

**Cite this article:** LeDoux JE. 2021 As soon as there was life, there was danger: the deep history of survival behaviours and the shallower history of consciousness. *Phil. Trans. R. Soc. B* **377**: 20210292. <https://doi.org/10.1098/rstb.2021.0292>

Received: 31 August 2021  
Accepted: 18 October 2021

One contribution of 16 to a theme issue  
'Systems neuroscience through the lens of  
evolutionary theory'.

**Subject Areas:**  
cognition


**Keywords:**  
fear, consciousness, evolution, survival circuits

**Author for correspondence:**  
Joseph E. LeDoux  
e-mail: ledoux@cns.nyu.edu

# As soon as there was life, there was danger: the deep history of survival behaviours and the shallower history of consciousness

Joseph E. LeDoux

Center for Neural Science, New York University, New York, NY 10003, USA

 JEL, 0000-0001-8518-132X

It is often said that fear is a universal innate emotion that we humans have inherited from our mammalian ancestors by virtue of having inherited conserved features of their nervous systems. Contrary to this common sense-based scientific point of view, I have argued that what we have inherited from our mammalian ancestors, and they from their distal vertebrate ancestors, and they from their chordate ancestors, and so forth, is not a fear circuit. It is, instead, a defensive survival circuit that detects threats, and in response, initiates defensive survival behaviours and supporting physiological adjustments. Seen in this light, the defensive survival circuits of humans and other mammals can be conceptualized as manifestations of an ancient survival function—the ability to detect danger and respond to it—that may in fact predate animals and their nervous systems, and perhaps may go back to the beginning of life. Fear, on the other hand, from my perspective, is a product of cortical cognitive circuits. This conception is not just of academic interest. It also has practical implications, offering clues as to why efforts to treat problems related to fear and anxiety are not more effective, and what might make them better.

This article is part of the theme issue 'Systems neuroscience through the lens of evolutionary theory'.

## 1. Introduction

That behaviour and evolution are interrelated is hardly a novel idea. Darwin emphasized it, as did pioneering ethologists such as Niko Tinbergen and Konrad Lorenz. The behaviourists who dominated psychology in the first half of the twentieth century paid little attention to evolution, but most contemporary psychologists and neuroscientists accept it as a key factor that must be accounted for to explain behaviour.

Efforts to understand the evolution of behaviour in neuroscience often compare closely related groups, such as humans and other mammals, or even other vertebrates. There are obvious reasons for doing so. For example, since the brain controls behaviour, studies of how brains evolved can help shed light on how behavioural repertoires evolved. But there are also reasons to look deeper [1,2].

I have spent much of my career working on the brain mechanisms of emotion, and especially how the brain detects and responds to danger. Roughly a decade ago, I began to see my work in a new light. This transformation is summed up by the title of the present article, which is a paraphrase of a quote from the American essayist and poet, Ralph Waldo Emerson [3]. In my version of Emerson's comment, I changed both instances of 'is' to 'was'. This stems from the realization that I had come to a few years ago—that danger must be as old as life itself. This epiphany may seem obvious to those who have thought long and hard about evolution. But it was an eye-opener to the mammalian neuroscientist in me, and prompted me to explore the natural

history of danger in more detail, which I did in my 2019 book, *The deep history of ourselves: the four-billion-year story of how we got conscious brains* [2]. Here, I summarize the path that took me to the science of danger in the first place, and why the latter led me to write *The deep history of ourselves*. I also consider the implications for our understanding of our emotions and their maladies.

## 2. Finding my way to danger

In the late 1970s, I completed my PhD studying consciousness in split-brain patients. From this research, my mentor, Michael Gazzaniga, and I concluded that an important feature of human consciousness is the maintenance of a sense of mental unity via cognitive interpretations that attribute meaning and cause to behaviours that are controlled non-consciously by the brain [4,5]. We speculated that emotional systems might be examples of systems that non-consciously generate behavioural responses, compelling some sort of interpretive narration to sustain a sense of mental unity. The fact that not much was known about emotions at the time made it an attractive option for me, given a comment made by another professor, the comparative anatomist, Harvey Karten, in a graduate seminar on the vertebrate brain. Karten suggested that students should avoid popular research topics, such as the role of the visual cortex in perception or the hippocampus in memory. We should, instead find an interesting but less well-travelled scientific road.

After graduate school, I therefore turned to the brain mechanisms of emotion. Because the tools available for studying the human brain were quite limited at the time, I chose to study how the brain controls emotional behaviours in rodents.

It's not as if the brain mechanisms of emotional behaviour had never been researched. Emotion was actually a popular scientific topic among physiologists in the first half of the twentieth century [6–11]. But when the field of neuroscience emerged as an independent discipline in the 1970s, cognition was the rage, and brain researchers were more interested in the mechanisms underlying processes such perception, attention and memory, than in emotion.

Although the earlier research had laid an important foundation for understanding the emotional brain, the knowledge was piecemeal, and mostly focused on what brain areas do, as if functions were literally localized in areas. The split-brain studies had taught me to think of information flow through the brain, and my ambition was to map the information flow underlying emotional behaviour from sensory receptors to motor outputs. Invertebrate researchers, such as Eric Kandel, had achieved this kind of sensory to motor mapping of behaviour in the invertebrate nervous system [12], and I thought that new research tools becoming available might make this doable in the brains of mammals as well.

## 3. Circuit busting fear conditioning

I adopted the behavioural procedure called Pavlovian fear conditioning [13] (this was also one of the tools Kandel had used). I was not particularly interested in fear, but the fear conditioning paradigm seemed to have some useful properties for studying emotional behaviour. For one thing, its associative underpinnings are the bread-and-butter mechanism by which organisms learn about stimuli that co-occur

with biologically significant events. The result is that the previously meaningless stimuli come to trigger species-typical behavioural and physiological responses, allowing the organism to respond to dangerous stimuli in advance of their harmful consequences. Also, the responses elicited in lower mammals are similar to those expressed in humans when in harm's way. Further, much was known about the behavioural principles underlying Pavlovian learning. And finally, consistent with Karten's advice, next to nothing was known about the neural mechanisms in mammals.

I submitted a grant proposal to map emotional behaviour pathways in the rat brain using the Pavlovian paradigm, but it was not funded because 'emotion is not a neuroscientific topic'. The way I ended up getting funding was by re-framing my work as being about the neural basis of learning.

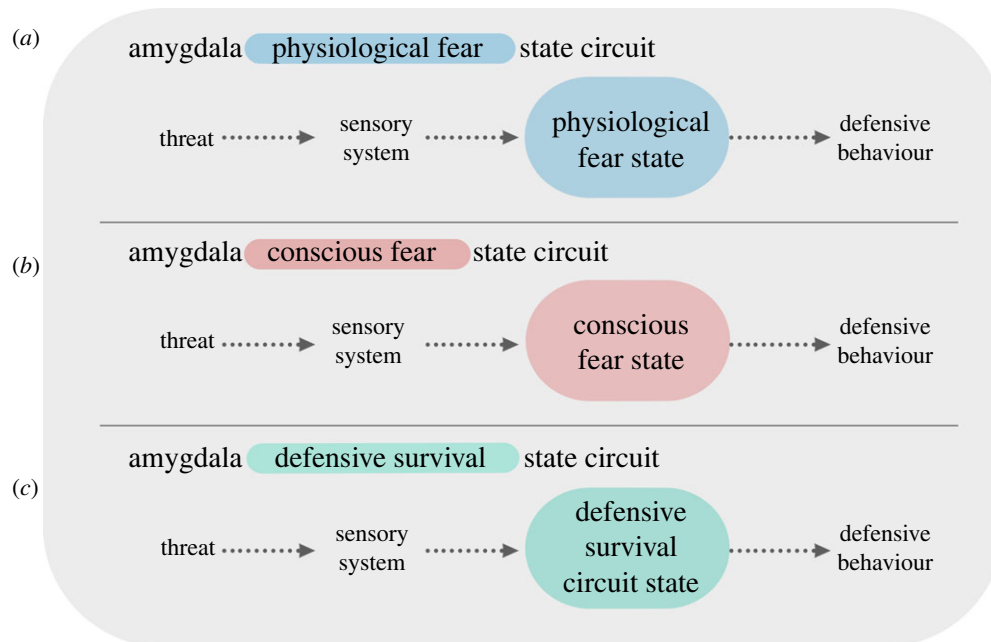
To map the connections of behavioural and physiological responses that were elicited by Pavlovian conditioned stimuli, I used axonal transport pathway tracing (which I had picked up while doing a rotation in Karten's lab), as well as some other methods that I learned 'on the job', such as chemical lesion approaches and electrophysiological recording techniques, especially in freely behaving animals.

I recall a Society for Neuroscience poster session in the mid-1980s where there were hundreds of presentations on the hippocampus and memory, and very few (maybe three or four) on the brain mechanisms of fear conditioning. Although our ranks were small [14–16], within a few years, we fear conditioning researchers had shown that sensory inputs to, and motor outputs from the amygdala are responsible for the behavioural and physiological expression of conditioned fear responses. And importantly, within the amygdala, a systematic pattern of synaptic connectivity from the input region (i.e. lateral nucleus) to the output region (i.e. central nucleus) was revealed. Further, while synaptic plasticity was found to occur throughout the amygdala circuitry, plasticity in the lateral amygdala seemed particularly important, based on the short latency of the neural changes and their necessity in supporting plasticity in other areas. Summaries of this body of work appeared in a 1994 article I wrote for *Scientific American* [17], in my 1996 book, *The emotional brain* [18] and in an *Annual Review of Neuroscience* paper in 2000 [19].

By the late 1990s, many researchers were becoming interested in the neural mechanisms of Pavlovian fear conditioning. With the basic circuits outlined, the effort was turning to the molecular mechanisms that might underlie learning. Some of the new researchers were from the usual areas of psychology and brain research, but others were from genetics, where the ability to alter genes was emerging. Over the next decade or so, accumulated findings implicated specific neurotransmitter receptors, ion channels, protein kinases and macromolecules in learning and plasticity in the amygdala of rodents [20–23]. Many of these findings were guided by earlier results about learning and memory storage in invertebrates [24,25], a point that will become relevant below.

## 4. Three views of fear

As I had hoped when I started this work, the fear conditioning procedure turned out to be an excellent way to elucidate how behavioural and physiological responses elicited by emotional stimuli are learned and controlled non-consciously by the brain. To emphasize my idea that amygdala circuits processed information non-consciously, I adopted the implicit/explicit



**Figure 1.** Amygdala fear versus defensive survival circuit views. (Online version in colour.)

distinction popular in learning and memory research [26]. Specifically, I referred to the amygdala as an implicit processor that controls fear behaviours non-consciously, with the explicit conscious experience of fear being a product of cortical cognitive circuits [18].

Findings obtained in studies of humans provided support for the idea that the amygdala non-consciously controls behavioural and physiological responses elicited by threats, turning one of my split-brain ideas into a reality. For example, subliminal presentation of threats resulted in amygdala activation and elicited physiological responses without the person having any awareness of the stimulus or reporting any feeling of fear [27,28]. Later studies showed that even when pressed, healthy participants did not report emotional feeling states in response to subliminal emotional stimuli [29]. A related line of research showed that in the so-called blindsight patients threats also elicited amygdala activity and body responses without conscious awareness of the stimulus [30]. And studies of patients with amygdala damage showed that they could still, under some conditions, have emotional experiences [31,32]. Fear, it seems, is not inextricably tied to the amygdala in humans. The amygdala, in other words might contribute to, but does not determine, the mental state of fear.

Most researchers working on fear conditioning came to this topic from a behaviourist intellectual background. I was once asked by one such colleague, 'Why do you talk about emotion?' He said 'we just study behavioural learning'. Given this, why did behaviourists ever call the procedure 'fear' conditioning?

Although the behaviourists had banned subjective states as explanations of behaviour, they did not eliminate the use of subjective state words in explaining behaviour. They referred to the task as 'fear' conditioning because fear for them was a description of the functional relationship between a dangerous stimulus and defensive behaviour. And when some behaviourists became physiological psychologists in the 1950s, the functional relation came to be thought of as a physiological state, which they called 'fear', that connected the threatening conditioned stimulus to the defensive conditioned response. The amygdala, of course, became the home of that that mindless fear state [14–16,33,34].

But some researchers rebelled against the sterile conception of behaviour offered by behaviourist approaches [35–40]. The emotion researchers among them promoted the idea that rodent fear was more or less like human fear, and therefore that animal studies could be used to understand human emotions. These anti-behaviourist brain researchers treated subcortical brain areas, like the amygdala and periaqueductal gray, as being responsible not just for defensive behaviours but also for the subjective feeling of fear in animals and humans alike [35].

Both groups of researchers, in other words, used the term 'fear' to describe functions subserved by subcortical circuits, but they used the term in different ways. The behaviourists–learning crowd did not bother to explain what they meant. For them, conscious fear was a useless construct that had been successfully purged from scientific discourse, and any serious scientist knows that 'fear' is simply a physiological state that controls behaviour. The anti-behaviourists, on the other hand, made it very clear that fear was conscious fear. The general public, and even other scientists who were not 'in the know', did not realize the difference, and simply assumed that when a scientist wrote or said 'fear' they meant conscious fear. By failing to explicitly define their terms, the behaviourist crew's message was lost.

I was caught in the middle [18,19,41–43]. I was aligned with my behaviourist-oriented colleagues in emphasizing that physiological states within amygdala circuits control defensive behaviour rather than make fearful feelings. But like the anti-behaviourists, I treated conscious fear as real and important. Yet, I viewed conscious fear as a product of cognitive processes instantiated in cortical circuits, rather than as a hard-wired function of subcortical areas [18,41,42] (figure 1).

## 5. Rethinking the emotional brain

Over the years, the amygdala went from being an obscure brain area of interest to only a handful of neuroscientists to cultural meme—the 'amygdala fear centre'. This meme appeared in novels, movies, songs, self-help books, cartoons and online

merchandise from tee-shirts to amulets to essential oils. Despite my efforts to draw a distinction between conscious fear and its behavioural and physiological correlates, my work was often used as scientific support for the meme, and I was often introduced at scientific and lay lectures as having discovered how conscious feelings of fear arise from the amygdala.

By 2012, it was clear to me that the loose, casual use of common language terms, especially psychological terms, as names for scientific constructs, can have a negative impact on our understanding of scientific findings and their implications. To be perfectly honest, though, I was not always as clear as I now realize I should have been. I resolved to find a better way to write and talk about fear and the brain, both to other scientists and to lay audiences.

I published scientific articles with titles like, 'Rethinking the emotional brain', 'Coming to terms with fear' and 'Semantics, surplus meaning and the science of fear' [44–46], and books called *Anxious* [47] and *The deep history of ourselves* [2]. In these publications, I argued that the use of mental state terms, like 'fear', to describe circuits that control behavioural and physiological responses conflates correlation with causation. For example, fear often occurs when we freeze or flee in a dangerous situation, but not because a feeling of fear that bubbles up out of the amygdala is the underlying cause of the responses. Although the subjective experience and the objective responses are both triggered by the same threat stimulus, I propose that they are separate consequences mediated by different outputs of the sensory system. One output takes the threat stimulus information to the amygdala to control behaviour and physiology, and the other takes it to cortical cognitive circuits as part of the process of assembling the conscious state we know of as fear.

In these publications, I emphasized that scientists should strive for conceptual clarity, and I offered a simple solution in the case of emotions. I suggested that we reserve the use of mental state terms, like 'fear', for the mental state that the word names, as opposed to using such terms to describe behavioural and physiological correlates of the conscious experience of fear. For example, rather than treating amygdala circuits activated by threats as 'fear' circuits of either kind (physiological or subjective), I proposed that they be called 'defensive survival circuits' (figure 1). After all, they control both defensive behaviours and the physiological responses that provide metabolic support for the behaviours in dangerous situations. I also suggested that the procedure called 'fear conditioning' be referred to as 'threat conditioning'.

The fact is that 'threat' and 'defense' were already commonly, though inconsistently, used in the field. All I did was suggest that we use the psychologically more neutral terms, instead of the one's with subjective implications, when referring to behaviour and body physiology, saving 'fear' for the mental state. There has been some progress, but also resistance. A colleague wrote a commentary about one of my papers [48] saying that following the ideas I have been promoting would take psychiatry back to the dark time when subjectivity ruled the field [49]. As I will argue below, though, the marginalization of subjective experience is why the treatment of mental disorders is not more effective.

## 6. The language of psychology

I am hardly the first to raise questions about scientific terminology of psychology [1,50–55]. Critics have noted that psychologists tend to be sloppy with words; that

category names can create illusions of understanding; and that the use of common language terms as names for scientific constructs can skew our understanding of the underlying processes. For example, calling the amygdala a fear generator [49], regardless of whether a physiological or subjective meaning of fear is intended, infects this brain area with subjective properties that it may well not deserve.

It is especially important to be careful when we extrapolate across species based on similarity in behaviours in humans and other animals. When we do this, we are using our introspections about the relation of our mental states to our behaviours to explain animal behaviour. The quote below from the pioneering ethologist, Nico Tinbergen, succinctly summarized the problem some 70 years ago:

'Hunger, like, anger, fear, and so forth, is a phenomenon that can be known only by introspection. When applied to another. . . species, it is merely a guess about the possible nature of the animal's subjective state'. [56, p. 5]

Being careful about the use of common language terms as names for scientific constructs is more important in the psychological sciences, including psychological aspects of neuroscience, than in other areas. For example, no one actually thinks that there is a hedgehog living in the gene of that name. But people do believe that fear lives in the amygdala 'fear centre'.

I agree with those who propose that the traditional psychological categories used to conceptualize behaviour can be problematic [1,57,58]. But I do not think this applies across the board. Problems especially arise when mental state words derived from human introspection [53] are used to talk about behaviours that do not depend on mental states [2,45–47], both in humans and other animals. In such cases, brain circuits that control such behaviours inherit the mental implications of the common language terms.

That said, there is a place for everyday words about mental states in the psychological and brain sciences. They are the coin of the mental realm, and are necessary when talking about our mental lives, even scientifically [51,57,59,60]. Summarizing such views about the special role of common language in psychology, I wrote this in my 2015 book *Anxious* [47, p. 40]:

Physicists, astronomers, and chemists don't need to take seriously commonsense ideas about nature because people's beliefs and attitudes about the stars, matter and energy, and chemical elements don't affect the subject under investigation. The fact that we commonly say (and some may actually believe) that 'the sun rises in the east' does not have any scientific bearing on the fact that sunrise is an illusion. But psychologists do have to pay attention to folk psychology because people's common beliefs about the mind influence their thoughts and actions in daily life and are thus an important part of what psychology is all about. Folk psychology is a window into the things that interest people and affect their lives.

## 7. Deep survival

Amygdala defensive survival circuits are present throughout vertebrate species [61–64]. One proposal is that this is a derived trait that emerged in fish tens of thousands of years after the first fish arrived [65,66]. Another possibility is that early vertebrates, in fact, had an amygdala homolog. That this may be the case is suggested by the work of Sten Grillner and colleagues on the jawless fish, lamprey, one of the oldest living vertebrates [67,68]. In fact, their findings show striking similarities between the entire forebrains of lamprey and all other vertebrates. Recent findings showing



that the vertebrate telencephalon evolved from a brain precursor region in ancestral chordates implies homology of these structures. Given this, it is possible that the amygdala, a region of telencephalon, may also have originated from this precursor region [69–71].

Most invertebrates, though, belong to a separate line of descent. Although their nervous systems are not equipped with an amygdala, they do behave defensively, and have their own defensive survival circuits. For example, circuits have been identified that control freezing in flies [72]. This finding was unfortunately described as being functionally similar to ‘fear’ circuits in mammals and, as such, suggested that it may be able to shed light on aspects of human ‘fear’. This kind of prose invites the conclusion by journalists that flies may be emotional beings with conscious feelings of fear similar to ours [73]. This is an excellent example of how using the survival circuit terminology could have helped avoid unintended misrepresentation in the press of otherwise important research.

Consistent with the idea of conserved survival circuitry in diverse animals is the fact that, as mentioned above, some genes and molecules involved in implicit learning about and control of behaviours in the presence of threat are conserved in vertebrates and invertebrates. This does not necessarily mean that having the genes shows that the function is conserved. But the correlation between the presence of the genes and the function is at least suggestive that similarities in implicit forms of learning in vertebrates and invertebrates may be due to a common set of synaptic plasticity-related genes inherited from a common basal ancestor. Research using modern genetic tools, in fact, supports this suggestion [74–76].

But the connection runs deeper. Some of these plasticity genes and molecules are present in radial organisms with diffuse, poorly centralized, nervous systems, like jellyfish and hydra, which express simple forms of learning [77]. Some of these genes and molecules also exist in sponges, which lack nervous systems, but may, long ago, have had one [78]. However, the focus on inherited features of nervous systems may be the wrong emphasis, as the trend predates neurons and synapses.

Single-cell protozoa share a common ancestor with animals. Being unicellular, protozoa could not have ever had a nervous system [78–81]. Yet they behave, approaching useful and avoiding harmful stimuli [82]. This is at least superficially similar to animal behaviour. Also like animals, they undergo simple forms of implicit learning about biologically significant stimuli [83], and possess some of the same plasticity genes and molecules that animals have [74–76]. And even bacteria approach useful and avoid harmful stimuli [84] and may even be capable of simple implicit learning [85], but this is not well established.

The key takeaway, though, is not about learning. It is about behaviour. Darwin’s protégé Romanes treated behaviour as an ambassador of the mind [86]. While this, of course, is true to some extent, behaviour is not fundamentally a psychological capacity [1,2]. It is simply one of the tools that organisms use to go about the day-to-day business of staying alive.

Bacteria and protozoa, for example, do not approach food out of hunger or ‘tumble’ away from harmful molecules out of ‘fear’. Such responses are defined by the physical fact that any movement of an object places it further from or closer to other objects in the environment, and by the physiological fact that such movements support survival in harmful and beneficial situations, respectively.

If withdrawal and approach are universal, life-sustaining capacities of all extant organisms [87], it follows that such behaviours may likely have existed as part of the homeostatic survival toolkit of early cells, which likely had to move around to manage their physiological viability in an ever-changing milieu—as soon as there was life, there was danger.

My inclination after reaching the above conclusions was to assume that animals inherited fundamental survival behaviours related to defense, feeding, fluid regulation and reproduction from unicellular microbes by virtue of genetic and molecular inheritance. But I realized there was another, less biologically daunting, way to think about this than to assume that complex defensive, feeding and reproductive behaviours of present-day animals were literally inherited from ancient microbes.

All organisms have to manage energy resources, regulate intracellular fluids, defend against harm and reproduce, but the way any one organism does these things depends on the kind of body its species evolved. For example, different mammalian species flee from danger by running, flying or swimming, depending on the kind of locomotory capacities that evolved with their bodies. In short, behaviours are body-plan-dependent (species-specific) survival implementations of universal (species-general) survival requirements. And survival circuits are the species-specific neural implementation by which the behavioural implementation of the universal requirement is achieved in each animal.

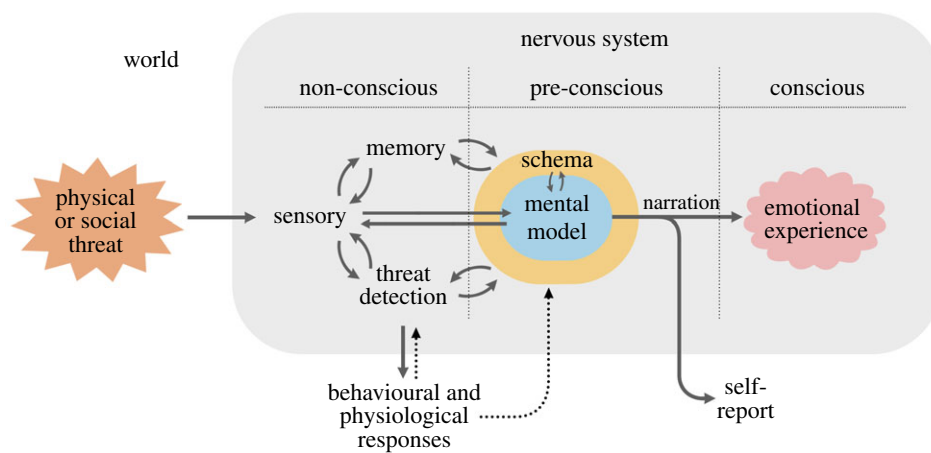
The point is, we should not be looking for how motor features of defensive or feeding behaviours were passed on from bacteria to archaea, and from them to protozoa and from them to animals. How animals manage survival motorically is essentially an artefact of their species’ body plan. Therefore, what we need to understand are the physiological substrata of survival functions that have been maintained across evolutionary transitions to ensure that animals possess body plan-appropriate, life-sustaining motoric responses that sustain individual life.

## 8. Human emotional experience

If the conscious mental state of fear in humans is not a product of the amygdala defensive survival circuit, how does it come about? I propose that it comes about much like any other conscious mental state. Although there are a number of theories of consciousness [88–93], I will focus on my cognitive account of mental state consciousness [2,94–98].

It is sometimes said that emotions like fear are universal. But I argue that it is danger, not fear, that is universal [19,94,96]. I define fear as a personal, schema-based, narrative-driven, culturally shaped, subjective experience that occurs in a biologically or psychologically significant situation (figure 2). Schema are collections of memories about specific kinds of situations in life that underlie our thoughts, feelings and actions in life.

The key idea is that conscious experience, whether emotional or not, is always preceded by non-conscious (pre-conscious) cognitive processing of sensations and memories, including schema, that results in a mental model of the situation. In an emotional situation, such as fear, brain and body consequences of threat-elicited survival circuit activation become part of the model. The output of the model is a non-conscious narration (which can be verbal and/or non-verbal). The narration then supplies the content of the



**Figure 2.** My cognitive model of the conscious experience of fear in humans. Reproduced from [96]. (Online version in colour.)

conscious experience, and separately, the content of verbal report. Because conscious experience and verbal report involve different downstream circuits, verbal report does not always perfectly reflect the person's experience.

Nevertheless, verbal report is the most reliable means, the gold standard, for assessing consciousness in humans [99–102]. The importance of verbal report was recognized in the late nineteenth century by the pioneering brain scientist David Ferrier, who lamented the absence of verbal report in monkeys and the limits that this lack imposed on his ability to study conscious perception in them [103]. The use of verbal report in studies of human consciousness will be discussed further below.

This perspective on emotion, of course, relates back the conclusion that Gazzaniga and I drew from our studies of split-brain patients—that an important feature of human consciousness is the maintenance of a sense of mental unity via cognitive interpretations that attribute meaning and cause to behaviours that are controlled non-consciously by the brain. In extending our ideas to emotional consciousness, we built on Schachter and Singer's cognitive theory of emotion [104], which I continue to do in spirit, if not substance, as do many other current cognitive theories of emotion [105–111].

A popular cognitive approach theory of emotion takes a constructionist perspective [105,112,113]. My approach falls roughly into this camp to the extent that it treats conscious emotions as cognitive conceptions that are assembled in the moment, rather than being elicited as innate mental states. I was on this track long before it carried the constructionist label.

The typical constructionist perspective views the conceptions underlying emotions as the result of interactions between two generalized activities in the brain, valence and arousal, and minimizes the importance of innate, species-typical processes. This reflects constructionists' battles with the so-called basic emotions theories, which treat core emotions like fear—including the conscious experience of fear—as innate products of subcortical circuits [35]. My model falls between basic and constructionist theories.

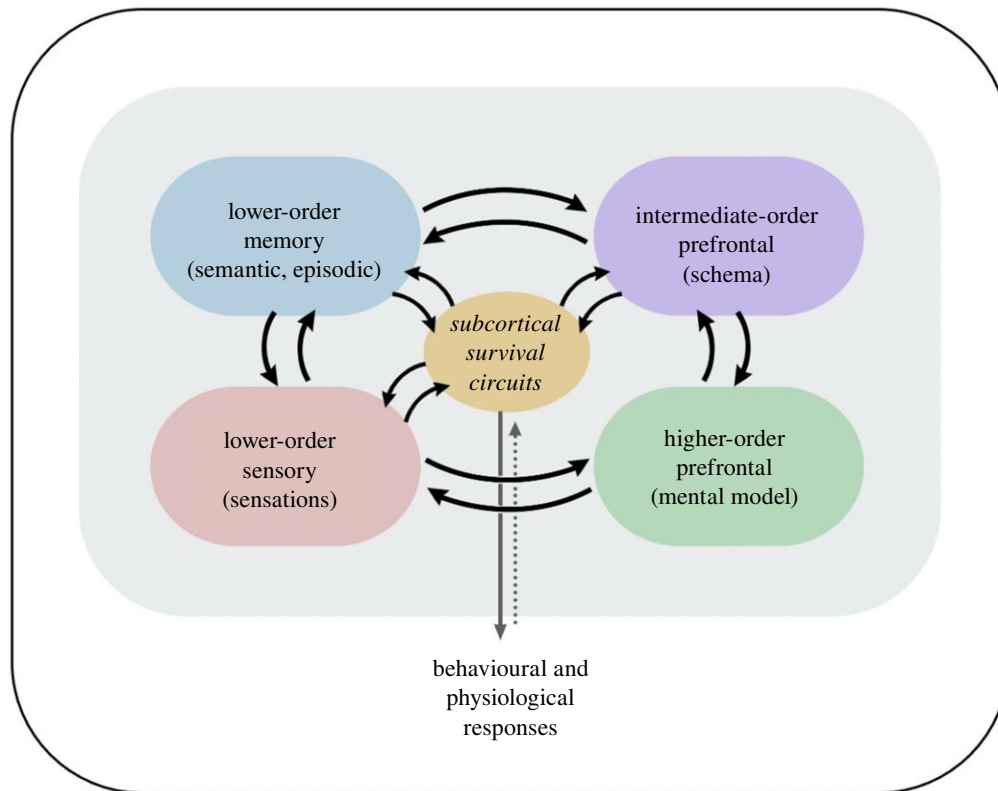
Like basic emotions theories, I emphasize innate circuits (in my case defensive survival circuits rather than fear circuits) that control behavioural and physiological body responses elicited by threats. But unlike basic emotionalists, and like constructionists, I view the conscious experience of fear as a product of cortical, cognitive activity. Unlike traditional constructionists, though, I think of arousal and valence as activities triggered by the particular innate

survival circuit that is active at the moment. For example, arousal, rather than being a purely generalized activity that occurs in any biologically significant situation, has elements that are tailored to current circumstances by the active survival circuit. Donald Pfaff has argued for something similar in relation to arousal—that it has specific and generalized components [114].

Another way my model differs from constructionist and other cognitive theories of emotion is that I have situated it within the higher-order framework of consciousness [88,90,115]. This class of theories proposes that conscious experiences result from the higher-order re-representation, hence conceptualization, of lower-order information. In the brain, the hypothetical higher-order network involves regions of lateral prefrontal cortex (PFC), including dorsal and ventral lateral PFC and the lateral frontal pole. For example, the conscious experience of an apple, in this theory, reflects the lateral PFC re-representation of lower-order visual cortex states.

Contemporary theories of consciousness, including higher-order theories, seldom consider the role of memory in consciousness. But without semantic memory, and especially memory-based schema [116,117], the sensory representation of an apple is not meaningful as an instance of the fruit of that name [118]. My multistate representation model makes memory an essential underpinning of higher-order consciousness, including emotional consciousness [94,96–98] (figure 3).

In the brain, for memory to impact visual perceptual consciousness, the activity of temporal lobe memory systems must be integrated with the sensory signals. A key way this occurs is by way of interactions between sensory cortex and temporal lobe areas, especially medial temporal lobe regions such as the hippocampus and perirhinal cortex [119]. Schema, in fact, are believed to be assembled and used by way of hippocampal interactions with medial PFC [117]. These sensory–memory representations, including schema, are then further represented by connections with prefrontal cortex areas, including the orbital, ventral medial and anterior cingulate areas located in the medial wall of the hemispheres, and the insula cortex buried deep in the Sylvian fissure. Each of these, in turn, interact with lateral PFC areas (dorsal and ventral lateral and frontal pole), by way of various intra-PFC connections [120], as discussed in detail elsewhere [2,90,94,96]. In the context of my higher-order theory, the medial wall and insula areas are considered 'intermediate' between sensory/memory circuits and lateral PFC areas.



**Figure 3.** Multi-state representation in higher-order awareness. Reproduced from [96]. (Online version in colour.)

These intermediate areas have been implicated not just in memory re-representation, but also in the encoding of body state signals, stimulus and response values, and self [94]. In the view presented here, such representations are not sufficient, on their own, and must be further represented by lateral PFC for the kind of higher-order awareness we humans experience. But this is hypothetical, and a higher-order account based on the intermediate areas is also imaginable.

If human emotions are cognitive processes, what makes them different from non-emotional cognitive states? Quite simply, the higher-order network re-represents additional lower-order information when the stimulus is an emotional one. For example, in fear, the consequences of amygdala defensive survival circuit activation in the brain (i.e. arousal) and body (visceral and somatic signals that feed back to the brain) are often also re-represented in the medial/insula PFC areas mentioned above. Aspects of the various prefrontal representations are then further integrated in the lateral PFC as a higher-order re-representation of diverse lower-order states. The emotional experience, therefore, depends not just on PFC but also on the entire cascade of non-conscious sensory, mnemonic/conceptual and survival circuit processing that is antecedent to PFC areas. It bears repeating that every conscious state is, until its last moment of fruition, non-conscious.

Other primates have similar, though less developed, lateral and medial/insula areas, while non-primate mammals mainly possess the medial/insula areas [2,96,121]. This anatomical situation may limit the kinds of conscious emotional states each can assemble relative to humans [96], as elaborated later.

The reason that amygdala survival circuits are so often associated with fear is because of their prominent role in controlling defensive responses when humans or other animals

are in danger of bodily harm from predators. As we have seen, this brain function runs deep in the vertebrate lineage. However, predator danger is only one of the many kinds of situations that can cause fear. If you are trapped on a mountain top without food, water, or shelter, body signals indicating low reservoirs of nutritional/energy supplies or fluids, or low body temperature, can lead to fear of death by starvation, dehydration or hypothermia. In humans, fear can also arise from being near someone with an infectious disease, or the news that you may have a life-threatening condition, or from family stress, social abuse, job insecurity, political instability or just about anything imaginable—threats do not have to be real to make one feel afraid or anxious. The only way to account for all of these under a single concept of fear is by way of cognitive interpretation of the situation in which you find yourself, or that you imagine you could be in. Innate states just would not do the trick.

## 9. The problem of animal consciousness

The question of how emotions come about in the brain of animals is subset of the thorny question of animal consciousness. I have written extensively about this [2,42,44,46,47,96] and can only touch on some of the key points here.

The question of whether non-human organisms are conscious, including mammals, lower vertebrates and invertebrates, and if so how, is quite contentious. The issues are far more complex than trying to find evolutionary roots of behaviours. We can directly measure behaviour in present-day organisms to obtain clues about what kinds of behaviours may have been present in their early representatives, and in the common ancestors they share with other organisms, in much the same as we do with brain structures. Consciousness,

though, is more challenging. Whether even our closest primate relatives are conscious is debated [122,123]. The problem is that behaviour and physiology have limits as measures of consciousness [124], especially when only non-verbal behaviour is available, as discussed further below.

In the late nineteenth century, animal psychology tried to use intelligent and emotional behaviour as marks of consciousness [86]. This, in part, is what led to behaviourism. With the behaviourists gone, some animal psychologists have revived such efforts. While the behaviourists certainly went too far, they were quite effective, if too dogmatic, in managing the language of psychology. And with them gone, we are back to the semantic 'wild west' of the late nineteenth century.

Part of the problem is that the evaluation of consciousness in animals, whether in mammals [35] or molluscs [125], is often based on intuitions and beliefs derived from one's understanding of their own behaviour, mind and/or brain. When the claims match common sense and lore, they feel correct, and when they are repeated authoritatively in scientific or lay communities, they come to be assumed as indisputable facts. This does not mean that other animals are unconscious robots. It simply means the behavioural data held up as proof of mental states are often not. The widespread assumption that innate defensive behaviours are a fool-proof reflection of conscious feelings of fear is a case in point.

According to J. S. Kennedy, author of *The new anthropomorphism*, 'Anthropomorphic thinking ... is built into us ... It is dinned into us culturally from earliest childhood. It has presumably also been 'pre-programmed' into our hereditary make-up by natural selection, perhaps because it proved to be useful for predicting and controlling the behavior of animals' [126, p. 5]. Our language, he says, is inherently anthropomorphic, and, as a result, our concepts and thoughts tend to lean in this direction as well. Kennedy concludes that anthropomorphism is part of human nature. This is perhaps why we all see human-like emotions in our pets. Me included. We do not have to be rigorous scientists every moment of waking life—but we must when we are being scientists.

At the turn of the twentieth century, when anthropomorphic approaches were commonplace in comparative psychology, Herbert Spencer Jennings, who studied the behavioural repertoire of single-cell protozoa, put it this way: 'If *Ameoba* were a large animal, so as to come within the everyday experience of human beings, its behavior would at once call forth the attribution to it of states of pleasure and pain, of hunger, desire and the like' [82, p. 336]. But beyond such general tendencies toward anthropomorphism, each scientist also comes to his or her career with personal biases and dispositions. As the philosopher Bertrand Russell once noted, 'All the animals that have been carefully observed have behaved so as to confirm the philosophy in which the observer believed before his observations began' [127, pp. 29–30].

Often, proponents of animal consciousness argue by analogy with human behaviour. This has been criticized by many on the grounds that it tends to ignore 'leaner' alternatives that explain behaviour without recourse to consciousness [122,123]. The critics do so, not to deny consciousness in non-human animals, but in the spirit of fostering a more rigorous scientific evaluation of what the data allow us to conclude about consciousness on the basis of behaviour alone.

Problems especially arise when the claims, rather than being treated as opinions or hypotheses, are presented as

indisputable facts that are presumed to be so obvious that no reasonable person could question them. And those who show interpretative restraint are denigrated as 'deniers' [128], a term that has a variety of negative connotations in contemporary culture, and that is not suitable in a scientific debate.

Similar issues exist when evaluating consciousness in pre-verbal children. Just because an infant human, or a rat, bird, lamprey, amphioxus or octopus, behaves to an environmental stimulus in a way that an adult human might when they are conscious of a similar kind of stimulus does not *necessarily* mean that the child or animal is having a conscious experience, a mental state, similar to what the adult human has. They *may*, and in some cases, likely, have some kind of conscious experience. But how can we truly know if they do, and if they do, what it is like for them?

To get around this conundrum one often has to make compromises about what is acceptable as scientific support for the hypothesis. At the end of the day, claims of consciousness in animals and human infants often rest more on intuitions and beliefs than on data, since there are often equally compelling non-conscious explanations of the behaviour [122,123].

Why are we on any firmer ground studying consciousness in human adults than in adult animals? Two points are crucial. One is that barring a congenital brain disorder, all humans have brains with same basic structural components and functional capacities. Given that, if I am conscious of my mental states and actions, I can with some confidence assume that you have this capacity as well. Like any species trait, consciousness will vary across individuals, but will be present, or at least is potentially present, in all members of our species.

The second point is that we humans can report on our conscious states verbally or non-verbally, but can only report on non-conscious states non-verbally. Other animals, and human infants, lacking the ability to verbally report, only have non-verbal means of responding to both conscious and non-conscious states. There is, therefore, no easy way to distinguish conscious and non-conscious states using non-verbal behaviour alone. Given the points I made above about the shortcomings of behaviour as an index of consciousness, even in humans, we have a significant methodological barrier. Obviously, the importance of verbal report as a methodological tool should not be taken to mean that human infants or language-impaired adults or non-human animals lack consciousness.

## 10. An empirical approach to animal consciousness

To be clear, even if repetitive, I do not deny consciousness in other animals. I am perfectly comfortable saying that some animals likely have conscious experiences, despite the methodological barriers that prevent clear proof. Yet, even if they are conscious, they would not be conscious in the ways we are. After all, human bodies differ in unmistakable ways from the bodies of other animals, including the bodies of our closest primate relatives. Similarly, our brains differ in important ways from the brains of other primates cousins [129–132], and theirs from other mammals [121,133]. It should, therefore, not be controversial to suggest that their mental states, including their emotional states, may also differ from ours.

Yet, in the light of the methodological issues, how would we ever know what the experiences of other animals



**Table 1.** Partitioning consciousness.

|   |
|---|
| Creature Consciousness (the state of being alive, awake, and behaviourally responsive to sensory stimuli)         |
| Mental States Consciousness (inner awareness of sensations, thoughts, feelings, and/or actions)                   |
| — Autonoetic Consciousness (cognitive awareness of oneself as an entity with a past, present and possible future) |
| — Noetic Consciousness (cognitive awareness of facts and/or ideas about the world and oneself)                    |
| — Anoetic Consciousness (non-cognitive, fringe awareness of external stimuli or internal states)                  |

are like? One way might be to use understanding of conscious experiences and their representation in the human brain as a basis of understanding what kind of consciousness other animals *might*, and *might not*, possess, given the ways that their brains are like, and different from, ours. I explored this idea in relation to other primates and other mammals in a recent publication [96] and will summarize and extend these suggestions below.

It is important to start by pointing out that the word ‘consciousness’ is often used as if it has a single referent. But it does not, and this has caused much confusion. At a minimum, it is important to distinguish ‘creature consciousness’ from ‘mental state consciousness’ [88] (table 1). The former refers to the condition of being awake and behaviourally responsive to environmental stimuli. All bilateral animals have creature consciousness. Mental state consciousness, on the other hand, is the condition of actually experiencing the world and one’s relation to it using mental models. It is crucial to keep this distinction in mind when evaluating claims of consciousness in animals, as some claims are primarily about creature consciousness.

For example, a much cited paper by Merker [134] defined consciousness as the state of being awake and responsive to sensory stimuli (i.e. creature consciousness). He used findings from patients with limited cortex due to hydrocephaly, and decorticated rats, as core components of his evidence. In these instances, and others he cites, consciousness is judged by sensory–motor responsiveness, not by evidence of mental states. In response to commentaries on his views, he says he is not interested in adult human consciousness (presumably cognitive-based mental state consciousness) but something more primitive. This implies that here he is referring to some kind of primitive mental state rather than creature conscious. But what would this be? Perhaps he meant a form of sensory sentience, which Lacalli proposed may have even existed in the invertebrate chordate ancestors of vertebrates [70,71]. I will build on a partition offered by Endel Tulving [135] that I have used in recent publications [2,94,96–98]. It may be able to account for what Merker might have had in mind, but in relation to mental state consciousness.

Tulving distinguished between three forms of mental state consciousness in humans (table 1). These are autonoetic (explicit self-awareness of one’s existence over time), noetic (explicit awareness of facts and concepts about the world or one’s self) and anoetic (implicit awareness of the world)

[2,94,96–98,135,136]. The first two are cognitive (i.e. based on episodic and semantic explicit memory) and the third is procedural (i.e. based on implicit learning). Another way of saying this is that autonoetic and noetic consciousness depend on the use of internal representations and mental models, but anoetic consciousness does not [96].

Because anoetic consciousness is not an intuitive idea, two clarifications are in order.

First, although ‘procedural’ and ‘implicit’ traditionally imply ‘non-conscious’ states, in the case of anoesis, non-conscious does not mean completely unconscious. Anoetic states reside on the ‘fringe’ or ‘penumbra’ of the ‘stream of consciousness’, to use William James’ terms [137], where the line between conscious and unconscious is fuzzy. In humans, these states are typically overshadowed by cognitive consciousness and go unnoticed [136]. But they are nevertheless present, and noticeable when cognitively attended to. They co-occur with cognitive states of consciousness and give these a feeling of ‘warmth’ and ‘intimacy’, again borrowing from James. They underlie your ability to know that your mental and body states are yours, without you ever having to explicitly affirm this [96,98,138]. The second clarification is that Tulving treated anoetic consciousness as a primitive awareness about the external world. However, I [96] and others [136] have proposed that penumbral/fringe anoetic states also represent the internal milieu (i.e. physiological conditions of the body and brain). Such anoetic states may be what Merker had in mind when he referred to a primitive kind of consciousness that is different from adult human consciousness.

The reason I am devoting so much discussion to Tulving’s partitions of conscious mental states is because I believe that understanding their representation in the brains of humans might give us insights into the processes the underlie mental state consciousness in other animals, especially those in our evolutionary past [2,96]. In particular, although we cannot directly measure consciousness in other animals, Tulving’s three states offer an indirect way in—because the three states depend on behaviourally measurable episodic, semantic and procedural memory processes, these provide pre-conscious proxies that can be studied in non-human species.

Researchers have used behavioural measures of episodic memory to ask whether primates, rodents and birds might have this capacity [119,139–142], which is so prominent in humans. Such behavioural studies can thus test the cognitive underpinnings of autonoetic consciousness in animals, but cannot measure autonoetic consciousness itself. To acknowledge this, the term ‘episodic-like’ memory is often used [140]. The key difference is that episodic/episodic-like memory is about what happened and when and where it happened, while autonoetic consciousness is about reflective awareness of one’s self over time in such episodes. Episodic-like memory, in short, takes researchers all the way to the finish line of consciousness. This is a significant achievement, despite not being able to cross the line.

Similarly, semantic memory has been well researched in primates and other mammals. For example, studies of object recognition in relation to instrumental-goal directed behaviours test what animals cognitively know factually and conceptually [119,143–146]. Procedural memory has also been extensively studied in primates and mammals [26,119,147]. As with episodic memory, studies of semantic and procedural memory tell us about the foundation of

noetic and anoetic awareness, but do not demonstrate that the awareness itself exists.

Tulving's partitions thus offer a top-down (reverse engineering) approach that starts with human consciousness, as opposed to bottom-up (forward engineering) evolutionary approaches to animal consciousness that start in the distant past [35,70,71,125,136]. An advantage of the top-down approach is that it explicitly identifies how consciousness might differ in different animals in our evolutionary past relative to our kinds of consciousness, based on similarities and differences in their brains and ours. In other words, our kind of consciousness tells us what to look for. Clearly, the further back we try to look, the harder it gets to use human consciousness to understand possible states of consciousness in other animals. But given the striking similarities in the detailed organization of the forebrain in all vertebrates [67,68,148], and the known involvement of homologous circuits in the relevant learning processes, we have some basis for speculating about the underpinnings of consciousness.

In humans, much is known about the neural basis of explicit (episodic and semantic) and implicit (procedural) memory [26,119,149–152]. The brain areas involved in the encoding of episodic memories include regions of the hippocampal formation, regions of medial PFC (anterior cingulate and ventromedial PFC) and regions of lateral PFC (dorsolateral and polar PFC). For semantic memories, including schema, the lateral prefrontal areas involved are the same as in episodic memory, but some areas of the hippocampal formation differ (parahippocampal in episodic; perihinal in semantic), and some areas of medial PFC also differ (anterior cingulate in episodic; medial orbital in semantic). The ventromedial PFC then connects both anterior cingulate and medial orbital to lateral PFC areas. Note that these are not simply one-way cascades, as there are reciprocal connections at each step.

Implicit procedural memory does not involve a single system, but instead is stored in circuits that process stimuli and/or control responses. For example, plasticity in sensory and motor cortex in sensory–motor learning; the amygdala in Pavlovian threat conditioning; and the basal ganglia in instrumental habit learning [19,26,153–157].

Non-human primates have homologs of many the areas implicated in episodic and semantic memory, with the possible exception of the lateral polar PFC [130,158,159]. At a minimum, this allows primates to use non-verbal episodic and semantic memory (including schema) to create mental models and use these to guide responses to recognized objects. Similarly, procedural circuits are quite similar in non-human primates and humans.

Other mammals possess homologues of the medial areas involved in non-verbal episodic and semantic memory present in primates and humans, but lack homologs of lateral PFC areas [121,133]. They thus must depend on hippocampal formation interactions with medial PFC for memory re-representations. Procedural circuits in mammals are largely homologous with those in non-human primates and humans.

The interim conclusion from this discussion is that the three kinds of memory/learning processes discussed reflect phylogenetic elaborations of the PFC over the evolutionary history of mammals. The implication is that the three kinds of conscious states associated with these memory processes may have followed the same evolutionary course, either arising in tandem with the memory/learning circuits, or because

of their existence. However, extension to consciousness remains hypothetical.

Can we take these ideas to lower vertebrates? An entry point for this may be the suggestion by Shepherd and others that, in mammals, olfactory paleocortex, via connections with medial wall cortical areas, especially orbitofrontal cortex, plays a role in olfactory consciousness [160–162]. But what kind of consciousness might that be? Following Tulving's scheme, it can be called anoetic olfactory consciousness. Also of note is that olfactory paleocortex and orbitofrontal cortex are both connected with the amygdala in mammals. This suggests the possibility that orbitofrontal re-representation of olfactory-triggered amygdala states might constitute anoetic experiences about the biological significance of olfactory stimuli, and of the bodily consequences that co-occur with these significant stimuli. But diurnal mammals also rely on other senses, and these also activate the amygdala, allowing a wider range of anoetic states via cortical re-representation.

Early vertebrates possessed a medial pallium, a primitive homologue of the mammalian hippocampus, that received olfactory and visual inputs and created and stored spatial representations that guided navigation and foraging [148,163–165]. If the amygdala homologue in early vertebrates was also connected with the medial pallium, re-representation of sensory and amygdala activity by the medial pallium may have supported both olfactory-based anoetic sensory experience, and amygdala-based anoetic experience of biologically significant stimuli. Given that they lack medial PFC homologues, medial pallium re-representation of sensory stimuli would have been the sole basis for sensory and amygdala informed experiences. As a result, such a state would likely have been even more primitive than anoetic experiences in mammals.

If the above is roughly correct, it would have an important implication. Namely, the medial pallium's capacity to acquire and retain spatial maps may have paved the way in mammals for additional functions of the hippocampus. In particular, I have in mind its capacity to form and store complex semantic memories and schema, and, via re-representation in medial PFC, to create mental models that underlie non-verbal noetic conscious awareness of the meaning of stimuli and responses in light of their value in survival situations. In non-human primates, and the evolution of lateral PFC, more elaborate executive functions may have enabled more complex forms of non-verbal behavioural control and noetic awareness. Further expansion of lateral PFC in humans, including enhanced executive functions and top-down control, and the invention of language, may have ushered in verbal noesis and autonoesis.

So far, I have described how different kinds of consciousness may have evolved in vertebrates. But this discussion begs the question of what might be the functional advantage of consciousness? The most general answer is that consciousness opens up novel forms of behavioural control [166]. But let me be more specific. Following the evolutionary thread just discussed, in early vertebrates, non-cognitive anoetic consciousness may have provided important advantages in navigation and foraging, including novel processing capacities for sensing the world, learning relations between sensory processing and biologically significant stimuli, and selection of actions based on past consequences [167]. With cortical expansion in mammals, and then primates, additional survival advantages related to behavioural control may have come

with more complex anoetic and novel cognitive (noetic) consciousness, with the latter involving the use of mental models. And in humans, with still further cortical expansions and the addition of language, verbal-based cognitive noesis and auto-noesis have allowed us to know ourselves as entities with a past and present, and also to anticipate possible futures. Regardless of whether actual states of consciousness can be experienced in non-human animals, the evidence described shows the kinds of neural underpinnings that may underlie the evolutionary past of anoetic, noetic and auto-noetic consciousness in humans, and perhaps in some non-human animals as well.

These ideas are, of course, highly speculative. But by linking the speculations to facts about the structure and functions of the human brain, and following the evolutionary history of the relevant structures underlying different kinds of conscious experiences in humans back in history, we may have at least some empirical grounding for such speculations.

## 11. Mental disorders put things in perspective

One of the factors that has motivated my efforts to reconceptualize fear is the much-discussed failure of the pharmaceutical industry to find better treatments for problems related to fear and anxiety [168,169]. The typical approach to drug discovery is to test animals, often rodents, using 'fear-like' or 'anxiety-like' behavioural tasks and administer compounds to them. The assumption is that because we inherited our 'fear' circuit from our mammalian ancestors, and because this circuit both assembles feelings of fear and controls behavioural and physiological responses in danger, it should be the case that drugs that make rats less timid behaviourally and less physiologically aroused should make people less fearful or anxious.

Yet, despite decades of research, new, better treatments were seldom found [170]; drugs that showed promise in animals often flopped in human clinical trials [171], and while those that made it into circulation may have had fewer side effects, they were at best only as effective as the older ones. As a result, Big-Pharma has been withdrawing funding for research on fear and anxiety, and also depression [168,169].

Given the various points made earlier, it seems unrealistic to have expected that drugs that change the behavioural or physiological responses in rodents would be the solution to the problem of how to make people feel subjectively less fearful or anxious. The more realistic expectation is that behavioural (avoidance) and physiological (hyperarousal) symptoms might be dampened, since that's what we know with confidence the drugs did in the laboratory studies. But if an anti-anxiety medication makes you less jittery and avoidant, but you still feel anxious, the treatment is not living up to the promise of its name.

To be fair, some drugs give some people some psychological relief. But the question is, does this result from the medication turning off a fear or anxiety circuit in the brain, or is it because it has indirect effects on feelings of fear or anxiety? Specifically, the psychological relief, when it occurs, could be secondary to other changes that affect fearful and anxious feelings. For example, some relief could come from alteration of defensive survival circuits that control behavioural and physiological symptoms, and/or alteration of somatic and visceral motoric control circuits in the brainstem,

the spinal cord or the periphery that are the final paths to the symptoms.

Clearly, such indirect effects are useful to the patient. And if changing such symptoms was the expectation, the industry, treatment community and patients might view the medications as a rousing success, since they would be understood as doing what they were designed to do in the animal studies, rather than not doing what they were incorrectly advertised to do. Another possible account of positive effects, at least for medications like benzodiazepines, on anxious feelings is general emotional blunting, rather than a specific reduction in anxiety, resulting from sedative effects consequent to enhancing GABAergic inhibitory neural activity in widespread areas of the brain [172].

It matters that we figure out how treatments work. Otherwise, we will never advance the cause of finding the best treatment, given the patient's particular problems. The fact is, what people really want from any kind of therapy is to feel better subjectively. The best way to improve this likelihood would be to put the mental back into mental disorders [48,173,174].

## 12. Conclusion

The survival circuits of mammals are neural implementations of ancient and necessarily persistent physiological survival requirements that have allowed organisms throughout the history of life to move in ways that extended their life beyond the present moment in biologically significant situations. The consequences of survival circuit activation can indirectly modulate conscious feelings. But they do not, on their own, define the conscious content of feelings. Feelings, instead, are cognitive interpretations of significant situations that we encounter in life. Although I had some inkling of these conclusions decades ago from studying split-brain patients, it was my decades of work on rodents that allowed me to see the field the way I do today. The approach I offer here, based on a distinction between ancient defensive survival circuits that control behavioural and supporting physiological responses, and more recent cognitive circuits that assemble conscious feelings of fear, seems at least worth considering as a possible way forward.

The bottom-line general conclusion I would like to end with is that our understanding of a psychological process in the brain is only as good as our conceptualization of what that process is. If we do not know what we are looking for, we will never find it. And as we search for knowledge about psychological processes in the brain, we should only use mental state words to refer to mental states. Additionally, we should refrain from attributing behavioural control to mental states, whether in humans or other animals, without compelling evidence. It is of course acceptable to speculate about the mental lives of other animals, which I have done my share of in this article. But when we speculate, we should be clear that we are doing so. I hope that I have at least achieved this.

**Data accessibility.** This article does not contain any additional data.

**Competing interests.** I declare I have no competing interests.

**Funding.** I received no funding for work on this paper.



## References

- Cisek P. 2019 Resynthesizing behavior through phylogenetic refinement. *Atten. Percept. Psychophys.* **81**, 2265–2287. (doi:10.3758/s13414-019-01760-1)
- LeDoux J. 2019 *The deep history of ourselves: the four-billion-year story of how we got conscious brains*. New York, NY: Viking.
- Emerson RW. 1870 *Society and solitude*. Boston, MA: Fields, Osgood & Co.
- LeDoux JE, Wilson DH, Gazzaniga MS. 1977 A divided mind: observations on the conscious properties of the separated hemispheres. *Annu. Neurol.* **2**, 417–421. (doi:10.1002/ana.410020513)
- Gazzaniga MS, LeDoux JE. 1978 *The integrated mind*. New York, NY: Plenum.
- Cannon WB. 1929 *Bodily changes in pain, hunger, fear, and rage*, 2nd edn. New York, NY: Appleton.
- Hess WR, Brugger M. 1943 Das subkortikale Zentrum der affektiven Abwehrreaktion. *HelvPhysiolPharmacolActa* **1**, 33–52.
- de Molina A F, Hunsperger RW. 1959 Central representation of affective reactions in forebrain and brain stem: electrical stimulation of amygdala, stria terminalis, and adjacent structures. *J. Physiol.* **145**, 251–265. (doi:10.1113/jphysiol.1959.sp006140)
- Hilton SM, Zbrozyna AW. 1963 Amygdaloid region for defense reactions and its efferent pathway to the brainstem. *J. Physiol.* **165**, 160–173. (doi:10.1113/jphysiol.1963.sp007049)
- MacLean PD. 1949 Psychosomatic disease and the "visceral brain": recent developments bearing on the Papez theory of emotion. *Psychosom. Med.* **11**, 338–353. (doi:10.1097/00006842-194911000-00003)
- MacLean PD. 1952 Some psychiatric implications of physiological studies on frontotemporal portion of limbic system (visceral brain). *Electroencephalogr. Clin. Neurophysiol.* **4**, 407–418. (doi:10.1016/0013-4694(52)90073-4)
- Kandel ER. 1976 *Cellular basis of behavior: an introduction to behavioral neurobiology*. San Francisco, CA: W.H. Freeman and Company.
- Miller NE. 1948 Studies of fear as an acquirable drive: I. Fear as motivation and fear reduction as reinforcement in the learning of new responses. *J. Exp. Psychol.* **38**, 89–101. (doi:10.1037/h0058455)
- Davis M. 1992 The role of the amygdala in conditioned fear. In *The amygdala: neurobiological aspects of emotion, memory, and mental dysfunction* (ed. JP Aggleton), pp. 255–306. New York, NY: Wiley-Liss, Inc.
- Kapp BS, Whalen PJ, Supple WF, Pascoe JP. 1992 Amygdaloid contributions to conditioned arousal and sensory information processing. In *The amygdala: neurobiological aspects of emotion, memory, and mental dysfunction* (ed. JP Aggleton), pp. 229–254. New York, NY: Wiley-Liss.
- LeDoux JE. 1992 Emotion and the amygdala. In *The amygdala: neurobiological aspects of emotion, memory, and mental dysfunction* (ed. JP Aggleton), pp. 339–351. New York, NY: Wiley-Liss, Inc.
- LeDoux JE. 1994 Emotion, memory and the brain. *Sci. Am.* **270**, 50–57. (doi:10.1038/scientificamerican0694-50)
- LeDoux JE. 1996 *The emotional brain*. New York, NY: Simon and Schuster.
- LeDoux JE. 2000 Emotion circuits in the brain. *Annu. Rev. Neurosci.* **23**, 155–184. (doi:10.1146/annurev.neuro.23.1.155)
- Blair HT, Schafer GE, Bauer EP, Rodrigues SM, LeDoux JE. 2001 Synaptic plasticity in the lateral amygdala: a cellular hypothesis of fear conditioning. *Learn. Mem.* **8**, 229–242. (doi:10.1101/lm.30901)
- Rodrigues SM, Schafer GE, LeDoux JE. 2004 Molecular mechanisms underlying emotional learning and memory in the lateral amygdala. *Neuron* **44**, 75–91. (doi:10.1016/j.neuron.2004.09.014)
- Lamprecht R, LeDoux J. 2004 Structural plasticity and memory. *Nat. Rev. Neurosci.* **5**, 45–54. (doi:10.1038/nrn1301)
- Won J, Silva AJ. 2008 Molecular and cellular mechanisms of memory allocation in neuronetworks. *Neurobiol. Learn. Mem.* **89**, 285–292. (doi:10.1016/j.nlm.2007.08.017)
- Kandel ER. 2006 *Search of memory: the emergence of a new science of mind*. New York, NY: W.W. Norton.
- Glanzman DL. 2010 Common mechanisms of synaptic plasticity in vertebrates and invertebrates. *Curr. Biol.* **20**, R31–R36. (doi:10.1016/j.cub.2009.10.023)
- Squire L. 1987 *Memory and brain*. New York, NY: Oxford University Press.
- Ohman A. 2005 The role of the amygdala in human fear: automatic detection of threat. *Psychoneuroendocrinology* **30**, 953–958. (doi:10.1016/j.psyneuen.2005.03.019)
- Morris JS, Ohman A, Dolan RJ. 1999 A subcortical pathway to the right amygdala mediating 'unseen' fear. *Proc. Natl Acad. Sci. USA* **96**, 1680–1685. (doi:10.1073/pnas.96.4.1680)
- Bornemann B, Winkelman P, van der Meer E. 2012 Can you feel what you do not see? Using internal feedback to detect briefly presented emotional stimuli. *Int. J. Psychophysiol.* **85**, 116–124. (doi:10.1016/j.ijpsycho.2011.04.007)
- de Gelder B, Vroomen J, Pourtois G, Weiskrantz L. 2000 Affective blindsight: are we blindly led by emotions? Response to Heywood and Kentridge (2000). *Trends Cogn. Sci.* **4**, 126–127. (doi:10.1016/S1364-6613(00)01470-4)
- Anderson AK, Phelps EA. 2002 Is the human amygdala critical for the subjective experience of emotion? Evidence of intact dispositional affect in patients with amygdala lesions. *J. Cogn. Neurosci.* **14**, 709–720. (doi:10.1162/08989290260138618)
- Feinstein JS, Buzza C, Hurlmann R, Follmer RL, Dahdaleh NS, Coryell WH, Welsh MJ, Tranel D, Wemmie JA. 2013 Fear and panic in humans with bilateral amygdala damage. *Nat. Neurosci.* **16**, 270–272. (doi:10.1038/nn.3323)
- Blanchard DC, Blanchard RJ. 1972 Innate and conditioned reactions to threat in rats with amygdaloid lesions. *J. Comp. Physiol. Psych.* **81**, 281–290. (doi:10.1037/h0033521)
- Fanselow MS, Poulos AM. 2005 The neuroscience of mammalian associative learning. *Annu. Rev. Psychol.* **56**, 207–234. (doi:10.1146/annurev.psych.56.091103.070213)
- Panksepp J. 1998 *Affective neuroscience*. New York, NY: Oxford University Press.
- Panksepp J. 2005 Affective consciousness: core emotional feelings in animals and humans. *Conscious Cogn.* **14**, 30–80. (doi:10.1016/j.concog.2004.10.004)
- Watt DF. 2000 Emotion and consciousness: Part II. *J. Consciousness Stud.* **7**, 72–84.
- Burghardt G. 1991 Cognitive ethology and critical anthropomorphism: a snake with two heads and hognose snakes that play dead. In *Cognitive ethology: the minds of other animals: essays in honor of Donald R Griffin* (ed. CA Ristau), pp. 53–90. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bekoff M. 2000 Animal emotions: exploring passionate natures: current interdisciplinary research provides compelling evidence that many animals experience such emotions as joy, fear, love, despair, and grief—we are not alone. *BioScience* **50**, 861–870. (doi:10.1641/0006-3568(2000)050[0861:AEENPN]2.0.CO;2)
- Griffin DR. 2001 *Animal minds: beyond cognition to consciousness*. Chicago, IL: University of Chicago Press.
- LeDoux JE. 2002 *Synaptic self: how our brains become who we are*. New York, NY: Viking.
- LeDoux JE. 2008 Emotional colouration of consciousness: how feelings come about. In *Frontiers of consciousness: chichele lectures* (eds L Weiskrantz, M Davies), pp. 69–130. Oxford, UK: Oxford University Press.
- LeDoux JE. 1990 Information flow from sensation to emotion: plasticity in the neural computation of stimulus value. In *Learning and computational neuroscience: foundations of adaptive networks. 1* (eds M Gabriel, J Moore), pp. 3–52. Cambridge, UK: MIT Press.
- LeDoux J. 2012 Rethinking the emotional brain. *Neuron* **73**, 653–676. (doi:10.1016/j.neuron.2012.02.004)
- LeDoux JE. 2014 Coming to terms with fear. *Proc. Natl Acad. Sci. USA* **111**, 2871–2878. (doi:10.1073/pnas.1400335111)
- LeDoux JE. 2017 Semantics, surplus meaning, and the science of fear. *Trends Cogn. Sci.* **21**, 303–306. (doi:10.1016/j.tics.2017.02.004)
- LeDoux JE. 2015 *Anxious: using the brain to understand and treat fear and anxiety*. New York, NY: Viking.



48. LeDoux JE, Pine DS. 2016 Using neuroscience to help understand fear and anxiety: a two-system framework. *Am. J. Psychiatry*. **173**, 1083–1093. (doi:10.1176/appi.ajp.2016.16030353)
49. Fanselow MS, Pennington ZT. 2017 The danger of LeDoux and Pine's two-system framework for fear. *Am. J. Psychiatry*. **174**, 1120–1121. (doi:10.1176/appi.ajp.2017.17070818)
50. Marx MH. 1951 Intervening variable or hypothetical construct? *Psychol. Rev.* **58**, 235–247. (doi:10.1037/h0054854)
51. Mandler G, Kessen W. 1964 *The language of psychology*, 2nd edn. New York, NY: John Wiley & Sons, Inc.
52. Block J. 1995 A contrarian view of the five-factor approach to personality description. *Psychol. Bull.* **117**, 187–215. (doi:10.1037/0033-2909.117.2.187)
53. Danziger K. 1997 *Naming the mind: how psychology found its language*. London, UK: Sage Publications, Ltd.
54. Taylor C, Dewsbury BM. 2018 On the problem and promise of metaphor use in science and science communication. *J. Microbiol. Biol. Educ.* **19**, 19. (doi:10.1128/jmbe.v19i1.1538)
55. Brick C, Hood B, Ekroll V, de-Wit L. 2021 Illusory essences: a bias holding back theorizing in psychological science. *Perspect. Psychol. Sci.* 1745691621991838. (doi:10.1177/1745691621991838)
56. Tinbergen N. 1951 *The study of instinct*. New York, NY: Oxford University Press.
57. Cisek P. 2021 Evolution of behavioural control from chordates to primates. *Phil. Trans. R. Soc. B* **377**, 20200522. (doi:10.1098/rstb.2020.0522)
58. Pessoa L, Medina L, Desfilis E. 2021 Refocusing neuroscience: moving away from mental categories and toward complex behaviors. *Phil. Trans. R. Soc. B* **377**, 20200534. (doi:10.1098/rstb.2020.0534)
59. Fletcher GJO. 1995 Two uses of folk psychology: implications for psychological science. *Philos. Psychol.* **8**, 221–238. (doi:10.1080/09515089508573155)
60. Kelley HH. 1992 Common-sense psychology and scientific psychology. *Annu. Rev. Psychol.* **43**, 1–24. (doi:10.1146/annurev.ps.43.020192.000245)
61. Martínez-García F, Lanuza E. 2018 Evolution of vertebrate survival circuits. *Curr. Opin. Behav. Sci.* **24**, 113–123. (doi:10.1016/j.cobeha.2018.06.012)
62. Butler AB. 2000 Topography and topology of the teleost telencephalon: a paradox resolved. *Neurosci. Lett.* **293**, 95–98. (doi:10.1016/S0304-3940(00)01497-X)
63. Portavella M, Vargas JP, Torres B, Salas C. 2002 The effects of telencephalic pallial lesions on spatial, temporal, and emotional learning in goldfish. *Brain Res. Bull.* **57**, 397–399. (doi:10.1016/S0361-9230(01)00699-2)
64. Lal P *et al.* 2018 Identification of a neuronal population in the telencephalon essential for fear conditioning in zebrafish. *BMC Biol.* **16**, 45. (doi:10.1186/s12915-018-0502-y)
65. Biechl D, Tietje K, Ryu S, Grothe B, Gerlach G, Wullmann MF. 2017 Identification of accessory olfactory system and medial amygdala in the zebrafish. *Sci. Rep.* **7**, 44295. (doi:10.1038/srep44295)
66. Maximino C, Lima MG, Oliveira KR, Batista Ede J, Herculano AM. 2013 "Limbic associative" and "autonomic" amygdala in teleosts: a review of the evidence. *J. Chem. Neuroanat.* **48–49**, 1–13. (doi:10.1016/j.jchemneu.2012.10.001)
67. Suryanarayana SM, Perez-Fernandez J, Robertson B, Grillner S. 2021 Olfaction in lamprey pallium revisited—dual projections of mitral and tufted cells. *Cell Rep.* **34**, 108596. (doi:10.1016/j.celrep.2020.108596)
68. Suryanarayana SM, Robertson B, Grillner S. 2021 The neural bases of vertebrate motor behaviour through the lens of evolution. *Phil. Trans. R. Soc. B* **377**, 20200521. (doi:10.1098/rstb.2020.0521)
69. Benito-Gutierrez E, Gattoni G, Stemmer M, Rohr SD, Schuhmacher LN, Tang J, Marconi A, Jékely G, Arendt D. 2021 The dorsoanterior brain of adult amphioxus shares similarities in expression profile and neuronal composition with the vertebrate telencephalon. *BMC Biol.* **19**, 110. (doi:10.1186/s12915-021-01045-w)
70. Lacalli T. 2021 Innovation through heterochrony: an amphioxus perspective on telencephalon origin and function. *Front. Ecol. Evol.* **9**, 318. (doi:10.3389/fevo.2021.666722)
71. Lacalli TC. 2021 An evolutionary perspective on chordate brain organization and function: insights from amphioxus, and the problem of sentience. *Phil. Trans. R. Soc. B* **377**, 20200520. (doi:10.1098/rstb.2020.0520)
72. Gibson WT *et al.* 2015 Behavioral responses to a repetitive visual threat stimulus express a persistent state of defensive arousal in *Drosophila*. *Curr. Biol.* **25**, 1401–1415. (doi:10.1016/j.cub.2015.03.058)
73. Viegas J. 2015 Flies have feelings: fear and maybe more. See <http://www.abc.net.au/science/articles/2015/05/15/4236373.htm>, retrieved May 11, 2018 2015.
74. Ryan TJ, Grant SG. 2009 The origin and evolution of synapses. *Nat. Rev. Neurosci.* **10**, 701–712. (doi:10.1038/nrn2717)
75. Conaco C, Bassett DS, Zhou H, Arcila ML, Degnan SM, Degnan BM, Kosik KS. 2013 Functionalization of a protosynaptic gene expression network. In *In the light of evolution: vol. VI: brain and behavior* (eds GF Striedter, JC Avise, FJ Ayala), pp. 3–19. Washington, DC: The National Academies Press.
76. Sakarya O, Armstrong KA, Adamska M, Adamski M, Wang IF, Tidor B, Degnan BM, Oakley TH, Kosik KS. 2007 A post-synaptic scaffold at the origin of the animal kingdom. *PLoS ONE* **2**, e506. (doi:10.1371/journal.pone.0000506)
77. Greenspan RJ. 2007 *An introduction to nervous systems*. New York, NY: Cold Spring Harbor Laboratory Press.
78. Arendt D, Tosches MA, Marlow H. 2016 From nerve net to nerve ring, nerve cord and brain—evolution of the nervous system. *Nat. Rev. Neurosci.* **17**, 61–72. (doi:10.1038/nrn.2015.15)
79. Jékely G. 2011 Origin and early evolution of neural circuits for the control of ciliary locomotion. *Proc. R. Soc. B* **278**, 914–922. (doi:10.1098/rspb.2010.2027)
80. Holland LZ. 2000 Body-plan evolution in the Bilateria: early antero-posterior patterning and the deuterostome–protostome dichotomy. *Curr. Opin. Genet. Dev.* **10**, 434–442. (doi:10.1016/S0959-437X(00)00109-X)
81. Holland ND. 2003 Early central nervous system evolution: an era of skin brains? *Nat. Rev. Neurosci.* **4**, 617–627. (doi:10.1038/nrn1175)
82. Jennings HS. 1906 *Behavior of the lower organisms*. New York, NY: Columbia University Press.
83. Hennessey TM, Rucker WB, McDiarmid CG. 1979 Classical conditioning in paramecia. *Anim. Learn. Behav.* **7**, 417–423. (doi:10.3758/BF03209695)
84. Vladimirov N, Sourjik V. 2009 Chemotaxis: how bacteria use memory. *Biol. Chem.* **390**, 1097–1104. (doi:10.1515/BC.2009.130)
85. Zhang H *et al.* 2014 Programming a Pavlovian-like conditioning circuit in *Escherichia coli*. *Nat. Commun.* **5**, 3102. (doi:10.1038/ncomms4102)
86. Keller FS. 1973 *The definition of psychology*. New York, NY: Appleton-Century-Crofts.
87. Schneirla TC. 1959 An evolutionary theory of biphasic processes underlying approach and withdrawal. In *Nebraska symposium on motivation* (ed. MR Jones), pp. 1–42. Lincoln, NB: University of Nebraska Press.
88. Rosenthal DM. 2005 *Consciousness and mind*. Oxford, UK: Oxford University Press.
89. Dennett DC. 1991 *Consciousness explained*. Boston, MA: Little, Brown and Company.
90. Brown R, Lau H, LeDoux JE. 2019 Understanding the higher-order approach to consciousness. *Trends Cogn. Sci.* **23**, 754–768. (doi:10.1016/j.tics.2019.06.009)
91. Graziano MSA, Guterstam A, Bio BJ, Wilterson AI. 2020 Toward a standard model of consciousness: reconciling the attention schema, global workspace, higher-order thought, and illusionist theories. *Cogn. Neuropsychol.* **37**, 155–172. (doi:10.1080/02643294.2019.1670630)
92. Tononi G, Koch C. 2015 Consciousness: here, there and everywhere? *Phil. Trans. R. Soc. B* **370**, 20140167. (doi:10.1098/rstb.2014.0167)
93. Mashour GA, Roelfsema P, Changeux JP, Dehaene S. 2020 Conscious processing and the global neuronal workspace hypothesis. *Neuron* **105**, 776–798. (doi:10.1016/j.neuron.2020.01.026)
94. LeDoux JE. 2020 Thoughtful feelings. *Curr. Biol.* **30**, R619–R623. (doi:10.1016/j.cub.2020.04.012)
95. LeDoux JE. 2020 How does the non-conscious become conscious? *Curr. Biol.* **30**, R196–R199. (doi:10.1016/j.cub.2020.01.033)
96. LeDoux JE. 2021 What emotions might be like in other animals. *Curr. Biol.* **31**, R824–R829. (doi:10.1016/j.cub.2021.05.005)
97. LeDoux JE, Brown R. 2017 A higher-order theory of emotional consciousness. *Proc. Natl Acad. Sci. USA* **114**, E2016–E2025. (doi:10.1073/pnas.1619316114)

98. LeDoux JE, Lau H. 2020 Seeing consciousness through the lens of memory. *Curr. Biol.* **30**, R1018–R1022. (doi:10.1016/j.cub.2020.08.008)
99. Rosenthal D. 2019 Consciousness and confidence. *Neuropsychologia* **128**, 255–265. (doi:10.1016/j.neuropsychologia.2018.01.018)
100. Weiskrantz L. 1997 *Consciousness lost and found: a neuropsychological exploration*. New York, NY: Oxford University Press.
101. Dehaene S, Sergent C, Changeux JP. 2003 A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proc. Natl Acad. Sci. USA* **100**, 8520–8525. (doi:10.1073/pnas.1332574100)
102. Koch C, Tsuchiya N. 2007 Phenomenology without conscious access is a form of consciousness without top-down attention. *Behav. Brain Sci.* **30**, 509–510. (doi:10.1017/S0140525X07002907)
103. Ferrier D. 1876 *The functions of the brain*. London, UK: Smith, Elder and Co.
104. Schachter S, Singer JE. 1962 Cognitive, social, and physiological determinants of emotional state. *Psychol. Rev.* **69**, 379–399. (doi:10.1037/h0046234)
105. Barrett LF. 2017 *How emotions are made*. New York, NY: Houghton Mifflin Harcourt.
106. Russell JA. 1980 A circumplex model of affect. *J. Pers. Soc. Psychol.* **39**, 1161–1178. (doi:10.1037/h0077714)
107. Ortony A, Clore GL, Collins A. 1988 *The cognitive structure of emotions*. Cambridge, UK: Cambridge University Press.
108. Frijda N. 1986 *The emotions*. Cambridge, UK: Cambridge University Press.
109. Mesquita B, Boiger M, De Leersnyder J. 2016 The cultural construction of emotions. *Curr. Opin. Psychol.* **8**, 31–36. (doi:10.1016/j.copsyc.2015.09.015)
110. Scherer KR. 1984 On the nature and function of emotion: a component process approach. In *Approaches to emotion* (eds KR Scherer, P Ekman), pp. 293–317. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
111. Flanagan O. 2021 *How to do things with emotions: anger and shame across cultures*. Princeton, NJ: Princeton University Press.
112. Barrett LF, Russell J (eds) 2015 *The psychological construction of emotion*. New York, NY: Guilford Press.
113. Russell JA. 2014 The greater constructionist project for emotion. In *The psychological construction of emotion* (eds LF Barrett, JA Russell), pp. 429–447. New York, NY: Guilford Press.
114. Pfaff D. 2019 *How brain arousal mechanisms work: paths toward consciousness*. New York, NY: Cambridge University Press.
115. Lau H, Rosenthal D. 2011 Empirical support for higher-order theories of conscious awareness. *Trends Cogn. Sci.* **15**, 365–373. (doi:10.1016/j.tics.2011.05.009)
116. Gilboa A, Marlatte H. 2017 Neurobiology of schemas and schema-mediated memory. *Trends Cogn. Sci.* **21**, 618–631. (doi:10.1016/j.tics.2017.04.013)
117. van Kesteren MT, Fernandez G, Norris DG, Hermans EJ. 2010 Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proc. Natl Acad. Sci. USA* **107**, 7550–7555. (doi:10.1073/pnas.0914892107)
118. Gregory RL. 1997 Knowledge in perception and illusion. *Phil. Trans. R Soc. Lond. B* **352**, 1121–1127. (doi:10.1098/rstb.1997.0095)
119. Murray EA, Wise SP, Graham KS. 2017 *The evolution of memory systems: ancestors, anatomy, and adaptations*. Oxford, UK: Oxford University Press.
120. Yeterian EH, Pandya DN, Tomaiuolo F, Petrides M. 2012 The cortical connectivity of the prefrontal cortex in the monkey brain. *Cortex* **48**, 58–81. (doi:10.1016/j.cortex.2011.03.004)
121. Preuss TM. 1995 Do rats have prefrontal cortex? The Rose-Woolsey-Akert program reconsidered. *J. Cogn. Neurosci.* **7**, 1–24. (doi:10.1162/jocn.1995.7.1.1)
122. Heyes C. 2008 Beast machines? Questions of animal consciousness. In *Frontiers of consciousness: Chichelle lectures* (eds L Weiskrantz, M Davies), pp. 259–274. Oxford, UK: Oxford University Press.
123. Povinelli DJ, Bering JM, Giambone S. 2000 Toward a science of other minds: escaping the argument by analogy. *Cogn. Sci.* **24**, 509–541. (doi:10.1207/s15516709cog2403\_7)
124. Richer J. 2021 The evolution of consciousness—the shortcomings of: the ancient origins of consciousness by Todd Feinberg and Jon Mallatt, and an alternative approach. *Hum. Ethol.* **36**, 27–35. (doi:10.22330/he/36/027-035)
125. Feinberg TE, Mallatt JM. 2016 *The ancient origins of consciousness*. Cambridge, MA: MIT Press.
126. Kennedy JS. 1992 *The new anthropomorphism*. New York, NY: Cambridge University Press.
127. Russell B. 1927 *Philosophy*. New York, NY: Norton.
128. de Waal FBM. 1999 Anthropomorphism and anthropodenial: consistency in our thinking about humans and other animals. *Philos. Top.* **27**, 255–280. (doi:10.5840/PHILTOPICS199927122)
129. Preuss TM. 2011 The human brain: rewired and running hot. *Ann. N Y Acad. Sci.* **1225**(Suppl. 1), E182–E191. (doi:10.1111/j.1749-6632.2011.06001.x)
130. Semendeferi K, Teffer K, Buxhoeveden DP, Park MS, Bludau S, Amunts K, Travis K, Buckwalter J 2011 Spatial organization of neurons in the frontal pole sets humans apart from great apes. *Cereb. Cortex* **21**, 1485–1497. (doi:10.1093/cercor/bhq191)
131. Koechlin E, Basso G, Pietrini P, Panzer S, Grafman J. 1999 The role of the anterior prefrontal cortex in human cognition. *Nature* **399**, 148–151. (doi:10.1038/20178)
132. Kaas JH, Herculano-Houzel S. 2017 What makes the human brain special: key features of brain and neocortex. In *The physics of the mind and brain disorders: integrated neural circuits supporting the emergence of mind* (eds I Opris, MF Casanova), pp. 3–22. Cham, Switzerland: Springer International Publishing.
133. Kaas JH. 2019 The origin and evolution of neocortex: from early mammals to modern humans. *Prog. Brain Res.* **250**, 61–81. (doi:10.1016/bs.pbr.2019.03.017)
134. Merker B. 2007 Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. *Behav. Brain Sci.* **30**, 63–81; discussion -134. (doi:10.1017/S0140525X07000891)
135. Tulving E. 2005 Episodic memory and autonoesis: uniquely human? In *The missing link in cognition* (eds HS Terrace, J Metcalfe), pp. 4–56. New York, NY: Oxford University Press.
136. Vandekerckhove M, Panksepp J. 2011 A neurocognitive theory of higher mental emergence: from anoetic affective experiences to noetic knowledge and autonoetic awareness. *Neurosci. Biobehav. Rev.* **35**, 2017–2025. (doi:10.1016/j.neubiorev.2011.04.001)
137. James W. 1890 *Principles of psychology*. New York, NY: Holt.
138. Klein SB. 2015 The feeling of personal ownership of one's mental states: a conceptual argument and empirical evidence for an essential, but underappreciated, mechanism of mind. *Psychol. Consciousness Theory Res. Practice* **2**, 355–376. (doi:10.1037/cns0000052)
139. Eichenbaum H, Fortin NJ, Ergorur C, Wright SP, Agster KL. 2005 Episodic recollection in animals: 'If it walks like a duck and quacks like a duck...'. *Learn. Motivation* **36**, 190–207. (doi:10.1016/j.lmot.2005.02.006)
140. Clayton NS, Bussey TJ, Dickinson A. 2003 Can animals recall the past and plan for the future? *Nat. Rev. Neurosci.* **4**, 685–691. (doi:10.1038/nrn1180)
141. Dickinson A. 2008 Why a rat is not a beast machine. In *Frontiers of consciousness* (eds L Weiskrantz, M Davies), pp. 275–288. Oxford, UK: Oxford University Press.
142. Balleine BW. 2019 The meaning of behavior: discriminating reflex and volition in the brain. *Neuron* **104**, 47–62. (doi:10.1016/j.neuron.2019.09.024)
143. Devito LM, Kanter BR, Eichenbaum H. 2010 The hippocampus contributes to memory expression during transitive inference in mice. *Hippocampus* **20**, 208–217.
144. DeVito LM, Lykken C, Kanter BR, Eichenbaum H. 2010 Prefrontal cortex: role in acquisition of overlapping associations and transitive inference. *Learn. Mem.* **17**, 161–167. (doi:10.1101/lm.1685710)
145. Clayton N, Dickinson A. 2006 Rational rats. *Nat. Neurosci.* **9**, 472–474. (doi:10.1038/nn0406-472)
146. Dickinson A, Balleine B. 2002 The role of learning in the operation of motivational systems. In *Steven's handbook of experimental psychology: learning, motivation and emotion*. 3 (ed. CR Gallistel), pp. 497–534, 3 edn. New York, NY: John Wiley & Sons.
147. Eichenbaum H, Cohen NJ. 2001 *From conditioning to conscious recollection*. New York, NY: Oxford University Press.
148. Striedter GF, Northcutt RG. 2020 *Brains through time: a natural history of vertebrates*. New York, NY: Oxford University Press.

149. Ritchey M, Libby LA, Ranganath C. 2015 Cortico-hippocampal systems involved in memory and cognition: the PMAT framework. *Prog. Brain Res.* **219**, 45–64. (doi:10.1016/bs.pbr.2015.04.001)
150. Schacter DL, Addis DR. 2007 On the constructive episodic simulation of past and future events. *Behav. Brain Sci.* **30**, 331–332. (doi:10.1017/S0140525X07002178)
151. Wheeler MA, Stuss DT, Tulving E. 1997 Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. *Psychol. Bull.* **121**, 331–354. (doi:10.1037/0033-2909.121.3.331)
152. Northoff G, Bermpohl F. 2004 Cortical midline structures and the self. *Trends Cogn. Sci.* **8**, 102–107. (doi:10.1016/j.tics.2004.01.004)
153. Robbins TW, Costa RM. 2017 Habits. *Curr. Biol.* **27**, R1200–R1206. (doi:10.1016/j.cub.2017.09.060)
154. Graybiel AM. 2008 Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.* **31**, 359–387. (doi:10.1146/annurev.neuro.29.051605.112851)
155. Balleine BW, O'Doherty JP. 2010 Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* **35**, 48–69. (doi:10.1038/npp.2009.131)
156. Weinberger NM. 1995 Retuning the brain by fear conditioning. In *The cognitive neurosciences* (ed. MS Gazzaniga), pp. 1071–1090. Cambridge, MA: The MIT Press.
157. Merzenich MM, Jenkins WM. 1995 Cortical plasticity, learning, and learning dysfunction. In *Maturational windows and adult cortical plasticity* (eds B Julesz, I Kovacs), pp. 247–272. Reading, MA: Addison-Wesley.
158. Neubert FX, Mars RB, Thomas AG, Sallet J, Rushworth MF. 2014 Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. *Neuron* **81**, 700–713. (doi:10.1016/j.neuron.2013.11.012)
159. Koehlin E. 2011 Frontal pole function: what is specifically human? *Trends Cogn. Sci.* **15**, 241. (doi:10.1016/j.tics.2011.04.005)
160. Shepherd GM. 2007 Perspectives on olfactory processing, conscious perception, and orbitofrontal cortex. *Ann. N Y Acad. Sci.* **1121**, 87–101. (doi:10.1196/annals.1401.032)
161. Merrick C, Godwin CA, Geisler MW, Morsella E. 2014 The olfactory system as the gateway to the neural correlates of consciousness. *Front. Psychol.* **4**, 1011. (doi:10.3389/fpsyg.2013.01011)
162. de Vries J, Ward LM. 2016 An "ecological" action-based synthesis. *Behav. Brain Sci.* **39**, e173. (doi:10.1017/S0140525X15002046)
163. Murray EA, Wise SP, Graham KS. 2018 Representational specializations of the hippocampus in phylogenetic perspective. *Neurosci. Lett.* **680**, 4–12. (doi:10.1016/j.neulet.2017.04.065)
164. Suryanarayana SM, Pérez-Fernández J, Robertson B, Grillner S. 2021 The Lamprey Forebrain - evolutionary implications. *Brain Behav. Evol.* **30**, 1–16. (doi:10.1159/000517492)
165. Bennett MS. 2021 Five breakthroughs: a first approximation of brain evolution from early bilaterians to humans. *Front. Neuroanat.* **15**, 693346. (doi:10.3389/fnana.2021.693346)
166. LeDoux J, Daw ND. 2018 Surviving threats: neural circuit and computational implications of a new taxonomy of defensive behaviour. *Nat. Rev. Neurosci.* **19**, 269–282. (doi:10.1038/nrn.2018.22)
167. Stephenson-Jones M, Samuelsson E, Ericsson J, Robertson B, Grillner S. 2011 Evolutionary conservation of the basal ganglia as a common vertebrate mechanism for action selection. *Curr. Biol.* **21**, 1081–1091. (doi:10.1016/j.cub.2011.05.001)
168. Miller G. 2010 Is pharma running out of brainy ideas? *Science* **329**, 502–504. (doi:10.1126/science.329.5991.502)
169. Hyman SE. 2013 Psychiatric drug development: diagnosing a crisis. *Cerebrum* **2013**, 5.
170. Griebel G, Holmes A. 2013 50 years of hurdles and hope in anxiolytic drug discovery. *Nat. Rev. Drug Discovery* **12**, 667–687. (doi:10.1038/nrd4075)
171. Kehne JH, Cain CK. 2010 Therapeutic utility of non-peptidic CRF1 receptor antagonists in anxiety, depression, and stress-related disorders: evidence from animal models. *Pharmacol. Ther.* **128**, 460–487. (doi:10.1016/j.pharmthera.2010.08.011)
172. Gray JA, McNaughton N. 2000 *The neuropsychology of anxiety*, 2nd edn. Oxford, UK: Oxford University Press.
173. LeDoux JE, Hofmann SG. 2018 The subjective experience of emotion: a fearful view. *Curr. Opin. Behav. Sci.* **19**, 67–72. (doi:10.1016/j.cobeha.2017.09.011)
174. Taschereau-Dumouchel V, Michel M, Lau H, Hofmann S, LeDoux JE. In press. Subjective fear, implicit metacognition, and the rehabilitation of consciousness as a major factor in the conceptualization and treatment of mental disorders. *Mol. Psychiatry*. (doi:10.1038/s41380-021-01395-5)