

The Transition to Experiencing: II. The Evolution of Associative Learning Based on Feelings

Simona Ginsburg

Department of Natural Science
The Open University of Israel
Raanana, Israel
simona@openu.ac.il

Eva Jablonka

Cohn Institute for the History and Philosophy of Science and Ideas
Tel Aviv University
Tel Aviv, Israel
jablonka@post.tau.ac.il

Abstract

We discuss the evolutionary transition from animals with limited experiencing to animals with unlimited experiencing and basic consciousness. This transition was, we suggest, intimately linked with the evolution of associative learning and with flexible reward systems based on, and modifiable by, learning. During associative learning, new pathways relating stimuli and effects are formed within a highly integrated and continuously active nervous system. We argue that the memory traces left by such new stimulus–effect relations form dynamic, flexible, and varied global sensory states, which we call *categorizing sensory states* (CSSs). These CSSs acquired a function: they came to act as internal “evaluators” and led to positive and negative reinforcement of new behavior. They are therefore the simplest, distinct, first-person motivational states that an animal can have. They constitute what we call basic consciousness, and are the hallmark of animals that can experience. Since associative learning has been found in many invertebrate taxa that first appeared during the Cambrian era, we propose that the processes underlying basic consciousness are phylogenetically ancient, and that their emergence may have fueled the Cambrian explosion.

Keywords

associative learning, basic consciousness, Cambrian explosion, categorizing sensory state (CSS), motivation

In the companion paper (Ginsburg and Jablonka 2007), we suggested that the raw material from which experiencing was molded by natural selection was a global sensory state that is a by-product of the incessant activity of the highly interconnected nervous system that characterizes all neural animals. We argued that overall sensory states became persistent (present-extending) when they became associated with the short- and long-term memory mechanisms involved in the simplest kinds of nervous-system-mediated learning and selective stabilization. Limited experiencing involves a small number of overall-sensory states that are global, integrated, unitary, and private. They are side effects of neural activity and do not play a causal role in learning. Limited experiencing is, however, the foundation on which what we call *unlimited experiencing*, or basic consciousness, arose.

In this article we suggest that limited experiencing evolved into feelings that function as flexible, ontogenetically constructed evaluators and discriminators of new inputs, and generators of new adaptive outputs, when open-ended, associative learning evolved. Of course, when looked at from a computational point of view, associative learning can be quite simple and we do not claim that associative learning *in the computational sense alone* is a sufficient condition for consciousness. The ability of robots or neural networks to learn by association, even when computationally sophisticated, captures only limited and fairly simple aspects of the process in animals, just as a camera captures limited aspects of seeing. What is lacking is an understanding of the experiencing-based value systems that guide associative learning. It is precisely the evolution of these systems, which, as we argue, are dependent on the biological evolution of associative learning from the simpler systems described in the companion paper that is our focus here.

The idea that associative learning evolved from less flexible learning mechanisms was suggested several times. Razran (1971) proposed that associative learning evolved from modulations of sensitization and habituation, and Wells (1968) developed a detailed scenario for the evolution of associative learning from sensitization, suggesting that pseudo-conditioning evolved into increasingly more refined associative sensitization, culminating in associative learning. Our focus, however, is on the central role of new value systems in this evolution. Open-ended, flexible, associative learning requires memorizing *new* associations and attributing intrinsic, whole-organism “value” to them. Because the relations that are formed are new, assigning values cannot be determined by natural selection, through their long-term effects on reproductive success. Instead, the individual organism must have internal criteria for deciding whether a new association or a new behavior is appropriate: it must be able to evaluate, during its own lifetime, whether a response is generally beneficial or detrimental. It must have an internal, flexible, yet robust, evaluative system, which can assess *new* stimuli and responses in

a highly context-dependent, ontogeny-sensitive manner. We call such value systems *reward systems* and focus on their evolution.

The preconditions for consciousness that we suggest, and our focus on value systems, are similar to ideas discussed by Edelman (1987, 2003, 2005) and Edelman and Tononi (2000). Like them, we believe that a necessary condition for consciousness to emerge was a highly distributed, connected neural network that allowed rich reentrant signaling. Reentrant signaling led to ongoing neural reverberations throughout the system as the organism responded to internal and external stimulations, and enabled the selective stabilization of a practically unlimited number of connection patterns. We agree with Edelman and Tononi (2000) that integration of different types of sensory processing, associated with temporal persistence of the bound sensory patterns, was fundamental to the evolution of feelings. We propose that these structures and processes and their associated functions (learning through feelings) appeared very early in animal evolution and drove the Cambrian explosion, and were tied up with the evolution of associative learning. In other words, we suggest that the transition from animals that could not learn by association to those that could led to the emergence of basic consciousness. This transition involved changes in neural organization including cephalization, new types of memory involving new molecular mechanisms, and the formation of sensory states that functioned as motivational states, that is, proper feelings.

Associative Learning and Reward Systems

Associative learning is a behavioral modification involving a *novel association* between two sensory stimuli, or between sensory stimuli and responses, that is brought about by reinforcement. At the neural level, it involves the modification or formation of new connections among neural circuit elements linking sensors and effectors. The nature of the stimuli that enter into association can include “neutral” stimuli, biologically important stimuli such as those that are typically linked to the maintenance of basic homeostatic and reproductive functions, the animal’s own responses, and the contexts in which particular stimuli and responses occur (Macphail and Bolhuis 2001). The crucial thing is that during associative learning a new relation is constructed, remembered, and can be recalled (Rose 1993). Inherent in the notion of associative learning is the assumption that the reinforcement of new behavior depends on internal evaluation—on internal value systems. Usually, the implicit assumption is that the value system is based on feelings of pleasure, pain, fear, and so on. However, generally speaking, the existence (and evolution) of such reward systems is taken for granted, and is not itself a subject of study.

Associative learning has been conclusively documented only in organisms with a central nervous system (CNS). We

found no evidence for it in sponges and the placozoan *Trichoplax*, and little evidence for it in cnidarians and ctenophores (Razran 1971; Rushforth 1973; G. E. Mackie, G. Matsumoto, personal communications, 1 November 2006). On the other hand, evidence for associative learning has been found in nematodes (Rankin 2004), platyhelminthes, crustaceans, arthropods, annelids, mollusks and, of course, chordates (Corning et al. 1973; Abramson 1994). It seems that a CNS is a necessary, though possibly not a sufficient, condition for the evolution of associative learning in animals.

The simplest types of associative learning are through associative sensitization and the modulation of habituation known as “protection from habituation” (Grau and Joynes 2005). Protection from habituation occurs when pairing an additional cue with the habituating one decreases habituation, while an unpaired cue does not. In associative sensitization, pairing two stimuli, in whatever order, sensitizes the neural pathways involved. Although no new stimulus–response link is created, the association between the two stimuli does have an impact on the response. This is different from pseudo-conditioning, another modification of sensitization, where the application of an *unpaired* stimulus (itself inadequate for eliciting the specific response) sensitizes the reaction, and the animal reacts to the original eliciting stimulus more readily.

It is customary to distinguish between two major types of associative learning—classical (Pavlovian) conditioning and instrumental or operant conditioning. Classical conditioning is a modification of behavior in which a new “neutral” stimulus is paired with a stimulus that already elicits a particular response (either because it is innate or because it was learnt at an earlier stage). An organism exposed repeatedly to pairs of the “neutral” and original stimuli will eventually respond to the “neutral” stimulus alone. For example, a dog normally salivates (unconditioned response, UR) at the smell of food (the unconditioned stimulus, US). But if the dog hears a bell (new “neutral” conditioned stimulus, CS) just before smelling the food, it will learn to associate the sound with being fed, and salivate when it hears the bell (conditioned response, CR). A simple and very widespread type of conditioning is inhibitory (punishment) conditioning, where the CS, when paired with the normally eliciting US, inhibits the UR (Razran 1971). Many types of classical conditioning and many combinations of different types of learnt associations have been described in the learning literature (Abramson 1994). However, although an animal can learn many associations, not all associations are equal: it is recognized that the effect of different CS–US pairings depends on the type of stimuli applied and on how “dominant” the reaction to each is, how predictive a CS is of the US, and the relation between the timings of the paired stimuli and the intervals between consecutive trainings (Rescorla 1988). In terms of attractors and reward systems, the attractor in classical conditioning is the new relation formed between

the new stimulus and the preexisting response. The reward systems are those neural/affective systems that stabilize this relation on the basis of their evaluation as pleasant or aversive.

Instrumental (or operant) conditioning is a form of associative learning in which the *actions* of the organism are reinforced by their consequences. For example, a rat placed in a box with a lever accidentally presses the lever and receives a pellet of food; if this happens a few times to a hungry rat, the rat will press the lever more and more often. A temporal association is thus formed between the rat’s response (lever pressing) and a special consequence of this response (food, the reinforcement). As with classical conditioning, there are many types of instrumental learning. They are commonly classified according to the type of reinforcement (desirable or aversive) and the effect of the response (prevents or enhances a desirable or an aversive response). The sensory biases and the motor constraints of the particular species, as well as the specific organization of its nervous system, are important factors in determining the rate and complexity of learning.

With instrumental conditioning, the motor act can be fixed and rigid, but the context in which it is performed can be completely new. For example, pecking behavior may be unlearned and stereotyped, but a bird can learn *what* to peck if the pecked grains are tasty, and this type of learning by consequence is a clear example of instrumental conditioning. It is also possible for the motor behavior to be novel and nonstereotyped (e.g., pressing a lever or jumping through a burning hoop), that is, it is *not* part of the animal’s preexisting behavioral repertoire. In such cases, when the animal learns to reorganize and recombine its motor acts in a new way, conditioning is called operant rather than instrumental. For instrumental and operant conditioning, the attractor is the new relation between the reinforcer and the motor behavior of the animal, and the reward systems are the neural/affective systems underlying reinforcement.

It is sometimes difficult to judge the novelty of a motor behavior and to separate between classical and instrumental/operant conditioning in natural situations. Consider feeding: pecking in a chick, which preexists at the time of training, could nevertheless undergo modifications according to the size and difficulty of handling the food item, and the handling of different items can be learnt. In nature, motor behaviors such as feeding and escaping are inherently flexible: they depend on, and are tailored to, local, somewhat variable, conditions. For example, some birds know that when there is fire (original stimulus) insects come out of their hiding places, so they fly to the vicinity of the burning region (which, of course, varies on different occasions) to feast upon them (Cody 1974). The birds learnt to associate smoke with fire, so smoke alone (the new stimulus) elicits the flexible orienting and motor behavior. The responses of an unrestrained animal in natural conditions are usually plastic, depending on interacting multiple stimuli and responses. The animal’s

learning involves both classical and instrumental conditioning, with different degrees of neutrality of associated inputs and different degrees of motor plasticity, so separating the types of conditioning in such situations is almost impossible (Brembs 2000). We believe that simple classical and instrumental conditioning were always intertwined in evolution (although they can be separated in laboratory conditions). Consequently, we disagree with Razran (1971), who suggested that classical conditioning preceded instrumental conditioning in evolution. Razran based his proposal on the greater memorability of instrumentally learnt behavior, on data suggesting the greater prevalence of classical conditioning in lower invertebrates, and on the greater novelty of instrumentally learnt reactions. The greater memorability of instrumentally learnt responses is not an argument in favor of later evolutionary origin, however, and neither is the argument that instrumental behavior leads to more novelty, since the number of perceptual stimuli and associations seems almost unlimited. Recent data from invertebrate taxa show that the distribution of classical and instrumental conditioning is equally broad (Abramson 1994).

Associative learning was, of course, modulated and sophisticated during evolution, leading to new types of learning and of value systems. In this article, however, we focus on the simplest learning and reward systems, so we do not discuss the evolution of these more complex learning mechanisms.

The importance of associative learning in ontogeny and evolution has not escaped the notice of philosophers (Dennett 1995; Beisecker 1999; Dretske 1999), and psychologists consider associative learning to be the most fundamental form of learning. Macphail (1987) and Macphail and Bolhuis (2001) have argued that learning in all vertebrate species, from goldfish to chimpanzee, can be explained by a general mechanism of associative learning. Although we disagree with this view, the generality and commonality of basic mechanisms of associative learning in both vertebrate and invertebrate taxa are not in doubt, and we accept that these mechanisms form the foundation for other species-specific modulations in learning ability. Moreover, as we argue below, the evolution of *long-term memory for new associations* led to the formation of a value system based on feelings.

Constructing Associative Learning

We assume that associative learning evolved on the basis of simpler mechanisms of neural learning. Although we do not rule out the possibility that a limited number of new associations may form in animals without a highly centralized nervous system, we think that the evolution of flexible and varied experiencing and learning required the interrelated evolution of cephalization, new memory mechanisms, and new value systems.

Cephalization

The evolution of greater centralization of the nervous system was probably related to the dawn of bilateral symmetry. A change in body shape from radial to bilateral need not involve drastic changes in neural machinery, but it inevitably leads to an anatomical reorganization of the nervous system, with far-reaching consequences for learning and memory. Once a single, forward direction of locomotion was defined, the anterior parts of animals became the first to meet or seek various stimuli in the environment. This resulted in the concentration of sensory cells in the front end of the animal—its head—and of other neural elements along the body. Such differentiation between sensory and motor regions allowed integration of sensory inputs and coordination of motor outputs *within* regions as well as coordinated interactions *between* regions.

Flatworms of the phylum Platyhelminthes are some of the earliest organisms with definite external bilateral symmetry. The planarian *Dugesia*, for example, has a nervous system consisting of a CNS and a peripheral nervous system (PNS). The CNS is made up of two ganglia in the head region (the brain) and a pair of ventral longitudinal nerve cords extending along the anterior–posterior axis. The brain has a rich wiring, with two-way anatomical projections leading to and from the sensory organs, converging input wiring from different sensors, and some direct connections to and from the nerve cords (Sarnat and Netsky 1985; Okamoto et al. 2005). Through their organization, various external signals can be integrated in the brain, and the outputs can be transmitted to the whole body. The PNS consists of sensory neurons located along the entire body and lateral nerves arranged in rings and plexuses perpendicular to the longitudinal nerve cords. This organization allows efficient motor control of parts of the body lying in close proximity and integration of the sensory input from these parts.

The type of rudimentary division of labor seen in the bilateral nervous system of flatworms paved the way for further specialization in the segmented worms. In segmented invertebrates, such as the annelids, in addition to the composite head ganglia, each body segment is equipped with a distinct ganglion. A segmental ganglion serves the reflex functions of the segment it occupies, as well as one or more adjacent body segments. Nerve fiber trunks—connectives—join the ganglia of successive body segments. Such an arrangement is a major evolutionary advance, because it enables extensive interconnections among neurons, with an economy of fiber branching.

In principle, centralization, which is a necessary condition for sensory integration and fine motor control, could be achieved by connecting neurons in a variety of ways. However, grouping individual neurons in ganglia, the fusion of head ganglia into a brain, and the formation of some long connections between nerve clusters minimize the combination of total wire length and the number of processing steps along

pathways. Neural networks arranged in this manner thus seem to optimize both wiring and processing efficiency (Kaiser and Hilgetag 2006). An innovation involving the growth of new connections between neurons is also very important: it allows shortcuts among active neurons, which enable better integration and more efficient signal transduction.

The Evolution of Memory

We suggest that generalized long-term memory evolved in two main stages. The first, which we discussed in the companion paper, was the evolution of mechanisms of short- and long-term habituation and sensitization—the temporal, quantitative modulation of existing reflex paths. This stage also involved the *transient* selective stabilization of *newly used* neural pathways. The second stage involved the evolution of *long-term* stabilization of new associations: memory traces persisted even in the absence of the activating stimulus. We suggest that the molecular memory mechanisms underlying long-term sensitization and habituation were recruited and further elaborated as associative learning evolved.

An example of simple associative learning in which long-term sensitization mechanisms are recruited and modulated is conditioning of the siphon withdrawal reflex in *Aplysia* (Antonov et al. 2001). Lightly tapping the siphon has a small effect on siphon withdrawal, whereas a shock to the tail leads to an enhanced withdrawal response. After repeatedly pairing siphon tapping with a tail shock, mechanical tapping on its own becomes more effective in eliciting siphon withdrawal. Two cellular mechanisms contribute to this conditioning. The first is activity-dependent facilitation: the tail shock activates sensory neurons in the tail, and these synapse onto interneurons that produce presynaptic facilitation of siphon sensory neurons (which synapse on the motor neurons) as well as interneurons that excite the motor neurons. Compared to unpaired training, paired training results in a greater increase of evoked firing from the motor neurons and greater facilitation of the monosynaptic postsynaptic potentials. This is correlated with the enhanced siphon withdrawal. The second cellular process contributing to conditioning is long-term potentiation (LTP) (Murphy and Glanzman 1997, 1999; Antonov et al. 2003), which is mediated by N-methyl-d-aspartate (NMDA) receptors—a subclass of glutamate receptors. The LTP mechanism acts as the detector of the temporal association between the paired stimuli. Thus, activity-dependent facilitation and LTP both contribute to conditioning, and the two processes may not be independent. The overall picture is that conditioning in *Aplysia* involves both presynaptic and postsynaptic associative mechanisms, acting together.

The evolution of the long-term memory underlying associative learning probably involved other new mechanisms, like the back-propagating action potentials found in mammalian hippocampal neurons (Markram et al. 1997; Paulsen

and Sejnowski 2000). It also relied on the growth and retraction of synaptic connections and other structural synaptic changes (e.g., Kim et al. 2003).

The simplest general neural model for Pavlovian associative learning assumes that sensory pathways for two independent stimuli (CS and US) converge anatomically (Martin et al. 2000). Synaptic strength is increased by coincident stimulation by the CS and US affecting the locus of convergence. Mechanisms such as LTP bring about changes in synaptic efficacy between the CS sensory pathway and connected (follower) neurons, ultimately ending in motor output. Thus, subsequent behavioral responses to the CS alone are strengthened. Other, more global memory mechanisms have also evolved. Several models suggest that memory processing is a system-level phenomenon in which patterns of firing by neuronal ensembles play a major role (Dubnau et al. 2003). Reentrant signaling, that is, the back and forth transmission of signals along reciprocal connections between groups of neurons (Edelman 1987), may enable finer and more reliable relations between specific stimuli and responses, and also make the signaling between nerve groups more persistent (Edelman and Tononi 2000).

In addition, the evolution of more diverse systems of neurotransmitters and neuromodulators must have been crucial to the ability to integrate and coordinate increasingly more sophisticated pathways between stimuli and responses. It is known that different categories of responses employ shared chemicals; for example, the ratio of serotonin to octopamine controls whether the behavior of lobsters is submissive or aggressive (Kravitz 1990), and serotonin is involved in the neuromodulation of various chemosensory circuits in *Caenorhabditis elegans*, including those associated with food aversion (Chao et al. 2004). Thus, even in invertebrates, the combination of diverse neurotransmitters and neuromodulators has further elaborated memory mechanisms and given enormous scope for associative learning.

The Evolution of Feelings: From Value Systems to Reward Systems

The evolution of feelings was closely allied with that of memory, and we suggest that it, too, went through two main stages. The foundation of feelings is the integrated, incessant activity present in the interconnected nervous system—the very fuzzy, constantly changing, overall sensation. As discussed in the companion paper, the first stage in the evolution of feelings was associated with the distinct, integrated sensory signatures conferred on the overall sensation when particular reflex paths became activated and could be modulated, and with the transient selective stabilization of neural sensory-motor pathways. Although still associated with innate reflex paths, many newly used, transient, distinct connections could contribute to overall sensation. However, these overall sensory states did not play any role in learning. They simply accompanied

transient selective stabilization of new neural pathways. They did not serve as criteria or motivational states for motor activities. We suggest that animals with this type of sensory state have “limited experiencing” and that the mechanisms underlying limited experiencing were the raw material from which open-ended learning and feelings evolved.

The second stage, the focus of this section, was a facet of the evolution of memory for newly formed associations, which left memory traces that lasted after the stimulus was no longer present. Such memory made associative learning possible and probably depended on cephalization. Since with associative learning the animal always and inevitably learns something new, the value systems that enable reinforcement and make learning possible have to accommodate novelty—they must be plastic, adjusted to the specific ontogeny of the individual animal. Distinct overall sensations, the by-products of neural activities, became not only more numerous, but they also began to play a causal role in the learning process. So how did a *state of the system*—accompanying feelings, become a *state for the system*—motivating feelings?

In outline, we suggest that the overall sensation that is elicited when an animal *first* encounters a new situation is still merely a state of the system, a by-product accompanying an unlearned response. However, when the animal learns a new relation, its neural effects are integrated within the original overall sensation, and this sensation—the feeling—can then be triggered by the newly learnt cue, and guide the corresponding behavior upon subsequent encounters. We will therefore argue that the function of feeling is to alert and “inform” the animal about its present general state and to trigger adaptive behavior.

Even in a relatively simple animal (e.g., an annelid worm), many different neural trajectories are involved in producing, for example, a food-seeking response. Because the neurons are highly interconnected and closely packed and because the animal can memorize associations, following behavioral and neural exploration, new trajectories become inevitably linked to the reflex network. Every encounter with a stimulus that elicits food seeking activates parts of this network, and neural connections that are repeatedly activated by encounters of the same type are strengthened, stabilized, and made persistent. This is the basis of both the ability to learn by association and the extension of the present.

The long-term stabilization of neural connections that are repeatedly activated when a particular response is elicited occurs in a ceaselessly active nervous system, against the background of overall sensation. It confers on the nervous system a particular distinct and persistent signature, a sensory state that is the consequence of the animal’s learning history. We call any activated, distinct, history-dependent, persistent overall sensory state a CSS. At the organism level, the CSS is a global sensory neurophysiological state—a feeling, an experiencing.

Every CSS is built around a basic reflex trajectory. However, this fixed trajectory can be thought of as a neurophysiological scaffold, which may become redundant after its initial activation, so that a particular CSS in the adult may have little similarity to the original, fixed-reflex scaffold. In the same individual at different stages of its life, as well as in different individuals, the CSS leading to food-seeking behavior, for example, may be different in detail, because the learning history and memory of the encounters with the food-seeking stimuli are different. The patterns of connections of the neurons that are activated during a given type of response have family resemblances, but are not identical. They share many connections, but no single connection is obligatory. The effects of any new stimulus or response can become part of a particular CSS; they can belong to the CSS that elicits food seeking, or withdrawal, or partner seeking. Thus, although CSSs are constrained to a certain domain of the animal’s actions, they have great flexibility within this domain, and yet are robust. Hence, each CSS is both typical (constrained, belonging to a type) and idiosyncratic because of the uniqueness of each individual’s ontogeny. Moreover, once CSSs are formed, the neural effects of a single incoming stimulus are produced in the context of other, preexisting, and previously formed neural trajectories. The effect of any additional incoming stimulus is dependent on this integration. The integrated internal state, rather than the effect of a single stimulus, becomes the cause of the behavior.

During associative learning, new paths of exploration lead to an “attractor,” the dynamic end state at which the animal’s physiology “settles.” For example, a sensory state associated with a need for food initiates and then constrains and guides the exploration process toward the stabilized attractor state, satiation. A new learnt cue associated with the food that is found will have effects on the CSS and trigger its future activation. In other words, the animal learns that the cue is specifically predictive (in this case that it predicts food), so when encountered again, the CSS will guide the food-related behavior of the animal.

CSSs are therefore flexible *value systems*: they are the neural dynamic structures that are necessary for the animal to modify its behavior according to the salience of a new cue. They guide the selective stabilization processes that occur in the nervous system and underlie the behavioral modifications taking place during and following learning. Since they are defined as patterns of *overall sensation*, their reinforcement-enabling effects are based on their global, yet specific and integrated sensory pattern, which is what we refer to as feeling. They are *reward systems*, because they are value systems based on feelings.

Every activated CSS has a value, which we refer to as either “positive” or “negative.” A CSS that the organism maintains or enhances is “positive”; one that it alters or reverses is “negative.” As we pointed out earlier, the selective stabilization

processes that lead to either the persistence or the modification of neural activity states (the precursors of CSSs) operate around preexisting, evolved, reflex trajectories. With CSSs, the general basis for a given “value” is founded on these precursor dynamic states, and is therefore established by past evolution. However, the effects of selective stabilization processes are memorized, and CSSs are therefore flexibly modified by ontogenetic experience. The adaptive state toward which the system tends is the attractor, and we can assign to the CSS that leads to it a positive or a negative value, which has a corresponding sensory signature. CSS maintenance is termed “positive” because activities that lead to its maintenance have usually been beneficial (over evolutionary history in general and over the animal’s ontogeny in particular). For example, upon detecting food, an organism may act in ways that *maintain* its existing sensory state, leading it to approach the food and consume it, so the CSS is “positive.” In the cases we term negative, only if activities *modify* the CSS is fitness enhanced, so the persistence of the CSS is usually detrimental. For example, when the animal has a CSS correlated with conditions leading to tissue damage (a fitness-reducing condition), it will act to *alter* the activity associated with this CSS, so the CSS is negative, and the organism acts in order to reach a different (positive) attractor CSS state. CSSs, therefore, not only evoke behavior but also *guide* behavior, by being positively or negatively loaded.

We stress that it is the internal dynamic sensory state—the CSS—that has a “value,” not the stimulus nor the actions of the organism. Once CSSs evolved, they became the major organizing causes of the animal’s actions. The effect of an external stimulus became merely a component of the system, contingently related to the direct organizing cause—the CSS. The same sensory or motor stimulus can activate different CSSs, either positive or negative, and hence elicit different behaviors. Moreover, internal stimulations of the network components of the CSS can activate it, so a response can be evoked without requiring an external stimulus. Although the overall activity in the whole nervous system contributes to every CSS, which sensory stimulation causes *which* change in behavior depends on the particular CSS in which it is incorporated.

The emergence of CSSs meant that an internal, flexible value system had evolved. A CSS categorizes as “reward” any action that relieves a “negative” or reinforces a “positive” sensory state. Change from a “negative” (“food-requiring”) state and reinforcement of a new “positive” state (“food-consuming”) will occur during trial-and-error behavior, with the “attractor”—the state to which the system “strives”—being a fitness-enhancing end state (“satiation”). Many different combinations of new stimuli can therefore be categorized as either “rewards” or “punishments,” depending on whether they contribute to positive or negative CSSs respectively. Each CSS can thus function as a special type of motivational state. We suggest, then, a definition of basic feeling: *basic feeling is*

an evaluative, categorizing sensory state (CSS) that (usually) leads to adaptive behavior. We suggest that CSSs with their assigned values form basic consciousness.

We can now account for the way in which the global sensory state *of* a system became a state *for* the system during evolution. The function of CSSs is to evaluate inputs and generate outputs: the effects of an input (an external stimulus or the animal’s own actions) (1) are incorporated into a preexisting CSS on the basis of the previous learning history of the animal and (2) generate actions by guiding selective stabilization processes, which maintain/reinforce or modify/reverse this CSS until an adaptive state is reached. A learnt external or internal input into the sensory region in the brain is first integrated within this region, and then evokes the activation of the appropriate neurons in motor regions. The individual-specific, learnt inputs that are received evoke a persistent, integrated global sensory state (a feeling—“fear,” “hunger”) that leads to a particular category of exploratory action (run away, seek food). These actions drive the animal until another integrated global and persistent sensory state, a feeling recognized as “good” and associated with an adaptive end state is reached. In James’s ([1890] 1950, vol. I: 140) words, consciousness (and experiencing) brings “more or less constant pressure to bear in favor of *those* of its [the animal’s] performances which make for the most permanent interest of the brain’s owner.” Consciousness is a “*fighter for ends*, of which many, but for its presence, would not be ends at all” (p. 141). “Fighting for ends” occurs through associative learning: “blind” reflex-like actions cease to be “blind” following their first activations, because the neural effects of newly learnt salient cues become related to fixed neural circuits (see James [1950, vol. II: 390]). In our terms, CSSs are constructed, and they begin to guide the animal’s actions. For example, an animal that upon an encounter with a new cue had the inborn feeling of fear that accompanies the escape reflex, in later encounters will have the newly learnt cue evoke a very similar feeling of fear, which both precedes and guides its escape. Moreover, the CSSs themselves, the feelings themselves, just like the cues that elicited them and the motor actions that they elicit, are modulated through learning, and can act as ever more sophisticated elicitors of action.

Associative learning and CSSs evolved hand in hand, appearing in early Bilateria: the ability to memorize new associations and the formation of overall, yet distinct, integrated sensory states that function as reward systems, are interrelated. Although the memory of an animal at any particular stage of life may be very limited, once associative learning and CSSs had evolved, the number of *potential*, newly learnt associations was greatly expanded, because many new links between the effects of stimuli and responses could belong to the same CSS.

The evolutionary scheme we propose is illustrated in Figure 1. We start with a hypothetical ancestral metazoan that had localized neural connections, but no integrating nerve

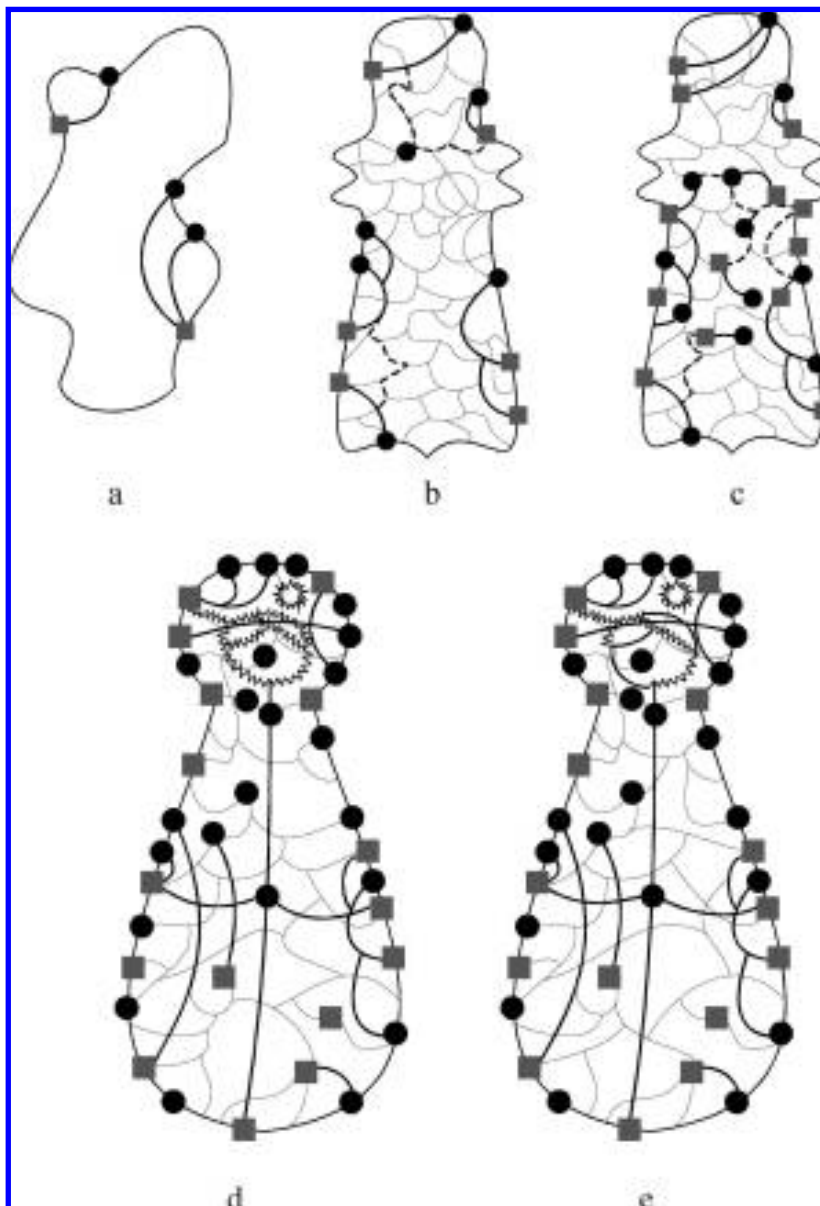


Figure 1. Stages in the evolution of associative learning. a: Hypothetical ancestor with only local reflexes (heavy continuous lines) that can be modified through habituation and sensitization. The circles denote sensors; the squares denote effectors. b: The first nerve net; the integrated net has several reflex (heavy lines) and many labile nonreflex neural connections (light lines): Some connections are transiently reinforced (broken lines). c: A more integrated nerve net, with some centralization and new, transiently reinforced connections; new associations are not yet remembered. d: Cephalization and new memory mechanisms enable the formation of CSSs; some new connections are persistently stabilized (skewed lines). e: Selection of genes that facilitate particular learnt responses lead to connections that are stabilized without learning (new heavy lines replacing some of the skewed lines in d).

net; it could learn through local habituation and sensitization (Figure 1a). Such an animal, if it ever existed, was probably evolutionarily short-lived. It was soon replaced by an animal with an interconnected nervous system, a nerve net, formed by connections between nerve cells (Figure 1b). Interconnectivity enabled selection for flexible responses built around the reflex circuits. A by-product of the nerve net was the first overall, white noise sensation. Natural selection for flexibility in this metazoan led to a more integrated system, where parts communicated via a more centralized interconnected nerve net

and new connections were transiently reinforced (Figure 1c). These new connections were associated with preexisting reflex trajectories and led to limited experiencing—to a limited partitioning of “sensory space.” There was, as yet, no long-term stabilization of *new* associations. The evolution of new anatomical features and mechanisms, such as increased cephalization and differentiation into sensory and motor brain regions, LTP, more neuromodulators, and reentrant signaling, led to memory for new associations (Figure 1d). Such memory involved sensory categorization of types of stimuli–response relations that

resulted from multiple-event encounters, and to the formation of CSSs that acted as motivators for action, guiding the animal to an attractor end state. When ontogenetically constructed, CSSs proved consistently adaptive, and when it was beneficial to learn more and more quickly, some neural connections that contributed to the CSS were genetically assimilated (see next section), and developed prior to actual experience (Figure 1e). As we argue in the next section, this was important for further evolution, as animals became adapted, both morphologically and cognitively, to new niches.

The simple affective systems that we described were organized around the reflex systems that enabled the animal to respond adaptively to changes in the world. These affective states were probably similar to the primordial emotions described by Denton (2006): a state akin to “hunger” related to food seeking, a state related to the pursuit of a sexual partner (“sexual urge”), withdrawal or freezing upon injury (“fear,” “panic”), and a few additional affective states related to existential needs (e.g., “thirst” and “suffocation” in land animals). In early Bilaterians, the variations in the neural networks of each CSS were limited, and each CSS had either a positive or a negative value, the basic “currencies” of reinforcement. It was only at later evolutionary stages that these affective systems were honed into more organized and complex affective categories. Only then did specialized pleasure and pain neural systems (Cabanac 1992) and more specialized neural structures directing what are called primary emotions evolve (Panksepp 1998, 2004).

Did Associative Learning Drive the Cambrian Explosion?

The transition to organisms with basic consciousness depended on the evolution of the neuron and the nerve net. No phylum older than the Ctenophora has a cell type that can be interpreted as a precursor of the neuron, although many of the molecular components of neurons, such as ion channels and receptor molecules, are found in protists, and the mechanisms of signal transduction in protozoans and metazoans are similar. Neurons and neural communication systems probably emerged quite suddenly. Stanley (1992) suggested that one of the intrinsic biological barriers that delayed the onset of the adaptive radiation of the metazoans may have been the difficulty of evolving the neuron. Before its evolution, the use of specialized contractile cells for feeding and locomotion was very limited, as is evident from the Porifera (sponges). But the advent of the neuron, with its ability to rapidly propagate electrical signals and its endless connectivity, opened up new possibilities. A dramatic realization of these possibilities occurred in what is known as the Cambrian explosion, a geologically very short period, 530 to 515 million years ago, when many metazoan phyla first appeared in the fossil record (Valentine 2002).

Although paleontologists all agree that very rapid morphological diversification occurred during this epoch, there are many controversies concerning the causes of this explosive radiation, and the relation between morphological and molecular evolution (see Wang et al. [1999] and Blair and Hedges [2005] for arguments for a pre-Cambrian origin of animal phyla, and Rokas et al. [2005] for the view that the rates of morphological and molecular evolution are compatible). Among the suggested external, environmental causes for the morphological and taxonomic radiation are changes in atmospheric free oxygen concentration, changes in the chemical constitution of the oceans, and drastic climatic changes. However, even if it is accepted that one or several external environmental changes were the trigger for the explosion, the nature of the *internal* biological changes that led to the patterns of diversification observed still has to be resolved. Some of the suggested internal causes are the reorganization of key regulatory genes in developmental networks, which led to new body plans; an ecological arms race driven by macroscopic predation (Conway Morris 2000); and the evolution of a new modality—sight (Parker 2003).

We suggest that the Cambrian explosion may have been fueled primarily by the emergence of associative learning and that it marks the ascendancy of learning and consciousness as driving forces in animal evolution. Hameroff (1998) has also suggested that the emergence of consciousness may have caused the Cambrian explosion, but he attributed it to emergent processes at the quantum level, which involved microtubule cytoskeletal structures. We propose that associative learning drove the evolution of new sensory modalities, of macroscopic predation, of new and more efficient feeding methods, and of new relations between ancient regulatory genes. The fossilized traces of Cambrian animals suggest that they were invading new niches and developing new behaviors such as active hunting, burrowing deeply into sediment, and making complex branching burrows (Conway Morris 2000, 2003).

The idea that learnt behaviors can drive physiological and morphological evolution is not new; it has been suggested repeatedly since the late 19th century. Hardy (1965, chs. 6 and 7) provided an excellent summary of these early ideas, and argued that what is known as the “Baldwin effect” had a major role in animal evolution. The Baldwin effect is the idea that evolutionary change is initiated by behaviorally learnt (or physiologically induced) ontogenetic adjustments to new environments, which foster and pave the way for hereditary changes that simulate the ontogenetic changes. Baldwin’s basic idea was later elaborated and empirically validated by Waddington, who showed how the reassortment and selection of combinations of existing genes can construct new genetic networks that produce inborn responses similar to those that were environmentally induced, a process he called “genetic assimilation” (Waddington 1957). A more inclusive term, “genetic accommodation,”

has been suggested by West-Eberhard (2003), who argues that any heritable variation in morphology, physiology, or behavior that contributes to the effectiveness of an adaptive ontogenetic response (making it more reliable, rapid, and precise, and/or ameliorating detrimental side-effects) will, through natural selection, be “accommodated” by the genetic system. Bateson (2005, 2006) suggested that adaptive plasticity in general, and learning in particular, guides genetic evolution, and called this guidance “the adaptability driver.”

Hardy believed that the dramatic adaptive radiations of reptiles, and later of mammals and birds, were driven by their enhanced learning ability. He argued that new learning strategies, which involved greater inquisitiveness and exploration, and in birds and mammals also social learning, led to the invasion of new niches and thus to new selection regimes, a process that was later called “behavioral drive” (Wyles et al. 1983). The selective regimes driven by new habits led both to neural adaptations and to morphological and physiological changes that enhanced the adaptedness of the learnt behaviors (Hardy 1965; Wyles et al. 1983). West-Eberhard (2003, ch. 28), too, suggests that developmental plasticity, including learning, had an important role in adaptive radiations.

Our suggestion that associative learning drove the Cambrian explosion is based on the same reasoning. Organisms that could learn flexibly by association had an enormous selective advantage. They could associate many *new* stimuli with old ones and many new acts with outcomes. They therefore had an increased capacity to discriminate and could anticipate the effects of environmental change and their own actions. Their ability to adapt ontogenetically to a variety of environments and to use new resources increased dramatically. These ontogenetic adaptations determined where animals looked for food and protection, how they sought mates and handled food, how they reacted to predation, and so on. They were fundamental to the construction of the niches that animals and their offspring inhabited (Avital and Jablonka 2000; Odling-Smee et al. 2003). For example, an animal that learnt that food sources are available in a particular area tended to stay and reproduce there; an animal that learnt that a hole in a rock affords protection against predators tended to hide in holes, and was likely to reproduce in or near them. Such behaviors introduce the offspring to the same learning environment and learning opportunities, and may lead to the genetic accommodation of any physiological, neurological, or morphological feature that enables more effective adaptation to this learning environment (including more sophisticated or more efficient learning ability). The explosion of new behaviors was therefore accompanied by an explosion of new congruent morphologies, which were probably based on regulatory modifications of genes in the existing developmental networks.

Learning may also have driven the directional evolution of particular sensors. For example, if the proximity to a food

source was frequently associated with a change in light intensity, the association could have driven the evolution of greater ability to discriminate visually. Rapid, directional evolution of vision and other modalities, such as touch, taste, and smell, is expected once associative learning has evolved. The evolution of the eye, which Parker (2003) suggested drove the Cambrian explosion, was, we believe, just one of the several important consequences of the evolution of associative learning (which is underlain by a highly interconnected nervous system, developing according to action-dependent processes), each with strong evolutionary domino effects.

In addition to directing the morphological and behavioral evolution of the lineage in which it evolved, learning had effects on other species’ evolution, because learning in one animal can exert strong selection pressure on the organisms with which it interacts. For example, as the learning ability of a predator species improves, there is strong selection for morphological and physiological adaptations (such as protective hard parts and escape reactions) in its prey species, and also for the prey’s ability to learn. Consequently, not only did associative learning lead to the radiation of the group in which it emerged, it also probably led to learning-guided morphological coevolution and a learning arms race in interacting species.

Discussion

We sketched the transition from an animal with limited experiencing and nonmotivational feelings to one with basic consciousness and motivations. We suggested that associative learning required the evolution of memory mechanisms that allow the long-term persistence of newly learnt and integrated cause–effect relations. In the framework of an interconnected, nervous system with overall sensation, the evolution of these memory mechanisms led to the formation of systemic CSSs, that is, motivating feelings. The function of CSSs is to act as internal guides and selectors of new relations and new behaviors. Their evolutionary emergence gave rise to the first motivations and the most basic types of individual experiences. Since these experiences are integrated, internally reverberating, whole-organism physiological effects, their accessibility to others is by definition (and nonmysteriously) limited. We see basic consciousness as the processes enabling the learning of new relations through the mediation of feelings. Intelligence could take off only in conjunction with feeling.

Surprisingly, few modern attempts have been made to describe the evolutionary origins of basic consciousness. Humphrey (1992, 2000, 2006) suggested that consciousness is founded on sensations (equivalent to what we call feelings) through which the animal knows what happens to it. He proposed that sensory activities developed gradually from local stimulation and responses, through local sensory stimulations

transmitted by nerves that feed back to the site of stimulation; finally, the sensory stimulation feeds back on itself in the sensory part of the brain, generating consciousness. Humphrey regards sensation as different in nature and processing from perception, through which the animal is informed about what happens in the external world. The prominence of sensation (rather than perception) is also central to Denton's view, which stresses the primary importance of internal sensors that monitor basic homeostatic functions (Denton 2006).

Our position differs from both Humphrey's and Denton's. First, we disagree with the categorical distinction between sensation and perception: an animal learns what happens in the external world by knowing (feeling) what happens to it. Although perceptions do not depend on associative learning and motivational feelings in extant perceiving animals, they are *evolutionarily* dependent on them. Second, unlike Humphrey and Denton, we link the evolution of consciousness to the evolution of learning and the dawn of the great metazoan radiation. Our view is that CSSs underlie the ability of organisms to feel and experience, and that they evolved in the context of selection for the memory involved in associative learning.

We argued that all the above processes are phylogenetically far more ancient than commonly assumed, and have come together through selection for associative learning. They do not require the vertebrate neural organization and brain areas that Edelman (1987, 2003, 2005), Crick (1994), Panksepp (1998), Damasio (1999), Koch (2004), Merker (2007), and others identified as generally necessary. We suggested that CSSs made their appearance in the first Bilateria that were able to learn by association, although precursors were present in ancient, cnidaria-like animals. However, the number of CSSs was initially very small, and there were strong constraints on the number of possible variants within a given CSS, reflecting the learning and memory limitations of the animals. As we see it, basic consciousness is not restricted to vertebrates, although we believe that vertebrates (and possibly other groups such as the cephalopods) did evolve qualitatively new aspects of consciousness, such as pain and pleasure, thoughts and beliefs, which are evolutionary derivatives of basic CSSs and reflect complex, evolved relations among them.

Our account raises the question of the "conscious status" of animals with a small number of neurons: Can animals with very few neurons—like *Caenorhabditis elegans*—that do learn by association be said to have motivating feelings and basic consciousness? We are here in the gray area where answers to such questions strongly depend on the extent of integration and the response richness one deems necessary. This dilemma is similar to that of researchers of the origin of life who study limited heredity systems. Can a short RNA molecule that replicates in a defined chemical milieu in laboratory conditions be said to be alive? We think that this question, like the question

about the status of consciousness in nematodes, is not very constructive and is not likely to advance research. When fundamental evolutionary transitions occur, there are likely to be gray areas between sharply defined biological organizations, and these gray areas evade clear definition. Our own position is that if an animal has integrated overall sensations, and these direct and guide its actions so that it has the potential to learn flexibly by association, it can be said to be conscious even when this (very basic) consciousness is very limited indeed.

Although feelings and associative learning are evolutionarily linked, in highly specialized nervous systems the ties between learning and feelings can be decoupled, leading to both normal and pathological conditions (Merker 2007). When learning occurs during dreaming, feeling is present, but inhibition of motor acts probably leads to the inhibition of some affective circuits. Learning under anesthesia (Garcia et al. 1974) involves the inhibition of probably all affective neural circuits, and with subliminal learning—learning relations between events without the subject being aware of the association presented to her (Wong et al. 2004)—some affective circuits are inhibited and attention is directed toward more salient events. Although instrumental learning involving a leg withdrawal reflex can occur in spinal rats (rats whose spinal cord has been severed), this learning is very limited (Grau and Joynes 2005); operant conditioning is impossible in spinal animals.

Learning by association without feeling reflects the ability of already specialized parts of a complex nervous system to become decoupled in certain natural and pathological/experimental conditions. Learning evolved in intact animals that needed to know about relations between salient events and to do something about them. Feeling-independent learning points to already specialized and evolutionarily derived features of a nervous system. Zombies are possible only as pathologies of previously conscious animals.

We are well aware that our account of the evolution of basic consciousness is very sketchy. We have only hinted at the role of systemic hormonal communication explored by LeDoux (1996) and have overlooked many other ingredients of the processes involved. However, we think that whatever the specific shortcomings of our scenario, the evolutionary framework and the emphasis we have suggested are useful and lead to some specific predictions and implications, which can be studied through experimental and comparative methods:

1. The reentry mechanisms suggested by Edelman will be found in the simplest Bilateria that are capable of learning by association.
2. The evolution of the eyes in cubozoans could have been driven by intense competition with already associatively learning Bilaterians, and if so the evolution of the eyes in cubozoans

is predicted to be post-Cambrian. Similarly, the evolution of new elaborate sensory and defense mechanisms in other cnidarian taxa lacking a CNS and flexible associative learning will be found to be associated with the Cambrian explosion and the appearance of Cambrian, associatively learning animals.

3. Comparative studies of the proteins and organization of the signal transduction pathways that underlie classical and instrumental conditioning in different groups will indicate whether they originated through diversification from one common ancestor or, as we think likely, through parallel evolution in different taxa.

4. We expect secondary radialization of body form, with its consequent reversion into a decentralized, radially organized nervous system, to be correlated with a significant reduction in the ability to learn by association (e.g., in the adult, but not the bilateral larval stage of sea urchins).

5. We expect that the neural mechanisms that underlie neural maturation and habit-driven behaviors on the one hand, and exploratory learning on the other, will differ. The absence of a behavioral exploratory phase in maturational triggering and habit-driven behavior and its obligatory presence during exploratory learning suggest that the latter requires additional neural mechanisms.

6. Since we argued that whatever its local function, the animal's "value" system, unlike the robot's, is based on feelings, associative learning cannot be a sufficient condition for consciousness (although it is a necessary and sufficient condition for basic consciousness in animals). Nevertheless, it should be possible to build feeling robots if (1) they are embodied and have a highly interconnected peripheral system linked to very rich arrays of external and internal sensors and effectors, and to a central system; (2) in addition to various algorithms, including conditional fixed-pattern and open-ended algorithms, the robot has a few but fundamental preinstalled value systems (such as "avoid mechanical and electrical damage" or "reconnect to a source of electricity every X hours"); (3) there is associative memory coupled with specific reentrant connections between elements of the system at different levels of hierarchical organization; and (4) there are parts of the system that allow highly efficient information integration at levels and on time scales which are equivalent to nervous, hormonal, and immune system integration and time scales.

First-person experiences need to be included in the scientific discussions of animal psychology. We should therefore add one more "cause" to the four that Tinbergen (1963) listed as necessary for biological explanations. In addition to his phylogenetic, functional, developmental, and immediate causes, which apply to all (neural and nonneural) living organisms, in metazoans with a CNS there is a fifth cause—intrinsic motivating experiencing—that is special to them and sets them apart.

Acknowledgments

We thank Konstantin Anokhin, Patrick Bateson, Daniel Dor, Iris Fry, Alex Gamma, Ehud Lamm, Denis Noble, Stephen Simpson, and Iddo Tavory for their useful and detailed comments on the ideas expressed in an earlier version of this article and Kim Sterelny for his helpful editorial and conceptual suggestions. We are very grateful to George Mackie and George Matsumoto for kindly answering our questions on cnidarians and to Charles Abramson for answering our questions on associative learning. Special thanks are due to Marion Lamb, who patiently read several versions of this article and suggested many improvements. We also thank Yael Givon, who drew the figure.

References

- Abramson CI (1994) *A Primer of Invertebrate Learning: The Behavioral Perspective*. Washington, DC: American Psychological Association.
- Antonov I, Antonova I, Kandel ER, Hawkins RD (2001) The contribution of activity-dependent synaptic plasticity to classical conditioning in *Aplysia*. *Journal of Neuroscience* 21: 6413–6422.
- Antonov I, Antonova I, Kandel E, Hawkins RD (2003) Activity-dependent presynaptic facilitation and Hebbian LTP are both required and interact during classical conditioning in *Aplysia*. *Neuron* 37: 135–147.
- Avital E, Jablonka E (2000) *Animal Traditions: Behavioural Inheritance in Evolution*. Cambridge: Cambridge University Press.
- Bateson P (2005) The return of the whole organism. *Journal of Biosciences* 30: 31–39.
- Bateson P (2006) The adaptability driver. *Biological Theory* 1: 342–345.
- Beisecker D (1999) The importance of being erroneous: prospects for animal intentionality. *Philosophical Topics* 27: 281–308.
- Blair JE, Hedges SB (2005) Molecular clocks do not support the Cambrian explosion. *Molecular Biology and Evolution* 22: 387–390.
- Brembs B (2000) *An Analysis of Associative Learning in Drosophila at the Flight Simulator*. Doctoral dissertation, Bayerische Julius Maximilians Universität Würzburg, Germany.
- Cabanac M (1992) Pleasure: The common currency. *Journal of Theoretical Biology* 155: 173–200.
- Chao MY, Komatsu H, Fukuto HS, Dionne HM, Hart AC (2004) Feeding status and serotonin rapidly and reversibly modulate a *C. elegans* chemosensory circuit. *Proceedings of the National Academy of Sciences of the USA* 101: 15512–15517.
- Cody M (1974) *Competition and the Structure of Bird Communities*. Princeton, NJ: Princeton University Press.
- Conway Morris S (2000) The Cambrian "explosion": Slow-fuse or megatonnage? *Proceedings of the National Academy of Sciences of the USA* 97: 4426–4429.
- Conway Morris S (2003) The Cambrian "explosion" of metazoans and molecular biology: Would Darwin be satisfied? *International Journal of Developmental Biology* 47: 505–515.
- Corning WC, Dyal JA, Willows AOD, eds (1973) *Invertebrate Learning*, Volumes 1–3. New York: Plenum Press.
- Crick F (1994) *The Astonishing Hypothesis: The Scientific Search for the Soul*. New York: Scribner's.
- Damasio AR (1999) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace.
- Dennett DC (1995) *Darwin's Dangerous Idea: Evolution and the Meaning of Life*. London: Penguin Books.
- Denton D (2006) *The Primordial Emotions: The Dawning of Consciousness*. Oxford: Oxford University Press.
- Dretske F (1999) Machines, plants and animals: The origins of agency. *Erkenntnis* 51: 19–31.

- Dubnau J, Chiang AS, Tully T (2003) Neural substrates of memory: From synapse to system. *Journal of Neurobiology* 54: 238–253.
- Edelman GM (1987) *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Basic Books.
- Edelman GM (2003) Naturalizing consciousness: A theoretical framework. *Proceedings of the National Academy of Sciences of the USA* 100: 5520–5524.
- Edelman GM (2005) *Wider than the Sky: A Revolutionary View of Consciousness*. London: Penguin Books.
- Edelman GM, Tononi G (2000) *The Universe of Consciousness: How Matter Becomes Imagination*. New York: Basic Books.
- Garcia J, Hankins WG, Rusiniak KW (1974) Behavioral regulation of the milieu interne in man and rat. *Science* 185: 824–831.
- Ginsburg S, Jablonka E (2007) The transition to experiencing: I. Limited learning and limited experiencing. *Biological Theory* 2: 218–230.
- Grau JW, Joynes RL (2005) A neuro-functional approach to learning. *International Journal of Comparative Psychology* 18: 1–22.
- Hameroff S (1998) Did consciousness cause the Cambrian explosion? In: *Towards a Science of Consciousness II: The 1996 Tucson Discussions and Debates* (Hameroff S, Kaszniak A, eds), 421–437. Cambridge, MA: MIT Press.
- Hardy A (1965) *The Living Stream: A Restatement of Evolution Theory and its Relation to the Spirit of Man*. London: Collins.
- Humphrey N (1992) *A History of the Mind*. London: Vintage.
- Humphrey N (2000) How to solve the mind-body problem. *Journal of Consciousness Studies* 7: 5–20.
- Humphrey N (2006) *Seeing Red: A Study in Consciousness*. Cambridge, MA: Harvard University Press.
- James W (1950) *The Principles of Psychology*. New York: Dover. Orig. 1890.
- Kaiser M, Hilgetag CC (2006) Nonoptimal component placement, but short processing, due to long-distance projections in neural systems. *PLoS Computational Biology* 2(7): e95.
- Kim JH, Udo H, Li HL, Youn TY, Chen M, Kandel ER, Bailey CH (2003) Presynaptic activation of silent synapses and growth of new synapses contribute to intermediate and long-term facilitation in *Aplysia*. *Neuron* 40: 151–165.
- Koch C (2004) *The Quest for Consciousness: A Neurobiological Approach*. Englewood, CO: Roberts.
- Kravitz EA (1990) Hormonal control of behavior: Amines as gain-setting elements that bias behavioral output in lobsters. *American Zoologist* 30: 595–608.
- LeDoux J (1996) *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon & Schuster.
- Macphail ME (1987) The comparative psychology of intelligence. *Behavioral and Brain Sciences* 10: 645–695.
- Macphail ME, Bolhuis JJ (2001) The evolution of intelligence: Adaptive specializations versus general process. *Biological Reviews* 76: 341–364.
- Markram H, Lübke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275: 213–215.
- Martin SJ, Grimwood PD, Morris RG (2000) Synaptic plasticity and memory: An evaluation of the hypothesis. *Annual Review of Neuroscience* 23: 649–711.
- Merker B (2007). Consciousness without a cerebral cortex: A challenge for neuroscience and medicine. *Behavioral and Brain Sciences* 30: 63–81.
- Murphy GG, Glanzman DL (1997) Mediation of classical conditioning in *Aplysia californica* by LTP of sensorimotor synapses. *Science* 278: 467–471.
- Murphy GG, Glanzman DL (1999) Cellular analog of differential classical conditioning in *Aplysia*: Disruption by the NMDA receptor-antagonist DL-2-Amino-5-phosphonovalerate. *Journal of Neuroscience* 19: 1062–1095.
- Odling-Smee FJ, Laland KN, Feldman MW (2003) *Niche Construction: The Neglected Process in Evolution*. Princeton, NJ: Princeton University Press.
- Okamoto K, Takeuchi K, Agata K (2005). Neural projections in planarian brain revealed by fluorescent dye tracing. *Zoological Science* 22: 535–546.
- Panksepp J (1998) *Affective Neuroscience: The Foundations of Human and Animal Emotions*. New York: Oxford University Press.
- Panksepp J (2004) Affective consciousness: core emotional feelings in animals and in humans. *Consciousness and Cognition* 14: 30–80.
- Parker A (2003) *In the Blink of an Eye: How Vision Kick-Started the Big Bang of Evolution*. Sydney: The Free Press.
- Paulsen O, Sejnowski T (2000) Natural patterns of activity and long-term synaptic plasticity. *Current Opinion in Neurobiology* 10: 172–179.
- Rankin CH (2004) Invertebrate learning: What can't a worm learn? *Current Biology* 14: R617–R618.
- Razran G (1971) *Mind in Evolution: An East-West Synthesis of Learnt Behavior and Cognition*. Boston: Houghton Mifflin.
- Rescorla RA (1988) Behavioral studies of Pavlovian conditioning. *Annual Review of Neuroscience* 11: 329–352.
- Rokas A, Kruger D, Carroll SB (2005) Animal evolution and the molecular signature of radiations compressed in time. *Science* 310: 1933–1937.
- Rose S (1993) *The Making of Memory: From Molecules to Mind*. London: Bantam Books.
- Rushforth ND (1973) Behavioral modifications in coelenterates. In: *Invertebrate Learning, Vol. 1, Protozoans Through Annelids* (Corning WC, Dyal JA, Willows AOD, eds), 123–169. New York: Plenum Press.
- Sarnat HB, Netsky MG (1985) The brain of the planarian as the ancestor of the human brain. *Canadian Journal of Neurological Science* 12: 296–302.
- Stanley SM (1992) Can neurons explain the Cambrian explosion? *Geological Society of America Abstracts with Programs* 24: A45.
- Tinbergen N (1963) On aims and methods of ethology. *Zeitschrift für Tierpsychologie* 20: 410–433.
- Valentine JW (2002) Prelude to the Cambrian explosion. *Annual Review of Earth and Planetary Sciences* 30: 285–306.
- Waddington CH (1957) *The Strategy of the Genes*. London: Allen and Unwin.
- Wang DY, Kumar S, Hedges SB (1999) Divergence time estimates for the early history of animal phyla and the origin of plants animals and fungi. *Proceedings of the Royal Society of London B* 266: 163–171.
- Wells MJ (1968) Sensitization and the evolution of associative learning. In: *Neurobiology of Invertebrates* (Salánki J, ed), 391–411. New York: Plenum Press.
- West-Eberhard MJ (2003) *Developmental Plasticity and Evolution*. Oxford: Oxford University Press.
- Wong PS, Bernal E, Snodgrass M, Shervin H (2004). Event-related brain correlates of associative learning without awareness. *International Journal of Psychology* 53: 217–231.
- Wyles JS, Kunkel JG, Wilson AC (1983) Birds, behavior, and anatomical evolution. *Proceedings of the National Academy of Sciences of the USA* 80: 4394–4397.