Dear Author,

Here are the proofs of your article.

- You can submit your corrections **online**, via **e-mail** or by **fax**.
- For **online** submission please insert your corrections in the online correction form. Always indicate the line number to which the correction refers.
- You can also insert your corrections in the proof PDF and **email** the annotated PDF.
- For fax submission, please ensure that your corrections are clearly legible. Use a fine black pen and write the correction in the margin, not too close to the edge of the page.
- Remember to note the **journal title**, **article number**, and **your name** when sending your response via e-mail or fax.
- **Check** the metadata sheet to make sure that the header information, especially author names and the corresponding affiliations are correctly shown.
- **Check** the questions that may have arisen during copy editing and insert your answers/ corrections.
- **Check** that the text is complete and that all figures, tables and their legends are included. Also check the accuracy of special characters, equations, and electronic supplementary material if applicable. If necessary refer to the *Edited manuscript*.
- The publication of inaccurate data such as dosages and units can have serious consequences. Please take particular care that all such details are correct.
- Please **do not** make changes that involve only matters of style. We have generally introduced forms that follow the journal's style.
  Substantial changes in content, e.g., new results, corrected values, title and authorship are not allowed without the approval of the responsible editor. In such a case, please contact the Editorial Office and return his/her consent together with the proof.
- If we do not receive your corrections **within 48 hours**, we will send you a reminder.
- Your article will be published **Online First** approximately one week after receipt of your corrected proofs. This is the **official first publication** citable with the DOI. **Further changes are, therefore, not possible.**
- The **printed version** will follow in a forthcoming issue.

Please note

After online publication, subscribers (personal/institutional) to this journal will have access to the complete article via the DOI using the URL: http://dx.doi.org/[DOI].
If you would like to know when your article has been published online, take advantage of our free alert service. For registration and further information go to: http://www.springerlink.com.

Due to the electronic nature of the procedure, the manuscript and the original figures will only be returned to you on special request. When you return your corrections, please inform us if you would like to have these documents returned.

Metadata of the article that will be visualized in OnlineFirst

ArticleTitle	The Evolutionary Rationale for Consciousness	
Article Sub-Title		
Article CopyRight	Konrad Lorenz Institute for Evolution and Cognition Research (This will be the copyright line in the final PDF)	
Journal Name	Biological Theory	
Corresponding Author	Family Name	**Grinde**
	Particle	
	Given Name	**Bjørn**
	Suffix	
	Division	Division of Mental Health
	Organization	Norwegian Institute of Public Health
	Address	Nydalen, P.O. Box 4404, Oslo, 0403, Norway
	Email	bjgr@fhi.no

	Received	16 April 2012
Schedule	Revised	
	Accepted	7 August 2012
Abstract	To answer the question of why we have consciousness, I propose the following evolutionary trajectory leading to this feature: Nervous systems appeared for the purpose of orchestrating behavior. As a rule of thumb the challenges facing an animal concern either approach or avoidance. These two options were originally hard-wired as reflexes. Improvements in adaptability of response came with an expansion of the computational aspect of the system and a concomitant shift from simple reflexes to instinctual behavior, learning, and eventually, feelings. The assessment of positive and negative feelings allows organisms to weigh various options, but for this to be a viable strategy, an awareness of hedonic value is required. This was presumably the first neural attribute to evolve that required awareness, and thus the key force in the evolution of consciousness. The attribute first appeared in the early amniotes (the phylogenetic group comprising reptiles, birds and mammals). Support for this model in current accounts of the neurobiology of feelings and consciousness is discussed.	
Keywords (separated by '-')	Amniotes - Consciousness - Emotions - Evolution - Mood modules - Self-awareness	
Footnote Information		

Biol Theory
DOI 10.1007/s13752-012-0061-3

LONG ARTICLE

The Evolutionary Rationale for Consciousness

Bjørn Grinde

Received: 16 April 2012 / Accepted: 7 August 2012

Abstract To answer the question of why we have consciousness, I propose the following evolutionary trajectory leading to this feature: Nervous systems appeared for the purpose of orchestrating behavior. As a rule of thumb the challenges facing an animal concern either approach or avoidance. These two options were originally hard-wired as reflexes. Improvements in adaptability of response came with an expansion of the computational aspect of the system and a concomitant shift from simple reflexes to instinctual behavior, learning, and eventually, feelings. The assessment of positive and negative feelings allows organisms to weigh various options, but for this to be a viable strategy, an awareness of hedonic value is required. This was presumably the first neural attribute to evolve that required awareness, and thus the key force in the evolution of consciousness. The attribute first appeared in the early amniotes (the phylogenetic group comprising reptiles, birds and mammals). Support for this model in current accounts of the neurobiology of feelings and consciousness is discussed.

Keywords Amniotes · Consciousness · Emotions · Evolution · Mood modules · Self-awareness

Introduction

Terminology

For the human species, consciousness is what life is about; yet presumably it is a trait lacking in the vast majority of organisms as it is difficult to envisage this attribute in the absence of an advanced nervous system. At some point in our evolutionary history the trait evolved, and if we can understand the evolutionary rationale, i.e., the adaptive significance, behind this event, we stand a better chance of understanding what consciousness is about. I present a model for the evolution of consciousness suggesting that the feature first appeared for the purpose of experiencing feelings, and that the capacity to feel evolved as a strategy toward a more flexible and adaptive way of evaluating behavioral options.

Few topics in science have a more extensive, and varied, depiction than the phenomenon referred to as consciousness. In order to present a coherent model it is pertinent to first discuss a few key terms. The following outline reflects what is useful for the present purpose, a general overview of the literature is beyond the scope of this article.

Consciousness implies an ability to be aware of sensory input and thus be in a position to monitor aspects of both the external and internal environment. Besides the ability to experience life, this attribute entails a neurobiological flexibility that can be used to drive a variety of behavioral outputs. In an animal capable of consciousness, some types of behavior are driven by motivation based on feelings—rather than on more hard-wired responses such as fixed action patterns and innate or learned behavioral patterns. In the terminology of Edelman (2004), *primary consciousness* (i.e., sensory consciousness or awareness) can be defined as the ability to integrate observed events with memory to create awareness of the present and immediate past; while *secondary consciousness* includes additional features such as self-awareness and reflective thoughts and thus allows for "being conscious of being conscious." Primary consciousness is sufficient to turn key components of brain activity into a cohesive 'film of life.'

A1 B. Grinde (✉)
A2 Division of Mental Health, Norwegian Institute of Public Health,
A3 Nydalen, P.O. Box 4404, 0403 Oslo, Norway
A4 e-mail: bjgr@fhi.no

Self-awareness (or self-recognition) implies an understanding of the "self" as a unique entity in the environment. The term suggests that the organism knows "who it is", i.e., it has concepts of 'agent' and 'agency.' Self-awareness is generally assayed with the mirror test (or related methods), and appears to be restricted primarily to humans and apes (Kitchen et al. 1996); although other mammals, such as cetaceans (Reiss and Marino 2001), as well as certain birds (Prior et al. 2008) may possess rudimentary forms of self-awareness.

Feelings imply brain activity causing affect. They include emotions as well as any sensation that are made available to (or impact on) conscious experience and that tend to have a positive or negative connotation; i.e., pleasure or displeasure. Emotions typically have social (or extrovert) components, while sensations concern primarily oneself. The term feelings consequently include affect caused by, e.g., physical pain and hunger, which are often not considered to be emotions. The parts of the brain involved in generating pleasure or displeasure may be referred to as *mood modules* (Grinde 2012). Feelings are, per definition, the conscious output of these modules. *Hedonic value* refers to the positive or negative aspect of feelings, as opposed to the particular type of sensation.

It should be noted that when employing words originally coined to describe human conditions in the characterization of animals, the question of appropriate use is necessarily somewhat arbitrary. Some people will, for example, claim that dogs have a nose, while others may say they do not possess a nose, but rather a snout. All living organisms have features in common with humans, but the features are unique to each species in their detailed structure and function. The snout and the nose are evolutionary homologous entities, but have evolved along different trajectories for a considerable amount of time. Similarly, the consciousness experienced by a dog is most likely different from that experienced by a human; but the two forms of consciousness are derived from a shared ancestor, which makes it reasonable to refer to them by the same term. In most cases, including consciousness, there is a somewhat arbitrary cutoff as to when the attribute possessed by an organism has the required similarity to the homologous (or analogous) human attribute to warrant the use of the same term.

Attributes of Consciousness

The feature of consciousness is one of many modules, or functions, that have been added to the mammalian brain over the course of evolution. It involves a select fraction of the processes taking place in the awake brain. The processes not brought to conscious awareness are referred to as *subconscious*. The subconscious activity has the capacity to direct the attention of the conscious brain in a fashion analogous to what, in the language of business, is referred to as "information given on a need to know basis." Thus, even information that impacts on our emotional life is not necessarily brought to conscious attention (Tamietto and de Gelder 2010). Presumably, consciousness is costly to operate, and only capable of handling one experience at the time; thus conveying too much information to the part of the brain responsible for consciousness could cause dangerous distractions.

Consciousness can be turned on or off, either by the power of control vested in the subconscious (as when falling asleep), by external means (as in anesthesia), or by damage to the brain (e.g., coma caused by a stroke). The various situations in which consciousness is off may be collectively referred to as *unconsciousness* (used here in a physiological, rather than Freudian, sense). The natural form of unconsciousness (sleep) is, however, different from anesthesia and coma in its capacity to generate dreams, and in that the subconscious retains the power to turn on awareness when needed, as when external stimuli suggest danger.

In humans, "accurate report" (e.g., in response to queries about a sensation or experience) may be used as a sign of consciousness (Seth et al. 2005), but in order to probe for a homologous feature in animals, we need to identify other defining qualities. A variety of neurobiological and behavioral correlates have been suggested, including: the presence of a thalamocortical complex, extensive "cross-talk" between dispersed nerve circuitry within this complex, a "default mode network" involving core activity in prefrontal and medial parietal regions of cortex, distinct sleep-wake cycling, behavioral flexibility (or behavior indicative of choice based on motivation rather than on hard-wired patterns), play behavior, signs of emotions or feelings, advanced communication, skill acquisition, and cultural transmission (for reviews, see (Rossano 2003; Crick and Koch 2003; Butler 2008; Cabanac et al. 2009; Edelman and Seth 2009). The position taken here is that several of these features, but not necessarily all, should be present in order to ascribe consciousness to an organism within the vertebrate lineage.

There is reasonable evidence indicating the existence of primary forms of consciousness in mammals and birds (Butler and Cotterill 2006; Edelman and Seth 2009), and possibly in reptiles as well (Cabanac et al. 2009). Taken together, these observations suggest that the trait first evolved in the common ancestor of these three classes, collectively referred to as amniotes, some 300 million years ago. Excluding the reptiles would mean that it evolved independently in birds and mammals; and a model not requiring convergent evolution is, arguably, more parsimonious.

All amniotes have a complex behavioral repertoire, and at least birds and mammals appear to have cultural transmission (Laland and Galef 2009). Moreover, the amniotes (but apparently neither fish nor amphibians) display signs of emotion, such as tachycardia and fever upon handling, an increase in brain dopamine activity (the neurotransmitter most closely associated with reward oriented behavior), and an apparent capacity to feel pain (Cabanac et al. 2009; Mosley 2011). Compared to lower vertebrates, amniotes have larger brains, and are thus presumably capable of a more complex response to the challenges of life. While it might be tempting to explain consciousness as an evolutionary strategy aimed at facilitating computational brainpower, or as a by-product of a sophisticated brain (Rosenthal 2008), advanced behavior—for example, communication in social insects—apparently does not require consciousness (Gould and Grant-Gould 1995) (and, one might add, neither do computers). Either presumed non-conscious species such as insects and fish do not possess a sufficiently sophisticated brain, or other factors beyond mere intricacy of response are required in order to explain the origin of consciousness. I shall argue in favor of the latter.

Amniotes were the first vertebrates to adapt to life on land. It has been discussed whether the complexity of terrestrial environments spurred the emergence of more complex behavior and consciousness (Cabanac et al. 2009). One would expect, however, that early terrestrial environments were a lot simpler, harboring a considerably lower diversity of life forms, compared to the oceans. Moreover, non-vertebrate animals, including annelids, arthropods and mollusks, colonized dry land at about the same time, or shortly after, without a similar expansion of the nervous system.

Interestingly, two of the most impressive escalations of brain capacity, i.e., in the molluscan class *Cephalopoda* (Edelman and Seth 2009) and the mammalian order *Cetacea* (Marino 2007), occurred in the ocean. In fact, cephalopods are the foremost candidates for consciousness in invertebrate animals (Mather 2008; Edelman and Seth 2009). The brains of these invertebrates are profoundly different as to neuroanatomical structures compared to amniotes. To the extent that they display signs of consciousness, a closer examination may therefore suggest general principles as to the underlying circuitry. Nevertheless, the presence of anything resembling consciousness in invertebrates would require convergent evolution, and has consequently limited relevance as to delineating the evolutionary trajectory leading to consciousness in humans. The present discussion will therefore focus on vertebrates.

A capacity for feelings, or emotions, are typically listed among the defining features of consciousness; however, even if consciousness were to be defined solely by other qualities, the current evidence suggests that the two features evolved concurrently (Cabanac et al. 2009; Denton et al. 2009; Mosley 2011). This observation may offer a more fruitful starting point for explaining the evolutionary scenario leading to vertebrate consciousness.

Evolution of Consciousness

The Rationale for Nervous Systems

The more primitive, decentralized nervous systems (e.g., in jellyfish and other members of the phylum *Cnidaria*) serve two functions: first, to collect information about the environment; and second, to initiate a response by activating muscles or glands. In more advanced, bilateral animals, nerve cells aggregate in ganglia or other centralized structures such as the vertebrate brain. These structures evolved for the purpose of a third function: to perform processing and evaluation of the information obtained prior to response.

While the most primitive nervous systems operate entirely on reflexes, or fixed action patterns, the expansion of processing implied a gradual shift toward learning and cognition. Nevertheless, even in humans, several types of external stimuli, and perhaps a majority of internal needs, are cared for by reflexive (subconscious) processing, exemplified by the adjustment of pupils in response to light and the heartbeat, respectively.

Nervous systems are associated with the management of behavior, and behavior is primarily a question of movement. Macroscopic plants are generally sedentary, and consequently have not evolved a nervous system. In the metazoans, however, nerve cells and their accompanying behavioral outputs were an evolutionary success. This success is founded on two pillars: One, neuronal circuitry allowed the organism to approach opportunities (e.g., nutrients and potential mates); and two, they made it possible to escape danger (e.g., toxins, inappropriate environments, and predators). This dichotomy—i.e., the pursuit of opportunities and the avoidance of aversive or dangerous conditions—is a core feature of all nervous systems.

Behavior that appears to be intelligent does not necessarily require consciousness, as (presumably) in the case of communication among social insects; and a response to sensual stimuli does not imply the sensation of feelings, as exemplified by the curling up of an earthworm in response to being poked. In other words, one should be careful about making assumptions as to the attributes of nervous systems based on observations of behavior alone.

The Rationale for Consciousness

Reflexes do not require extensive centralized processing. Brain power evolved gradually, presumably due to the advantages of integrating more information before executing a response, and being able to base that response on previous experience. The latter quality implies the ability to learn; but even relatively primitive animals, such as nematodes, may have this capacity (Zhang et al. 2005). Eventually, the response to the challenges of life was no longer simply an issue of whether or not to approach or escape, but rather, a matter of weighing the pros and cons in a decision-making process allowing a large number of finely tuned alternatives. The advantage came in the form of flexibility in dealing with the environment; i.e., behavior that adapts to variable conditions.

In order to implement an advanced form of behavioral response, a nervous system would need to evaluate the survival value of various expected outcomes. Early nervous systems must have already been tuned to the approach-or-avoid dichotomy of most situations; i.e., the outcome tends to be either positive or negative. However, in the case of more advanced animals, in order to effectively assess various alternatives a strategy for comparison was required. For example, how much pain, or risk, is it worth to try to lay down a prey? In order to respond optimally to this type of situation, the organisms needed a 'common currency' for positive and negative, or 'good' and 'bad' (McFarland and Sibly 1975). The chosen currency is what I refer to as the hedonic value component of feelings, and is implicit in the terms 'reward' and 'punishment.' The amniote brain considers the net outcome of various actions (the sum of positive and negative expectations), and hence the presumed optimal survival outcome (Cabanac 1992).

In other words, feelings presumably originated as a further elaboration of the neurobiological processing taking place between the sensory system and the executive branch of the nervous system. Feelings add value to the information obtained. The value is positive (pleasant) in the case where approach behavior is appropriate, and negative (painful or otherwise unpleasant) if avoidance is called for. The strength and duration of the expected feelings determine their worth. The score given to various options is based on a composite of innate tendencies and previous experience; e.g., humans may have an innate propensity to fear snakes (which implies a punitive feeling), though we can learn that certain snakes do not harm us.

Note that this strategy requires two attributes of the brain: one, to weigh alternatives based on hedonic value; and two, to translate the inference into action by generating motivation based on pleasure maximization (Cabanac 1992). It may be possible to conceive of ways to achieve similar performance without the use of feelings, but feelings appear to be a rational choice; moreover, it was presumably the choice opted for by evolution in the case of the amniotic lineage.

As to the present discussion, the core point is that for feelings to work, or make any sense as a currency to respond to, a capacity to assess (and hence experience) their positive and negative value is required. Most invertebrates respond to sensory input, but presumably not by weighing hedonic value. It is difficult to envision how feelings could function as "a currency for decision making" without an awareness component. *I surmise that the requirement for that awareness was the cue that engendered the emergence of consciousness.*

Why the Amniotes?

Evolution has moved in the direction of radically increased complexity of nervous systems primarily in three phyla: *Chordata* (vertebrates), *Arthropoda* and *Mollusca*. The development can be seen as a consequence of an evolutionary 'arms race:' if one species improves its fitness by evolving more elaborate, or more flexible, behavior; interacting species needs to follow suit in order to survive. Feelings, compared to a more innate response, increased the flexibility and adaptability of behavior, though in the early stages not necessarily its complexity.

Although complex behavior is evident in present arthropods and mollusks, evolution may not have introduced consciousness. In other words, it seems a bit simplistic to assume that consciousness is a consequence of an increased computational capacity. The following presumed features of the early amniotes and their environment may help explain why the attribute emerged in terrestrial vertebrates:

1. At the time, the amniotes most likely possessed a more sophisticated central nervous system compared to the other two phyla mentioned above, and thus a better starting point for further elaborations.
2. They were relatively large animals with long generation times and small litters, which implies that they evolved slowly. Consequently, adaptation to novel environments relied to a greater extent on individual adaptability, rather than on genetic modification.
3. They evolved (advanced) lungs as an adaptation to terrestrial life. The brain is an expensive organ to operate, requiring a substantial portion of the energy (and oxygen) an organism can procure (Mink et al. 1981). Introducing a system of feelings, and concomitant awareness, as a strategy for complex decision-making presumably provided a considerable push in the direction of a larger and more demanding brain. The concentration of oxygen is much higher in air

compared to water, but of the terrestrial animals only amniotes developed an efficient breathing organ.

4. Although it seems likely that early terrestrial ecosystems were less complex than contemporary marine environments, the situation may have *changed* more rapidly on land, implying a greater selection pressure for behavioral flexibility.

I believe these four points may have contributed to the emergence of consciousness, but one should also take into account a possible stochastic element in the evolutionary process. As mentioned earlier, elaborate *non-conscious* behavior is possible, and a decision-making strategy employing feelings was almost certainly due, in part, to chance—or providence.

Further Elaborations of Consciousness

For the early amniotes, awareness presumably hinged on the assessment of behavioral options based on their hedonic value. The strategy proved to be successful, and evolution has since elaborated extensively on this first, simple version of (primary) consciousness. The elaborations adapted to the requirements of the various species; for example, olfactory signals play a prominent role in the conscious life of a dog.

In the human lineage, attributes such as self-awareness, culture, language, advanced cognitive power, and the curious sense of free will enhanced the experience of life. Concomitantly, the original function became less obvious, as the conscious brain evolved into a partly independent unit with ‘a life of its own.’ The subconscious presumably directed ever more information to the conscious brain, as more information would imply a better foundation for decision-making—limited primarily by available brain capacity. The original pleasure or displeasure dichotomy became obscured, as today the human experience of life is based on a smear of sensory input mixed with memories and thoughts that have none, or limited, hedonic value. Consciousness is often active even in the apparent absence of any (obvious) emotional valence. The integration of various sensory and cognitive information appears more important, and decisions are to a larger extent based on cognition, taking long term objectives into account, rather than on feelings alone.

Yet, consciousness has its shortcomings, which may explain why a substantial portion of the brain’s processing capacity is retained by the unconscious. For example, only select parts of the sensory input meant to monitor internal and external environments are sent to the conscious brain, most of the signals received by sensory organs are filtered away. The constraints on consciousness also explain why following intuition sometimes (for example in the case of solving problems and in fine-tuned muscle movements) is a better strategy than exercising deliberate control.

The main drawbacks of consciousness are as follows:

1. It is a relatively slow process. Conscious perception of a stimulus requires 100-200 ms (Crick and Koch 2003).
2. It is probably energy intensive; thus, a more automated response conserves nutrients and oxygen in cases, such as the regulation of heartbeat, where awareness cannot add any meaningful input to the response.
3. While the subconscious most likely works as a parallel processor, organizing several tasks simultaneously (e.g., heartbeat and temperature regulation), the conscious brain can only handle one task at a time; if additional tasks require conscious input, it is necessary to shift back and forth between them (Baars 1997).
4. According to the present model, consciousness evolved for decision making, not for execution; thus the process does not have the power to deal with tasks such as how to orchestrate optimal performance of legs and arms.
5. Cognition is vulnerable to the whims of the individual. For the sake of the genes, flexibility comes with the price of uncertainty.
6. Feelings and awareness were only generated in cases where they made evolutionary sense. For example, we do not feel a tumor unless it happens to press on nerve cells installed for other purposes; because during our evolutionary history, being aware of a tumor would not have helped.

Due to these limitations, consciousness is not the sole, or even prime, ‘mover’ of behavior; instead most human behavior stems from a mixture of conscious and subconscious processing (Kunde et al. 2003; Cabanac and Bonniot-Cabanac 2007; Pessiglione et al. 2008; Baumeister et al. 2011).

Neurobiological Support for the Present Model

Neural Correlates of Consciousness

The evolutionary scenario presented above suggests the following predictions: One, consciousness and feelings have related neurobiological features (as to neuroanatomy and/or neurochemistry), as they appeared at the same time and for a shared purpose; and two, if they evolved to care for the basal process of approach or avoidance, the core circuitry involved might be situated in the more ancient parts of the brain. It is worthwhile to consider whether these implications are supported by data.

The neurobiology of consciousness is elusive, presumably because it relies on constant communication between widely dispersed nerve circuits, rather than on the localized ‘off-or-on’ activity of a particular center. One view that has gained broad acceptance is that the main anatomical components are within the thalamocortical complex, which may include the basal ganglia and possibly other parts of the forebrain (Crick and Koch 2003; Edelman and Seth 2009; Cabanac et al. 2009; Ward 2011). In this view, consciousness depends on the continuous chattering of circuits within the thalamocortical complex (Alkire et al. 2008; Noirhomme et al. 2010). More specifically, our experiences may reflect perturbations on a background of more regular, spontaneous activity (Buzsaki 2007); and attention may be a question of which of a variety of nascent, perturbation-causing nerve cell coalitions gain dominance at any given moment (Crick and Koch 2003).

Somewhat surprising is the recent suggestion that consciousness may be independent of either intact cortex or thalamus. Hydranencephalic children, i.e., humans born without cortex (or with minimal remnants thereof), appear to be conscious (Merker 2007; Beshkar 2008), as do animals in which cortex or (possibly) thalamus are removed (Panksepp et al. 1994; Alkire et al. 2008). In such instances, it is conceivable that remaining structures of the forebrain—particularly components of the basal ganglia such as the nucleus accumbens, ventral pallidum, and striatum—are sufficient for generating primary conscious states. Alternatively, the brain might compensate for the absence or loss of cortex by delegating functions to available nervous tissue.

A reasonable model based on the above discourse is that the functions regulating consciousness are associated with subcortical structures, perhaps in particular the intralaminar nuclei of the thalamus (Alkire and Miller 2005; Jones 2001). Here, direct injections of agonists to the generally inhibitory neurotransmitter GABA cause rapid sedation in rats (Miller and Ferrendelli 1990), a patient in minimal conscious state for 6 years improved drastically after stimulation of these nuclei by electrodes Schiff and Fins 2007), and thalamic damage in humans can result in a vegetative state, while restoration of consciousness is associated with restoration of functional connectivity between thalamus and (cingulate) cortex (Alkire et al. 2008). Moreover, the associated thalamic reticular nucleus has been implicated in schizophrenia, a disturbance of consciousness (Ferrarelli and Tononi 2011); and related structures in the hypothalamus apparently plays a similar, central role in the regulation of sleep (Szymusiak and McGinty 2008; Gvilia 2010).

The cortex presumably adds substance and content, not only to conscious experience, but also to dreams (Nir and Tononi 2010). Other structures, such as the claustrum (Crick and Koch 2005), may help in the process of gathering and integrating information from different parts of the brain.

Neural Correlates of Feelings

As for feelings, it seems reasonable to divide the parts of the brain involved in generating hedonic value into three main modules: one for punishment, and two for rewards, i.e., seeking (or wanting) and liking (or consuming) (Panksepp 1998; Kringelbach and Berridge 2009). Recent data suggest that these three modules to a large extent use the same brain structures; that is, all types of punishment and reward—whether from food, sex, burns, social relations, etc.—converge on shared neural substrates for the generation of hedonic value (Leknes and Tracey 2008; Tabibnia et al. 2008; Lieberman and Eisenberger 2009; Takahashi et al. 2009; Berridge and Kringelbach 2011). Again, particular regions of the cortex (e.g., prefrontal, orbitofrontal, and insular cortex) may act as a sort of dashboard to add ‘flavor’ and distinctiveness to various rewards and punishments, while subcortical structures—including areas associated with the basal ganglia, the amygdala, and thalamus—act more like a ‘motor,’ generating the hedonic quality (reviewed in (Grinde 2012)). The main ‘hedonic hotspots,’ in which direct stimulation can cause activation (in the form of enhanced pleasure) upon relevant stimulation (via electrodes or local injection of neurotransmitter modulators), are found in nucleus accumbens and pallidum (Pecina 2008; Smith et al. 2010); while stimulation of certain areas of the thalamus can inhibit pain (Bittar et al. 2005).

Comparison

Dopamine appears to play a central role in the seeking type of rewards (Barbano and Cador 2007; Leknes and Tracey 2008) as well as in consciousness (Lou 2011; Palmiter 2011). The considerable increase in telencephalic (the major part of the forebrain) dopamine receptors in reptiles compared to amphibians is taken as a further indication that consciousness first evolved in the amniotes (Cabanac et al. 2009).

The foregoing observations suggests that the core, or regulatory, circuitry for both feelings and consciousness is situated in basal parts of the brain; moreover, they point toward a considerable overlap between the neurobiology of emotions and that of consciousness, which accords with the notion that the two evolved together. Specifically, the neurobiology of the two converge in the basal ganglia, perhaps in the diencephalon (the minor part of the forebrain), in the function of dopamine, and, of course, in the use of the cortex for elaborating the experience.

Journal : Large 13752
Dispatch : 20-8-2012
Pages : 10
Article No. : 61
☐ LE
☐ TYPESET
MS Code : BTHE-D-12-00007
☑ CP
☑ DISK

It should be noted that even if the emergence of feelings spurred awareness, the two features have been molded by evolution for some 300 million years, which is ample time for considerable divergence in neurobiology. Moreover, a shared neurobiology in the early amniotes, while expected, is not required. Still it is interesting to note that, although cognitive capacity is lacking in hydranencephalic children, they do seem to experience feelings, including pleasure and pain (Merker 2007; Beshkar 2008). This observation lends credence to the idea of shared neurobiological features for feelings and consciousness; although one should consider that the presence of feelings is one of the criteria used to assess consciousness, which limits the validity of the argument. It is not obvious to what extent these children retain the basal ganglia, which may prove more crucial than either cortex or thalamus.

Discussion

Consciousness Evolved

I have outlined a model for the evolution of consciousness suggesting that the feature first appeared for the purpose of experiencing feelings. The capacity to feel evolved as a strategy toward a more flexible and adaptive way of evaluating behavioral options. The model is based on the following considerations:

1. The core function of a brain is to make behavioral decisions, and these were, in most of our evolutionary history, primarily a matter of either approach or avoidance. This dichotomy is a characteristic feature of all nervous systems.
2. In order to compare the survival value of various approach and avoidance options, a “common currency” for positive and negative salience, i.e., hedonic value, is required (McFarland and Sibly 1975). The ensuing assessment allows for a more flexible and sophisticated response compared to what innate or learned patterns of behavior can deliver.
3. Feelings, in the form of positive and negative incitements (e.g., reward and punishment), seem to be a reasonable choice of currency. The two are weighed against each other in order to create the right motivation, implying that the brain will motivate the individual to act according to the principle of pleasure maximization (Cabanac 1992).
4. For feelings to make any sense, an awareness of good and bad, pleasure and displeasure, is required. There seems to be no other obvious requirement for conscious experience in (early) amniote evolution, and consciousness is apparently not required for complex behavior.
5. Sensory input provides the primary source of relevant information for behavioral decisions, and would therefore be expected to play a dominant role in delivering reward and punishment, and in the conscious experience of life. On the other hand, only select sensations, those pertinent for advanced decision making, engage the mood modules. Adding hedonic value is not required for a sense organ to trigger behavior, as exemplified by reflexes.
6. Various lines of evidence suggest that awareness and feelings evolved concurrently in early amniotes.
7. The amniote form of awareness, or primary consciousness, has been further elaborated by the evolutionary process into the more advanced, secondary conscious experiences of humans.
8. Consciousness and feelings have neurobiological features in common, as expected if the two evolved together for a shared purpose.
9. The core, regulatory circuitry appears to be situated in the basal, sub-cortical parts of the brain; as would be expected for an evolutionary expansion of the core function of nervous systems—i.e., to make decisions about approach or avoidance. Expansion of the cortical mantle presumably occurred later, and caused enrichment of content as to both feelings and consciousness.

Starting with early vertebrates, it is theoretically possible to envision the evolution of advanced, human-like behavior without introducing feelings as a currency for weighing alternatives—a purely cognitive assessment of options would, for example, be conceivable. Evolution did not follow this trajectory, perhaps because: For one, cognition was not sufficiently advanced to make this a viable strategy; and two, moving from fixed action patterns to learned behavior, and then on to motivation based on feelings, is a more probable evolutionary scenario. This scenario is in line with how evolution is known to work; i.e., changes in the genome typically reflect indirect means to direct the body the genes reside into promote their propagation. For example, in mammals the sexual drive, rather than a desire to have children, is sufficient to ensure fertilization.

The evolutionary trajectory leading to the human brain may be considered providential in that it offers us an experience of life (Baars 1997), and a capacity for happiness (Grinde, 2012). Fish and amphibians presumably lack this capacity; they respond to sensory stimuli, but may not *feel* pleasure or pain (the issue is discussed in (Rose 2007; Cabanac et al. 2009; Sneddon 2009)).

Human Consciousness

Dating back to the ancient Greek philosophers, there has been numerous ways to categorize and describe human

Journal : Large 13752 | Dispatch : 20-8-2012 | Pages : 10
Article No. : 61 | ☐ LE | ☐ TYPESET
MS Code : BTHE-D-12-00007 | ☑ CP | ☑ DISK

consciousness. I have mentioned the distinction between primary and secondary consciousness (Edelman 2004), as these terms are useful for the present model. Damasio (1999) prefers the term *self* as (partly) synonymous with secondary consciousness; i.e., as the personal experiences, thoughts and memories of an organism with the capacity for self-awareness. The self is further divided into core self and extended self; respectively, a stable representation of an individual's life, and the autobiographic information that accumulates in the mind. In the present biological model, however, this distinction seems somewhat arbitrary.

The following list is an attempt to use the present biological model to categorize the types of brain processes that are delegated to, and cared for by, the conscious part of the human brain:

1. Feelings (which here include the activity of the mood modules, as engaged by either emotions, sensations or cognition—i.e., all forms of affect).
2. Sensing (input from sensory organs that may or may not activate mood modules).
3. Cognition (thinking and related mental activity that may or may not activate mood modules).
4. Motivation and volition (initiators of actions based on the above three).

According to the present model, type 1 was the instigating rationale for the evolution of consciousness. The additional information deriving from sensory organs, i.e., type 2, may be brought to the conscious brain in order to secure that all relevant information is available for scrutiny. The subconscious does filter away the vast majority of signals reaching, for example, eyes and ears; but it would be difficult to install a filter that only left information of obvious relevance for making decisions, particularly as the conscious brain may be the best judge as to what constitute relevant information. Cognition, type 3, evolved gradually to improve the process of decision making, but eventually, in the human lineage, took the shape of an 'independent' feature of the brain. In fact, it has evolved to the point where decisions are made partly in the absence of, or in disregard of, the hedonic value of various options. The final type, number 4, is required as a link between feelings/cognition and actual behavior; but the relevant activity is not always brought to conscious attention.

It has been suggested that consciousness is simply an epiphenomenon, i.e., an incidental byproduct of selection in the direction of cognition and a more advanced brain (see, for example (Rosenthal 2008)). In my mind, the epiphenomenon model is less attractive for the following reasons: One, consciousness is a rather distinct and noticeable feature of the brain, and conspicuous features are in general unlikely to appear unless selected for; and two, it is possible to outline a scenario that depicts why evolution opted for consciousness (as exemplified by the present text). With a reasonable evolutionary rationale available, selection offers a more compelling explanation compared to a model describing the feature as an epiphenomenon. These arguments, however, do not rule out the possibility that the evolutionary trajectory leading to consciousness was in part characterized by coincidental events.

Most bodily features, somatic as well as mental, evolve to various states of sophistication in different lineages. Their final complexity is primarily a question of survival value. The nose, for example, is considerably more advanced in dogs compared to humans; while both consciousness and emotions presumably display their most elaborate forms in humans. I have previously proposed that, if so, humans may have the propensity to be the most happy (and most miserable) of any animal (Grinde 2012).

One factor hampering our efforts to understand consciousness may simply be that the human version of the feature has progressed far beyond the original state. So much information has been added to our 'film of life' that we do not easily sense the dichotomy of positive and negative feelings, which presumably dominated in the early amniotes. The neurobiology of the human brain reflects this advancement, making it difficult to identify the anatomical and neurochemical correlates of human consciousness. In this regard, reptilian brains may provide clues as to the nature of incipient substrates for early forms of primary consciousness.

Other Forms of Consciousness?

The success of combining feelings and consciousness in a strategy to generate sophisticated behavior begs the question of whether evolution may have opted for this combination more than once. The core elements of the nervous system—including the use of sensory cells, processing units, and muscles as effector organs—are present in most animals; thus convergent evolution in the direction of consciousness seems plausible. The three most successful metazoan phyla (*Chordata* (vertebrates), *Arthropoda* and *Mollusca*) all have sophisticated nervous systems and complex behavior. In fact, convergent evolution of advanced features is possible; eyes, for example, evolved independently (presumably from the shared starting point of light sensitive patches of skin) in these three phyla as a consequence of the obvious advantages of vision (Land and Nilsson 2002).

Among the invertebrates, the coleoid cephalopods (octopuses, squid, and cuttlefish) are considered prime candidates for consciousness (for reviews, see (Mather 2008; Edelman and Seth 2009)). These animals display advanced behavior, such as learning based on reward-like stimuli (Borrelli and Fiorito 2008), navigating mazes

(Moriyama and Gunji 1997), and possibly learning based on the observation of other members of the species (Fiorito and Scotto 1992). They can recognize a variety of objects and have considerable capacity for memory (Borrelli and Fiorito 2008; Hochner et al. 2006). In other words, their brains seem capable of a degree of processing and flexibility of behavior well beyond what one might expect from a collection of mere innate or learned behavioral patterns. Apparently they have evolved a level of sophistication, in terms of evaluating options, similar to that of amniotes. The key question as to whether they have anything resembling conscious experience may be whether evolution opted for the strategy of using feelings as a means to assess behavioral opportunities. Feelings seem to be a compelling choice, but there may be other options that are difficult for a human to conceive. If these creatures do possess an analogue to human consciousness, their 'film of life' must be quite different from what we experience.

Final Comment

Smith (2010) notes that while we have made considerable progress in understanding most aspects of the natural sciences, when it comes to understanding consciousness, we are no closer today than at the time of Darwin. I believe we do have a better grasp today, but one problem may be in communicating what we know to a wider audience. The issue of human consciousness is easily distorted by emotional sentiments, including metaphysical or religious ideas. Pope John Paul II, for example, has supposedly claimed that while scientists may have the brain, the mind belongs to God (Lane 2009). Biological explanations face not only challenges from the clergy, but also the problem of disseminating ideas effectively to disparate scientific and cultural traditions such as philosophy and the social sciences.

Acknowledgments The author would like to thank Michel Cabanac (Laval University, Quebec, Canada) and David B. Edelman (The Neurosciences Institute, San Diego, CA) for critical reading of the manuscript and for numerous highly valuable comments.

References

Alkire MT, Miller J (2005) General anesthesia and the neural correlates of consciousness. Prog Brain Res 150:229–244

Alkire MT, Hudetz AG, Tononi G (2008) Consciousness and anesthesia. Science 322:876–880

Baars B (1997) In the theather of consciousness. Oxford University Press, Oxford

Barbano MF, Cador M (2007) Opioids for hedonic experience and dopamine to get ready for it. Psychopharm 191:497–506

Baumeister RF, Masicampo EJ, Vohs KD (2011) Do conscious thoughts cause behavior? Ann Rev Psych 62:331–361

Berridge KC, Kringelbach ML (2011) Building a neuroscience of pleasure and well-being. Psych Well-Being 1:3

Beshkar M (2008) The presence of consciousness in the absence of the cerebral cortex. Synapse 62:553–556

Bittar RG, Kar-Purkayastha I, Owen SL, Bear RE, Green A, Wang S, Aziz TZ (2005) Deep brain stimulation for pain relief: a meta-analysis. J Clin Neurosci 12:515–519

Borrelli L, Fiorito G (2008) Behavioral analysis of learning and memory in cephalopods. In: Menzel R (ed) Learning theory and behavior. Elsevier, New York, pp 605–628

Butler AB (2008) Evolution of brains, cognition, and consciousness. Brain Res Bul 75:442–449

Butler AB, Cotterill RM (2006) Mammalian and avian neuroanatomy and the question of consciousness in birds. Biol Bul 211:106–127

Buzsaki G (2007) The structure of consciousness. Nature 446:267

Cabanac M (1992) Pleasure: the common currency. J Theor Biol 155:173–200

Cabanac M, Bonniot-Cabanac MC (2007) Decision making: rational or hedonic? Beh Brain Funct 3:45

Cabanac M, Cabanac AJ, Parent A (2009) The emergence of consciousness in phylogeny. Beh Brain Res 198:267–272

Crick F, Koch C (2003) A framework for consciousness. Nat Neurosci 6:119–126

Crick F, Koch C (2005) What is the function of the claustrum? Phil Trans Royal Soc London B 360:1271–1279

Damasio A (1999) The feeling of what happens: body and emotion in the making of consciousness. Harcourt, New York

Denton DA, McKinley MJ, Farrell M, Egan GF (2009) The role of primordial emotions in the evolutionary origin of consciousness. Conscious Cogn 18:500–514

Edelman G (2004) Wider than the sky: the phenomenal gift of consciousness. Yale University Press, Yale

Edelman DB, Seth AK (2009) Animal consciousness: a synthetic approach. Trends Neurosc 32:476–484

Ferrarelli F, Tononi G (2011) The thalamic reticular nucleus and schizophrenia. Schizophrenia Bul 37:306–315

Fiorito G, Scotto P (1992) Observational learning in octopus vulgaris. Science 256:545–547

Gould JL, Grant-Gould C (1995) The honey bee. Scientific American Library, New York

Grinde B (2012) The biology of happiness. Springer, Dordrecht

Gvilia I (2010) Underlying brain mechanisms that regulate sleep-wakefulness cycles. Int Rev Neurobio 93:1–21

Hochner B, Shomrat T, Fiorito G (2006) The octopus: a model for a comparative analysis of the evolution of learning and memory mechanisms. Biol Bul 210:308–317

Jones EG (2001) The thalamic matrix and thalamocortical synchrony. Trends Neurosc 24:595–601

Kitchen A, Denton D, Brent L (1996) Self-recognition and abstraction abilities in the common chimpanzee studied with distorting mirrors. Proc Natl Acad Sci USA 93:7405–7408

Kringelbach ML, Berridge KC (2009) Towards a functional neuroanatomy of pleasure and happiness. Trends Cogn Sci 13:479–487

Kunde W, Kiesel A, Hoffmann J (2003) Conscious control over the content of unconscious cognition. Cognition 88:223–242

Laland KN, Galef BG (2009) The question of animal culture. Harvard University Press, Harvard

Land M, Nilsson D-E (2002) Animal eyes. Oxford University Press, Oxford

Lane N (2009) Life ascending: the ten great inventions of evolution. WW Norton, New York

Leknes S, Tracey I (2008) A common neurobiology for pain and pleasure. Nature Rev Neurosc 9:314–320

Lieberman MD, Eisenberger NI (2009) Neuroscience pains and pleasures of social life. Science 323:890–891

Lou HC (2011) Paradigm shift in consciousness research: the child's self-awareness and abnormalities in autism, ADHD and schizophrenia. Acta Paediat 101:112–119

Marino L (2007) Cetacean brains: how aquatic are they? Anat Rec 290:694–700

Mather JA (2008) Cephalopod consciousness: behavioural evidence. Conscious Cogn 17:37–48

McFarland DJ, Sibly RM (1975) The behavioural final common path. Philos Trans Royal Soc London B 270:265–293

Merker B (2007) Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. Beh Brain Sci 30:63–81

Miller JW, Ferrendelli JA (1990) Characterization of GABAergic seizure regulation in the midline thalamus. Neuropharma 29:649–655

Mink JW, Blumenschine RJ, Adams DB (1981) Ratio of central nervous system to body metabolism in vertebrates: its constancy and functional basis. Am J Phys 241:R203–R212

Moriyama T, Gunji YP (1997) Autonomous learning in maze solution by octopus. Ethology 103:499–513

Mosley C (2011) Pain and nociception in reptiles. Vet Clin N Am Exotic Ani Prac 14:45–60

Nir Y, Tononi G (2010) Dreaming and the brain: from phenomenology to neurophysiology. Trends Cogn Sci 14:88–100

Noirhomme Q, Soddu A, Lehembre R, Vanhaudenhuyse A, Boveroux P, Boly M, Laureys S (2010) Brain connectivity in pathological and pharmacological coma. Front Syst Neurosci 4:160

Palmiter RD (2011) Dopamine signaling as a neural correlate of consciousness. Neurosci 198:213–220

Panksepp J (1998) Affective neuroscience. Oxford University Press, Oxford

Panksepp J, Normansell L, Cox JF, Siviy SM (1994) Effects of neonatal decortication on the social play of juvenile rats. Phys Behav 56:429–443

Pecina S (2008) Opioid reward 'liking' and 'wanting' in the nucleus accumbens. Phys Behav 94:675–680

Pessiglione M, Petrovic P, Daunizeau J, Palminteri S, Dolan RJ, Frith CD (2008) Subliminal instrumental conditioning demonstrated in the human brain. Neuron 59:561–567

Prior H, Schwarz A, Gunturkun O (2008) Mirror-induced behavior in the magpie (Pica pica): evidence of self-recognition. PLoS Biol 6:e202

Reiss D, Marino L (2001) Mirror self-recognition in the bottlenose dolphin: a case of cognitive convergence. Proc Natl Acad Sci USA 98:5937–5942

Rose JD (2007) Anthropomorphism and 'mental welfare' of fishes. Dis Aqua Org 75:139–154

Rosenthal DM (2008) Consciousness and its function. Neuropsychologia 46:829–840

Rossano MJ (2003) Expertise and the evolution of consciousness. Cognition 89:207–236

Schiff ND, Fins JJ (2007) Deep brain stimulation and cognition: moving from animal to patient. Cur Opin Neurol 20:638–642

Seth AK, Baars BJ, Edelman DB (2005) Criteria for consciousness in humans and other mammals. Conscious Cogn 14:119–139

Smith CU (2010) Darwin's unsolved problem: the place of consciousness in an evolutionary world. J Hist Neurosci 19:105–120

Smith K, Mahler S, Pecina S, Berridge K (2010) Hedonic hotspots: generating sensory pleasure in the brain. In: Kringelbach M, Berridge K (eds) Pleasures in the brain. Oxford University Press, Oxford, pp 27–49

Sneddon LU (2009) Pain perception in fish: indicators and endpoints. ILAR J 50:338–342

Szymusiak R, McGinty D (2008) Hypothalamic regulation of sleep and arousal. Ann New York Acad Sci 1129:275–286

Tabibnia G, Satpute AB, Lieberman MD (2008) The sunny side of fairness—preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). Psychol Sci 19:339–347

Takahashi H, Kato M, Matsuura M, Mobbs D, Suhara T, Okubo Y (2009) When your gain is my pain and your pain is my gain: neural correlates of envy and schadenfreude. Science 323:937–939

Tamietto M, de Gelder B (2010) Neural bases of the non-conscious perception of emotional signals. Nat Rev Neurosci 11:697–709

Ward LM (2011) The thalamic dynamic core theory of conscious experience. Conscious Cogn 20:464–486

Zhang Y, Lu H, Bargmann CI (2005) Pathogenic bacteria induce aversive olfactory learning in *Caenorhabditis elegans*. Nature 438:179–184

Journal : **Large 13752**
Dispatch : **20-8-2012**
Pages : **10**
Article No. : **61**
☐ LE
☐ TYPESET
MS Code : **BTHE-D-12-00007**
☑ CP
☑ DISK