Controversy in Consciousness: Is the Brainstem Sufficient for Sentience?

Shadia Kawkabani  
*University of Akron, sk129@zips.uakron.edu*

Kevin P. Kaut  
*University of Akron, kpk@uakron.edu*

Follow this and additional works at: [https://ideaexchange.uakron.edu/jnbp](https://ideaexchange.uakron.edu/jnbp)

Part of the Bioethics and Medical Ethics Commons, Nervous System Diseases Commons, and the Other Mental and Social Health Commons

Please take a moment to share how this work helps you through this survey. Your feedback will be important as we plan further development of our repository.

Recommended Citation

Available at: [https://ideaexchange.uakron.edu/jnbp/vol1/iss1/2](https://ideaexchange.uakron.edu/jnbp/vol1/iss1/2)

This Article is brought to you for free and open access by IdeaExchange@UAkron, the institutional repository of The University of Akron in Akron, Ohio, USA. It has been accepted for inclusion in Journal of Neuropsychology and Behavioral Processes by an authorized administrator of IdeaExchange@UAkron. For more information, please contact mjon@uakron.edu, uapress@uakron.edu.
Controversy in Consciousness: Is the Brainstem Sufficient for Sentience?

Shadia Kawkabani and Kevin P. Kaut
Department of Psychology, The University of Akron, Akron, Ohio

Abstract

Historically, scientists and physicians have taken a corticocentric view of consciousness, emphasizing the need for a cortex in producing the conscious experience. The preserved consciousness observed in hydranencephalic children and decorticated rats suggests that some form of consciousness may be produced by a subcortical network. The brainstem, a phylogenetically conserved brain system, could serve as the major integrative network to produce this form of consciousness—referred to as ‘affective consciousness’, the evolutionary antecedent to the reflective consciousness allowing humans to reflect on experiences. The functional integration of the brainstem with the amygdala, motor network, and other subcortical structures provides the architecture necessary to support an affective state of consciousness by which instinctual-emotional goal-directed behavior is produced. This subcortical system operates by what Merker (2007) calls the selection triangle—an interface between bodily actions (action selection), the world (target selection), and personal motivation—thus producing action through integration. Accordingly, a form of consciousness may persist in the absence of the cortex, such as that observed in the persistent vegetative state (PVS). As such, it is helpful to distinguish between affective and reflective consciousness, inasmuch as this distinction would have ethical implications in treatment of those with disorders of consciousness.

Jahi McMath was just 13 years old when she became the unfortunate focus of a national debate regarding the legal definition of death. McMath had undergone a routine adeno-tonsillectomy, uvulopalatopharyngioplasty, and submucous resection of the bilateral inferior turbinates on December 9, 2013 for the treatment of her obstructive sleep apnea (Luce, 2015). While recovering in the pediatric ICU, McMath began bleeding from her nose and mouth then suffered cardiopulmonary arrest before circulation was restored via cardiopulmonary resuscitation and subsequent placement on a mechanical ventilator and IV fluids (Luce, 2015). McMath never regained consciousness and on December 12, 2013, she was declared brain dead by the physicians at Oakland Children’s Hospital and Research Center (Luce, 2015).

By California state law and under the Uniform Determination of Death Act (UDDA), McMath was, by all accounts, considered dead, and on December 15, 2013, her family was notified that she would be removed from the ventilator, which would ultimately culminate in somatic death (Luce, 2015). McMath’s family adamantly rejected a determination of death by neurologic criteria refusing to believe that Jahi was truly dead, but under California state law, they had no legal precedence or grounds for keeping Jahi on life support (Luce, 2015). The family filed a motion in which they argued that California Health and Safety Code Section 7180 (the state’s UDDA, which defines death as permanent whole brain death) allowing death to be determined by neurologic criteria, violated their religious freedom to believe that Jahi was still alive (Luce, 2015). The court ruled that Jahi’s determination of death was legitimate and on January 5, 2014, the county coroner issued a death certificate with time of death dated December 12, 2013 – when she was initially declared dead (Luce,
The family was able to transfer Jahi to a private care facility until she could be transferred to a hospital in New Jersey, one of only two states in the country that allows a family to reject a determination of death by neurologic criteria on the basis of their religious freedom (Luce, 2015).

On September 30, 2014, the family’s lawyer requested a hearing to present new evidence in support of the family’s claim that Jahi was still alive (Luce, 2015). The evidence presented included videos of Jahi moving her hands and feet, seemingly in response to her mother’s commands, as well as MRI and EEG scans demonstrating continual brain activity (Luce, 2015). The imaging studies were regarded as irrelevant by the physician on the case, Dr. Fisher, and her movements, he said, were reflexive only (Luce, 2015). Those videos were also publicly released. Although he has not examined Jahi in person, after reviewing her case and the video evidence, Dr. Alan Shewmon, a pediatric neurologist noteworthy for his extensive research into consciousness, believes that she is still alive (Lupkin, 2014). Another neurologist, Dr. Wei Xiong at University Hospitals in Cleveland, Ohio, states that the hand movement executed by Jahi in these videos is less consistent with posturing (a reflexive movement exhibited in brain death) and more consistent with a complex motion, which would be unusual for someone who is brain dead (Lupkin, 2014). Jahi has been out of a long-term care facility since August of 2014 and was relocated to her mother’s home in New Jersey where she receives 24 hour nursing care (Lupkin, 2014). Now, Jahi has gone through puberty and has reportedly had menstrual cycles, a developmental process that should not be possible in a brain dead patient (Lupkin, 2014). Since Jahi is still considered dead in California, her mother’s next step will be getting Jahi’s death certificate reversed so she and her daughter can relocate back to their home in California (Lupkin, 2014).

In such a case as Jahi’s, we are forced to call into question the legitimacy of determining death by neurologic criteria, along with the ethical and legal consequences of using this standard. The subject of brain death cannot be addressed without first discussing consciousness and its neural correlates, an issue that is yet to be settled by the scientific community (see Eisenberg, 20xx). The traditional view of consciousness is a corticocentric one, which asserts that no aspect of consciousness can exist in the absence of a functional cortex. While evidence points to a critical role of the cortex in establishing integrated thought and cognition as well as complex reflective consciousness, there is no compelling evidence that a primary form of what is referred to as “affective consciousness” cannot exist without the cortex. This form of consciousness could be generated by a network of subcortical structures integrating and converging on the brainstem where primary consciousness ultimately arises. In this and the following paper (this volume), we discuss prevailing theories of consciousness and suspected neural correlates, as well as the legal implications of the definition, before delving into what we believe constitute plausible neural correlates. We also discuss the importance of establishing a universal definition of consciousness in which we distinguish between affective (primary) and reflective (secondary) consciousness, which are supported by different neural networks and structures. This will be accomplished by investigating disorders of consciousness, the evolutionary origin of consciousness, and challenging the corticocentric view of conscious awareness.
Defining and Differentiating Types of Consciousness

Until recently, discussions of consciousness were attributed largely to the field of philosophy. The scientific community has been somewhat late ‘coming to the table’ for this issue (i.e., arguably within the last 15-20 years) driven by methodologies permitting greater attention to the neural networks putatively involved in consciousness (e.g., PET scanning; fMRI; evoked potentials). As might be expected, neuroscience has emphasized the corticocentric view of consciousness, insisting that the conscious experience is principally a function of the cortex. In the section that follows here, we briefly review the evidence that points to this conclusion, and in subsequent sections consider also subcortical structures and their potential role in producing what is referred to as unreflective, affective states of consciousness.

At the outset, we must define consciousness. Consciousness is typically regarded as a process inherent to the individual, although there is often little consideration given to the various levels of consciousness which exist. The assumption that the cortex is necessary to produce consciousness is likely the result of our tendency to conflate consciousness with cognition. Importantly, a functional evolutionary view of consciousness considers that multiple levels of consciousness exist (Panksepp et al., 2007). Several argue that without the cortex, individuals continue to possess a more robust primary form of consciousness called affective consciousness or pre-reflective self-awareness (Izard, 2007; Northoff, 2007; Panksepp et al., 2007). This unreflective state is characterized by basic, core emotions, from which reflective consciousness arose and which are evolutionarily advantageous in the way they drive our instinctual behavior (Denton, McKinley, Farrell, & Egan, 2009; Izard, 2007; Panksepp et al., 2007). Any tendency to disregard this multifaceted nature of consciousness may be partially responsible for the predominantly corticocentric view of consciousness. It could be argued that the empirical evidence for a singular corticocentric view is lacking, unable to completely account for the conscious experience with its insufficient generalizability and its failure to consider subcortical structures.

For the purposes of our discussion, we will incorporate the operational definitions of reflective (secondary) and primary (affective) consciousness according to Izard (2007):

Reflective consciousness is characterized by symbolic processes, memory, and ultimately, the capacity for awareness of self and others and for monitoring one’s own behavior...primary consciousness is characterized by sensory processes that generate subjective feelings (cf. James 1890/1950; Izard 1990), especially emotion feelings, and also includes awareness of and responsiveness to objects in the environment. (p. 97)

Although this paper is mainly concerned with affective consciousness and the neural correlates that allow for its generation, reflective consciousness is obviously more widely recognized. Affective consciousness is generally overlooked (or rather, less recognized) when diagnosing disorders of consciousness. The foundational logic upon which the Uniform Determination of Death Act (UDDA) is established presumably rests on either the absence or presence of reflective consciousness, and does not include any discussion of basic affective consciousness; the assumption is that absolutely no aspect of consciousness survives functional or radical decortication. The traditional definition of consciousness assumes that only one form exists. In the sections that follow, we stress the importance of distinguishing this definition, which is consistent with reflective consciousness, from affective consciousness.
One view holds that consciousness is a state of being characterized by three core elements: awareness, attention, and sense of self (see Figure 2; Garrett, 2015).

Awareness refers to knowledge that a situation is happening. In discussing this abstract concept, we must note that there is a threshold of detection which, when surpassed, brings information to the forefront of conscious awareness. For example, conscious awareness is often measured by assessing visual consciousness, where awareness of some external visual stimuli can be measured via neuroimaging techniques in which a series of words are rapidly flashed on a screen before a subject (Garrett, 2015). The words are flashed at rates either at or below the threshold of detection, so when a word is not consciously perceived, only the visual cortex is stimulated; alternatively, both the parietal and prefrontal cortex are activated in conjunction with the visual cortex when someone becomes aware of having seen the word (Garrett, 2015). The results of such studies demonstrate the involvement of a prefrontal-parietal network in creating conscious awareness of a particular stimulus or situation. Awareness is not, however, contingent upon prefrontal-parietal activity, but on the connectivity between this network and the thalamus – an absence of which is observed among patients in a vegetative state (Laureys, 2005). Without awareness, these patients also cannot demonstrate attention.

When discussing attention, we make it sound transactional: we ‘pay’ attention when we take notice of a particular stimulus or situation, allocating neural resources towards focusing on some neural input while excluding others (Garrett, 2015). Every day, we are flooded with multitudes of sensory stimuli competing with one another for our attention. It is important for a neurological mechanism to be in place to determine which of these stimuli we attend to. The thalamus, or specifically the reticular nucleus of the thalamus, may play a large role in choosing between competing stimuli (Crick & Koch, 2003). The winning stimulus will be the one to which an individual consciously attends. There are two networks involved in attention, and they work in conjunction with one another: a dorsal network, which focuses attention towards a particular stimulus with the intention of satisfying a specific goal, and a ventral network, which responds to the attentional demands of the stimulus (Garrett, 2015). The anterior cingulate cortex is thought to play an executive role in allocating attentional resources by modifying its firing rate to meet the demands of a particular stimulus or task (i.e. it will increase its firing rate in response to a cognitively demanding task and decrease its firing rate when the task does not call for much attention) (Garrett, 2015). Both attention and awareness of the fact that one is attending to a particular stimulus/situation are essential, but not independently sufficient, to determine reflective consciousness. In order to be considered conscious, by its traditional definition, one must also possess a sense of self.

This sense of self is a central component to identity and is an important element of reflective consciousness. “Self-awareness’ refers to the capacity to become the object of one’s own attention” (Morin, 2006, p. 359). An animal possesses self-awareness when it can distinguish itself from its environment, is cognizant of its own experience as a separate entity expressing volitional behavior, and has awareness of itself as an entity which continues to remain that same entity for the duration of its existence (Morin, 2006). It becomes a “reflective observer” (Morin, 2006, p. 359) processing information as it pertains to the self in relation to the external environment. Numerous studies have investigated the neural correlates of self-awareness. For example, with use of the rubber hand illusion and positron emission...
tomography, Tsakiris, Hess, Boy, Haggard, and Fink (2007) demonstrated the importance of the right posterior insula and right frontal operculum in generating a sense of body ownership. When participants of the study did not perceive the rubber hand to be their own, the somatosensory cortex and parietal cortex were active instead (Tsakiris et al., 2007). As illustrated in Figure 3, this evidence suggests that important to self-consciousness/self-awareness is a network formed by the parietal cortex, somatosensory cortex, posterior insula, and frontal operculum, which operate to connect a conscious percept with one's own body (Tsakiris et al., 2007). What, then, happens when one or more of these elements of consciousness are not present in an individual?

**The Persistent Vegetative State**

Disorders of consciousness are often categorized as coma, persistent vegetative state (PVS), and minimally conscious state (see also Panksepp et al., 2007, including akinetic mutism, hyperkinetic mutism, and delirium). We will limit our discussion to the first two, and give most attention to the PVS. Coma and PVS are both characterized by a complete loss of consciousness, and are distinguished by the duration of this altered state of consciousness (Panksepp et al., 2007). Coma is a temporary loss that is eventually recovered, whereas PVS is considered permanent and unlikely to ever resolve, once a period of 12 months or more has elapsed (Panksepp et al., 2007). In PVS, individuals continue to undergo normal sleep-wake cycles and have variable cranial nerve and spinal reflexes, but they have a complete unawareness of self and the environment, incontinence, no language comprehension, and no evidence of sustained, reproducible, purposeful, or voluntary behavioral responses – they are in a state of “wakefulness without awareness” (Garrett, 2015; Laureys, 2005, p. 556). The Multi-Society Task Force (1994) on PVS was assembled to provide guidelines for the diagnosis and treatment of PVS and set forth the following seven diagnostic criteria:

1) No evidence of awareness of self or environment and an inability to interact with others; 2) no evidence of sustained, reproducible, purposeful, or voluntary behavioral responses to visual, auditory, tactile, or noxious stimuli; 3) no evidence of language comprehension or expression; 4) intermittent wakefulness manifested by the presence of sleep-wake cycles; 5) sufficiently preserved hypothalamic and brain-stem autonomic functions to permit survival with medical and nursing care; 6) bowel and bladder incontinence; and 7) variably preserved cranial-nerve reflexes (pupillary, oculocephalic, corneal, vestibulo-ocular, and gag) and spinal reflexes. (p. 1500)

Having established the medical standard for defining PVS, we would advocate for an extended definition which considers a more multidimensional and nuanced nature of consciousness. Accordingly, PVS in particular might be characterized by a complete loss of reflective consciousness with preserved affective consciousness; alternatively, a minimally conscious state (MCS) might be characterized by partial loss of reflective consciousness and preserved affective consciousness.
The vegetative state can be assessed using neuroimaging techniques as well as the patient’s clinical presentation. The neuropathology of PVS is typically reflected by a state of ‘functional decortication’ (Panksepp et al., 2007), which can be seen with imaging. Without the assistance of neuroimaging, one cannot accurately determine an individual’s level of awareness; they may be minimally or entirely aware, remaining cognitively intact but unable to verbally communicate or behaviorally respond, as in locked-in syndrome (Garrett, 2015). With the use of fMRI during mental imagery tasks, it is possible to assess a patient’s level of awareness by giving commands and identifying regions of the brain that correspond to a particular command with the activity visualized on the fMRI scan. Theoretically, those who possess some level of reflective consciousness should be able to volitionally modify their brain activity while undergoing a mental imagery task.

In a study conducted by Monti et al. (2010), brain activity was monitored via fMRI as patients were asked to imagine playing tennis. In one case, a patient was believed to be in a permanent vegetative state for five years before fMRI revealed she was minimally conscious (Monti et al., 2010). Unfortunately, most physicians rely on the patient’s clinical presentation at the bedside rather than exhausting medical resources to assess the patient with imaging. Because of the reliance on neurobehavioral testing, it is not unlikely that numerous others have been misdiagnosed with regard to their state of consciousness. For this reason, it is important to assist neurobehavioral testing with imaging studies.

It is critical to note that the aforementioned mental imagery tasks should only be used to assert the presence of reflective consciousness, considering that even in the absence of cortical activity, affective consciousness may persist (i.e., reflected by intact subcortical function). Consciousness is an abstract concept and its generation is not only dependent upon the presence of certain structures and physiological patterns of activity, but on the integration of processes conducted across structures. It is our contention that the prevailing perspective of consciousness as solely reflective may need elaboration and clarification. Recognizing the broader neural correlates of consciousness as they pertain to diagnosing PVS (in particular) and other disorders of consciousness can significantly impact the way we conceptualize, diagnose, and treat such disorders.

**Prevailing Theories to Explain the Conscious Experience**

When discussing consciousness, it is necessary to guard against the notion of a ‘sentience homunculus’ within the brain, acting as a discrete center for conscious control, making executive decisions influencing actions and thoughts. Contemporary perspectives of consciousness advocate for distributed modular outputs integrated across cortical and subcortical space. Accordingly, Tononi (2008) equates consciousness with integrated information. He states that the quantity and quality of consciousness is dependent upon the degree to which information generated by a distributed neuronal network is integrated, as well as the extent of the relationships between components of that network, or complex. This is referred to as the integrated information theory (IIT) (Tononi, 2008). Integrated information (i.e., identified as ‘I’) is the sum of information produced by the neuronal network converging together. ‘Qualia space (i.e., ‘Q’) is a space where each axis represents a possible state of the complex, each point is a probability distribution of its states, and arrows between points represent the informational relationships among its elements generated by causal mechanisms (connections)” (Tononi, 2008, p. 216). A specific experience is perceived, or brought into conscious awareness, by the integration of these informational relationships which, when connecting various points in the Q, generate a specific shape within the Q, creating a distinct conscious experience. This framework is largely consistent with the theory
proposed by Crick and Koch (2003) which falls somewhere between proposing a distributed network theory of consciousness and a localized one.

According to Crick and Koch (2003), the cortex functions via a highly interconnected network of coalitions between neurons which compete with and support one another, where the winning coalition determines the external stimuli to which we attend and of which we become conscious. A coalition is composed of a network of interconnected nodes, each responsible for generating an output for a particular feature of a percept; an isolated node cannot reach the necessary threshold for generating consciousness, but a network of nodes, or a coalition, can (Crick & Koch, 2003). This is evidenced by numerous studies of minimally conscious patients who show a conservation of widespread cerebral networks, while their completely unconscious counterparts do not (Schiff et al., 2007). Again, this assumes only one form of consciousness, i.e., reflective consciousness. A stimulus is brought into consciousness when the features of that stimulus or object are bound together. Binding refers to the process in which the various features of an object or situation are integrated with one another to form a particular percept (Crick & Koch, 2003).

Crick and Koch (2003) believe that the specific neural correlates of consciousness may be localized to a small set of neurons in the posterior region of the cortex, but that the events leading up to a percept being brought into conscious awareness are widely distributed. As such, consciousness cannot be produced without the widespread distribution of nodes within a coalition binding together their outputs and yielding a particular conscious experience. Again, the network emphasis for Crick and Koch (2003), much like that of Tononi (2008), reinforces the broadly accepted perspective of consciousness as a function of cortico-cortical integration. And, as we have noted, this is unequivocally central to acceptable theories of consciousness. Indeed, the computational processing specializations of the cerebral cortex underlying much of our modern neuropsychology underscores modular specialization, and highlights the functional significance of distributed processing (i.e., linking cross-cortical processing domains) for conscious behavior. Nevertheless, it stands to reason that in a highly conserved central nervous system, a more ‘primordial’ or ancestral vestige of consciousness may exist, albeit overshadowed by the emergent properties of the thinking and reasoning brain, i.e., ‘cortex’.

The notion that consciousness can exist without the cortex is not a novel one. In 1954, and in the midst of an increasingly corticocentric view of consciousness, Penfield and Jasper put forth a radical proposal, centered around their observations of epileptic patients (as cited in Merker, 2007). They (Penfield and Jasper) suggested that the highest level of integration is not accomplished by the cortex, but by the upper brainstem onto which neural activity converges to produce the conscious experience. They called this system the centrencephalic system (as cited in Merker, 2007). Their theory lays the foundation for Merker’s (2007) mesodiencephalic system, which is composed of primary integrative machinery housed at the level of the midbrain and basal diencephalon, and supplied by the telencephalon, which provides it with the contents of the conscious experience.

**Penfield and Jasper’s Epileptic Studies Lay The Foundation For a Subcortical Model**

Penfield and Jasper constructed their centrencephalic system from their observations of epileptic patients. As discussed previously, the PVS is characterized by a state of wakefulness in the absence of awareness. A similar dissociation between wakefulness and...
awareness is demonstrated by patients experiencing absence seizure – a disorder characterized by a period of unresponsiveness while awake (Laureys, 2005). Interestingly enough, when Penfield and Jasper stimulated the cortex of their subjects, they were able to produce all varieties of seizure except absence seizure (as cited in Merker, 2007). Functional MRI scans (fMRI) taken while a patient experiences an absence seizure (during which an absence of conscious awareness is demonstrated) reveal a loss of function in the bilateral frontoparietal cortices (Laureys, 2005). While treating epileptic patients, Penfield and Jasper, then, were surprised to see a preservation of consciousness despite the surgical extrication of relatively large portions of the cortex – considered the region of highest integrative function (as cited in Merker, 2007). Moreover, Penfield and Jasper observed that there was no disruption in the consciousness of a patient who was kept under conscious sedation while regions of the cortex were actively being removed (as cited in Merker, 2007).

The groundbreaking work of Penfield and Jasper suggested that damage to the cortex does not always result in a diminished state of consciousness. Meanwhile, lesions of the brainstem (Moruzzi & Magoun, 1949) and suppression of subcortical arousal systems (involving the upper brainstem reticular formation and the intralaminar thalamus) during limbic seizure (Blumenfeld, 2012; Kundishora et al., 2017) result in states resembling coma, the persistent vegetative state, or pathological somnolence. Though deficits in select higher order cognitive processes are ultimately observed with cortical lesioning (i.e., depending on location and extent), consciousness of self and environment may remain intact, challenging the assertion that the highest level of integrative functioning takes place in the cortex.

Subcortical Model of Consciousness

While the higher order cortices are necessary for the execution of more integrated and coordinated cognition and thought, as well as the ability to self-reflect on one’s own thought processes, they may not be intrinsic to consciousness. The overemphasis on the cortex may inadvertently limit our understanding of consciousness by prematurely limiting the debate around its neural correlates and the extent of cortical activity needed for consciousness to persist. Merker (2007) proposes a mesodiencephalic system [a system involving the midbrain and diencephalon] in which the highest level of integrative function takes place in the brainstem. Accordingly, we must not assume that the primitive nature of the brainstem and other subcortical structures excludes them from consideration.

It seems quite reasonable to concede that the cortex is necessary for reflective consciousness. However, the immediate unreflective experience of consciousness—also referred to as primary or affective consciousness— requires only subcortical regions of the brain (Merker, 2007). Northoff (2007) calls this form of consciousness, which allows us to “immediately and unreflectively experience our self,” (p. 101) pre-reflective self-awareness, and proposes a subcortical mediation of this type of consciousness by self-related processing. Self-related processing refers to our ability to discriminate between the self and environment by examining interoceptive stimuli (i.e., internally generated) in relation to exteroceptive stimuli (i.e., externally generated; Northoff, 2007). The ventral striatum/ nucleus accumbens, as well as the “tectum, periaqueductal gray, dorsal medial thalamus, and colliculi have been identified by Schneider et al. (submitted)” (Northoff, 2007, p. 101) as subcortical structures involved in self-related processing (refer to Figure 5). The circuitry of afferent and efferent sensorimotor projections to and from these structures led Northoff (2007) to conclude that
they may be critical in relating the sensorimotor with the self, thus giving rise to an immediate experience of the self through self-relatedness to one's immediate environment. This type of consciousness reflects what one is experiencing is not simply an external stimulus, but oneself as related to that stimulus. This, and other numerous factors, are involved in producing consciousness.

Consciousness emerges through the confluence of enabling factors and specific factors (Denton et al., 2009). Enabling factors refer to the particular regions and circuits of the brain that give rise to any degree of consciousness, and specific factors refer to those elements that are responsible for creating a specific conscious percept (e.g., sensory input; Denton et al., 2009). A cluster of nuclei in the brainstem, basal diencephalon, and the thalamus regulate the cortex and give humans the ability to experience consciousness (Denton et al., 2009). Together, "the ascending fibers [of these nuclei] create the essential conditions for any consciousness to occur" (Denton et al., 2009, p. 506) – namely, wakefulness, arousal, and sleep, which are regulated, in part, by cholinergic release from these nuclei. These essential conditions, thus, are generated by a brainstem–forebrain system innervated by a highly distributed network whose activities converge on the brainstem, where, according to Merker (2007), the information is integrated for action. In this system, conscious awareness is predicated upon a functioning brainstem receiving input from other regions of the brain. Cortical functioning contributes to the brainstem's supply of information regarding the external environment, but is not necessary for consciousness to exist.

Though an important source of input to the upper brainstem is supplied by the cortex, it is necessary to identify the presence of both efferent and afferent projections and note that the manner in which the cortex interacts and integrates with the mesodiencephalic network is really suggestive of the upper brainstem’s role in generating consciousness (Merker, 2007). Some of the first attempts at identifying the neural correlates of consciousness led to the discovery that nearly 80% of PVS patients have abnormalities in the thalamus and that a combination of lesions in both the corpus callosum and dorsolateral brainstem were a frequent occurrence in PVS as well (Panksepp et al., 2007). Giacino and Whyte identified two major patterns of brain damage in the PVS: 1) diffuse axonal Injury, which is typically related to thalamic malfunction resulting in ischemia and 2) focal brainstem lesioning (as cited in

Figure 5. Mesial view of right hemisphere. The mesencephalon (midbrain) contains the periaqueductal gray, and the superior and inferior colliculi (i.e., the tectum [T]). The diencephalon contains the thalamus [TH], and the cross-hatched area anterior to the diencephalon is the location of the deeper-situated ventral striatum and nucleus accumbens. From: https://anatomytool.org/content/sobotta-1914-fig549-arteries-mesial-surface-brain-no-labels
Panksepp et al., 2007). These results would suggest that the disruption in consciousness observed in the PVS may be more likely the result of damage to the mesodiencephalic system.

In discussions of consciousness, the literature repeatedly points to the posterior parietal cortex, the prefrontal cortex, and the anterior cingulate cortex in mediating the three central components of consciousness: attention, awareness, and the sense of self. These three regions of the cortex are intersected by the thalamus. According to Merker, they all have a “mutual interface in the nuclei of the dorsal thalamus... The latter [the nuclei] can be divided into first-order (largely sensory relay) and higher-order (“association”) thalamic nuclei (Sherman & Guillery, 2001), and it is with the latter, higher-order nuclei, that the mesodiencephalic system maintains an intimate and complex relationship” (Merker, 2007, p. 74). The thalamus is an important mediator in regulating states of arousal (Schiff et al., 2007). During periods of wakefulness, the thalamus, in conjunction with the brainstem, maintains the brain’s metabolic rate at the threshold for access to conscious awareness (Schiff et al., 2007). In disorders of consciousness, there is typically a reduction of global metabolic activity to approximately 50% of normal levels (Schiff et al., 2007; Laureys, 2005). However, in some cases of PVS, there is not a significant reduction of global metabolic activity, but in these cases, patients did eventually recover consciousness (Laureys, 2005). Diminished states of arousal exhibited during focal limbic seizure follow from suppression of the upper brainstem activating system (to which the reticular formation is essential) and the intralaminar thalamus (Blumenfeld, 2012; Kundishora et al., 2017). In a study by Kundishora et al. (2017), some consciousness was restored during limbic seizure with the use of deep brain stimulation of the thalamus. The results of these studies further support a consciousness system comprising the diencephalon and the midbrain with the functional convergence of the thalamus and reticular activating system of the brainstem playing an integral role in states of arousal.

From the results of their 1954 study on absence seizures, Penfield and Jasper conclude that it is the brainstem and not the cortex that plays the highest role in integrating the outputs of a distributed network which converge at the brainstem to carry out purposeful action (as cited in Merker, 2007). The importance of the brainstem in generating consciousness is demonstrated by the restorative ability of DBS treatment in anesthetized animals. Increased activity of the EEG from synchronized discharge to low voltage fast activity and improved behavior (generating an arousal response) were resultant of stimulation of the reticular formation of the brainstem (Moruzzi & Magoun, 1949). “The distribution of the excitable area... includes the central core of the brain stem, extending from the bulbar reticular formation forward through the pontile and mesencephalic tegmentum into the caudal diencephalon” (Moruzzi & Magoun, 1949, p. 252). Their experiment demonstrates the necessary role of this subcortical network in generating consciousness, especially wakefulness and arousal of a subject to alertness.

Penfield and Jasper conclude that it is the brainstem and not the cortex that plays the highest role in integrating outputs of a distributed network which converge on the brainstem....

Merker (2007) asserts that the centrencephalic system of Penfield and Jasper involves the reticular formation extending into the midline, and intralaminar reticular thalamic nuclei. The brainstem, innervated by higher order cortices, ultimately integrates...
incoming cortical input together to yield consciousness (Merker, 2007) (refer here to Figure 6). Also showing integrative function – bringing together multiple sensory modalities important for spatial orientation -- is the superior colliculus at the roof the midbrain (i.e., tectum), which receives its input from the substantia nigra of the basal ganglia and projects its integrated output through a multitude of efferents to the thalamus and brainstem (Merker, 2007). Ventral to the superior colliculus is the reticular formation, thalamus, periaqueductal gray, and substantia nigra, and at the bottom of the midbrain lies the hypothalamus, which integrates multiple categories of goal oriented behavior (Merker, 2007). Merker (2007) proposes that the synencephalon (a region of the brainstem between the diencephalon and midbrain) is perfectly situated to create a bottleneck in which distributed cortical outputs project through the integrative mechanisms of the midbrain and converge at the brainstem where the “massively parallel and distributed information capacity of the cerebral hemispheres is further integrated] into the limited-capacity, sequential mode of operation” (p. 63) necessary for the execution of volitional behavior. The system operates by what Merker (2007) calls a selection triangle by which the brainstem carries out integration for action by connecting three components of behavior: motivation, target selection, and action selection.

**Figure 6. Centrencephalic network model supporting conscious behavior** (adapted from Merker, 2007). The diencephalon consists of the thalamus and the hypothalamus. The mesencephalon (region of the midbrain) includes the superior and inferior colliculi (collectively referred to as the tectum), the periaqueductal gray (region of white dashed circle) and the substantia nigra (indicated by dark-filled circle).

*Mesial view of brain taken from:*
http://biodrawing.com/Neurology_modules/NervousSystemSite/Neuroanatomy/Brainstem/Introduction.html

**Merker asserts that the centrencephalic system of Penfield and Jasper involves the reticular formation extending into the midline, and intralaminar reticular thalamic nuclei.**
There are three main contributions to behavior indicated by the selection triangle. It considers that a conscious organism has a particular motivation (the motivational) for carrying out a specific action (action selection) based upon the assessment of its environment (target selection); action selection is mediated by the resources and environmental states available to that organism (the target states of the external environment) as well as consideration into the amount of energy that would be expended in each of its possible choices of action (the action states of the body) (Merker, 2007). In order for the organism to select an optimal action state based upon the available target states of its external environment and in accordance with its motivation for producing an action, it must have a sense of the world around it, and this is accomplished through sensory inputs and/or via the action of the superior colliculus executing self-related processing (Merker, 2007; Northoff, 2007). This necessitates that the organism have an awareness of its own physical existence, as well as emotionally and motivationally guided behavior to accomplish a particular goal within an external environment with which it is capable of interacting (Merker, 2007). The interaction between bodily action (the body/self), target selection (the world), and motivation (the motivational) described by the selection triangle is consistent with Northoff’s (2007) pre-reflective self awareness through self-related processing, which emphasizes one’s ability to discriminate between the self and the environment and subsequently relate oneself to one’s surroundings for the immediate, unreflective experience of consciousness to arise.

The acquisition of information about the environment is obtained through use of the senses. While the sensory organs do enrich our conscious experience, it should be noted that consciousness is not contingent upon one’s ability to perceive each of the various modalities. Someone with impaired vision or hearing (e.g., blind or deaf) still possesses consciousness. For simplicity’s sake, however, let us trace the process of conscious perception first starting at the level of the cortex. Various sense organs (e.g., vision/retina; hearing/cochlea) transduce the respective sensory stimuli. Transduced signals (i.e., action potential electrophysiology) project to the corresponding primary and secondary sensory cortices and then towards the midbrain and forebrain in a parallel fashion, to produce bodily action through movement. In the absence of cortical input and/or the senses, however, a primary form of consciousness would still be supported by the mesodiencephalic system proposed by Merker (2007) in combination with that proposed by Izard (2007) and in accordance with Tononi’s (2008) IIT in which he equates consciousness with integrated information. We will refer to this explicitly as the subcortical model of consciousness – the foundations for the unreflective process.

This unreflective state of awareness may be also be a function, in part, of the basal ganglia (i.e., caudate, putamen, globus pallidus, substantia nigra of midbrain) and other influences including the thalamus and brainstem structures (Izard, 2007; see also Afrasiabi et al., 2021). These structures may be viewed as part of a distributed subcortical network by which information integration takes place across each structure of the network and, as Merker suggests (2007), ultimately converges on the brainstem where it is integrated further, giving rise to consciousness. The basal ganglia project to the ventral anterior and ventral lateral nuclei of the thalamus as well as the superior colliculus. When there is intact cortical function, the thalamus relays information back to the cortex which would support reflective consciousness. With or without the cortex, however, the substantia nigra of the basal ganglia also projects its output towards the superior colliculus, which integrates various sources of information and provides a sense of spatial orientation of the body in relation to its environment as well as a necessary distinction between the self and the environment, providing us the neural substrate possibly responsible for target selection and pre-reflective self awareness (Merker, 2007; Northoff, 2007). Also projecting its output to the superior colliculus (via direct innervation and indirect projections through the periaqueductal gray) is the hypothalamus (refer again to Figure 6),
which plays an integral role in "regulating and integrating motivational states related to goal-directed behaviors" (Merker, 2007, p. 71).

It is here then, between the hypothalamus and periaqueductal gray, that the superior colliculus provides the necessary machinery to “implement a three-way interface” between motivation, target selection, and action selection (Merker, 2007, p. 71). Once information is integrated by the superior colliculus, the output is projected to the thalamus which works in conjunction with the reticular activating system of the brainstem in mediating arousal states (Schiff et al., 2007; Morruzi & Magoun, 1949). The integrated information by the superior colliculus also projects to the brainstem via colicural efferents where it is combined with the distributed and parallel outputs of the cortex (if present) as well as the outputs of the integrative network of the mesodiencephalon with the functional convergence of the limbic system and amygdala -- all converging on the brainstem to generate affective consciousness (Merker, 2007; Denton et al., 2007). If this subcortical model of consciousness holds true, it will be absolutely necessary to emphasize whole brain loss of functioning in diagnosing brain death, for any amount of brain activity – namely in the subcortical regions of the brain - could indicate that some form of consciousness may still exist.¹ Ultimately, further research and biomedical ethics will need to address the possible dimensions of consciousness and what the functional implications might be regarding the distinction between reflective and non-reflective (i.e., subcortical) aspects of conscious experience. Toward this end, a review of the subcortical mechanisms of conscious experience continues to be of paramount importance, and it is here that perspectives of the evolutionary origins of consciousness are of value.

Transitions Across Time and Phylogeny

The emergence of consciousness is likely due to its adaptive advantage. In support of this evolutionary advantage, Merker’s (2007) selection triangle (see Figure 7) posits that with the advent of consciousness, animals could select the most optimal choice for action (action selection) based upon their evaluation of the present environment (target selection) to fulfill a particular goal or desire (motivation), such as remaining hydrated for survival (refer to Figure 8). When confronted with an environmental threat (e.g., lack of food/water availability, presence of a predator, etc.), an animal’s behavior is mediated by the primordial emotions (e.g., thirst, hunger, fear, etc.) in order to alleviate that threat. The primordial emotions are defined by Denton et al. (2009) as “the subjective element of the instincts which are the genetically programmed behaviour patterns which contrive homeostasis,” (p. 500) generated by subcortical regions and from which consciousness may have originated. The neural correlates of consciousness can be studied with respect to the primordial emotions by examining the immediate gratification that is achieved by satiating the particular primordial emotion causing an animal distress (Denton et al., 2009). Satiating the primordial emotion subsequently increases an organism’s odds of

¹ It is of particular importance to note that the mere presence of subcortical activity, even if determined to support affective consciousness, does not suggest this is – in itself – a manifestation of consciousness incompatible with a decision to terminate supportive maintenance of an otherwise seemingly ‘non-conscious’ patient. Indeed, quality of life and an assessment of functional awareness must be considered here as well.
survival by restoring homeostasis, thus giving organisms in possession of consciousness an important adaptive advantage.

The functional evolutionary view of consciousness echoes Merker's view of consciousness as an immediate and unreflective experience. It posits that an ancient form of primary, affective consciousness gave rise to the more sophisticated self-reflective consciousness, which enables higher order cognitive processes, such as the ability to consciously reflect on one's own experience and possess awareness of one's own awareness (Panksepp et al., 2007). Abundant research in animal models supports the theory that there is an affective consciousness which preceded the secondary (reflective) form of consciousness that gives humans the presumably unique ability to conceptualize and discuss our own conscious experiences. Mammals possess epicenters for affective experiences within the medial and ventral regions of the brain, which endow them with the ability to experience primordial emotions such as thirst, hunger, and suffocation in instances of a disturbed homeostatic state (Denton et al., 2009; Panksepp et al., 2007). Primordial emotions hold functional relevance for driving goal-oriented behavior that aims to restore homeostasis.

Baars (2005) suggests that consciousness itself is “biologically fundamental and phylogenetically ancient” (p. 7) and was likely present at least as far back as the earliest mammals—for alertness, sensory perception, memory, and motor function are all necessary components in producing goal-oriented behavior. Because of the highly conserved and phylogenetically ancient nature of the brainstem, it is probable that the origin of consciousness is tied to the brainstem as well as other subcortical regions of the brain. Goal-oriented behavior is ultimately driven by the core primordial emotions, which are produced by diencephalic, mesencephalic, and medullary mechanisms (Denton et al., 2009). Therefore, the theory proposed by Denton et al. (2009) is consistent with Merker’s (2007) theory that consciousness preceded the development of the cortex and is implemented by the primitive lower regions of the brain.

Interestingly, as increasingly lower regions of the brain (such as the periaqueductal gray) are electrically stimulated, there is a corresponding increase in affective intensity (Panksepp, 1998). Panksepp (1998) envisions a system of emotional circuits coded by value, converging and interacting with:

“coherent brainstem representations of the body and nearby attentional/waking systems of the brain. Affective feelings may be caused by the neurodynamics of basic emotional circuits interacting with neural schema of bodily action plans. One key brain area where such interactions occur is found within centromedial diencephalic midbrain areas such as the periventricular and periaqueductal gray (PAG) and nearby tectal and tegmental zones. Here [...] a Simple Ego-type Life Form (a primitive SELF structure) is instantiated in those circuits. The ability of this ‘primal SELF’ to resonate with primitive emotional values may help yield the raw subjectively experienced feelings of pleasure, lust, anger, hunger, desire, fear, loneliness and so forth.” (p. 566)

The system he has conceptualized provides us with a neural dynamic process by which the primal subcortical circuitry may produce the primordial emotions and primary affective consciousness.

In discussing the primitive regions of the brain versus the neocortex, it is necessary to distinguish between higher-level processes as they pertain to degree of sophistication versus degree of control. As noted, we recognize the tendency to conflate consciousness with cognition. The two are distinct. Consciousness is not inherently sophisticated or complex. While less sophisticated and more primitive than the cortex, the brainstem executes a higher level of control over behavior than the cortex. Merker (2007) provides an example of this: in instances in which an individual is experiencing fear, hunger, or pain, subcortical regions of
the brain can supersede the executive control of the cortex to direct volitional behavior, essentially outcompeting the cortex for control over action. For example, when blood gas levels fall out of their normal range, automatic respiratory function (controlled by the brainstem) falls out of unconscious automaticity and into consciousness awareness as one is filled with an overwhelming sense that they are suffocating and subsequently must make a behavioral choice (action selection) about their circumstances (target state) to alleviate their state of respiratory distress (motivation) (Merker, 2007).

The above example suggests how a primal function such as respiration can motivate goal-directed behavior, highlighting this distinction between cognitive complexity and consciousness; indeed, conscious behavior is not inherently derived from complex cognition (Merker, 2007), but can potentially exist at a level 'beneath' that which is assumed to emerge from cortico-cortical mechanisms. This notion of conscious and behavioral complexity in humans is the subject of the second, follow-up article, in this sequence: Biomedical and Ethical Implications of a Subcortical System for Consciousness.

References


*Journal of Neuropsychology and Behavioral Processes, 1*, 11-26


