

# Evolution of Consciousness: Phylogeny, Ontogeny, and Emergence from General Anesthesia

**Author**

RAGHUNATH TIRUVAIPATI

**Date:**

2022-March-15

**Affiliations**

<sup>1</sup> \* Departments of Anesthesiology and Neurosurgery, Neuroscience Graduate Program, University of Michigan Medical School, Ann Arbor, MI 48109;

<sup>2</sup> † Veterans Administration Long Beach Healthcare System, Department of

Anesthesiology and Perioperative Care, Center for the Neurobiology of Learning and Memory, University of California, Irvine, CA 92697

Are animals conscious? If so, when did consciousness evolve? We address these long-standing and essential questions using a modern neuroscientific approach that draws on diverse fields such as consciousness studies, evolutionary neurobiology, animal psychology, and anesthesiology. We propose that the stepwise emergence from general anesthesia can serve as a reproducible model to study the evolution of consciousness across various species and use current data from anesthesiology to shed light on the phylogeny of consciousness. Ultimately, we conclude that the neurobiological structure of the vertebrate central nervous system is evolutionarily ancient and highly conserved across species and that the basic neurophysiologic mechanisms supporting consciousness in humans are found at the earliest points of vertebrate brain evolution. Thus, in

agreement with Darwin's insight and the recent “Cambridge Declaration on Consciousness in Non-Human Animals,” a review of modern scientific data suggests that the differences between species in terms of the ability to experience the world is one of degree and not kind.

Evolutionary biology forms a cornerstone of the life sciences and thus the neurosciences, yet the emergence of consciousness during the timeline of evolution remains opaque. As the theory of evolution began to eclipse both religious explanations and Enlightenment doctrines regarding the singularity of human consciousness, it became clear that consciousness must have a point of emergence during evolution and that point likely occurred before *Homo sapiens*. “How,” Darwin questioned, “does consciousness commence?” His post-*Beagle* research on this question evidently caused him violent headaches. One such headache can be expressed as the 20th century philosophical distinction of phenomenal consciousness and access consciousness (Block, 2007). Phenomenal

consciousness relates solely to subjective experience, whereas access consciousness includes (among other processes) the ability to report such experiences verbally (other distinctions related to consciousness can be found in [Table 3.1](#)). Thus, the scientist looking for objective indices of subjective events is primarily limited to humans manifesting access consciousness, an obstacle in studying the evolution of consciousness antecedent to our species. We could, however, take solace in the dictum that ontogeny recapitulates phylogeny and search for clues in developing humans. Unfortunately, Haeckel's theory of recapitulation is not scientifically sound and, even if applicable in this case, we would still be constrained by the high probability that babies develop phenomenal consciousness before access consciousness. To overcome the limitations in identifying the birth of consciousness, we need a reproducible experimental model in which (i) consciousness emerges from unconsciousness at a discrete and measurable point, (ii) phenomenal consciousness and access consciousness are closely

juxtaposed or collapsed, and (iii) assessment of neural structure and function is possible. In this chapter, we consider top-down and bottom-up approaches to consciousness, nonhuman consciousness, and the emergence of consciousness from general anesthesia as a model for the evolution of subjectivity.

## **TOP-DOWN AND BOTTOM-UP APPROACHES TO CONSCIOUSNESS**

To locate the birth of consciousness on the evolutionary timeline, it will be beneficial to consider the basic neural machinery that is thought to be involved in human consciousness (Crick, 1994; Damasio, 1999; Edelman and Tononi, 2000; Tononi, 2004; Baars, 2005; Denton, 2005; Blumenfeld, 2011). The distinction between phenomenal and access consciousness was noted, but phenomenal consciousness itself reflects the dissociable neurobiological processes of awareness and arousal (Paus, 2000; Schiff and Plum, 2000; Jones, 2003; Laureys, 2005; Lydic and

Baghdoyan, 2005) (Table 3.1). Awareness refers to the content of consciousness (red apple vs. blue sky), whereas arousal refers to brain activation and level of consciousness (alert vs. drowsy vs. asleep vs. anesthetized). A number of current theories about consciousness propose that the cortex is the primary site containing the neural correlates of awareness (Tononi and Edelman, 1998; Van der Werf et al., 2002; Crick and Koch, 2003; Seth et al., 2005; Franks, 2008; Brown EN et al., 2011), whereas midline subcortical brain structures provide ascending arousal influences to the cortex (Van der Werf et al., 2002; Franks, 2008; Brown EN et al., 2011). Thus, we can explore both top-down and bottom-up approaches to consciousness.

## **Top-Down Approach**

Seth et al. (2005) propose three main physiological reasons supporting the importance of the neocortex to the process of consciousness. First, the electroencephalogram of virtually all mammals and birds in the awake state is characterized by

desynchronized, high-frequency, and low-amplitude activity. This pattern changes to one of low-frequency, high-amplitude activity during depressed levels of consciousness such as nonrapid eye movement (NREM) sleep, minimally conscious states, and anesthesia. Thus, a state-dependent change in the electrical firing properties of the neurons across the neocortex varies with the level of arousal and strongly supports the idea that neuronal activity in the brain (and particularly in the neocortex) is a necessary requirement for consciousness (Revonsuo, 2006).

Second, consciousness appears to be linked more specifically with neural activity in the thalamocortical system. In this view, the midline brain structures of brainstem and midbrain are thought to be important for keeping the cortex in an aroused or awake state, whereas the cortical regions are thought to serve as specific cognitive modules contributing to the contents of conscious experience. The idea that certain brain regions are more important than others for generating the contents of

consciousness is further supported by a number of basic neurological facts. For instance, a person could suffer the loss of the cerebellum or large bilateral portions of the medial temporal lobes, including amygdala and hippocampus complex, and would not become unconscious. However, focal damage to specific areas of cortical tissue will change the contents of a person's consciousness in a way that matches the loss of function associated with the specific area damaged. Cortical lesions can thus result in such specific impairments of consciousness that one may no longer be able to speak, perceive color, or identify parts of themselves as their own (Aguirre et al., 1998). Damage to lower midline brain structures, on the other hand, will likely alter the level of consciousness (i.e., arousal) without necessarily changing its contents.

Thalamocortical oscillations have been posited to be of critical importance to consciousness because they help integrate functionally diverse and spatially distinct cognitive modules in the cortex (Saalman et al., 2012; Schmid et al., 2012). The interplay of



segregation and integration is a fundamental focus of the integrated information theory of consciousness (Tononi, 2004, 2012). The capacity of the thalamocortical system to achieve both integration and differentiation is reflected in higher levels of Phi, a proposed metric for consciousness (Tononi, 2004). Phi reflects the amount of information generated by an integrated system beyond the information contained within the components of the system. In principle, this measure captures the emergent property of the system (consciousness) that cannot be causally reduced to individual subsystems (particular brain regions). Phi is predicted to decrease during sleep and seizures; preliminary evidence suggests it also decreases during anesthesia (Lee et al., 2009b), possibly due to impaired long-range coupling of neural spike activity (Lewis et al., 2012). Although the integrated information theory of consciousness has yet to be definitively demonstrated, it is a guiding paradigm that can inform the evolution of consciousness from the network perspective. Creatures with brain network

systems that are capable of generating high values of Phi are more likely to be conscious (Edlund et al., 2011).

Third, widespread brain activity appears correlated with conscious activity. Sensory input spreads quickly from sensory cortex to parietal, temporal, and prefrontal areas (Dehaene et al., 2003). This spread of cortical activity is also associated with recurrent local feedback occurring along the way, followed shortly thereafter by long-range feedback from anterior to posterior structures (Lamme, 2006). These long-range connections are thought to be important for the experiential aspects of consciousness (i.e., awareness) (Singer, 1993) and appear to be preferentially suppressed during general anesthesia (Lewis et al., 2012; Schröter et al., 2012). In particular, there is strong evidence that networks across the frontal and parietal cortices are associated with awareness across multiple sensory modalities (Gaillard et al., 2006; Fahrenfort et al., 2008; Blumenfeld, 2012). The lateral frontoparietal network plays a role in mediating consciousness of

the environment, whereas the medial frontoparietal network plays a role in mediating internal conscious states such as dreaming and internally directed attention (Boly et al., 2007; Denton et al., 2009). It is becoming increasingly clear that the directionality of corticocortical network communication is relevant to conscious processing. Information processing from the caudal to rostral direction (feedforward) is associated with sensory processing that can occur in the absence of consciousness (e.g., general anesthesia, priming) (Imas et al., 2005; Gaillard et al., 2007). In contrast, information processing in the rostral-to-caudal direction (feedback or cortical reafference) is thought to be associated with experience itself and is preferentially inhibited by general anesthetics (Imas et al., 2005; Lee et al., 2009a; Ku et al., 2011).

The neocortical view of consciousness originates, in part, from early morphologic examination of brain differences across species that suggested the capacities of consciousness increased as brains evolved from more primitive reptilian organization,

to mammalian (or, with a limbic system, paleomammalian), and then neo-mammalian organization, characterized by an intricately folded neocortex. This conceptualization of brain evolution occurring in stages during which more “advanced” brains—along with their expanded behavioral repertoire—were built on the structure of earlier forms was popularized by Maclean as “the triune brain” (Maclean, 1990). Importantly, this view of brain evolution is now largely considered erroneous (Emery and Clayton, 2005; Jarvis et al., 2005). It did offer an easy conceptualization for relating brain structure with function and suggested evolutionary time points for when various behaviors would have emerged. Newer findings, however, strongly refute the model of a triune brain, especially the concept of a later-developing neocortex (Fig. 3.1) (Emery and Clayton, 2005). As it turns out, a precursor of the neocortex was actually present in the earliest evolving vertebrates, a claim based on some aspects of connectivity and homology of early transcription factor expression (Striedter, 2005). The basic

structural pattern of a brainstem, midbrain, and forebrain did not need to be completely reinvented as each new species emerged. Rather, as various ecological niches were exploited by various creatures, those brain regions best suited for enhancing survival in the local environment were emphasized for further development (Emery and Clayton, 2005).

## **Bottom-Up Approach**

Since the discovery of the ascending reticular activating system by Moruzzi and Magoun in the late 1940s (Moruzzi and Magoun, 1949), the fundamental and permissive role for arousal in generating conscious states has been well established. It is now clear how a number of specific nuclei and specific cell types within the brainstem, midbrain, basal forebrain, and diencephalon send long-range axons throughout the cortex to enhance arousal and generate a neurochemical environment in the cortex that is capable of supporting consciousness (Lydic and Baghdoyan, 2005). The

role of arousal in regulating overall levels of consciousness is clearly established in connection with depressed levels of consciousness as during sleep or coma (Laureys et al., 2004). How arousal machinery interacts with consciousness during more subtle cognitive and behavioral manipulations is the subject of much current research (Cahill and Alkire, 2003; Coull et al., 2004; Devilbiss et al., 2006). However, through the study of arousal as it relates to emotion (Paus, 2001; McGaugh, 2005), another link is made that puts some of Darwin's later investigations into a more modern light.

Darwin spent the later years of his career investigating the similarities and differences associated with emotional expression in man and animals (Darwin, 1872). He reasoned that if animals show emotion through behavioral expression, and man is an animal, then the behavioral expression of emotion in man must share a similar neurobiologic evolution with the other animals capable of expressing similar emotions. Put another way, years before behaviorism dominated neuroscience, Darwin

saw how commonalities in emotional expression across species likely reflected the occurrence of similar underlying states of mind that only made sense within a theory of evolution. Modern study into the emotional lives of animals now reveals how fundamentally similar the brain structures are that support affective reactions in animals and humans (Panksepp, 2011).

Consciousness may not have emerged from the need to make an internal representation of the outside world, but rather as an extension of very basic primitive or primordial emotional influences. Such emotional influences would generate an arousal response in an organism and prepare its brain for action. This hypothesis is well elaborated by Denton (2005) in his book on primordial emotions. It posits that the most basic instincts, such as thirst, hunger for air, hunger for salt and food, and the desire for sex are the defining starting points for evolving a conscious brain (Denton et al., 2009). This idea holds within it the concept of intention, desire, and action selection, where the basic intention of a

movement is in the service of fulfilling a desire. As noted by Darwin (1872), “So strongly are our intentions and movements associated together, that if we eagerly wish an object to move in any direction, we can hardly avoid moving our bodies in the same direction, although we may be perfectly aware that this can have no influence.”

The basic behavior of an organism is driven by a fundamental physiologic need to maintain homeostasis. Those cells and systems used for monitoring and maintaining the internal milieu are referred to as interoceptors (Denton, 2005). The basic behaviors driving homeostasis are evident as far back as the first multicellular organisms that needed a vascular system to provide nutrients to those cells no longer exposed directly to the environment. Creatures that could meet their basic homeostatic needs are the ones that survived; those that did not suffered extinction. The brain structures needed for generating arousal and primitive emotional responses are generally located in the brainstem, midbrain, and limbic system and are as



old as the vertebrate radiation itself (Jing et al., 2009).

Recent work on the lamprey, a jawless fish whose common ancestor forms the basis for all vertebrates more than 500 million years ago, has revealed just how ancient the neuroanatomy and neurochemistry needed for action selection is. Findings reveal that the lamprey's behavioral motor output system shows similar complexity to higher-level vertebrates who are capable of regulating behavior by both direct and indirect motor output pathways from the basal ganglia (Stephenson-Jones et al., 2012). In other words, the lamprey is capable of both selecting a motor output to perform and at the same time inhibiting the performance of other possible outputs. Thus, they are capable of making a choice depending on the situation with which they are confronted. This “reduction of uncertainty” (a classic definition of information) through action selection may be the precursor to the highly informative states of consciousness characteristic of humans. We address

the relationship of motoric behavior and consciousness in the next section.

More complex neocortical abilities offered a survival advantage to more complex brains by giving organisms a larger grasp of their surroundings, but these systems developed over time and used sensory information from the environment or exteroceptors (Denton, 2005). Denton illustrates his point with the example of a dehydrated frog placed next to a water source in the sun. The frog has only a limited capacity in its visual system and when placed next to a source of water it will usually die without moving, unless it stumbles on the water by accident. If, by chance, the frog finds the water, it will drink, which suggests functioning interoceptors. In contrast, the more highly evolved visual system of the lizard allows that creature to see the water and immediately drink, suggesting that its more evolved brain more successfully couples its exteroceptor-mediated perceptions with its interoceptor-mediated needs. This coupling of an internally based need system with an externally based situational awareness

system is likely the foundation for the emergence of consciousness, and it closely corresponds to the mental machinery seen in humans for generating awareness and arousal.

The brainstem arousal centers are, for the most part, juxtaposed with the sensory motor inputs and outputs of the cranial nerves that supply the head and neck with its ability to orient a creature to its environment and provide a stable platform for sensing its surroundings. The motor output of the cranial nerves is fundamentally linked with the expression of emotion in essentially all vertebrates, and this likely emanates from the oldest of the predator–prey relationships. In essence, an open mouth signifies a meal for the predator, and if the hunt is successful, it would likely be associated with internal sensations of goal/task completion that would serve to fulfill a basic need for food in the predator. This goal completion/desire fulfillment would likely have positive reinforcing value for an organism and might easily be hypothesized to lead to internal states comparable to a sense of pleasure

(Panksepp, 2011). For the prey, an open mouth heading toward it would certainly be cause for alarm, prompting an immediate escape response that, if successful, might be associated with an internal state of heightened arousal and fear (Panksepp, 2011). Thus, the most basic emotions and arousal states are associated with internal feedback networks that serve to guide an organism's behavior toward its best possible situational outcome. This emotional arousal machinery underlies essentially all behavioral choices in the vertebrate brain.

## **CONSCIOUSNESS IN NONHUMAN SPECIES**

If consciousness evolved in conjunction with cephalad development of the central nervous system, then its emergence should, in principle, be identifiable at a discrete point on the tree of evolution. Darwin reasoned that the cognitive differences between species must be one of degree and not kind. This conclusion is consistent with the recent Cambridge declaration that occurred on July 7, 2012, at the first annual Francis Crick memorial

conference on consciousness. A group of prominent scientists formally declared in a document entitled the “Cambridge Declaration on Consciousness in Non-Human Animals” that the neurobiological structures needed to support consciousness are not uniquely human (Low, 2012). This declaration essentially states that the capacity for consciousness likely emerged very early in evolutionary terms, and those processes that support consciousness in humans are likely characteristic of many living creatures. In fact, according to the declaration, based on a number of considerations from comparative brain anatomy and current knowledge about the neurobiology of consciousness, it would seem almost certain that some form of consciousness is present in all mammals and could have emerged on the evolutionary timeline at the branch point of amniotes.

However, long before the Cambridge declaration, some thinkers expressed serious concerns about attributing higher levels of consciousness to all life. Indeed, René Descartes, often considered the

philosophical father of the mind–body relationship, questioned whether a conscious self arose in the animal kingdom. He avoided ascribing a conscious self to a particular animal because by doing so he recognized that he might be compelled to ascribe a conscious self to all animals. This all-or-none approach did not reflect an evolutionary theory perspective, which raised the possibility of a conscious continuum. This continuum, however, also introduces difficulty. As pointed out by Gallup (1985), in discussing the emergence of consciousness in animals,

Where do we draw the line? On the one hand, we could decide not to draw a line. This would presume that all living things are sentient, conscious, and mindful. While the data are admittedly incomplete, the issue should be taken seriously. Life on this planet consists of several million different species. Most are microorganisms, plants, and insects. I doubt that there is a paramecium, a rosebush, or a termite alive today which is aware of its own

existence or has the capacity to become the object of its own attention.

With Gallup's statement, we begin to see the need for clarity in how or why we associate certain behaviors with subjective experience and the need for some operational definitions of the “consciousness” being studied.

To identify the origin of sentience along an evolutionary timeline, it is beneficial to consider a common element that might link consciousness across species, rather than focusing on the ostensibly unique qualities of human experience such as self-reflection. Furthermore, this common element should likely relate to a goal-directed behavior or response pattern that confers a survival advantage in a given environment. In line with philosophers such as Merleau-Ponty and neuroscientists such as Rudolfo Llinás and György Buzsáki, we support motility (also referred to in this context as motricity) as a strong candidate for the evolutionary anlage of consciousness (Cotterill, 2001; Goodrich, 2010).

Consider, for example, the unicellular paramecium, which is covered with several thousand cilia. These cilia can serve both the function of sensing environmental stimuli and initiating motility responses (e.g., attraction, avoidance) based on the nature of those stimuli. This preneural example of a single structure (i.e., cilia and their coordinated activity) mediating both sensation and response is intriguing but does not establish the primacy of motility as a kernel of consciousness. Perhaps a more compelling case is that of the sessile sea squirt, which possesses neural structures only transiently during a larval stage (Llinás, 2001). Neural ganglia and primordial sensory processing allow the sea squirt to find a suitable local environment and underwater surface for attachment. However, after this goal is achieved the neural tissue is digested, suggesting a role related exclusively to movement. Although it is unlikely that paramecia and sea squirts have phenomenal experience, these early examples of sensation in the service of motility lead us to start the search for the neurobiological origins of



consciousness in phylogenetically conserved structures.

## **WHAT IS THE NEURAL “CORE” OF CONSCIOUSNESS**

To identify which aspects of the mental machinery should be the focus of evolutionary consideration for consciousness, we need to identify the neural correlates of the most primitive core of human consciousness. The still emerging field of consciousness studies has been dominated in the last decade by a search for the neural correlates of consciousness, which have been defined as the specific and minimally adequate brain states that correspond to states of consciousness (Tononi and Koch, 2008). However, studies of the content of consciousness (e.g., the awareness of a red rose placed in your visual field) already assume a conscious brain; thus, the neural activity or structure identified in these paradigms corresponds to a specific content within a preexisting consciousness (Hohwy, 2009). Studying the level of consciousness

(e.g., arousal states) is also beset with difficulties. For example, the transition from a fully conscious to unconscious state will inform us primarily of correlates required for the full spectrum of waking human consciousness rather than the minimal or core requirements. Furthermore, we must also grapple with how to identify the true correlate (or substrate) of consciousness vs. neural prerequisites or neural consequences of consciousness (Aru et al., 2012; de Graaf et al., 2012).

To address some of these difficulties, a recent study explored the neural correlates of the primitive form of consciousness that arises during emergence from general anesthesia (Långsjö et al., 2012). With anesthesia, the level of consciousness can be manipulated as an experimental variable, and the resultant changes in brain activity can then be determined with various neuroimaging and neurophysiologic techniques. Numerous studies have now examined what happens to brain activity when consciousness is removed by anesthesia (Alkire et al., 2008; Brown et al., 2010); however, fewer

studies have investigated the correlates of consciousness associated with its return following a period of anesthesia (Friedman et al., 2010; Lee U et al., 2011; Xie et al., 2011; Bonhomme et al., 2012). In one recent study of healthy male volunteers, positron emission tomography (PET) was used to investigate the neural correlates of the recovery of consciousness from the i.v. anesthetics propofol and dexmedetomidine (Långsjö et al., 2012). The order of the state transition is important because the investigation of consciousness to unconsciousness may yield a variety of nonspecific deactivations due to drug effects that do not necessarily play a core role in consciousness. The emergence of consciousness (as judged by the return of a response to command) was correlated primarily with activity of the brainstem (locus coeruleus), hypothalamus, thalamus, and anterior cingulate (medial prefrontal area). Surprisingly, there was limited neocortical involvement that correlated with this primitive form of consciousness. Frontal–parietal connectivity appeared to be the key cortical response, which has

been confirmed by studies of consciousness and anesthesia using electroencephalography (Lee U et al., 2011). Similar findings were seen in another imaging study investigating the emergence of consciousness from sleep (Balkin et al., 2002). In the sleep study, midline arousal structures of the thalamus and brainstem also recovered function well before cortical connectivity resumed. Thus, the core of human consciousness appears to be associated primarily with phylogenetically ancient structures mediating arousal and activated by primitive emotions (Liotti et al., 2001; Denton et al., 2009), in conjunction with limited connectivity patterns in frontal–parietal networks (Merker, 2007; Brusseau, 2008) (Fig. 3.2).

The emergence from general anesthesia may be of particular interest to evolutionary biology, as it is observed clinically to progress from primitive homeostatic functions (such as breathing) to evidence of arousal (such as responsiveness to pain or eye opening) to consciousness of the environment (as evidenced by the ability to follow a command) to

higher cognitive function. Unlike the emergence of consciousness over millions of years in phylogeny or months during the gestational period in ontogeny, the emergence of consciousness from the anesthetized state is a reproducible model system that can be observed in real time over the course of hours. Multimodal investigation using neuroimaging and neurophysiology, in conjunction with clinical observation and cognitive evaluation, could uncover key shifts of neural activation or network organization that support conscious processing. For example, high-density electroencephalography could be used during recovery from general anesthesia to measure Phi to help delineate in humans the threshold for emerging consciousness. Such a threshold could then be compared with other species in the waking state to determine the relative value with reference to the neural core of human consciousness. This approach could be applied to any number of brain network properties, as assessed quantitatively through graph theoretical methods (Stam and van Straaten, 2012b).

Network approaches—which have broad applicability in mathematics, biology, computer science, and sociology—might be particularly attractive to test hypotheses across species, where functionally similar cognitive systems may arise from neurobiologically distinct structures. For example, the mammalian cortex and avian pallium are histologically distinct (Table 3.2) (Butler, 2008), but may subserve similar network functions that can be quantitatively assessed and compared with human findings. General anesthesia represents a way of turning back the evolutionary clock of cognitive function in humans and—depending on the “depth” and length of anesthetic exposure—allows investigators to observe the return of neural function in a way that could recapitulate phylogeny. Although not without difficulties (including the contamination of access consciousness, because language is involved in assessing return of consciousness after anesthesia), advantages of emergence from anesthesia as a model system for the evolution of consciousness include convenience, reproducibility,

real-time observation, possibility of subjective report of experiences (with experiments in humans), and amenability to neuroscientific investigation across multiple species.

## **WHEN DOES CONSCIOUSNESS OF THE WORLD ARISE**

The recent experiments with general anesthesia in humans suggest that phylogenetically ancient structures in the brainstem and diencephalon—with only limited neocortical involvement—are sufficient to support primitive consciousness. Where, then, does consciousness arise on the evolutionary timeline? One might be tempted to conclude that consciousness commenced as our mammalian ancestors evolved just beyond reptiles and their predominantly subcortical brains. However, paleontological findings suggest that the synapsid line that gave rise to mammals and the sauropsid line that gave rise to reptiles and birds both diverged from the primitive anapsid line at a single point ~315 million years ago (Warren et al., 2008). Furthermore,

there is significant evidence that avian species are capable of higher cognition and even consciousness itself (Butler and Cotterill, 2006). For example, birds demonstrate evidence of explicit episodic recall (i.e., conscious memory of an event) (Emery and Clayton, 2004) and theory of mind (i.e., attribution of subjective mental events to another being) (Emery and Clayton, 2001). Thus, it would be misguided to try to identify a single point at which consciousness emerged because evidence suggests that consciousness evolved along two independent lineages. As pointed out by Butler et al. (2005), birds and mammals share a number of homologous traits despite this evolutionary divergence, including a dramatic increase in their brain–body ratios (compared with reptiles), homeothermy, extended parental care of offspring, habitual bipedalism, distinct sleep stages, and complex social interactions. The neurobiology also reflects homologous advances, particularly in the mammalian neocortex and the avian pallium (Table 3.2). These advances include the emergence of recurrent or feedback



processing, which is not found in reptiles. Thus, both birds and early mammals are equipped with a neural substrate consistent with conscious processing: phylogenetically conserved brainstem, diencephalic structures such as thalamus and hypothalamus, and association neocortex (or equivalent) capable of recurrent processing. All of these structures appear to play a role as the neural core for primitive consciousness in humans, as evidenced by experiments with general anesthesia.

The critical role of subcortical structures in consciousness has been further argued based on clinical observations of hydranencephalic children, who are essentially devoid of neocortex and yet who still demonstrate some behavioral signs of consciousness (Merker, 2007). Others have attempted to link the arousal-related components of consciousness with the contents of consciousness by placing emphasis on the dynamic recurrent activity that occurs in the thalamus or through the thalamic reticular nucleus when consciousness is present (Min, 2010; Ward, 2011). As such, the PET study

showing that the emergence of consciousness is correlated with increased activity in “primitive” brain regions may reflect an arousal-related response to the test stimulus itself rather than a direct awareness of the stimuli that are occurring in the thalamus. In either event, the data clearly show that the neurocircuitry associated with arousal is fundamental to consciousness. A further recent study investigating long-term memory encoding also imaged the neural correlates of subjective emotional arousal. As shown in Fig. 3.2, the neural correlates for awareness of subjective arousal induced by viewing of emotional stimuli involve the same midbrain arousal structures seen with activation of primordial emotions (Hayama et al., 2012).

Regarding ontogeny of *H. sapiens*, peripheral sensory receptors are thought to be present from 20 weeks of gestation in utero. The developmental anlage of the thalamus is present from around day 22 or 23 postconception, and thalamocortical connections are thought to be formed by 26 weeks of gestation (Brusseau, 2008). Around the same time of

gestation (25–29 weeks), electrical activity from the cerebral hemispheres shifts from an isolated to a more continuous pattern, with sleep-wake distinctions appreciable from 30 weeks of gestation. Thus, both the structural and functional prerequisites for consciousness are in place by the third trimester, with implications for the experience of pain during in utero or neonatal surgery. It is of interest to note that the third trimester of human development is thought to be the period in which the maximal proportion of time spent in REM sleep occurs across the lifespan (Birnholz, 1981). This finding supports the ontogenetic theory of REM sleep as a process of internally driven neuronal activation that prepares the developing cortex for the coming influx of sensory stimuli at birth. The theory of REM sleep as a form of protoconsciousness has recently undergone further elaboration (Hobson, 2009).

## **WHEN DOES CONSCIOUSNESS OF THE SELF ARISE**

One component of consciousness that seems linked to higher cognitive abilities is awareness of the self rather than simply awareness of the environment. One way to test for this possibility is to use what is known as the mirror self-recognition (MSR) test (Keenan et al., 2003). In 1970, Gallup found that chimpanzees, but not monkeys, were able to pass the MSR test (Gallup, 1970). This test presupposes that the experimental subject has sufficient cognitive ability to be aware of itself as an entity that is distinct from another conspecific. This ability then defines one form of consciousness (i.e., the ability to have awareness of one's own awareness or self). In Gallup's well-controlled experiment, the animals were first allowed ample time with mirror exposure to allow social responses to their reflected images to diminish greatly. The number of social responses and the number of self-directed responses were measured before the animals had a mark covertly placed on their forehead or ear while they were briefly anesthetized. The animals were then allowed to recover from anesthesia, and some hours later a

mirror was reintroduced. On seeing themselves in the mirror, the marked chimpanzees—but not the marked monkeys—exhibited mark-directed responses by spending time investigating the area of the mark and examining their fingers after touching the mark. The findings led Gallup to conclude “insofar as self-recognition of one's mirror image implies a concept of self, these data would seem to qualify as the first experimental demonstration of a self-concept in a subhuman form.” Regarding the difference between chimpanzee and monkey, he further concluded, “Our data suggest that we may have found a qualitative psychological difference among primates, and that the capacity for self-recognition may not extend below man and the great apes.” The distinction among primates suggests that the qualitative nature of the conscious experience varies greatly across species and the introspective nature of human consciousness may be evolutionarily quite rare. The MSR test has now been used to examine the ability of other species to show evidence of self-awareness. Primates that have

passed the MSR test include chimpanzees, orangutans, and bonobos. The case for the gorilla is equivocal with mostly negative findings; several studies have suggested that more socialized gorillas might be able to pass the test. Humans begin to develop a sense of self and pass the MSR test starting around 18 months of age, and by 24–36 months, almost all western children will show a positive MSR response (Amsterdam, 1972). The distinction between great apes and monkeys would seem to provide a clear demarcation in the capacity for consciousness between species. Numerous studies have supported this demarcation, with multiple failed attempts to detect self-awareness in monkeys, despite one recent report to the contrary (Rajala et al., 2010). However, a number of methodological concerns limit enthusiasm for the one contrary study, and overall the data continue to suggest that macaques do not evidence MSR behavior (Anderson and Gallup, 2011). In evolutionary terms, if objective evidence of self-awareness can be taken as evidence for

consciousness, then consciousness as it occurs in the primate with their more fully developed cortex may have evolved ~5 million years ago, at around the time when great apes split off from the lesser apes.

Mirror self-recognition may not be limited to the relatively big-brained great apes. More recent work with other big-brained creatures suggests the possibility that dolphins, and at least one African elephant, may also be capable of this response (Delfour and Marten, 2001; Reiss and Marino, 2001; Plotnik et al., 2006). As apes, elephants, and cetaceans have a very remote common ancestor, these findings would seem to suggest that the mental machinery prerequisite for self-awareness must be at least as old as the development of the placental divide in mammals (Wildman et al., 2007).

However, we may be able to take this idea on another path in evolutionary time. As noted, the cognitive abilities of some birds are now thought to be comparable to the abilities of some primates (Emery and Clayton, 2004). Evidence suggests that the brain development of the bird, which evolved on

a different path from mammals, still has a conceptually similar thalamocortical structure that can be delineated (Jarvis et al., 2005). The cognitive abilities of various birds seem to correlate with the relative size of the analogous avian prefrontal cortex. Indeed, the crow-like Corvidae (crows, ravens, magpies, rooks, jackdaws, and jays) appear to have the most advanced behavioral repertoire, as well as the largest prefrontal cortex (pallium) (Emery, 2006). Importantly, a recent report shows magpies (having a relatively large prefrontal cortex) exhibit behavior consistent with MSR (Prior et al., 2008). This finding, coupled with the current understanding of avian neuroanatomy and its well-developed thalamocortical structure, suggests that the foundations required for both consciousness of the world and consciousness of the self may have formed as early as the amniote radiation (Warren et al., 2008).

From a cognitive perspective, the meaning of self-awareness behaviors in a mirror remains somewhat controversial (Morin, 2011). Some argue that the



mirror behavior could be more easily explained by simple knowledge of one's body. The neurobiology of having a body sense is something that is highly linked with a sense of consciousness (Damasio, 2003). Perhaps, as stated by Morin (2011), “all an organism requires to self-recognize is a mental representation of its own physical self; the organism matches the kinaesthetic representation of the body with the image seen in the mirror and infers that ‘it's me.’” A number of other arguments against overinterpreting MSR have been made, yet despite these relevant concerns, from an evolutionary point of view the presence or absence of a MSR response is at least a starting point for considering what having such a response might mean as a basis for the evolution of consciousness. The MSR response allows one to question what is functionally and structurally different about brains that can self-recognize vs. those that cannot.

**WHY IS HUMAN CONSCIOUSNESS UNIQUE**

We have argued that the brainstem, diencephalon, and limited association cortex capable of recurrent processing is consistent with a core or primitive consciousness. However, what accounts for the richness of human experience in contrast to those of early mammals or birds? Drawing on the integrated information theory of consciousness, the evolution of more complex brain networks capable of synthesizing the outputs of more functionally diverse modules would result in a higher capacity for consciousness. Indeed, integration of information appears to correlate positively with fitness in artificial agents (animats) ([Edlund et al., 2011](#)). It is unknown in biology, however, whether it is the level or quality of consciousness that differs across species. Although *H. sapiens* may have more advanced cognition, it is difficult to imagine that a sedentary human has a higher level of consciousness than a highly alert beast in pursuit of prey; the richness of conscious experience may be what differs. Alternatively, it is possible that advanced symbolic processing in human cognition eclipses the

subjective characteristics of experience. In other words, cognition may be potentially opposed to phenomenal consciousness. Despite these considerations, human consciousness—especially the capacity for self-consciousness and reflection/projection in time—seems unique. Although evidence suggests that the core of consciousness is rooted in phylogenetically older structures such as the brainstem and diencephalon (Merker, 2007), the evolution of that which is particular to human consciousness may be more closely associated with the development of the frontal cortex. The relative size of the frontal lobes with respect to the total neocortex is roughly the same in modern humans and great apes, but richer interconnectivity might account for advanced cognition in *H. sapiens* (Semendeferi et al., 2002). In particular, directed anterior-to-posterior connectivity has been associated with conscious perception and is dominant in humans (Ku et al., 2011) but not in rodents (Imas et al., 2005; Nieder, 2009) (Fig. 3.3). It has been suggested that the afferent information

from the periphery converging at the hub of the posterior parietal cortex becomes, with the expansion of the frontal cortex, dominated by a strong anterior-to-posterior reafference (Noack, 2012). Indeed, a recent neural mass model based on structural connectivity data from diffusion tensor imaging in humans predicts an information flow from the frontal to the posterior parietal cortex (Stam and van Straaten, 2012a). In essence, this information flow reversal suggests that human consciousness is more defined by internal dynamics than external stimuli. This level of information flow reversal may help explain, in part, those animals capable of an MSR response. According to one theory, human consciousness is a closed system or “oneiric” (dreamlike) state that is simply modulated by environmental input (Llinás and Ribary, 1993), a theory consistent with REM sleep as a building block for human consciousness. The relative independence from environmental determination of conscious content would potentially permit a greater diversity or richness of experience in comparison

with species without dominance of anterior-to-posterior flow. This independence would also facilitate the projection and simulation associated with future plans, of clear relevance to survival. It is important to note, however, that the role of information flow in consciousness is unclear at this time and requires further neuroscientific investigation.

## **CONCLUSION**

The emergence of consciousness on the evolutionary timeline has been scientifically considered at least since the time of Darwin. The emergence of consciousness from the anesthetized state may provide a practical and reproducible model for characterizing the real-time evolution of the core neural correlates required for consciousness of the world and of the self. Using recent data from general anesthesia in humans, we suggest that the arousal centers in the brainstem and diencephalon—in conjunction with even limited neocortical connectivity and recurrent processing—can result in

primitive phenomenal consciousness. By “reverse engineering,” we postulate that early mammals and birds possessing these structures (or their equivalents) are capable of phenomenal consciousness. However, the increased complexity of networks and a functionally dominant prefrontal cortex in the brain of *H. sapiens* likely accounts for the unique richness of the human experience.

## ACKNOWLEDGMENTS

G.A.M. is supported by National Institutes of Health Grant 1R01 [GM098578](#) and the James S. McDonnell Foundation.

## Figures

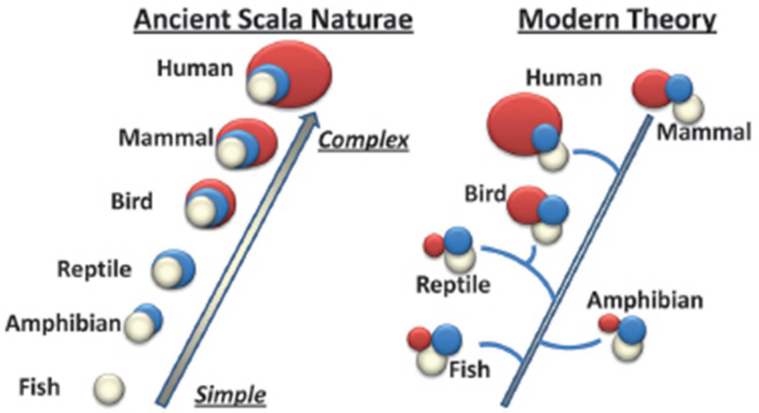
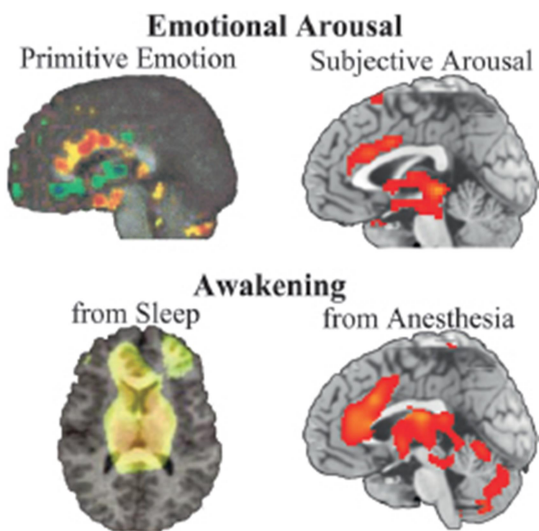


FIGURE 3.1

Theories of brain evolution. Ancient brain structure evolution theory of Scala Naturae showing brain development proceeding from simple to more complicated with the addition of new brain regions as evolution progressed. This erroneous view is compared with a modern understanding of brain structure evolution that reveals a basic common structure evolved in the vertebrate brain and various regions expanded to accommodate each specific animal's needs. Modified from Emery and Clayton (2005) with permission from Elsevier. [NOTE: Figure can be viewed in color in the PDF version of this volume on the National Academies Press website, [www.nap.edu/catalog.php?record\\_id18573](http://www.nap.edu/catalog.php?record_id18573).]



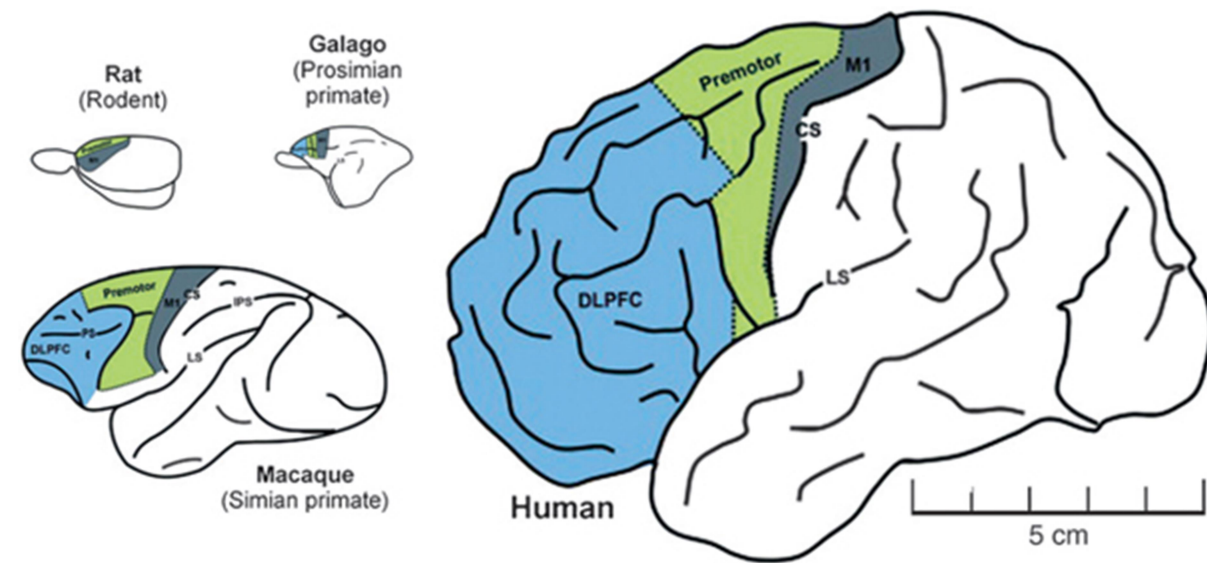
## FIGURE 3.2 Brain structures functionally related to primitive emotional arousal and the return of consciousness following sleep or anesthesia

The primitive emotional response of air hunger shows activations in brainstem and anterior cingulate regions; thalamic changes are also seen ([Liotti et al., 2001](#)). Subjective emotional arousal activates similar regions in an event-related functional MRI study of picture viewing. Reproduced with permission from [Hayama et al. \(2012\)](#). Midline thalamic and anterior cingulate arousal is seen with PET neuroimaging when consciousness first reemerges following sleep or anesthesia. Reproduced with permission from [Balkin et al. \(2002\)](#) and [Långsjö et al. \(2012\)](#). A common brainstem, thalamic, cingulate neuroanatomy associated with conscious brain activity is seen. Images used with permission.

[NOTE: Figure can be viewed in color in the PDF version of this volume on the National Academies Press



website, [www.nap.edu/catalog.php?record\\_id18573](http://www.nap.edu/catalog.php?record_id18573).  
]



### **FIGURE 3.3 Schematic showing relative size of frontal lobe across different species and the potential capacity for anterior-posterior information flow**

The DLPFC (blue) areas represent the prefrontal cortex, and the schematic shows how the prefrontal cortex proportionally increases in size with increasing brain size across species. Relative brain size is scaled to the human brain. Modified from Nieder (2009) with permission from Elsevier. [NOTE: Figure can be viewed in color in the PDF version of this volume on the National Academies

Press

website, [www.nap.edu/catalog.php?record\\_id18573](http://www.nap.edu/catalog.php?record_id18573).

]

## Tables

### TABLE 3.1 Definitions Relevant to Consciousness

| Terms                                  | Explanation   |
|--|---|
| Easy vs. hard problem of consciousness | This distinction was drawn by philosopher David Chalmers. “Easy” problems of consciousness (which are easy in principle only) include understanding the neural basis of feature detection, integration, verbal report, etc. The hard problem is the problem of experience; even if we understand everything about neural function, it is not clear how that would explain |

| <b>Terms</b>              | <b>Explanation</b>  |
|---------------------------|---|
|                           | subjectivity.   |
| Awareness                 | Cognitive neuroscientists and philosophers use the term “awareness” to mean only subjective experience. In clinical anesthesiology, the term awareness is (inaccurately) used to include both consciousness and explicit episodic memory. |
| Wakefulness vs. awareness | Wakefulness refers to brain arousal, which can be manifest by sleep-wake cycles and can occur even in pathologic conditions of unconsciousness such as vegetative states. Thus, being awake is dissociable from being aware.              |
| Phenomenal vs.            | Phenomenal consciousness is   |

| <b>Terms</b>                        | <b>Explanation</b>   |
|-------------------------------------|--|
| access consciousness                | subjective experience itself, whereas access consciousness is that which is available to other cognitive processes (such as working memory or verbal report).                              |
| External vs. internal consciousness | External consciousness is the experience of environmental stimuli (such as the sound of an orchestra), whereas internal consciousness is an endogenous experience (such as a dream state). |
| Consciousness vs. responsiveness    | An individual may fully experience a stimulus (such as the command “Open your eyes!”) but not be able to respond (as when a patient is paralyzed but conscious during                      |

| <b>Terms</b>  | <b>Explanation</b>   |
|---|--|
|   | surgery).  |
| Levels of consciousness vs. contents of consciousness | Levels of consciousness include distinctions such as alert vs. drowsy vs. anesthetized, whereas the contents of consciousness refer to particular phenomenal aspects such as a red rose vs. a blue ball. |

**TABLE 3.2 Comparison of Neocortex and Pallium with Respect to Requirements for Cell Assemblies**

| <b>Requirements for Hebbian Cell Assembly</b> | <b>Structure of Mammalian Neocortex</b> | <b>Structure of Avian Pallium</b> |
|---|---|-----------------------------------|
| Many neurons of the same kind                 | About 85% pyramidal cells               | High number of multipolar cells   |
| Connections with                              | Most synapses                           | Many                              |

| <b>Requirements for Hebbian Cell Assembly</b>                     | <b>Structure of Mammalian Neocortex</b>       | <b>Structure of Avian Pallium</b>    |
|---|---|--------------------------------------|
| each other  | are between pyramidal cells                   | synapses between multipolar cells    |
| Excitatory connections  | About 90% of synapses are type 1 (excitatory) | Many synapses excitatory             |
| Modifiable connections  | About 75% of synapses are on spines           | Dendrites are densely spiny          |
| Individual neurons connected to as many other neurons as possible | About 8,000 synapses per neuron               | Many synapses per neuron             |
| Distant connections across the network                            | Large amount of white matter                  | Axons more interspersed with neurons |

