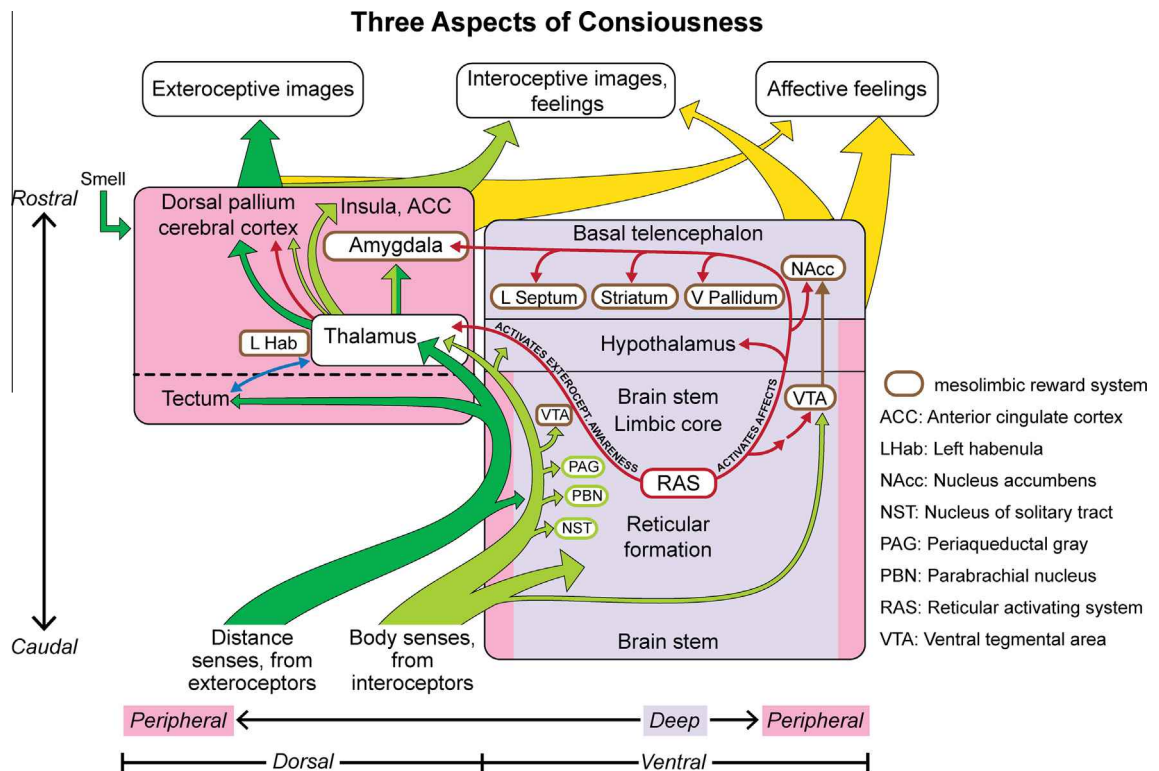


Fig. 1. How diverse structures in the vertebrate (mammalian) brain contribute to three different aspects of consciousness: exteroceptive, interoceptive, and affective. Large brain regions are represented by the two rectangular boxes. Notice that exteroceptive images (such as consciously seeing a flower) arise from mostly different brain regions than do affective feelings (such as liking or disliking the flower). Adapted from Feinberg and Mallatt (2016a).

Third is the case of birds differing from mammals. The behavioral evidence for consciousness existing in birds is as strong as it is for mammals in general (and many authorities now accept that cats, dogs, and horses are conscious) (Århem, Lindahl, Manger, & Butler, 2008; Boly et al., 2013; Butler, 2008; Butler, Manger, Lindahl, & Århem, 2005; Lefebvre, Nicolakakis, & Boire, 2002; Pepperberg, 2009). Yet birds' cerebral areas for consciousness differ considerably from those of mammals in structure and relative location. For example, the primary visual area is much farther forward in the bird cerebrum than in mammals. The differences are so substantial that only recently have enough commonalities been found to allow comparisons (Dugas-Ford, Rowell, & Ragsdale, 2012; Jarvis et al., 2013; Karten, 2013). Birds and mammals have experienced a lot of independent evolution in these cerebral structures since diverging from a common reptile-like ancestor a third of a billion years ago. This shows diversification of the brain structures responsible for consciousness.

As a fourth example of diversity, the brain regions for exteroceptive consciousness seem to differ in different groups of vertebrates. Fish and amphibians (“anamniotes”) differ from mammals and birds in lacking topographically organized sensory maps of the world in their small cerebrums. That is, anamniotes lack the maps that have been associated with conscious images in the cerebral cortex of mammals (see above). Most workers say that a large cerebral cortex (or the bird homologue) is required for consciousness (Boly et al., 2013; Butler, 2008; Edelman, Baars, & Seth, 2005; Edelman & Seth, 2009; Ribary, 2005; Rose et al., 2014; Seth, Baars, & Edelman, 2005), and would take the lack of cerebral sensory maps as evidence that anamniotes are not conscious. However, fish and amphibians do have such topographical maps in their subcortical midbrain, or optic tectum (Fig. 2). To us, this is evidence of tectal consciousness in lower vertebrates. Our reasoning is that these tectal maps would have no purpose unless used to produce images, to which fish and amphibians can refer for accurate behavioral interactions with their environment. Indeed, many workers who study the optic tectum use the terms of consciousness to describe its function in anamniotes: “perception,” “recognition,” (Dicke & Roth, 2009), and “object identification and location” (Wullimann & Vernier, 2009). Sensory consciousness in the anamniotes is consistent with their elaborated sensory organs (eyes that are more acute than those of most mammals, hearing ears, a lateral line, olfactory organs, many electroreceptors and mechanoreceptors). Consciousness is also consistent with the alert attention that fish pay to the abundant sensory stimuli they encounter, along with the complex behaviors they perform in response to the stimuli (Abbott, 2015; Bshary & Grutter, 2006; Griffin, 2001; Hotta et al., 2015; Kardong, 2012; Vindas et al., 2014). Our reasoning and evidence say that the cortex-based consciousness of mammals and birds differs from the predominantly tectum-based consciousness of fish and amphibians. That is diversity.

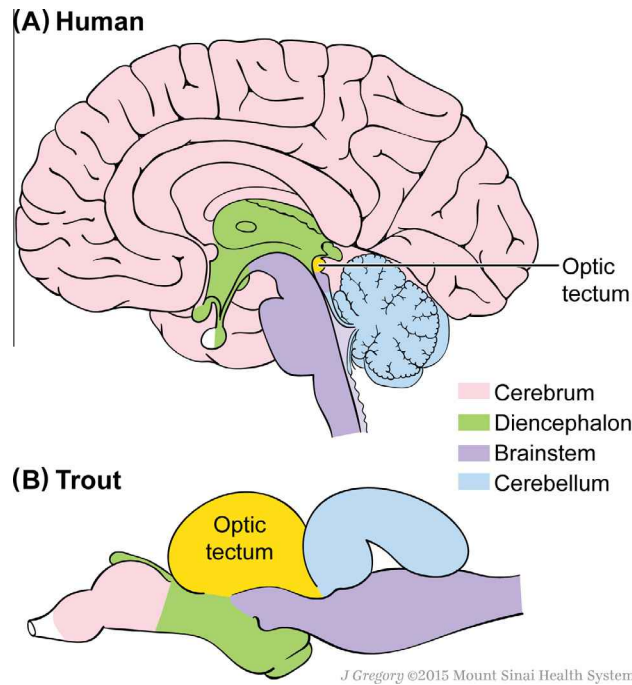


Fig. 2. Two different brain regions for conscious images among the vertebrates. (A) Cerebral cortex of a mammal (human). (B) Optic tectum of a fish.

Fifth, though least certain, some invertebrates may be conscious. After assembling a set of criteria for identifying consciousness in the vertebrates (Tables 2 and 3), we applied them to various invertebrates (Feinberg & Mallatt, 2016a). The nearest relatives of vertebrates, the invertebrate sea squirts and lancelets (amphioxus), have comparatively simple brains and neural circuits (Lacalli, 2008, 2016; Lacalli & Holland, 1998), so they failed our test. By contrast, arthropods (mainly insects and crabs) and cephalopod molluscs (mainly octopuses) met most of the criteria, including multiple, complex sensory organs and highly organized sensory pathways to the brain, behavioral preferences indicating remembered likes and dislikes, and more. For arthropods, the small sizes of their brains throw doubt on whether they could be conscious, and although the large-brained cephalopods exceed most of the requirements for which they have been tested, they have not been studied enough. Still, we have uncovered some evidence for arthropod and cephalopod consciousness. Because arthropods, vertebrates, and cephalopods are only distantly related, this would suggest the greatest diversity of all. Other investigators have likewise suggested early origins and some invertebrate consciousness, for reasons other than ours (Barron & Klein, 2016; Cabanac, 1996; Ginsburg & Jablonka, 2010a, 2010b, 2015; Huber, Panksepp, Nathaniel, Alcaro, & Panksepp, 2011; Packard & Delafield-Butt, 2014).

In summary, the striking diversity among the neural architectures indicates that there is not a *single* emergent process that can explain all aspects of primary consciousness, the “explanatory gaps,” and the “hard problem.” Thus, past studies of consciousness may have focused too much on single, dominant causes, such as the reciprocal corticothalamic interactions mentioned above, or primary drives from the core of the brain (Denton, 2006), or motor actions and rhythms (Linás, 2002; Merker, 2007). Instead, there are many different “emergences” that contribute to subjectivity. Based on these considerations, we next look for the *common factors* among the diverse neural architectures (both within and across species) that have primary consciousness.

4. Consciousness is so diverse that many biological features must characterize it: the general and special features

While the law of parsimony dictates that the simplest answer (i.e., with the fewest variables) is usually the best, it can lead to overly simplistic explanations when a process is multi-factorial in origin. In such cases more elaborate theories are better, and in the end are more parsimonious because only they explain *all* the observations.

By scrutinizing the many diverse anatomies that produce consciousness, we found features shared by all (Tables 2 and 3). These features are either observable structures or known natural processes. They indicate that consciousness is always created in hierarchical levels in complex systems, and all the levels are necessary. Here is a preliminary summary of the features in the tables: starting with the basic hierarchical systems of life, neurons were added and evolved into increasingly complex neural hierarchies.

Most basic are the *general biological features* that apply to all living organisms, even those without nervous systems (Table 2). These features were the first to evolve and they laid the foundations for life and consciousness. The second level

Table 2

The defining features of consciousness, Levels 1 and 2: General biological features and reflexes.

First level: General biological features , which apply to all living things	
1. <i>Life, embodiment, and process</i>	Life: use of energy to sustain self, responsiveness, reproduction, adaptiveness. All known life is cellular Embodiment: body with interior separate from the exterior. Has a boundary, such as a cell membrane or skin Processes: Life functions are processes not material things
2. <i>System and self-organization</i>	System: Entity considered as a whole, in which arrangements and interactions between the parts are important Self-organization: Interactions of the parts organize the patterns at global level of the whole
3. <i>Hierarchy, emergence, and constraint</i>	Hierarchy: Complex system with different interacting levels, organized from simpler to more complex: e.g., macromolecules to cells to organs to the organism. New, more elaborate features <i>emerge</i> at each higher level. Higher levels <i>constrain</i> lower levels for integration of the parts
4. <i>Teleonomy and adaptation</i>	Teleonomy: Biological structures perform programed, goal-directed functions Adaptation: A teleonomic structure or function as evolved by natural selection
Second level: Reflexes , which apply to all animals with nervous systems	
1. <i>Rates and connectivity</i>	Fast rates: Reflexes are <i>fast</i> , automatic responses to stimuli. Neural communication is rapid Connectivity: Simple reflex arcs are chains of several neurons connected at synapses. More complex arcs have more neurons in the chain (C) and in networks (N); they also have more neuronal interactions (I) and process more information (P). Further increase in CNIP was the royal road to complex nervous systems and consciousness
2. <i>Advanced examples include basic motor programs from central pattern generators (but still not conscious)</i>	

Table 3

The defining features of consciousness, Level 3: Special neurobiological features.

Third level: Special neurobiological features , which apply to animals with sensory consciousness	
1. <i>Elaborate sensory organs^a</i>	Image-forming eyes, multiple mechanoreceptors, olfactory and taste chemoreceptors. Plus high locomotory mobility, to travel and gather the sensory information
2. <i>Complex neural hierarchies</i>	Overall neural complexity: A brain, many neurons, ^b many neuronal subtypes Hierarchy complexity: For example, at least <i>four</i> successive levels of neurons before pre-motor centers in the conscious sensory pathways of human ^c
3. <i>Neural hierarchies create unique neural-neural interactions</i>	Interactions: Extensive reciprocal (reentrant, recurrent) communication within and between the hierarchies for the different senses Synchronized communication by gamma-frequency oscillations may be required, or else an “activated EEG” ^d
4. <i>Multisensory convergence^a</i>	Pathways of the different senses converge in the brain: allows unification of the senses into a single experience
5. <i>Neural hierarchies create isomorphic representations and mental images, and/or affective states^e</i>	Isomorphic representations: neurons arranged in topographic maps of the world or body structures Affective states ^e : Involve affect-associated neurotransmitters or neuromodulators such as dopamine and serotonin
6. <i>Unique combination of nested and non-nested hierarchical functions</i>	Nested function is the assembling of conscious unity by progressively unifying the sensory precepts Non-nested features include: some top-down control by higher brain; the physically separate parts of the conscious neurohierarchy; and topical convergence on grandmother neurons in mammals
7. <i>Attention</i>	Selective attention mechanisms in brain: for directing consciousness to salient objects in the environment. Related feature of <i>arousal</i> is also present
8. <i>Memory</i>	Memory regions in brain: needed for temporal continuity of experience, for providing learned reference-templates by which newly sensed objects can be recognized, etc.

^a Actually, our theory says that vision and visual consciousness evolved first, but this was consequently followed by elaboration of the other senses for olfactory, mechanosensory, electrosensory, taste, and auditory consciousness (Feinberg and Mallatt, 2016a). These followed so quickly as to be effectively simultaneous with vision.

^b Our rough guess here is that 20,000 neurons is not enough for consciousness (amount in gastropod *Aplysia*), but 10,000,000 neurons may be enough (amount in zebrafish): Table 9.2 in Feinberg and Mallatt (2016a).

^c Based on the number of levels in humans, to and including the primary sensory cerebral cortex. This marker is rough, because it does not consider the extensive neural interactions within the levels.

^d See Koch et al. (2016).

^e Non-structural, behavioral markers of affective consciousness are: operant conditioning based on learned positive and negative valences, behavioral trade-off between choices of different valence, frustration behavior, self-delivery of rewards, and conditioned place preference: Chapter 8 in Feinberg and Mallatt (2016a).

is reflexes that occur in animals that have the general biological features plus the added dimensions that nervous systems bring to an animal's life. But reflexes operate without creating sensory consciousness. Finally, only systems with consciousness possess the *special neurobiological features* (Table 3), which evolved through the elaboration of reflexes into neurohierarchies with more neurons, more levels, and more interactions between the neurons and levels (Feinberg & Mallatt, 2016a; see also Cornelis & Coop, 2014, and Ginsburg & Jablonka, 2010b). Although consciousness always evolved step by step just like any other emergent biological process, the special features – in association with the general features and reflexes – provided the advanced and unique properties that make primary consciousness possible.

5. The general and special features explain the subjective features of consciousness

The fossil record shows that the first, bacteria-like, life (with general features) evolved about 3.5 billion years ago, and the first multicellular animals that had nervous systems and reflexes arose around 580 million years ago. Then, according to our theory, with the appearance of the special features, consciousness was created in the earliest fishes (and perhaps the first arthropods), which evolved in the Cambrian Period about 540–520 million years ago (Buatois, Narbonne, Mángano, Carmona, & Myrow, 2014; Erwin & Valentine, 2013; Feinberg & Mallatt, 2013, 2016a; Schopf & Kudryavtsev, 2012).

Thus, we propose that critical to the nature and origin of consciousness is that with the addition of the special features to the general features, there occurred a *transition from reflexes to conscious mental images and affects*, and it was at that point that subjectivity, the philosophical “explanatory gaps,” and the hard problem were all naturally created (Fig. 3). In the following sections we explain how this occurred.

5.1. The transition to mental images: the general and special features and the “explanatory gaps”

We first consider how sensory mental images evolved and created “explanatory gaps” and along with them some aspects of the hard problem. We use the term “mental images” in the context of primary consciousness, to mean consciously experienced mental representations of things in the world or in the subject's body as these things are being sensed (Edelman, 1992). We do *not* use “images” to mean the capacity for imagining scenes in the absence of sensory information (Shepard, 1978).

We know that basal forms of sensory responsiveness were reflexive and innate, and they characterized the early, non-conscious bilaterian animals such as ancestral worms. These reflexes can be fully described in an objective way and do not entail any “gaps” in their biological explanation. But when the increasing brain complexity and the evolution of the special features first turned reflexive processing into “mental images” before 520 million years ago, many aspects of

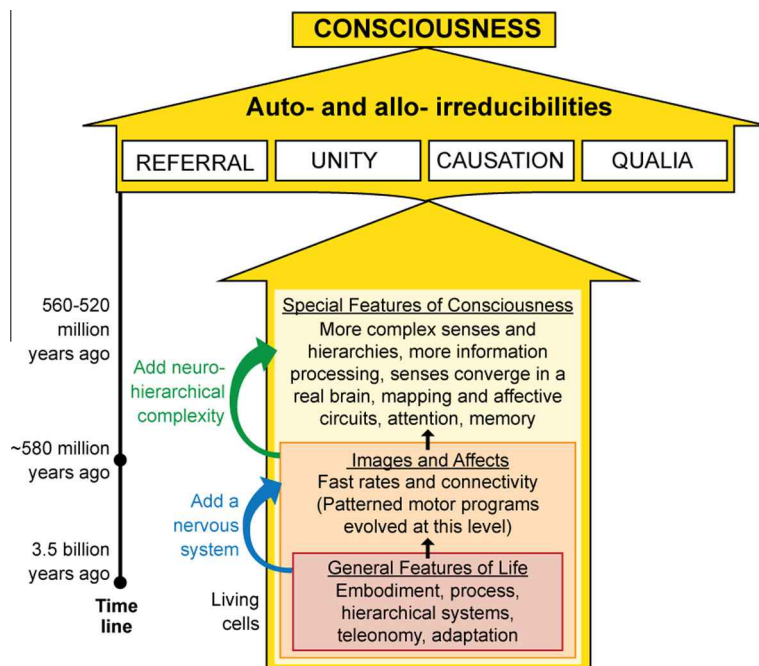


Fig. 3. The nested hierarchy of the general and special *objective* features explains the *subjective* features of consciousness (namely the four NSFCs, Referral, Unity, Causation, and Qualia). The “Special Features of Consciousness” are present in all vertebrates (and arguably in arthropods and cephalopods). The “General Features of Life” were retained when consciousness evolved.

ontological subjectivity were created. This is because mental images are referred, unified, mentally causal, and possess qualia (Table 1) – and thus images are one aspect of primary consciousness that entails all four explanatory gaps that characterize subjectivity.

In the first vertebrates, these sensory images were enabled by a great expansion of the senses for detecting things at a distance: detailed vision, hearing, olfaction, electroreception, and advanced mechanoreception using the lateral line of fish. Accompanying this sensory expansion was the evolution of special embryonic tissues named neural crest and ectodermal placodes, which develop into most of the receptors and sensory neurons for the distance senses of vertebrates (Feinberg & Mallatt, 2016a, chap. 5; Schlosser, 2014). These same kinds of distance senses arose in the first arthropods and cephalopods (Feinberg & Mallatt, 2016a, chap. 9; Mather, 2012; Strausfeld, 2013). Taking all this varied sensory information, then organizing it, sharpening it, attending to it, recording it (in memory), and joining it into a detailed image – that is what drove the evolution of the special neurobiological features, with their complex neuroprocessing hierarchies (Table 3).

The next question is how did the construction of mental images create the specific explanatory gap for each of the NSFC, referral, unity, causation and qualia.

5.1.1. Mental images and the transition to referral

Despite the vital contribution of the special features to consciousness, the general features remain operative and were critical to the creation of each NSFC (Fig. 3). Referral, for example, is a *teleonomic system feature of an embodied animal* (Table 2, First Level). The eye-blink reflex automatically protects the cornea from abrasive dust in the same way that a noxious pin-prick can still activate reflexive withdrawal in coma patient, or an earthworm uses its withdrawal reflex to retreat into the safety of its burrow when sensing vibrations from a predator in the external environment. In these primitive and non-conscious reflexes, the world-versus-organism and internal-versus-external relationships are already established, thus setting the stage for the later evolution of referred mental images.

With these reflexes serving as the neural scaffolding upon which referral is based, the addition of the special neurobiological features including elaborate sensory organs, complex neural hierarchies, and nested and non-nested sensory functions, led to the creation of the higher-order, differentiated, mental images that simulated the world and body but were not about the brain itself (Feinberg, 2000, 2012; Feinberg & Mallatt, 2013, 2016a, 2016b, chap. 1, 2016c, chap. 2; Table 3). Thus, referral away from the brain is a multisource process based on both the general and special features, the latter having been added to the former. The transition was uninterrupted due to the continuous evolutionary increase in neurohierarchical complexity.

5.1.2. Mental images and the transition to unity

Mental unity, like referral, is determined by the general and special features. A general feature of unity is that it is a *process*, as is life, and not a material thing that can be assigned to any one location. In other words, there is no single place where consciousness is *physically* unified (Dennett, 1991). That is, the many brain regions that contribute to mental unity are spread out in space, as exemplified by the different sensory pathways for the different senses and by the various brain-centers of multisensory convergence; and consciousness itself results from processes performed by these widespread neurons (James, 1904).

And as was the case with the evolution of referral, the creation of subjective mental unity occurred in the Cambrian at the critical transition between the general biological and special neurobiological features. As noted above, the special feature of *complex neural hierarchies* constructs mapped representations of different environmental objects from multiple types of sensory input and joins all these percepts into a unified conscious image. Consciousness arose as unified because using *fragmented* sensory maps to guide one's behavior would be inefficient or lethal. The feature of neurohierarchy is important to unity for another reason as well: constraint imposed by higher brain centers upon lower levels of the neurohierarchy allows the unification and nestedness of sensory systems (Feinberg, 2011).

Conscious unity may be assigned to neurobiological causes. Synchronized oscillatory patterns of spiking activity in the gamma frequency range occur across complex, intercommunicating networks of neurons, and these oscillations may contribute to unity by binding different percepts and different sensory modalities together (Zmigrod & Hommel, 2013). Such oscillations have been tied to many other aspects of consciousness as well, although they are much better studied in mammals than in more basally arising vertebrates like fishes (mammals: Buzsáki, 2006; Cabral-Calderin, Schmidt-Samoa, & Wilke, 2015; Engel, Fries, & Singer, 2001; Melloni et al., 2007; Northoff, 2013a, 2013b; Uhlhaas et al., 2009; fishes and amphibians: Bullock, 2002; Caudill, Eggebrecht, Gruberg, & Wessel, 2010; Northmore & Gallagher, 2003). Recently, doubt has been thrown on all this because the gamma oscillations seem to be absent in certain conscious states in humans; however, a similar phenomenon of integrated, low-voltage, fast waves (“activated EEG”) remains a consistent marker of consciousness (Koch et al., 2016).

As was the case with referral, we now see how the philosophical, neurobiological, and evolutionary points of view can be reconciled and integrated to explain mental unity. Due to the evolution and neurobiology of the special neurobiological features, mental unity – another “gap” between the brain and subjective consciousness (the hard problem) – was created. But, most importantly for the hard problem, this occurred without the addition of any new “mysterious” or “fundamental” processes.

5.1.3. Mental images and the transition to mental causation

Mental causation, in which subjective consciousness with mental images can “mysteriously” affect the “material world,” entails every general and special feature enumerated in [Tables 2 and 3](#). Among the general features that produce mental causation, *embodiment* is especially noteworthy. An individual brain is only capable of subjectively acting upon motor neural pathways that are embodied within that brain. I have no ability to directly control *your* actions through *my* thoughts, or vice versa. The general feature of *self-organization* also relates closely to mental causation because every living organism carries out programmed, teleonomic, goal-directed processes within the self-organization feature ([Mayr, 2004](#)) – and mental causation is certainly goal-directed. The groundwork for mental causation was laid by the non-conscious reflexive stage that we have modeled as a simple worm, because even reflexes are behaviors that affect the surroundings.

We can readily see how philosophical, neurobiological, and evolutionary features together explain the history of mental causation. Starting with an embodied organism (worm) that reflexively interacted with its environment to carry out its goal-directed survival processes, the subsequent addition of elaborate, non-reflexive, neural hierarchies led to complex and context-dependent behaviors. This was because the brain's new, mapped, images of the environment guided the behavioral actions to the right location and because the newly evolved affective feelings (see below) motivated and directed the behaviors toward “positive” and away from “negative” stimuli. This transition opened a seeming “gap” between subjective experience and drive, on the one hand, and the objective world that was affected by the behaviors, on the other. But, once again, there is nothing philosophically “mysterious” about this process when viewed within the evolutionary context and the general biological and special neurobiological features.

5.1.4. Mental images and the transition to qualia

Exteroceptive and interoceptive qualia arise when the brain represents sensory information from the subject's environment or body. As with the other three NSFC, the key factors are the combination of the general and special features, with qualia arising when the special features appeared.

The general feature of *adaptation* is paramount for explaining qualia because qualia act to distinguish among a wide range of sensory stimuli, as is needed for survival. To mate successfully and leave offspring, a female fish must be able to distinguish a brightly colored male from another, dully hued female. We propose that the first qualia were relatively simple, perhaps the crudest sensory assemblages that could form any image. Then, the qualia rapidly evolved more discriminations among viewed features (edges, shapes, shades, colors, depths, movements), among different odors, tastes, and mechanical stimuli. This increasing richness would have occurred during the explosive evolution of the distance senses in the Cambrian Period (see above). A strength of this proposal is that once the ability to make any conscious discriminations evolved – any at all – they could become increasingly fine-grained simply by adding more processing neurons to the existing neural hierarchies. The first, simple qualia evolved into millions of subtly differentiated qualia by a natural process not much more difficult than mathematical addition. The increasingly refined sensory discriminations provided survival benefits for finding food items and mates, avoiding predators, detecting dangers, and these benefits would apply in many different habitats and adaptive situations.

Qualia overlap the other three NSFC, so we have already explained much of this phenomenon and its natural, physical, and evolutionary origin. Thus, a subjective mental image is *qualia* assembled into a *unified* scene that *refers* to the outer world or body and guides behaviors and creates mental *causation* (for references, see [Table 1](#)). Logically, this means that because the mental images stem from *complex sensory hierarchies, isomorphic maps, and hierarchical nestedness*, so do the qualia. And, as we have just reviewed, both referral and unity are *system features of embodied brains*. In this respect qualia are no different.

So if one asks “why should there be qualia at all?” as [Chalmers wonders \(1995, 1996\)](#) we could answer “because they are highly adaptive.” But from a broader perspective, we could say there are *multiple parts to the answer*: Qualia are another unique consequence of the multiple general and special neurobiological features of complex nervous systems.

Finally, consider another aspect of consciousness that Chalmers called the *character of experience*. This is the perplexing question of why “red” subjectively feels exactly and uniquely the way red does. Or, why does the activation of the auditory pathway lead to subjectively heard sounds? Isn't that beyond scientific explanation? Chalmers states this problem of the character of experience as:

... why do individual experiences have their particular nature? When I open my eyes and look around my office, why do I have *this* sort of experience? At a more basic level, why is seeing red like *this*, rather than like *that*! It seems conceivable that when looking at red things, one might have had the sort of color experiences that one in fact has when looking at blue things. Why is the experience one way rather than the other? Why, for that matter, do we experience the reddish sensation that we do, rather than some entirely different sensation, like the sound of a trumpet?

[[Chalmers, 1996, p. 5](#)]

Again, we reply that the reason this problem is so “hard” is because it requires a multi-disciplinary answer that combines the neurobiological, neuroevolutionary, and neurophilosophical domains. If you ask, “why does red subjectively feel “red” and not as the note C sharp,” we first give the neurobiological answer that the neural pathways of color processing and those of sound processing are quite different, so they shouldn't and indeed couldn't feel the same. We can measure different people to tell whether red-coding neurons (light at 570 nm) ever get crossed with blue-coding neurons (light at 440 nm) or with auditory-path neurons, and can show that they seldom or never do so. Second, we would point out from an evolutionary perspective that color and sound stimuli usually signify different things, so there is strong selection pressure for a response

to a sound to be appropriate to the sound, rather than to an irrelevant color. Third, to answer “Why is it subjective?” we offer the philosophical solution that arises from systems and hierarchy theory (Fig. 3), as well as our philosophical analysis of the subjective/objective divide discussed below in Section 6.

5.2. Affective consciousness and the transitions to the subjective features

While sensory mental images can be neutral, the feelings of affective consciousness always have a *valence* – the aversiveness (negativity) or the attractiveness (positivity) of the stimulus or event. The most basal affects are the capacity to have experiences of a negative (noxious) or positive (pleasurable) valence.

As mentioned, affects do not have the map-like isomorphism that characterizes mental images. Without isomorphism to use as a marker, it is much more difficult for us to date and model the evolution of affects than of mental images, and to determine which living animals have affects. Therefore, the evolution of affects must be reasoned out step by step as follows. Investigators who study the relatively simple behaviors and nervous systems of sea slugs and other gastropod molluscs have identified a core circuit for affects (Gillette & Brown, 2015; Hirayama, Catanho, Brown, & Gillette, 2012; Hirayama & Gillette, 2012). They call this a sensory integrator circuit for incentives (Gillette & Brown, 2015), and it associates with neuronal ‘central pattern generators’ that control rhythmic movements and survival behaviors. The integrator circuit labels the sensory input it receives as either rewarding (+) or aversive (–) then it codes the motivations that dictate approach (+) versus avoidance (–) behaviors. It also potentiates *memories* of the salient stimuli, for associative learning. In gastropods, this incentive circuit is very simple, and may have too few neurons to produce true, affective, experiences in these invertebrates. Yet in all the vertebrates the corresponding set of structures is enormously complex, including parts of the brain’s mesencephalon, diencephalon, and telencephalon: specific parts are the ventral tegmental area or posterior tubercular nucleus, laterodorsal tegmental nucleus, habenula, amygdala, ventral striatum, and more (Butler & Hodos, 2005; O’Connell & Hofmann, 2011; Ryczko & Dubuc, 2013; Ryczko et al., 2013; Stephenson-Jones, Flores, Robertson, & Grillner, 2012).

We have concluded that in fishes these brain structures are similar enough to those in humans to indicate all vertebrates have affective consciousness (Feinberg & Mallatt, 2016a, chap. 8). These affective brain structures are at the top of complex neurohierarchies that have all the general and special features of consciousness listed in Tables 2 and 3. For example, affective neural circuits require the general features of an *embodied brain*, and they perform *teleonomic* processes that aid survival. Furthermore, they participate in the special features of *arousal* and forming *memories*. Affects in vertebrates also seem to depend on specific neurotransmitter and neuromodulatory molecules such as brain-generated dopamine and serotonin (Barron, Søvik, & Cornish, 2010; Gillette & Brown, 2015; Hikosaka, 2010; Naderi, Jamwal, Chivers, & Niyogi, 2016; Ramage-Healey, 2014; Waddell, 2013), and these neurochemicals are present in all vertebrates, fish and lampreys.

Behavioral evidence, based on global operant learning and other criteria, also indicates that all vertebrates experience affects (Feinberg & Mallatt, 2016a). Therefore, we date the evolutionary appearance of conscious affects to the origin of the brain complexity with the above-mentioned neurochemicals, around 540–520 million years ago. This date was also when vertebrates evolved the other aspect of consciousness, mental imagery. The same date may apply to the arthropods, as their brains and behaviors meet many of our criteria for affects (Perry & Barron, 2013; Strausfeld, 2013; Waddell, 2013).

Consistent with the proposals of Michel Cabanac and others, we conclude that affective consciousness benefited the first vertebrates (and arthropods?) by efficiently directing motor responses to salient stimuli. Affects do this by motivating the animals and telling them which such stimuli to approach and which to avoid (Cabanac, 1996; Gallese, 2013; Giske et al., 2013; Ohira, 2010; Packard & Delafield-Butt, 2014).

However, when compared to the consciousness of mental images, the affective aspects of sensory pain and pleasure, and the pure global affects such as fear, raise a question about the “explanatory gaps” they create. This question does not involve the three subjective features of unity, causality, or qualia. No confusion or new gaps to fill here, because affects, just like mental images, are unified, causal, and qualitative. Instead the question involves the other subjective feature, *referral*, which differs for affects versus mental images: Whereas the consciousness of mental images “refers” a stimulus to a specific place

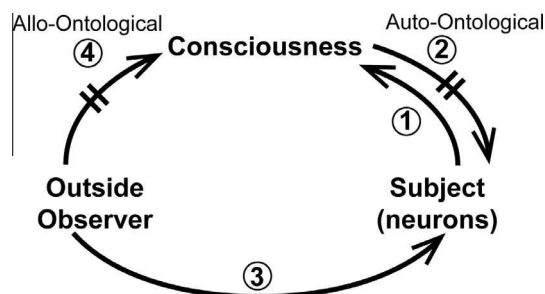


Fig. 4. Auto-ontological and allo-ontological irreducibilities: Fundamental barriers to reducing the objective to the subjective (and vice versa). (1) Subject has access to his or her conscious experiences. (2) Auto-ontological irreducibility: subject has no conscious access to his or her own objective neurons. (3) Observer has access to subject’s material neurons. (4) Allo-ontological irreducibility: observer has no access to subject’s conscious experiences.

in the external world or to the point in the body where the stimulus was applied, affective consciousness refers to the global, embodied self (“I” am scared.). However, this difference in referral sites proves irrelevant in the end because for mental images and affects alike *the brain never refers experience to the neural substrate of the brain that creates it*. We call this key feature of the conscious brain “auto-ontological irreducibility” (Fig. 4) and it provides more pieces to our understanding of the hard problem.

6. Auto-ontological and allo-ontological irreducibilities: the “experiential gaps

We have demonstrated how the neurobiological basis and evolutionary origins of the NSFC (the “explanatory gaps”) can be explained without any mysterious or missing ingredients. However, when subjectivity evolved from the general and special features over 520 mya, two barriers arose to uniting subjective awareness and objective science in a single conceptual framework (Fig. 4). We call these barriers *auto-* and *allo-ontological irreducibilities* (Feinberg, 2012).

Auto-ontological irreducibility means one’s subjective consciousness never experiences or “refers to” the objective neurons that create it. As Gordon Globus expressed it (1973, p. 1129), “It does not appear that the brain in any way codes or represents . . . its own structure. (The nervous system has no sensory apparatus directed to its own structure).” This irreducibility is explained in part by the way consciousness evolved: an organism’s survival depends on neural networks attending to the outer world and to the body, so it would be maladaptive if sensory neural networks evolved to consciously attend to their own functions. That would be redundant because neuronal activity is already optimized by physiological mechanisms that maintain the ionic constancy of body fluids bathing the neurons (Abbott, 2004; Hall, 2011, pp. 358–382). Additionally, supportive *glial cells* automatically help with this ionic homeostasis and they also attend to the other needs of neurons by regulating neurotransmission at synapses, energy metabolism, blood flow, and immune defense (Oberheim, Goldman, & Nedergaard, 2012; Verkhratsky & Parpura, 2014).

Neurons have operated efficiently like this since the pre-conscious, reflexive stage. For example, phylogenetic analysis indicates the ancestors of vertebrates, of arthropods, and of molluscs all had glial cells (Hartline, 2011). Besides being redundant, any conscious attention to neuronal maintenance would distract from the other important purposes of consciousness. Inefficiency and waste of effort are selected against, so no “objective” experiencing of the brain’s own neuronal signals ever evolved. This means the auto-ontological barrier was present ever since the dawn of consciousness.

Additionally, even if it were possible for someone to objectively observe his or her own neurons in the act of creating experience – via the use of a hypothetical “autocerebroscope” for example (Feigl, 1967) – these observations would amount to the same thing as an outsider’s third person observations, thus raising the same ontological barriers to one’s experience as is encountered by an outsider. This means there will always be an “*experiential gap*” between the subjective and objective points of view. There is no way that the subject can become objectively aware of his or her own neurons in the same way they are *subjectively experienced* from the embodied inside.

Allo-ontological irreducibility means that an outside observer has no access to a subject’s conscious experience (Fig. 4). It emerges not only from referral *but from all the NSFC*. This is because, by definition, all the NSFC are ontologically subjective and therefore cannot be experienced by anyone other than the subject. And, as we have already demonstrated, the neurobiological basis and evolutionary origins of the NSFC can be explained without any mysterious or missing ingredients. This point is not trivial because allo-ontological irreducibility is real, even absolute: Although two individuals can have similar or equal experiences (with both sensing the environmental temperature, light level, or a sound, for example), they naturally have different bodies, brains, viewpoints, different memories for reference, and incomplete intercommunication, so the allo-ontological barrier between them will always remain. Thus, the allo-ontological gap can be “bridged scientifically” or *naturalized* without needing to reduce the subjective to the objective points of view, or *visa versa*. The allo-ontological barrier exists and it contributes to the “experiential gap” between the objective and subjective, yet it poses no problem to science.

Therefore, despite the biological uniqueness of auto-ontological and allo-ontological irreducibilities, and the fact that they form absolute barriers between the objective and the subjective points of view, they have a perfectly natural explanation that is provided by the evolution and nature of the general and special features. The ontological subjectivity of consciousness is the result of its being an embodied, neurohierarchical, system-feature process, and all its unique subjective qualities follow from that.

In summary, for both mental images and affects, we can clearly understand the relationships among the philosophical issues of subjectivity and the evolution and neurobiology of complex nervous systems. Once nervous systems become sufficiently hierarchical, organized, and complex with the addition of the special features, there emerged the unique subjective features including auto- and allo-ontological irreducibilities, and the hard problem naturally arose without the addition of any mysterious features, new fundamental properties, or quantum factors. We emphasize that to solve the difficulties of subjective experience, it is necessary to bring together the philosophical, neurobiological, and evolutionary perspectives, with each informing the others (Feinberg & Mallatt, 2016a).

7. Neurobiological naturalism

Having updated our theory of neurobiological naturalism, we will summarize its unique approach to the problems of consciousness. Rather than considering subjectivity as a single perplexing “mystery,” we divided it into more manageable

Table 4

Summary of the main findings.

A. <i>Our logical steps to the solution of the hard problem</i>	
1.	Biology has no explanatory gaps but consciousness does, posing the hard problem
2.	Multiple explanatory gaps exist, not one (Table 1)
3.	Diversity of brain structures for consciousness means the gaps cannot have just one explanation
4.	Instead, many general and special features contribute (Tables 2 and 3)
5.	Evolutionary transitions from the general to the special features explain both aspects of subjectivity: mental images and affects. This closes the explanatory gaps
6.	Biological consideration of the allo-ontological and allo-ontological irreducibilities shows that the subjective/objective divide is real but can be explained by normal science. That bridges the experiential gaps
7.	The theory of <i>neurobiological naturalism</i> addresses the hard problem by dividing consciousness (subjectivity) into more parts and from more perspectives than do other theories
B. <i>Newly recognized division of the hard problem</i>	
1.	Explanatory gaps: gaps to explaining subjectivity; are solvable; now are explained so these gaps are closed
2.	Experiential gaps (Fig. 4): the objective/subjective divide; is real and cannot be closed, but now is bridged by scientific characterization

sub-problems. First, we broke down subjectivity into four “explanatory gaps”: referral, unity, qualia, and causation. We next divided and analyzed the diverse neurobiological sources of primary consciousness and identified the numerous general biological and special neurobiological features that create it. We then traced the evolutionary origins of primary consciousness and demonstrated how the general and special features naturally created subjectivity and the “explanatory gaps.” Finally, we showed how once subjectivity evolved, two additional features – auto- and allo-irreducibilities – naturally and “non-mysteriously” emerged and contributed to the hard problem. Ultimately, this approach of de-constructing and repeatedly dividing the problem of consciousness into its components allowed a natural solution that enabled us to see that three different disciplines – philosophy, neurobiology, and evolution – are needed to explain consciousness.

Although we are presenting a broad and unified resolution from a new perspective, we wish to point out that neurobiological naturalism is consistent with many other neurological theories of consciousness, in almost all ways except for precise evolutionary dating. That is, neurobiological naturalism reconciles with theories that focus on recurrent neuronal interactions and feedback loops, information integration, oscillatory binding, neural coding strategies, or other brain processes that contribute to the creation of consciousness (Baars, 2002; Baars, Franklin, & Ramsay, 2013; De Assis, 2016; Gennaro, 2012; Koch et al., 2016; Llinás, 2002; Min, 2010; Northoff, 2013a, 2013b; Ribary, 2005). However, each of those theories emphasizes a particular neurobiological aspect whereas neurobiological naturalism joins many such aspects and it brings in more philosophy and evolution.

Table 4 summarizes the logical steps and findings of this paper. We conclude that the theory of neurobiological naturalism “closes” the explanatory gaps between the biological and subjective features of consciousness, and *naturalizes* and “scientifically bridges” the ontological barriers and “experiential gaps” of auto-ontological and allo-ontological irreducibilities, thus providing a comprehensive solution to the hard problem of consciousness.

Acknowledgements

Thanks are extended to Jill K. Gregory of the Department of Academic Medical Illustration at the Icahn School of Medicine at Mount Sinai for her excellent work on the figure illustrations.

References

- Abbott, N. J. (2004). Evidence for bulk flow of brain interstitial fluid: Significance for physiology and pathology. *Neurochemistry International*, 45, 545–552.
- Abbott, A. (2015). Clever fish. *Nature*, 521, 413–414.
- Aleman, B., & Merker, B. (2014). Consciousness without cortex: A hydranencephaly family survey. *Acta Paediatrica*, 103, 1057–1065.
- Århem, P., Lindahl, B. I. B., Manger, P. R., & Butler, A. B. (2008). On the origin of consciousness—Some amniote scenarios. In H. Liljenstrom & P. Århem (Eds.), *Consciousness transitions: Phylogenetic, ontogenetic, and physiological aspects* (pp. 77–96). San Francisco, CA: Elsevier.
- Baars, B. J. (1988). *A cognitive theory of consciousness*. New York: Cambridge University Press.
- Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, 6, 47–52.
- Baars, B. J., Franklin, S., & Ramsay, T. Z. (2013). Global workspace dynamics: Cortical “binding and propagation” enables conscious contents. *Frontiers in Psychology*, 4.
- Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology*, 58, 373.
- Barron, A. B., & Klein, C. (2016). What insects can tell us about the origins of consciousness. *Proceedings of the National Academy of Sciences*, 113(18), 4900–4908.
- Barron, A. B., Søvik, E., & Cornish, J. L. (2010). The roles of dopamine and related compounds in reward-seeking behavior across animal phyla. *Frontiers in Behavioral Neuroscience*, 4.
- Bayne, T. (2010). *The unity of consciousness*. Oxford: Oxford University Press.
- Bayne, T., & Chalmers, D. J. (2003). What is the unity of consciousness? In A. Cleeremans (Ed.), *The unity of consciousness: Binding, integration, and dissociation* (pp. 23–58). New York: Oxford University Press.
- Berlin, H. (2013). The brainstem begs the question: *Petitio principii*. *Neuro-psychoanalysis*, 15(1), 25–29.
- Boly, M., Seth, A. K., Wilke, M., Ingmundson, P., Baars, B., Laureys, S., & Tsuchiya, N. (2013). Consciousness in humans and non-human animals: Recent advances and future directions. *Frontiers in Psychology*, 4.
- Brain, W. R. (1951). *Mind, perception and science*. Oxford: Blackwell Scientific Publications.

- Bshary, R., & Grutter, A. S. (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature*, 441, 975–978.
- Buatois, L. A., Narbonne, G. M., Mángano, M. G., Carmona, N. B., & Myrow, P. (2014). Ediacaran matground ecology persisted into the earliest Cambrian. *Nature Communications*, 5.
- Bullock, T. H. (2002). Biology of brain waves: Natural history and evolution of an information-rich sign of activity. In K. Arian & N. C. Moore (Eds.), *Advances in electrophysiology in clinical practice and research*. Istanbul: Kjellberg.
- Butler, A. B. (2008). Evolution of brains, cognition, and consciousness. *Brain Research Bulletin*, 75, 442–449.
- Butler, A. B., & Hodos, W. (2005). *Comparative vertebrate neuroanatomy* (2nd ed.). Hoboken, NJ: Wiley-Interscience.
- Butler, A. B., Manger, P. R., Lindahl, B. I. B., & Århem, P. (2005). Evolution of the neural basis of consciousness: A bird–mammal comparison. *BioEssays*, 27, 923–936.
- Buzsáki, G. (2006). *Rhythms of the brain*. New York: Oxford.
- Cabanac, M. (1996). On the origin of consciousness, a postulate and its corollary. *Neuroscience and Biobehavioral Reviews*, 20, 33–40.
- Cabral-Calderin, Y., Schmidt-Samoa, C., & Wilke, M. (2015). Rhythmic gamma stimulation affects bistable perception. *Journal of Cognitive Neuroscience*, 27, 1298–1307.
- Caudill, M. S., Eggebrecht, A. T., Gruber, E. R., & Wessel, R. (2010). Electrophysiological properties of isthmus neurons in frogs revealed by in vitro and in vivo studies. *Journal of Comparative Physiology A*, 196(4), 249–262.
- Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2, 200–219.
- Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. New York: Oxford University Press.
- Chalmers, D. (2010). *The character of consciousness*. New York: Oxford University Press.
- Churchland, P. M. (1985). Reduction, qualia and the direct introspection of brain states. *Journal of Philosophy*, 82, 8–28.
- Churchland, P. M., & Churchland, P. S. (1981). Functionalism, qualia and intentionality. *Philosophical Topics*, 12, 121–132.
- Cornelis, H., & Coop, A. D. (2014). Afference copy as a quantitative neurophysiological model for consciousness. *Journal of Integrative Neuroscience*, 13, 363–402.
- Craig, A. D. (2010). The sentient self. *Brain Structure & Function*, 214, 563–577.
- Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, 6, 119–126.
- Damasio, A., Damasio, H., & Tranel, D. (2012). Persistence of feelings and sentience after bilateral damage of the insula. *Cerebral Cortex*, 23, 833–846.
- Dardis, A. (2008). *Mental causation: The mind-body problem*. New York: Columbia University Press.
- Davidson, D. (1980). *Essays on actions and events*. Oxford: Clarendon Press.
- De Assis, L. P. (2016). Neural binding, consciousness, and mental disorders: Complexity as a common element. *ANS: The Journal for Neurocognitive Research*, 57, 110–120.
- Dennett, D. C. (1991). *Consciousness explained*. Boston: Little Brown.
- Denton, D. (2006). *The primordial emotions: The dawning of consciousness*. London: Oxford.
- Dicke, U., & Roth, G. (2009). Evolution of the visual system in amphibians. In M. D. Binder, N. Hirokawa, & U. Windhorst (Eds.), *Encyclopedia of neurosciences* (pp. 1455–1459). Berlin: Springer-Verlag.
- Dugas-Ford, J., Rowell, J. J., & Ragsdale, C. W. (2012). Cell-type homologies and the origins of the neocortex. *Proceedings of the National Academy of Sciences*, 109, 16974–16979.
- Edelman, G. M. (1989). *The remembered present: A biological theory of consciousness*. New York: Basic Books.
- Edelman, G. M. (1992). *Bright air, brilliant fire: On the matter of the mind*. New York: Basic Books.
- Edelman, G. M. (2004). *Wider than the sky: The phenomenal gift of consciousness*. New Haven: Yale University Press.
- Edelman, D. B., Baars, B. J., & Seth, A. K. (2005). Identifying hallmarks of consciousness in non-mammalian species. *Consciousness and Cognition*, 14, 169–187.
- Edelman, D. B., & Seth, A. K. (2009). Animal consciousness: A synthetic approach. *Trends in Neurosciences*, 32, 476–484.
- Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, 2, 704–716.
- Erwin, D. H., & Valentine, J. W. (2013). *The Cambrian explosion*. Greenwood Village, CO: Roberts and Company Publishers.
- Fabbro, F., Aglioti, S. M., Bergamasco, M., Clarici, A., & Panksepp, J. (2015). Evolutionary aspects of self- and world consciousness in vertebrates. *Frontiers in Human Neuroscience*, 9.
- Feigl, H. (1967). *The “mental” and the “physical”*. Minneapolis: University of Minnesota Press.
- Feinberg, T. E. (2000). The nested hierarchy of consciousness: A neurobiological solution to the problem of mental unity. *Neurocase*, 6, 75–81.
- Feinberg, T. E. (2011). The nested neural hierarchy and the self. *Consciousness and Cognition*, 20, 4–17.
- Feinberg, T. E. (2012). Neuroontology, neurobiological naturalism, and consciousness: A challenge to scientific reduction and a solution. *Physics of Life Reviews*, 9, 13–34.
- Feinberg, T. E., & Mallatt, J. (2013). The evolutionary and genetic origins of consciousness in the Cambrian Period over 500 million years ago. *Frontiers in Psychology*, 4.
- Feinberg, T. E., & Mallatt, J. (2016a). *The ancient origins of consciousness. How the brain created experience*. Cambridge, MA: MIT Press.
- Feinberg, T. E., & Mallatt, J. (2016b). Neurobiological naturalism. In R. R. Poznanski, J. Tuszynski, & T. E. Feinberg (Eds.), *Biophysics of consciousness: A foundational approach*. London: World Scientific Publishing Company.
- Feinberg, T. E., & Mallatt, J. (2016c). The evolutionary origins of consciousness. In R. R. Poznanski, J. Tuszynski, & T. E. Feinberg (Eds.), *Biophysics of consciousness: A foundational approach*. London: World Scientific Publishing Company.
- Fry, I. (2000). *The emergence of life on earth: A historical and scientific overview*. New Brunswick, NJ: Rutgers University Press.
- Gallese, V. (2013). Bodily self, affect, consciousness, and the cortex. *Neuropsychanalysis*, 15, 42–45.
- Gennaro, R. J. (2012). *The consciousness paradox: Consciousness, concepts, and higher-order thoughts*. Cambridge, MA: MIT Press.
- Gillette, R., & Brown, J. W. (2015). The sea slug, *Pleurobranchaea californica*: A signpost species in the evolution of complex nervous systems and behavior. *Integrative and Comparative Biology*, icv081, 1–12.
- Ginsburg, S., & Jablonka, E. (2010a). The evolution of associative learning: A factor in the Cambrian explosion. *Journal of Theoretical Biology*, 266(1), 11–20.
- Ginsburg, S., & Jablonka, E. (2010b). Experiencing: a Jamesian approach. *Journal of Consciousness Studies*, 17(5–6), 102–124.
- Ginsburg, S., & Jablonka, E. (2015). The teleological transitions in evolution: A Gántian view. *Journal of Theoretical Biology*, 381, 55–60.
- Giske, J., Eliassen, S., Fiksen, Ø., Jakobsen, P. J., Aksnes, D. L., Jørgensen, C., & Mangel, M. (2013). Effects of the emotion system on adaptive behavior. *American Naturalist*, 182, 689–703.
- Globus, G. G. (1973). Unexpected symmetries in the “world knot”. *Science*, 180, 1129–1136.
- Gottfried, J. A. (2010). Central mechanisms of odour object perception. *Nature Reviews Neuroscience*, 11, 628–641.
- Griffin, D. R. (2001). *Animal minds: Beyond cognition to consciousness*. Chicago: University of Chicago Press.
- Hall, J. (2011). *Gayton and Hall textbook of medical physiology* (12th ed.). Philadelphia: Saunders.
- Hartline, D. K. (2011). The evolutionary origins of glia. *Glia*, 59, 1215–1236.
- Heil, J., & Mele, A. (Eds.). (1993). *Mental causation*. Oxford: Clarendon Press.
- Hikosaka, O. (2010). The habenula: From stress evasion to value-based decision-making. *Nature Reviews: Neuroscience*, 11, 503–513.
- Hirayama, K., Catanho, M., Brown, J. W., & Gillette, R. (2012). A core circuit module for cost/benefit decision. *Frontiers in Neuroscience*, 6.
- Hirayama, K., & Gillette, R. (2012). A neuronal network switch for approach/avoidance toggled by appetitive state. *Current Biology*, 22, 118–123.
- Hotta, T., Takeyama, T., Heg, D., Awata, S., Jordan, L. A., & Kohda, M. (2015). The use of multiple sources of social information in contest behavior: Testing the social cognitive abilities of a cichlid fish. *Frontiers in Ecology and Evolution*, 3, 85.
- Huber, R., Panksepp, J. B., Nathaniel, T., Alcaro, A., & Panksepp, J. (2011). Drug-sensitive reward in crayfish: An invertebrate model system for the study of seeking, reward, addiction, and withdrawal. *Neuroscience and Biobehavioral Reviews*, 35(9), 1847–1853.

- Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, 32, 127–136.
- James, W. (1904). Does “consciousness” exist? *The Journal of Philosophy, Psychology and Scientific Methods*, 1(18), 477–491.
- Jarvis, E. D., Yu, J., Rivas, M. V., Horita, H., Feenders, G., Whitney, O., ... Siang-Bakshi, C. (2013). Global view of the functional molecular organization of the avian cerebrum: Mirror images and functional columns. *Journal of Comparative Neurology*, 521(16), 3614–3665.
- Kaas, J. H. (1997). Topographic maps are fundamental to sensory processing. *Brain Research Bulletin*, 44, 107–112.
- Kardong, K. (2012). *Vertebrates: Comparative anatomy, function, evolution* (6th ed.). Dubuque, IA: McGraw-Hill Higher Education.
- Karten, H. J. (2013). Neocortical evolution: Neuronal circuits arise independently of lamination. *Current Biology*, 23, R12–R15.
- Kim, J. (1998). *Mind in a physical world. An essay on the mind–body problem and mental causation*. Cambridge, MA: MIT Press.
- Kirk, R. (1994). *Raw feeling*. Cambridge, MA: MIT Press.
- Koch, C. (2012). *Consciousness: Confessions of a romantic reductionist*. Cambridge, MA: MIT Press.
- Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, 17(5), 307–321.
- Lacalli, T. C. (2008). Basic features of the ancestral chordate brain: A protochordate perspective. *Brain Research Bulletin*, 75, 319–323.
- Lacalli, T. C. (2016). The origin of vertebrate neural organization. In A. Schmidt-Rhaesa, S. Harszsch, & G. Purschke (Eds.), *Structure and evolution of invertebrate nervous systems*. Oxford: Oxford University Press.
- Lacalli, T. C., & Holland, L. Z. (1998). The developing dorsal ganglion of the salp *Thalia democratica*, and the nature of the ancestral chordate brain. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 353, 1943–1967.
- Lefebvre, L., Nicolakakis, N., & Boire, D. (2002). Tools and brains in birds. *Behaviour*, 139, 939–973.
- Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, 64, 354–361.
- Llinás, R. R. (2002). *I of the vortex: From neurons to self*. Cambridge, MA: MIT Press.
- Mather, J. (2012). Cephalopod intelligence. In J. Vonk & T. K. Shackelford (Eds.), *The Oxford handbook of comparative evolutionary psychology* (pp. 118–128). Oxford: Oxford University Press.
- Mayr, E. (2004). *What makes biology unique? Considerations on the autonomy of a scientific discipline*. Cambridge: Cambridge University Press.
- Meehl, P. (1966). The compleat autocerebroscopist: A thought experiment on Professor Feigl's mind/body identity thesis. In P. K. Feyerabend & G. Maxwell (Eds.), *Mind, matter and method* (pp. 103–180). Minneapolis: University of Minnesota Press.
- Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., & Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *The Journal of Neuroscience*, 27(11), 2858–2865.
- Merker, B. (2007). Consciousness without a cerebral cortex: A challenge for neuroscience and medicine. *Behavioral and Brain Sciences*, 30(01), 63–81.
- Metzinger, T. (2003). *Being no one. The self-model theory of subjectivity*. Cambridge, MA: MIT Press.
- Min, B. K. (2010). A thalamic reticular networking model of consciousness. *Theoretical Biology and Medical Modeling*, 7(10).
- Mori, K., Manabe, H., Narikiyo, K., & Onisawa, N. (2013). Olfactory consciousness and gamma oscillation couplings across the olfactory bulb, olfactory cortex, and orbitofrontal cortex. *Frontiers in Psychology*, 4, 743.
- Naderi, M., Jamwal, A., Chivers, D. P., & Niyogi, S. (2016). Modulatory effects of dopamine receptors on associative learning performance in zebrafish (*Danio rerio*). *Behavioural Brain Research*, 303, 109–119.
- Nagel, T. (1989). *Mind and cosmos: Why the materialistic neo-Darwinian conception of nature is almost certainly wrong*. Oxford: Oxford University Press.
- Northmore, D. P., & Gallagher, S. P. (2003). Functional relationship between nucleus isthmi and tectum in teleosts: Synchrony but no topography. *Visual Neuroscience*, 20(03), 335–348.
- Northoff, G. (2013a). *Unlocking the brain: Consciousness* (Vol. 2). New York: Oxford University Press.
- Northoff, G. (2013b). What the brain's intrinsic activity can tell us about consciousness? A tri-dimensional view. *Neuroscience & Biobehavioral Reviews*, 37, 726–738.
- O'Connell, L. A., & Hofmann, H. A. (2011). The vertebrate mesolimbic reward system and social behavior network: A comparative synthesis. *Journal of Comparative Neurology*, 519, 3599–3639.
- Oberheim, N. A., Goldman, S. A., & Nedergaard, M. (2012). Heterogeneity of astrocytic form and function. In R. Milner (Ed.), *Astrocytes: Methods and protocols, Methods in molecular biology* (Vol. 814, pp. 23–45). New York: Springer Science + Business Media.
- Ohira, H. (2010). The somatic marker revisited: Brain and body in emotional decision making. *Emotion Review*, 2, 245.
- Packard, A., & Delafield-Butt, J. T. (2014). Feelings as agents of selection: Putting Charles Darwin back into (extended neo-) Darwinism. *Biological Journal of the Linnean Society*, 112, 332–353.
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.
- Panksepp, J. (2011). Cross-species affective neuroscience decoding of the primal affective experiences of humans and related animals. *PLoS One*, 6(9), e21236.
- Panksepp, J., & Panksepp, J. B. (2013). Toward a cross-species understanding of empathy. *Trends in Neurosciences*, 36, 489–496.
- Pepperberg, I. (2009). *Alex and me*. New York: HarperCollins.
- Perry, C. J., & Barron, A. B. (2013). Neural mechanisms of reward in insects. *Annual Review of Entomology*, 58, 543–562.
- Popper, K. R., & Eccles, J. C. (1997). *The self and its brain*. New York: Springer.
- Remage-Healey, L. (2014). Frank Beach Award Winner: Steroids as neuromodulators of brain circuits and behavior. *Hormones and Behavior*, 66, 552–560.
- Revonsuo, A. (2006). *Inner presence: Consciousness as a biological phenomenon*. Cambridge, MA: MIT Press.
- Revonsuo, A. (2010). *Consciousness: The science of subjectivity*. Hove, UK: Psychology Press.
- Ribary, U. (2005). Dynamics of thalamo-cortical network oscillations and human perception. *Progress in Brain Research*, 150, 127–142.
- Risi, S., & Stanley, K. O. (2014). Guided self-organization in indirectly encoded and evolving topographic maps. In *Proceedings of the 2014 conference on genetic and evolutionary computation, July* (pp. 713–720). New York: ACM.
- Rose, J. D., Arlinghaus, R., Cooke, S. J., Diggles, B. K., Sawynok, W., Stevens, E. D., & Wynne, C. D. L. (2014). Can fish really feel pain? *Fish and Fisheries*, 15, 97–133.
- Ryczko, D., & Dubuc, R. (2013). The multifunctional mesencephalic locomotor region. *Current Pharmaceutical Design*, 19, 4448–4470.
- Ryczko, D., Grätsch, S., Auclair, F., Dubé, C., Bergeron, S., Alpert, M. H., ... Dubuc, R. (2013). Forebrain dopamine neurons project down to a brainstem region controlling locomotion. *Proceedings of the National Academy of Sciences*, 110, E3235–E3242.
- Schlosser, G. (2014). Development and evolution of vertebrate cranial placodes. *Developmental Biology*, 389, 1.
- Schopf, J. W., & Kudryavtsev, A. B. (2012). Biogenicity of Earth's earliest fossils: A resolution of the controversy. *Gondwana Research*, 22, 761–771.
- Schrödinger, E. (1967). *What is life? The physical aspects of living cell with mind and matter and autobiographical sketches*. Cambridge: Cambridge University Press.
- Searle, J. R. (1984). *Minds, brains and science*. Cambridge, MA: Harvard University Press.
- Searle, J. (1992). *The rediscovery of the mind*. Cambridge, MA: The MIT Press.
- Searle, J. R. (1997). *The mystery of consciousness*. New York: New York Review of Books.
- Searle, J. R. (2007). Dualism revisited. *Journal of Physiology – Paris*, 101, 169–178.
- Sellars, W. (1963). *Science, perception and reality*. London: Routledge and Kegan Paul.
- Seth, A. K., Baars, B. J., & Edelman, D. B. (2005). Criteria for consciousness in humans and other mammals. *Consciousness and Cognition*, 14, 119–139.
- Shepard, R. N. (1978). The mental image. *American Psychologist*, 33, 125–137.
- Shepherd, G. M. (2007). Perspectives on olfactory processing, conscious perception, and orbitofrontal cortex. *Annals of the New York Academy of Sciences*, 1121, 87–101.
- Sherrington, C. S. (1947). *The integrative action of the nervous system*. New Haven, CN: Yale University Press.

- Sperry, R. W. (1977). Forebrain commissurotomy and conscious awareness. *Journal of Medicine and Philosophy*, 2(2), 101–126.
- Stephenson-Jones, M., Floros, O., Robertson, B., & Grillner, S. (2012). Evolutionary conservation of the habenular nuclei and their circuitry controlling the dopamine and 5-hydroxytryptophan (5-HT) systems. *Proceedings of the National Academy of Sciences of the United States of America*, 109, E164–E173.
- Strausfeld, N. J. (2013). *Arthropod brains: Evolution, functional elegance, and historical significance*. Cambridge, MA: Harvard University Press.
- Teller, P. (1992). Subjectivity and knowing what it's like. In A. Berckermann, H. Flohr, & J. Kim (Eds.), *Emergence or reduction? Essays on the prospects of nonreductive physicalism* (pp. 180–200). Berlin/New York: Walter de Gruyter.
- Tye, M. (2000). *Consciousness, color, and content*. Cambridge, MA: MIT Press.
- Uhlhaas, P., Pipa, G., Lima, B., Melloni, L., Neuenschwander, S., Nikolić, D., & Singer, W. (2009). Neural synchrony in cortical networks: History, concept and current status. *Frontiers in Integrative Neuroscience*, 3, 1–19.
- Velmans, M. (2000). *Understanding consciousness*. London: Routledge.
- Velmans, M., & Schneider, S. (2007). *The Blackwell companion to consciousness*. New York: Wiley-Blackwell.
- Verkhatsky, A., & Parpura, V. (2014). Introduction to neuroglia. *Colloquium series on neuroglia in biology and medicine: From physiology to disease* (Vol. 1, No. 1, pp. 1–74). Morgan & Claypool Life Sciences (February).
- Vindas, M. A., Johansen, I. B., Vela-Avitua, S., Nørstrud, K. S., Aalgaard, M., Braastad, B. O., ... Øverli, Ø. (2014). Frustrative reward omission increases aggressive behaviour of inferior fighters. *Proceedings of the Royal Society of London B: Biological Sciences*, 281, 1–8.
- Waddell, S. (2013). Reinforcement signalling in *Drosophila*; dopamine does it all after all. *Current Opinion in Neurobiology*, 23, 324–329.
- Walter, S., & Heckmann, H.-D. (Eds.). (2003). *Physicalism and mental causation*. Exeter, England: Imprint Academic.
- Wullimann, M. F., & Vernier, P. (2009). Evolution of the brain in fishes. In M. D. Binder, N. Hirokawa, & U. Windhorst (Eds.), *Encyclopedia of neurosciences* (pp. 1318–1326). Berlin: Springer-Verlag.
- Zmigrod, S., & Hommel, B. (2013). Feature integration across multimodal perception and action: A review. *Multisensory Research*, 26(1–2), 143–157.