What Behavioral Abilities Emerged at Key Milestones in Human Brain Evolution? 13 Hypotheses on the 600-Million-Year Phylogenetic History of Human Intelligence

Max S. Bennett*

Independent Researcher, New York, NY, United States

This paper presents 13 hypotheses regarding the specific behavioral abilities that emerged at key milestones during the 600-million-year phylogenetic history from early bilaterians to extant humans. The behavioral, intellectual, and cognitive faculties of humans are complex and varied: we have abilities as diverse as map-based navigation, theory of mind, counterfactual learning, episodic memory, and language. But these faculties, which emerge from the complex human brain, are likely to have evolved from simpler prototypes in the simpler brains of our ancestors. Understanding the order in which behavioral abilities evolved can shed light on how and why our brains evolved. To propose these hypotheses, I review the available data from comparative psychology and evolutionary neuroscience.

Keywords: evolutionary neuroscience, evolutionary psychology, evolution of theory of mind, evolution of episodic memory, primate intelligence

INTRODUCTION

Humans have an incredibly diverse suite of intellectual faculties. We can build cognitive maps, infer intentions of others, remember specific historical events, communicate with each other using language, learn motor skills through observation, and more. But all these varied faculties, which emerge from the complex human brain, are likely to have evolved from simpler prototypes in the inevitably simpler brains of our ancestors. This general idea of progressive complexification of behavior from simpler roots has been elegantly articulated in Paul Cisek's theory of "phylogenetic refinement," whereby an extant animal's behavioral repertoire is interpreted as a consequence of evolutionary refinement from more basic phylogenetic building blocks (Cisek, 2019). A challenge to interpreting human behavioral, intellectual, and cognitive faculties through the lens of phylogenetic refinement is in identifying the faculties present in our ancestors, as these were the building blocks upon which the process of phylogenetic refinement operated.

To aid this tracking of the phylogenetic refinement of behavior, this paper presents 13 hypotheses regarding the specific behavioral abilities that emerged at key milestones during the 600-million-year phylogenetic history from early bilaterians to extant humans. Given the breadth of this topic, the scope of this paper is narrowed in three ways. Firstly, it develops hypotheses only on phylogenetic history. Secondly, it develops hypotheses only on the *human lineage* from early

bilaterians to extant homo sapiens. Thirdly, it develops hypotheses regarding the subset of behaviors that are frequently considered as "intelligent." I will briefly review each of these three refinements to clarify the scope of the analysis herein.

Focus on Phylogenetic History

Tinbergen's Four Questions (Tinbergen, 1963), provide a useful tool for categorizing the levels at which behavior can be explained. His four questions are:

(1) Phylogeny: what were the evolutionary *steps* by which this behavior came to be? What function *did* the behavior serve in the environment in which the behavior originally emerged?
(2) Function: what was the *current* function that the behavior serves, as measured by reproductive and survival success? Note that the original function of a behavior is not necessarily the same as its current function.
(3) Mechanism: what are the underlying (neural, hormonal, biomechanical, etc.) mechanisms by which this behavior is implemented?
(4) Ontogeny: how does this behavior emerge in the development of an individual organism?

The scope of this paper is to investigate the phylogenetic history of behavioral abilities, and as such sets its focus on only one of Tinbergen's questions: phylogeny. Of course, these questions not entirely separable; mechanisms, functions, and ontogeny provide essential clues to the phylogenetic origins of behaviors. As such, mechanism, function, and ontogeny will be invoked as evidence in favor or against various speculations regarding the phylogenetics. However, the hypotheses themselves stake claims only on the phylogenetic history of a behavior, not its mechanism, present function, or ontogeny.

Focus on the Human Lineage

The scope of this paper is intentionally anthropocentric – it seeks to chronicle the phylogenetic history of behavioral abilities in the *human* lineage from early bilaterians and extant *Homo sapiens*. This requires an essential caveat to the hypotheses presented here. Proposing a hypothesis regarding the emergence of abilities along the evolutionary lineage from early bilaterians to humans is not the same thing as proposing a hypothesis regarding a unique ability of humans relative to other extant animals alive today. For example, the hypothesis that episodic memory emerged in early mammals is not the same as a hypothesis that *only mammals* exhibit episodic memory. Convergent evolution is not the exception, but the rule. Flying evolved independently multiple times (Ben-Hamo et al., 2016). Lens-based eyes evolved independently multiple times (Ogura, 2004). Alas, as we will see, the evidence is quite strong that episodic memory also evolved independently numerous times – amongst cephalopods, birds, as well as in mammals. Further, extant animals today independently evolved abilities that have never been present in the human lineage (such as the electroreception of certain fish and echolocation of bats). As such, the hypotheses in this paper should not be used to make comparisons between *Homo sapiens* and other extant species.
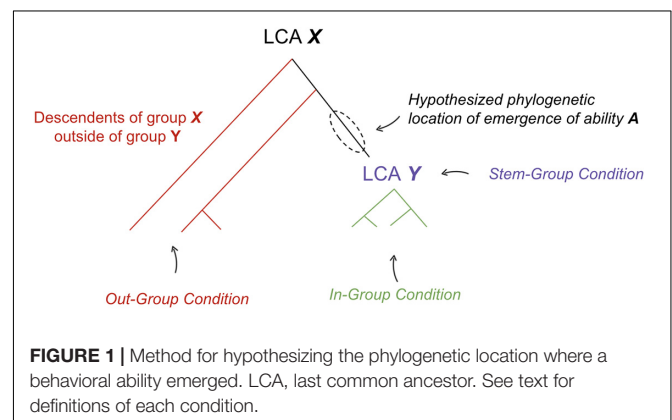
Focus on "Behavioral Abilities"

The scope of this paper will attempt to focus on the phylogenetic history of what I will call "behavioral abilities." I define a "behavioral ability" as an intellectual or cognitive faculty that animals are capable of invoking. I use this term for two reasons. Firstly, I use "behavioral ability" instead of "behavior" because the scope is not intended to review the entire behavioral repertoire of our ancestors. Further, the term "ability" is meant to focus the analysis to the realm of purported animal intelligence – attempting to understand the capacities afforded by the brains of our ancestors. Secondly, I use the term "behavioral ability" instead of "cognitive capacity" or "intellectual capacity" to avoid being restrained by pre-existing definitions of cognition or intelligence, and instead remain agnostic as to the differences in the many purported definitions of each. I use this novel term "behavioral ability" while acknowledging this refinement of scope will be imperfect. What gets categorized as a faculty deemed intelligent is ripe with an unavoidable terminological quagmire and anthropocentrism. Despite this challenge, the chronicling of such behaviors is still fruitful and illuminating as to the function and mechanisms of brains.

The Logic of Hypothesizing the Emergence of a Behavioral Ability

The hypotheses presented here all take the following form: "behavioral ability *A* emerged at some point amongst the early stem members of a phylogenetic group *Y*" **Figure 1**.

To support such a claim, evidence is presented in the form of three conditions:

(1) In-Group Condition: Diverse groups of early diverging species across group *Y* contain ability *A*, implemented in homologous neural mechanisms with shared ontogeny.
(2) Out-Group Condition: Evidence is supportive of one of the following three claims: (a) descendants of earlier diverging phylogenetic group *X* outside of group *Y* do not contain ability *A* or (b) ability *A* is implemented in non-homologous neural mechanisms with different

FIGURE 1 | Method for hypothesizing the phylogenetic location where a behavioral ability emerged. LCA, last common ancestor. See text for definitions of each condition.

ontogeny in earlier diverging group *X* outside of *Y*, relative to species within *Y* or (c) principle of parsimony suggests a homologous neural substrate was repurposed for new ability *A* within early members of group *Y*. The principle of parsimony states that if evidence is equally supportive of multiple phylogenetic relationships, we prefer the evolutionary tree that requires the fewest number of evolutionary changes (Fitch, 1971).

(3) Stem-Group Condition: Ability *A* would have been adaptive within the purported ecological niche of early members of group *Y* (i.e., the now extinct "stem-group" of *Y*, especially those likely to be the last common ancestor of group *Y*).

If the preponderance of evidence is supportive of the in-group, out-group, and stem-group conditions, then this is considered meaningful evidence in support of the hypothesis that behavioral ability *A* emerged at some point amongst the early members of a phylogenetic group *Y* and was thereby inherited by many of its descendants.

The relevance of criteria #2c is subtle but important. There are cases where shared neural structures are *independently* repurposed for similar functions. A simple example of this is that of wings. Both birds and bats repurposed the same structure (front legs) for a similar ability (flying). Although bird and bat wings share homology, the last common ancestor of birds and bats did not have wings. It is reasonable to see why this happens: given the constraints of evolution, different species that find themselves in ecological niches with similar selection pressures may end up similarly repurposing older structures for the same function. In such cases, the true phylogenetic origin of a behavioral ability can only be deciphered either through closer look at the actual substrates (identifying revealing differences in bird and bat wings) or through the principle of parsimony. Most non-mammal amniotes do not have wings, hence by principle of parsimony one would argue that it is more likely that wings were independently gained in the bird and bat lineage, as opposed to being independently lost across all other non-mammal amniotes (which would have required comparably more evolutionary changes). In the case of wings, fossil records are a useful adjudicator, but when it comes to brain structures, fossils are much less informative and hence we must rely on these other strategies to deduce the neural structures and the behaviors they enabled.

I am intentionally broad with respect to the specific evolutionary timing, focusing on the major divergences (e.g., comparing early members of Vertebrata with early members of Mammalia), and am less specific on more detailed timing (e.g., comparing early members of Mammalia with early members of Placentalia or Boreoeutheria). With the foundation of these broad hypotheses, however, further work can add additional detail to the steps of phylogenetic refinement by which these abilities emerged.

BEHAVIORAL ABILITIES THAT EMERGED IN EARLY BILATERIANS

See **Figure 2** for cladogram of bilaterian-cnidarian divergence.

Hypothesis #1: "Taxis Navigation" Emerged in Early Bilaterians

Taxis Navigation in Bilaterians (In-Group Condition)

The predominant navigational and hunting strategy across bilaterians is to navigate toward food and away from danger. Taxis navigation is the navigational strategy of simply turning *toward* or *away* from specific stimuli (e.g., chemotaxis, phototaxis, and thermotaxis). Even early diverging bilaterians,

FIGURE 2 | Cladogram of bilaterians and cnidarians.
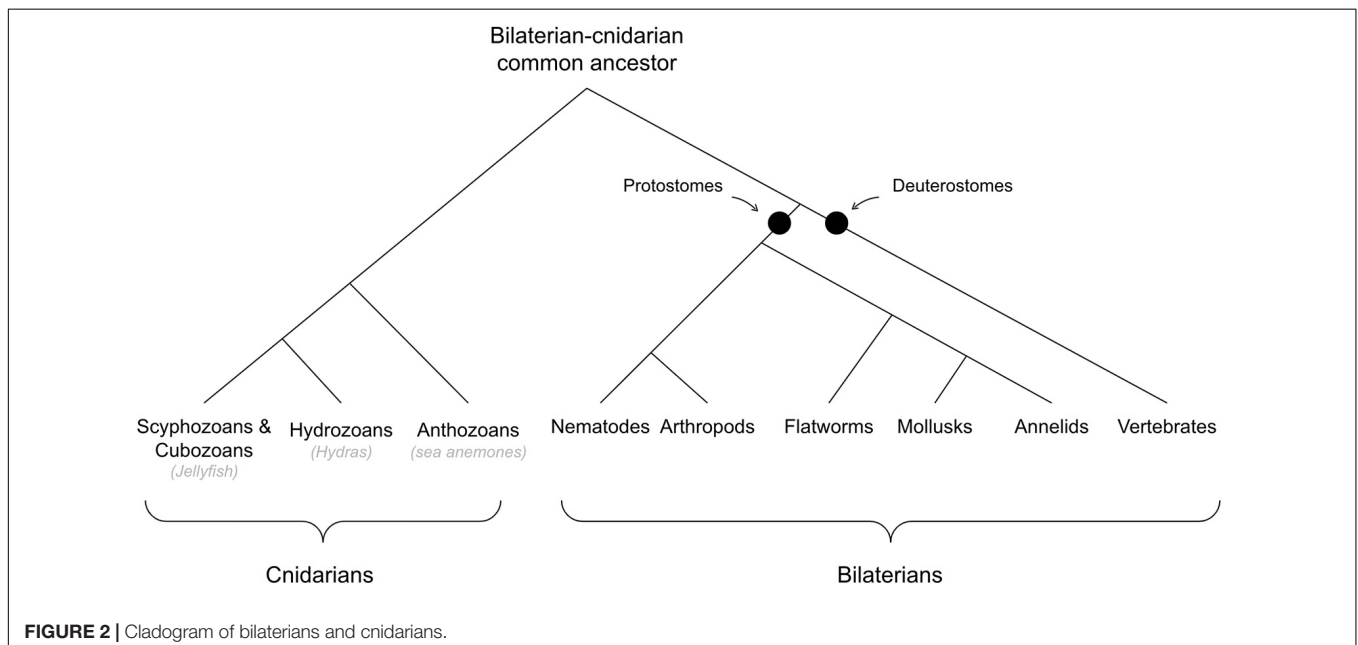
including those thought to be model organisms for urbilateria, such as *Caenorhabditis elegans* and flatworms, demonstrate taxis navigation. Both species climb sensory gradients to approach or avoid various stimuli (Pearl, 1903; Larsch et al., 2015).

There are two notable features of such taxis navigation in these model organisms for urbilateria: cross-modal integration and valence. Cross-modal integration describes the ability to integrate competing input across modalities to decide whether to approach or avoid a cue. For example, the decision by *C. elegans* to cross a copper barrier (which is aversive) to get to food on the other side is sensitive to the strength of the food smell relative to the concentration of copper (Hobert, 2002; Ishihara et al., 2002). Further, flatworms will navigate toward a food cue even if there is aversive light at the same source. But if made bright enough, flatworms will no longer travel all the way to the food source; and the brighter the light, the further from the food source they will end up. This type of integrated decision making was also shown across numerous modalities such as mechano-sensation and thermo-sensation (Inoue et al., 2015).

Valence refers to the ability to change decisions based on internal states, hence representing an ability to make decisions depending on what an organism needs. The concept of valence modulated by internal states is a foundational component of motivation and learning theories of behavior (Balleine, 2001, 2018; Berridge, 2019). For example, In *C. elegans,* food smells that trigger approach when hungry often have no effect when well fed (Davis et al., 2017). Further, some cues like carbon dioxide, which can signal both food as well as predators, shift from attractive when hungry to aversive when well fed (Rengarajan et al., 2019).

The neural substrates of taxis navigation, valence, and cross-model decision making are relatively well understood in *C. elegans*. The sensory neurons within *C. elegans* seem to encode valence directly: many of the sensory neurons in *C. elegans* are directly modulated by neuropeptides that signal internal states such as hunger or stress, and thereby provide a likely substrate by which behavior is sensitive to internal states (Chao et al., 2004; Chalasani et al., 2010; Sengupta, 2013; Guillermin et al., 2017; Rengarajan et al., 2019). In this sense, these sensory cells can be interpreted as primitive versions of valence neurons, with some responsive to positive valence stimuli (positive valence neurons), and others responsive negative valence stimuli (negative valence neurons). Cross-modal integration seems to occur through input to common neural circuits that control locomotion: there are command interneurons in the brain of *C. elegans*, which control turning by modulating locomotor central pattern generators (Garrity et al., 2010). One set of command interneurons biases movement toward *forward* locomotion whereas the other seems to bias movement toward *turning*. Different groups of sensory cells selectively target different sets of these command interneurons. The positive valence neurons in *C. elegans* that activate in response to the onset of food smells stimulate interneurons that trigger forward locomotion and suppress turning (Chalasani et al., 2007). On the other hand, negative valence neurons that activate in response to aversive stimuli such as copper, predator smells, or noxious heat, stimulate interneurons that trigger turning (Kimata et al., 2012). In general, these forward

and turning neurons exhibit lateral inhibition of each other (Wakabayashi et al., 2004; Li et al., 2014).

Taxis Navigation in Non-bilaterian Metazoans (Out-Group Condition)

Sea anemones are the extant cnidarians most likely to resemble early cnidarians: they are believed to have diverged earliest in cnidarian evolution, resembling early cnidarians in the fossil record (Yuan et al., 2011), and they resemble the developmental polyp stage that all cnidarians go through (Harris, 1990; Hinde, 1998). The hunting strategy of sea anemones is primarily one of waiting for food to come to them, and catching prey with their tentacles (Ruppert et al., 2004).

Medusae cnidarians (such as jelly fish) likely evolved later than polyp cnidarians (such as sea anemones) and admittedly do spend most of their time moving. However, even medusae do not show active hunting toward food. Rather, they have various reflexive movement routines that drive locomotion in general. For example, medusae swimming seems to be simply driven by orienting their heads in the direction of current to ensure staying in the same general location (Fossette et al., 2015). While chemical cues such as those for food in medusae change the overall speed of locomotion (Matanoski et al., 2001), and the sensitivity of the feeding reflex (Pantin, 1935; Batham and Pantin, 1950a; Ewer and Fox, 2009), it doesn't seem to broadly drive medusa navigation. Medusae do indeed swim up and down vertically in a Lévy walk-like manner, (Hays et al., 2011) but this functions to maximize the likelihood of encountering food, not to actively pursue it.

Most cnidarians do not even show immediate reflexive escape responses. Cnidarians often have a retraction reflex, but this doesn't drive locomotion in any specific direction (Batham and Pantin, 1950a). Some jellyfish do exhibit escape swimming, but most do not, which implies such behavior evolved independently and is not a general feature of early Cnidaria (Mackie, 2004). Instead of rapid escape responses, repeated aversive stimuli seem to toggle cnidarians between different broad phases of reflexive locomotion. Repeated aversive stimuli applied to the sea anemone will increase their likelihood to begin crawling and change location (Batham and Pantin, 1950a). But this is often triggered some time later after the stimulus, sometimes even within the next night cycle many hours later (Batham and Pantin, 1950a). Further, it is also unclear whether movement is reliably in the opposite direction of the aversive stimulus (Batham and Pantin, 1950b). In fact, often the direction of relocation seems to be driven by factors irrelevant to the aversive stimulus, such as simply moving in the direction of increased elevation (Batham and Pantin, 1950b). Sea anemones have also been shown to change location in this undirected manner if temperatures rise too high (Sund, 1958).

Even sexual reproduction in cnidarians is, in general, not one that requires navigating toward mates or even interacting directly with them. Instead, pheromone and light cues trigger a coordinated "spawning" whereby gametes are released into the sea, and hence mating between nearby cnidarians is made possible (Hinde, 1998).

Taken together, it seems to be the case that although stimuli can impact overall movement arousal in cnidarians, they do not, in general, exhibit directed taxis navigation. That said, there are a few exceptions where cnidarians do show directed navigation toward or away from stimuli. For example, hydra specifically orient their bodies and navigate directly toward light sources (Wilson, 1891; Haug, 1933). And box jellyfish can use eye spots on their head to avoid obstacles when swimming through a tank (Garm et al., 2007). However, this obstacle avoidance seems to use lens-like eyes that are generally agreed to have evolved independently, and is unlikely to represent navigational strategies of the early cnidarian-bilaterian common ancestor (Nilsson, 2013; Bosch et al., 2017). Sea anemones have also been found to move specifically in the direction of light (Parker, 1916), but this has been shown to be independent of their own visual apparatus and driven by cues of nearby amoebae – sea anemones without these symbiotic amoebae fail to navigate toward or away from light (Pearse, 1974; Foo et al., 2019).

The neural mechanisms for navigation are also different between Cnidaria and Bilateria. Locomotion in Cnidaria seems to operate via a completely independent reflexive circuit. For example, if you remove the oral disk of a sea anemone, and leave only the bottom of the animal, they still navigate as regularly and normally (Batham and Pantin, 1950b).

Larvae of earlier diverging metazoans, such as sponges, show taxis navigation (Wapstra and van Soest, 1987; Woollacott, 1993; Leys and Degnan, 2001; Leys et al., 2002), even though their adult forms show no such behavior. However, this taxis navigation uses completely different mechanisms from that of bilaterians: ciliated cells along the back of the larvae become rigid in response to various cues as with single cellular taxis navigation. The mechanism for how this response is coordinated across cells is not yet clear but may simply leverage independent photoreceptors within each of these ciliated cells directly.

Adult forms of ctenophores (comb jellies) move through ciliated pumping as well and may represent an adult form metazoan with taxis navigation coordinated through neurons. However, how well extant ctenophores represent early metazoans is unclear, with meaningful evidence indicating that ctenophores independently evolved many features of nervous systems (Ryan, 2014; Moroz, 2015; Moroz and Kohn, 2016; Liebeskind et al., 2017). If true, their basis of taxis navigation would not be indicative of early metazoans before the cnidarian-bilaterian last common ancestor.

Possible Adaptive Function(s) of Taxis Navigation in Early Bilaterians (Stem-Group Condition)

Fossil records have found evidence of small worm-like bilaterians navigating the microbial mats of the sea floor in the Ediacaran period (Chen et al., 2013). Such taxis navigation would likely have enabled these early bilaterians to navigate chemical gradients more efficiently and effectively on the Ediacaran seafloor in search of carcasses or microbial patches.

Taxis navigation is present within single-celled organisms and is clearly a fundamental feature of cellular navigation. But evidence suggests that taxis-navigation in early bilaterians was the

first time that taxis-navigation was implemented in the substrate of neurons and muscles, as opposed to cellular cilia. The cilia-based navigation of sponge or cnidarian larvae may have been limited in its ability to integrate competing inputs, limited in their ability to incorporate global need states, and may not have scaled well to larger organisms. In contrast, this neuron-based implementation of taxis-navigation in early bilaterians came with the crucial features of integration (balancing competing stimuli to make a single cross-modal decision) as well as valence (modulating those decisions by internal states), while potentially also enabling taxis navigation in larger morphologies.

Conclusion

Taken together, evidence across the in-group, out-group, and stem-group conditions is generally supportive of the hypothesis that taxis navigation (with neurons and muscles) emerged in early bilaterians, and was not present, at least with the same faculties, in the cnidaria-bilaterian last common ancestor. First, many, if not all, early diverging bilaterians show taxis navigation using neurons and muscles with the features of valence and cross-modal integration (in-group condition). Second, there is sparse evidence of taxis navigation in non-bilaterian eumetazoans, with many showing no such behavior (out-group condition). And third, the ecological strategy of early bilaterians was likely to actively pursue food, whereas for earlier diverging eumetazoans, possibly resembling sea anemones, it seems more likely to have been one of waiting for food to come to them – the latters strategy being one where taxis-navigation would have been much less adaptive (stem-group condition).

Hypothesis #2: Associative Learning Emerged in Early Bilaterians
Associative Conditioning in Bilaterians (In-Group Condition)

Associative learning, both classical and instrumental, has been shown across Bilateria, even those that diverged very early such as mollusks (Hawkins et al., 1989), flatworms (Prados et al., 2012), and nematodes (Ardiel and Rankin, 2010).

Consistent with the idea that associative learning across Bilateria have common evolutionary roots, associative learning across protostomes and deuterostomes shares a broad set of common features - including latent inhibition, overshadowing, blocking, second order conditioning, and trace conditioning. Latent inhibition has been shown in honeybees (Abramson and Bitterman, 1986; Chandra et al., 2000, 2010; Sandoz et al., 2000; Ferguson et al., 2001; Fernández et al., 2009; Fernandez et al., 2012), mollusks (Loy et al., 2006), fish (Mitchell et al., 2011), goats (Lubow and Moore, 1959), and rats (Ackil et al., 1969; Boughner and Papini, 2006). Overshadowing and blocking have been observed in flatworms (Prados et al., 2012), honeybees (Couvillon and Bitterman, 1989; Smith and Cobey, 1994; Couvillon et al., 1997; Smith, 1997), mollusks (Sahley et al., 1981; Colwill et al., 1988; Loy et al., 2006; Acebes et al., 2009), rats (Kamin, 1968; Prados et al., 2013), humans (Arcediano et al., 1997; Prados et al., 2013), rabbits (Merchant and Moore, 1973), and monkeys (Cook and Mineka, 1987). Second order conditioning has been observed in mollusks (Hawkins et al., 1998; Loy

et al., 2006), rats (Lay et al., 2018), and monkeys (Cook and Mineka, 1987). Further evidence for the idea that such learning features are foundational to old neural circuits, and not more advanced ones, these features of associative learning can be seen even in spinalized rats (rats that only contain a spinal cord) (Illich et al., 1994).

Further, the biological mechanisms for associative learning across bilaterians is incredibly similar. All bilaterians use very similar presynaptic and postsynaptic mechanisms for learning. For example, both invertebrates and vertebrates use cAMP as well as NMDA and AMPA receptors in learning processes (Kandel, 2001, 2006; Dubnau et al., 2002; Glanzman, 2010; Hawkins and Byrne, 2015). Neuromodulators such as dopamine and serotonin are involved in gating both presynaptic and postsynaptic learning processes across both vertebrates and invertebrates. The importance of such neuromodulators in learning has been shown in invertebrates such as crickets (Hammer and Menzel, 1998; Farooqui et al., 2003; Vergoz et al., 2007), fruit flies (Burke et al., 2012; Liu et al., 2012), honeybees (Hammer and Menzel, 1998; Farooqui et al., 2003; Vergoz et al., 2007), and C. elegans (Kusayama and Watanabe, 2000; Qin and Wheeler, 2007). Further, mechanisms for some forms of trace conditioning also seem to have common neurobiology whereby traces decay after synaptic discharge, and subsequent third factors such as dopamine signaling trigger synaptic weight changes (Cassenaer and Laurent, 2012; Dylla et al., 2013).

Associative Conditioning in Non-bilaterian Metazoans (Out-Group Condition)

Associative learning has for the most part not been observed in non-bilaterian metazoans. Attempts to show associative learning in cnidarians have shown negative results (Rushforth, 1973; Torley, 2009). I am aware of only one report of associative learning in Cnidaria, where a sea anemone was reported to learn to contract tentacles in response to a light that predicted shock (Haralson, 1975). Subsequent results have not replicated this, and others who have reviewed the available literature have similarly concluded that Cnidaria do not exhibit associative learning (Ginsburg and Jablonka, 2019).

Possible Adaptive Function(s) of Associative Conditioning in Early Bilaterians (Stem-Group Condition)

The adaptive benefits of associative conditioning are straightforward to imagine in the ecological niche of an early bilaterian animal of the Ediacaran. The ability to learn to associate certain cues, such as light, with food or predators, would have been useful in biasing taxis navigation toward safer and more food-rich areas.

Conclusion

Taken together, evidence across the in-group, out-group, and stem-group conditions is supportive of the hypothesis that associative learning emerged in early bilaterians. First, associative learning is observed even in very early diverging bilaterians (in-group condition). Second, there are numerous negative results of associative learning in non-bilaterians eumetazoans (out-group condition). And third, if the ecological niche of early bilaterians was to actively pursue food, while for earlier eumetazoans it was to wait for food to come to them, then associative learning would have been uniquely adaptive for early bilaterians, who were regularly making important navigational decisions (stem-group condition). The hypothesis that associative learning emerged in early bilatearans has also been proposed by others (Ginsburg and Jablonka, 2021).

BEHAVIORAL ABILITIES THAT EMERGED IN EARLY VERTEBRATES

See **Figure 3** for cladogram of vertebrate-invertebrate divergence.

Hypothesis #3: Map-Based Navigation Emerged in Early Vertebrates
Map-Based Navigation in Vertebrates (In-Group Condition)

Many diverse vertebrates, including those that diverged early such as fish (Burt de Perera et al., 2016), reptiles (Wilkinson and Huber, 2012; Broglio et al., 2015), turtles (López et al., 2001), amphibians (Phillips et al., 1995), and tortoises (Wilkinson et al., 2007) show incredibly sophisticated mapping abilities (Rodríguez et al., 2002a) – capable of learning un-cued locations and capable of flexibly generating new navigation routes. For example, fish can remember specific locations in 3-dimensional space (Karnik and Gerlai, 2012; Lucon-Xiccato and Bisazza, 2017; Wallach et al., 2018) and can generate a correct novel path to specific goal locations from many different starting places (Brown, 2015). Fish can learn a place preference for locations that avoid pain, which persists even when visual cues are switched and changed (as long as background place cues are kept constant), demonstrating a stable representation of a location (Valente et al., 2012). Fish can also use spatial maps for navigation even when they are at odds with taxon-based cues (Katie, 2015). Fish can latently learn a map of a maze where no locations have been rewarded, and one-shot generate the correct path to an observed goal location (Gómez-Laplaza and Gerlai, 2010). Reptiles can generalize learning about navigating around detours quite well too (Wilkinson and Huber, 2012).

As further evidence that spatial maps emerged in early vertebrates, spatial memory across vertebrates use similar circuitry, and that circuitry is unique to vertebrates. It is known that the hippocampus is the locus for spatial maps in mammals (O'Keefe and Dostrovsky, 1971; Ranck, 1973). Hippocampal lesions lead to failures in many tests of map-based navigation, while leaving cue learning intact (Broadbent et al., 2004; Clark et al., 2005). Further, recording studies have shown grid cells, head direction cells, and place cells within various regions of the hippocampal complex, demonstrating a very sophisticated model of allocentric space within the hippocampus (O'Keefe, 1976; Fyhn et al., 2004; Hafting et al., 2005; Sargolini et al., 2006).

Evidence suggests that the pallium of non-mammalian vertebrates contains a homologous region to the hippocampus, sharing similar circuitry, embryonic origins, and genetic markers

FIGURE 3 | Cladogram of vertebrate-invertebrate divergence.

(Rodríguez et al., 2002a,b; Gupta S. et al., 2012; Abellan et al., 2014; Tosches et al., 2018). In reptiles the hippocampal homolog is the medial pallium, while in teleost fish it is the dorsolateral telencephalon (Rodríguez et al., 2002a,b). Lesioning these areas in goldfish or turtles similarly prevents allocentric spatial learning while leaving cue learning intact (López et al., 2003; Durán et al., 2010; Broglio et al., 2015). Recording studies of these hippocampal homologs in non-mammalian vertebrates such as fish similarly show such head direction cells, edge detection cells, velocity cells, and place cells (Vinepinsky et al., 2018, 2020). And lastly, the circuitry of these hippocampal homologs in non-mammal vertebrates is similar to the circuitry of the hippocampus in mammals (Giassi et al., 2012; Fotowat et al., 2019).

Some evidence suggests that the basal ganglia, a structure that emerged in the first vertebrates (Grillner and Robertson, 2016), is also necessary for map-based navigation (Goodroe et al., 2018). Interactions between the hippocampus and basal ganglia seem to be necessary for the development of conditioning place preference or conditioned place aversion (Ito et al., 2008). Some have suggested that there may have precursors of the basal ganglia in the bilaterian common ancestor, given possible homology to arthropod central complex (Strausfeld and Hirth, 2013), but this is currently unclear.

Map-Based Navigation in Invertebrates (Out-Group Condition)

Note that I will use the label invertebrate as a shorthand for non-vertebrate bilaterians.

Taxon navigation refers to learning to take specific actions at specific cues. Map-based navigation, on the other hand, is when an animal learns a spatial map of an environment enabling them to calculate new routes that have never been tried before to get to a familiar location, and to learn about locations, even if individual cues are no longer present or have moved. Many reviews of navigation amongst even sophisticated invertebrates have concluded that their navigational strategy is exclusively one of taxon navigation, not of map-based navigation (Wehner et al., 1996; Walker, 1997; Benhamou et al., 1990).

Invertebrates fail at tests designed to evaluate their ability to generate maps. For example, if ants are put in an experimental condition where they must have an outbound route away from their nest to get to food, and a different inbound route back to the nest, they will readily learn how to follow this exact route. The question is – are they learning simply to turn at specific cues, or do they understand the map of their environment? A simple test is to place an ant trying to get back to its nest on the outbound path; if the ant understood a map of the environment, it would simply turn around and go right back (this shows flexible calculation of new routes to a target location) - but the ant doesn't do this, it becomes lost (Wehner et al., 2006). Further, if you change a single cue in an environment as an ant navigates (but leave other cues indicative of the map of the environment intact), it can completely disrupt their navigation (Cheng, 2012).

In another experiment, it was shown that a bee can learn to push a cap in the middle of a plate to find food underneath. If you then put a different shaped cap in the center with food underneath, and put the old cap on the far end of the plate, the bee will simply go to the original cap. It must completely re-learn to push the new cap: the bee never learned to associate a location with the reward (Abramson et al., 2016). In contrast, honeybees can learn to navigate through

a very complex mazes using taxon navigation, such as by learning to turn right or left at specific-colored cues (Zhang et al., 2000). Early diverging bilaterians such as *C. elegans* and planarians also do not navigate using spatial maps and instead use basic taxis mechanisms (Pearl, 1903; Luersen et al., 2014; Larsch et al., 2015).

There have been reports and suggests that a few select invertebrates have been shown to build map-like memories (Boles and Lohmann, 2003; Menzel et al., 2005, 2011). But this is still with considerable controversy, with others claiming that the ability to navigate from novel locations in these invertebrates is in fact just a sophisticated path integration mechanism and includes no map-like representations (Wehner et al., 1996; Lehrer, 1997a,b). This is further complicated by the fact that many impressive abilities of insects seem to emerge from mushroom bodies (Perry et al., 2013; Cope et al., 2018), a cortex-like structure which is believed to have evolved independently (Farris, 2008). As such, any behavioral abilities emerging from the mushroom bodies are more likely a case of independent evolution and not representative of the protostome-deuterostome last common ancestor.

Possible Adaptive Function(s) of Map-Based Navigation in Early Vertebrates (Stem-Group Condition)

Fossils of *Haikouichthys* spp., believed to be some of the earliest vertebrate species (Shu, 2003; Shu et al., 2003, 2009), have been dated to the early Cambrian (Zhang et al., 2001; Shu et al., 2003). The ecological niche of these early vertebrates is unknown, but relative to the huge fossils of presumably predatory arthropods from the Cambrian, we can speculate that there was strong pressure to avoid predation. Map-based navigation may have been a technique that these early vertebrates used to quickly swim to safety in response to predators as well as how to avoid dangerous locations, both useful adaptations to avoid predation. Some evidence suggests that early vertebrates may have evolved in shallow waters (Griffith, 1994; Sallan et al., 2018), which would have contained many more landmarks than in the open sea, perhaps making map-based navigation uniquely useful.

Conclusion

Taken together, evidence across the in-group, out-group, and stem-group conditions is generally supportive of the hypothesis that map-based navigation emerged in early vertebrates. First, map-based navigation is observed even in early diverging vertebrates and the neural substrates of map-based navigation in vertebrates seem to be structures that emerged only in early vertebrates (in-group condition). Second, there are numerous negative results of map-based navigation in invertebrates, especially those thought to be model organisms for the protostome-deuterostome last common ancestor such as *C. elegans* or planarians; and those invertebrates that do show some potential evidence of map-based navigation, such as arthropods, seem to do so in independently evolved substrates, such as the mushroom body (out-group condition). And third, map-based navigation would have been adaptive in the predatory shallow water environment of early vertebrates.

Hypothesis #4: "Interval Timing" Emerged in Early Vertebrates

Interval Timing in Vertebrates (In-Group Condition)

Even vertebrates that diverged very early, such as fish, show a remarkable ability to learn the timing of events. For example, goldfish and zebrafish can remember the exact time interval between a cue and a shock occurring and will selectively speed up to escape right before the shock occurs (Drew et al., 2005; Lee et al., 2010; Meck et al., 2012). Such interval timing is shown across vertebrate phyla, including in fish (Sumbre et al., 2008), birds (Bateson and Kacelnik, 1997; Ohyama et al., 1999; Buhusi et al., 2002), non-human primates (Gribova et al., 2002), and mice (Roberts and Church, 1978; Gallistel et al., 2004; Buhusi et al., 2005). Although some vertebrates do struggle with such tasks (Grossmann, 1973; Kleniginna and Currie, 1979; Laurent and Lejeune, 1985; Lejeune and Wearden, 1991).

Interval timing also seems to be implemented by uniquely vertebrate brain regions, namely, the striatum, hippocampus, and cerebellum. Many models of how interval timing is implemented in the brain place the striatum as the locus of interval timing (Matell and Meck, 2000, 2004; Meck and Benson, 2002). Additional models also incorporate the hippocampus into these timing mechanisms (MacDonald et al., 2011; Oprisan and Buhusi, 2013; Oprisan et al., 2018; Rolls and Mills, 2019; Shimbo et al., 2021). Lesions of the hippocampus in humans and rats disrupt learning on interval timing tasks (Meck et al., 1984, 1987; Melgire et al., 2005; Balci et al., 2009; Yin and Meck, 2014). Damage to the striatum in the basal ganglia leads to even more severe interval timing performance (Malapani et al., 1998, 2002). The results on lesions in both structures are specific to the timescale of hundreds of milliseconds to seconds. When it comes to millisecond-based timing, and perhaps also absolute timing in general, the cerebellum, seems to be essential (Ivry and Spencer, 2004).

Interval Timing in Invertebrates (Out-Group Condition)

As noted in other reviews of invertebrate learning and behavior, invertebrates consistently show an inferior perception of time to vertebrates, if any perception of time at all (reviewed in Abramson and Wells, 2018). Invertebrates struggle to learn and predict the specific times at which events will occur and differentiate between different timings (Abramson and Feinman, 1990; Balci, 2015). Honeybees show an inability to tell the difference between stimuli separated by 15, 30, 60, or 120 s (Craig et al., 2014). They also do not increase responses as the time of stimulus presentation approaches, suggesting they do not anticipate an event based on its timing (Abramson and Boyd, 2001; Craig et al., 2014). Similar results have been shown for crabs (Abramson et al., 1988; Abramson and Feinman, 1990; Balci, 2015). There are some studies that have shown evidence of interval timing in bees (Boisvert and Sherry, 2006), but the methodology of analysis has been questioned because it only demonstrated group-average differences and did not report on individual performance (Abramson and Wells, 2018).

Possible Adaptive Function(s) of Interval Timing in Early Vertebrates (Stem-Group Condition)

If early vertebrates experienced strong predation pressures from Cambrian arthropods and lived near shores with many landmarks of underwater rocks, plants, and even corals, then interval timing may have yielded certain adaptive benefits. For example, it would have been adaptive to learn the timing between faraway predator cues and their actual arrival. Further, as seen in the next section, it is possible that interval timing was a prerequisite for omission learning, which has many benefits for flexible and robust predictions, which would have been useful in avoiding predation.

Conclusion

Taken together, evidence across the in-group, out-group, and stem-group conditions is supportive of the hypothesis that the ability to learn the specific timing between events emerged with early vertebrates. First, interval timing has been observed in early diverging vertebrates and the neural substrates of interval timing seem to be structures that emerged in early vertebrates (in-group condition). Second, there are numerous negative results of interval timing in invertebrates (out-group condition). And third, it is straightforward to imagine the adaptive benefits of interval timing in the predatory environment in which early vertebrates evolved.

Hypothesis #5: Omission Learning Emerged in Early Vertebrates

Omission Learning in Vertebrates (In-Group Condition)

Another difference in the observed behavioral abilities between extant vertebrates and invertebrates is omission learning. The standard paradigm wherein omission learning is evaluated is as follows. First, a cue (e.g., a light) is paired with a shock, via classical conditioning. Second, trained animals are given the option to prevent the shock after the cue is presented, such as by moving to one side of the cage or by pushing a lever. These actions either terminate the cue and prevent the shock or prevent the shock without terminating the cue. Three levels of understanding can be tested in this paradigm. Level one is Pavlovian learning: animals can learn the predictive relationship between the cue and the shock. Level two is offset learning: animals can learn to repeat behaviors that terminate the cue that has been paired with pain. Level three is omission learning: in response to a cue that reliably predicts pain, animals can learn to repeat behaviors that lead to the omission of predicted pain, even if the cue is *not* terminated.

Many vertebrates, including dogs (Cole and Wahlsten, 1968), mice (Kamin, 1957; Avcu et al., 2014), and fish (Portavella, 2004; Vindas et al., 2012, 2014) demonstrate the ability to learn from omission. A key neural substrate of omission learning in vertebrates is dopamine reward prediction errors: tonic dopamine can pause during negative valence and burst during positive valence. A leading model of how this works is that striosomes in the vertebrate striatum learn to predict and time activations of dopamine and directly inhibit dopamine neurons
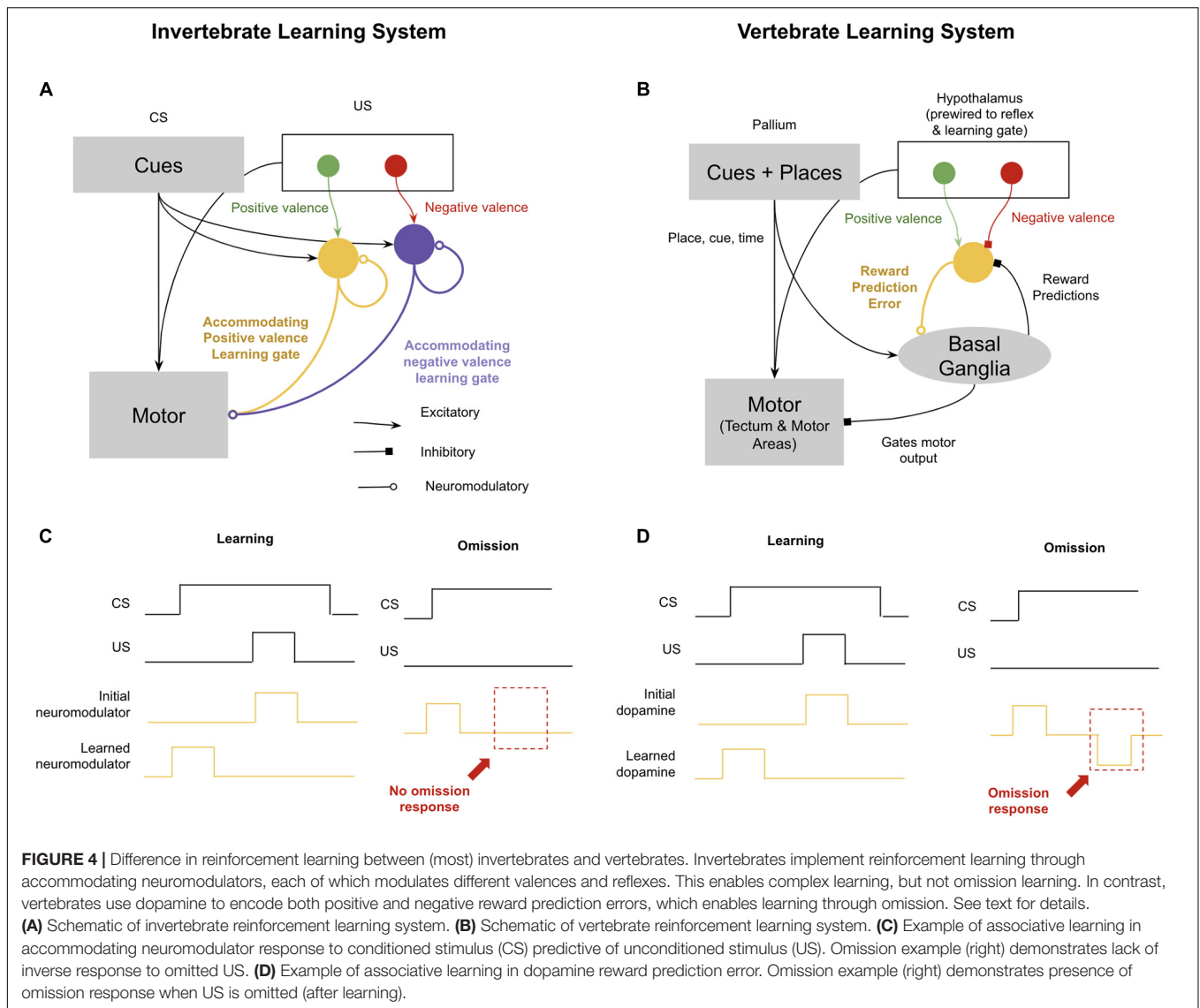
(Brown et al., 1999). In other words, striosomes filter out predicted dopamine activations, and hence trigger opposing dopamine activations when an unconditioned stimulus is omitted (e.g., positive valence when pain is omitted, or negative valence when food is omitted). Evidence suggests that this omission learning circuitry is shared by all vertebrates, including fish. In zebrafish for example, there are reward prediction errors during omission throughout its brain (Li, 2012; Cheng et al., 2014). And further, if you inactivate the habenula in zebrafish during avoidance learning, it strongly biases them from avoiding, and they become stuck only freezing, as you would expect if avoidance learning were driven by the rewarding aspects of omitted pain (Agetsuma et al., 2010).

Omission Learning in Invertebrates (Out-Group Condition)

Invertebrates seem, for the most part, to only learn from conditioned cue offsets, and seem incapable of learning from solely the omission of an unconditioned stimulus (reviewed in Abramson and Wells, 2018). In other words, invertebrates seem to only operate at "level two" in the above description. The key test to tell the difference between offset and omission learning is whether an animal's behavior is reinforced if it results in the omission of an aversive unconditioned stimulus, even when the conditioned cue (which normally predicts the unconditioned stimulus) is not terminated by said behavior. With omission learning this omission without termination will still be reinforcing, with only offset learning it will not. This general type of avoidance test has been performed across many invertebrates, including crabs, ants, and honeybees (Abramson et al., 1988; Abramson and Wells, 2018), all consistently demonstrating the inability to learn from omission. Consistent with this, flies have been shown to struggle to respond appropriately when expected cues are omitted (Wenner and Wells, 1990; Sanderson et al., 2013).

The observation that invertebrates fail at omission learning is somewhat perplexing because many observed features of learning in invertebrates, such as blocking, are best explained by the Rescorla–Wagner (RW) model, whereby learning is gated by surprise (Rescorla and Wagner, 1972). If invertebrates gate learning by surprise, then one would expect an explicit prediction error signal, which should enable invertebrates to learn from omission. However, a more detailed examination of the neural substrates of associative learning in invertebrates and vertebrates helps illuminate why this difference exists, and why although both vertebrates and invertebrates learn using an Rescorla–Wagner learning rule, invertebrates still can't learn from omissions.

The standard invertebrate circuits for conditioning are those described elegantly in *Aplysia* spp. by Kandel (1979, shown in **Figure 4**). During the pairing of a conditioned stimulus and an unconditioned stimulus, sensory neurons for the conditioned stimulus potentiate their synapses with neuromodulatory neurons. Next time the conditioned stimulus occurs, they can activate neuromodulatory neurons directly. It is then believed that these neuromodulatory neurons are accommodating, meaning that if they first burst sufficiently in

FIGURE 4 | Difference in reinforcement learning between (most) invertebrates and vertebrates. Invertebrates implement reinforcement learning through accommodating neuromodulators, each of which modulates different valences and reflexes. This enables complex learning, but not omission learning. In contrast, vertebrates use dopamine to encode both positive and negative reward prediction errors, which enables learning through omission. See text for details. **(A)** Schematic of invertebrate reinforcement learning system. **(B)** Schematic of vertebrate reinforcement learning system. **(C)** Example of associative learning in accommodating neuromodulator response to conditioned stimulus (CS) predictive of unconditioned stimulus (US). Omission example (right) demonstrates lack of inverse response to omitted US. **(D)** Example of associative learning in dopamine reward prediction error. Omission example (right) demonstrates presence of omission response when US is omitted (after learning).

response to the conditioned stimulus, they will be unresponsive to an unconditioned stimulus occurring shortly thereafter. This means that during an animal's experience with a conditioned stimulus and then an unconditioned stimulus, the neuromodulatory burst will shift forward to the conditioned stimulus and away from the unconditioned stimulus. The consequence of this is a simple version of the Rescorla–Wagner rule that will generate the effects of blocking, second-order conditioning, and overshadowing (Hawkins and Kandel, 1984; Hawkins et al., 1998). This circuit can even enable offset learning, if you assume that negative valence conditioned stimulus' inhibit positive valence neuromodulators, and their offset drives rebound excitation. But what this primitive circuit will fail to learn is when something is predicted but never occurs. There is no notion of timing in this circuit, and hence if a conditioned stimulus is active, predicting a shock 5 s later, and the conditioned stimulus remains but the shock never occurs, nothing in this circuit will encode a

prediction error in that omission. The relationship between omission learning and the emergence of time perception is likely not a coincidence. Without the ability to accurately encode interval timing, learning from omission would be very difficult: to assess when something has been omitted, an animal must not only predict *what* will happen, but also *when* it will happen.

Possible Adaptive Function(s) of Omission Learning in Early Vertebrates (Stem-Group Condition)

As shown in temporal difference learning, the ability to learn from predicted events that do not occur leads to much more robust and flexible learning (Sutton, 1988). Hence, it is reasonable to speculate that omission learning enabled early vertebrates to learn more effectively to predict food locations and predator behaviors.

Conclusion

Taken together, evidence across the in-group, out-group, and stem-group conditions is supportive of the hypothesis that the ability to learn from omissions emerged in early vertebrates and did not exist in brains beforehand. First, omission learning has been demonstrated across vertebrates and seems to be implemented in neural structures that are emerged in early vertebrates (in-group condition). Second, multiple negative results have been demonstrated in invertebrates (out-group condition). And third, omission learning would have offered many adaptive benefits to early vertebrates, as this ability is a key feature of temporal difference reinforcement learning.

BEHAVIORAL ABILITIES THAT EMERGED IN EARLY MAMMALS

See **Figure 5** for cladogram of mammal divergence.

Hypothesis #6: Vicarious Trial and Error Emerged in Early Mammals

Vicarious Trial and Error in Mammals (In-Group Condition)

It has long been observed that mice navigating a maze will occasionally pause at choice points and toggle their head back and forth – this behavior has been called vicarious trial and error (VTE) and it has been interpreted as the mice considering their options (Muenzinger and Gentry, 1931; Tolman, 1939, 1948; Tolman and Minium, 1942). VTE shows several interesting features. First, animals perform VTE selectively when decisions are hard, such as when in the early stages of learning about a maze (Tolman and Minium, 1942), when the difference between the outcomes is small (Tolman, 1939; van der Meer et al., 2010), when starting in different places (Gardner et al., 2013; Schmidt et al., 2013), when reward contingencies change (Johnson and Redish,

2007; Steiner and Redish, 2012; Regier et al., 2015), or when there is conflict between evidence (Schmidt et al., 2013). When decisions are not difficult, such as when an animal can learn simple procedural strategies, such as always turn left at this cue, then VTE goes away (Gardner et al., 2013). This VTE behavior has also been shown across the mammalian taxa, including in non-human primates and humans (reviewed in Redish, 2016). Such head turning behavior in humans has been shown to be predictive of superior performance (van der Meer et al., 2010).

Three structures are highly implicated in VTE behavior: the neocortex, the hippocampus, and the ventral striatum (VS). When rats perform VTE, the hippocampus replays the same path sequences of place cells that represent the routes to each goal, demonstrating that the animal is first simulating the path toward one goal, and then simulating the path toward the other, without actually moving (Johnson and Redish, 2007; Gupta A. S. et al., 2012; Wikenheiser and Redish, 2015). The preplay events are identified as sharp wave ripples in recording studies. These sharp wave ripples are not found in non-mammal vertebrates (Van Twyver and Allison, 1972; Hough and Bingman, 2004; Kahn et al., 2008; Rattenborg et al., 2011; Ben-Yishay et al., 2020; Vinepinsky et al., 2020) with the exception of birds, where they have been found in the avian hippocampus (a functional homolog of mammalian hippocampus) during sleep (Payne et al., 2020).

It has been hypothesized that this hippocampal preplay is initiated by the prefrontal cortex (Redish, 2016). Evidence for this can be seen in the fact that hippocampal disruption increases VTE behavior (Robbe et al., 2006; Bett et al., 2015), as if the animal is trying to simulate actions but struggling to successfully do so without an intact hippocampus. Consistent with this, hippocampal damage makes rats more impulsive in delayed gratification tasks (Cheung and Cardinal, 2005), as if unable to simulate the benefit of waiting. The prefrontal cortex is known to highly influence such goal-based decisions (Killcross and Coutureau, 2003; Sharpe and Killcross, 2015) and be required for goal related activity in the hippocampus (Ito et al., 2015;

FIGURE 5 | Cladogram of mammal divergence.
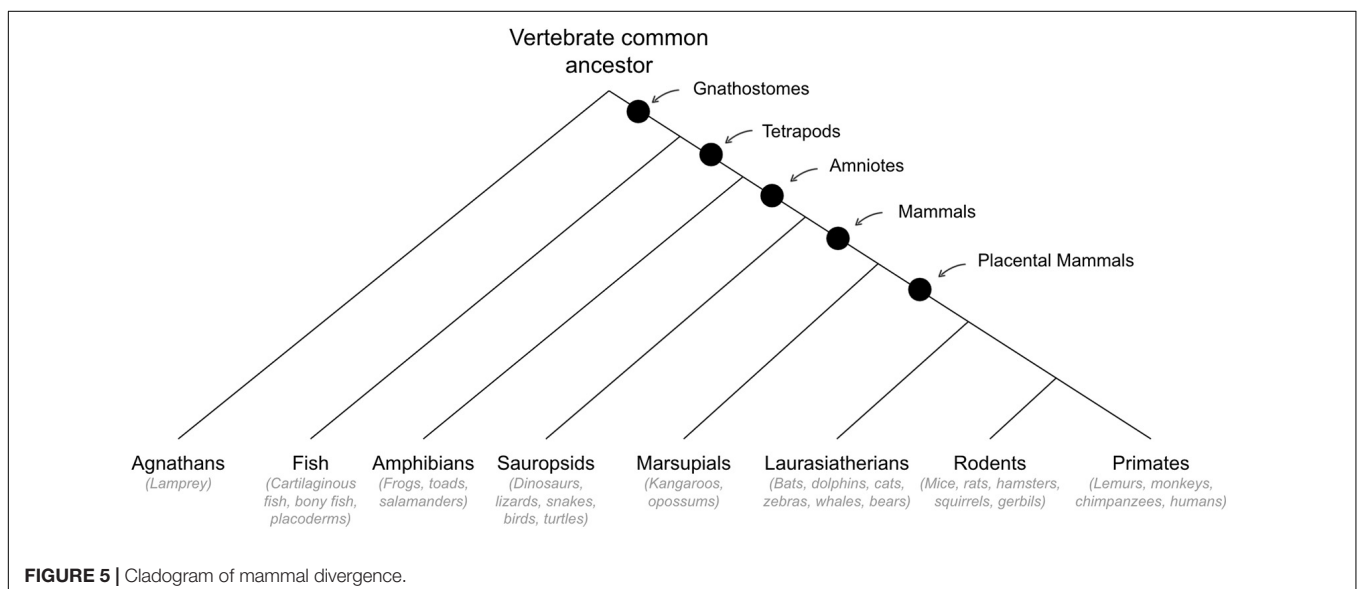
Spellman et al., 2015). Prefrontal cortex disruption impairs an animal's ability to make hard decisions at choice points, implying disruption to VTE (Ragozzino et al., 1999). Specifically at choice points in mazes, rats exhibit entrainment between oscillatory activity in the hippocampus and prefrontal cortex (Benchenane et al., 2010; O'Neill et al., 2013; Spellman et al., 2015). Also supportive of the idea of prefrontal cortex involvement in VTE, mice with PFC lesions are still able to solve the Morris water maze task and navigate spatial maps normally (Poucet, 1989, 1990; Granon and Poucet, 1995), but show impaired behavior when the task is made exceptionally hard by making rats start from completely novel locations (Granon and Poucet, 1995). Lastly, prefrontal cortex is known to be engaged during planning in general (Redish, 2016).

Evidence suggests that the ventral striatum (VS) is the neural substrate whereby the outcome of the simulated options are evaluated. During VTE behavior, cells in the VS encode reward values of the goal outcome of each option (van der Meer and Redish, 2009; Steiner and Redish, 2014; Stott and Redish, 2014). Importantly, these are the same cells that become active during the receipt of actual rewards (Tremblay and Schultz, 1999; Nicola et al., 2004; Roitman et al., 2005; Padoa-Schioppa and Assad, 2006; Stott and Redish, 2014).

Vicarious Trial and Error in Non-mammal Vertebrates (Out-Group Condition)

To my knowledge, there have been no published evaluations of specifically VTE outside of mammals. Some suggestive evidence has been seen in birds, who have demonstrated the ability to plan their foraging paths ahead of time (Sulikowski and Burke, 2015), which is suggestive that birds can also engage in a form of VTE. However, for reasons that will be discussed more thoroughly in hypothesis #8, birds are poor model organisms for the brains of the amniote common ancestor; bird brains independently underwent substantial modification since the amniote common ancestor and as such are unlikely to be representative of the brains of early amniotes.

Possible Adaptive Function(s) of Vicarious Trial and Error in Early Mammals (Stem-Group Condition)

The early mammals of the Mesozoic period are believed to have been small nocturnal (Polyak, 1957; Jerison, 1973; Gerkema et al., 2013; Wu et al., 2017), arboreal (Fröbisch and Reisz, 2009; Luo et al., 2015; Meng et al., 2015) insectivores (Kielan-Jaworowska et al., 1979; Lofgren et al., 2004). They would have likely been under extreme predation pressure from the massive archosaurs of the Mesozoic. Their ecological niche was likely hiding in trees and burrows, only to emerge for the purpose of quickly hunting food. VTE may have been useful in deciding which path to take across branches to either get to nearby insects or to avoid nearby predators. Navigating tree branches with far eyesight presents unique challenges and evolutionary pressures that may not have been previously experienced to the same degree: namely, irreversible choices. As a small animal living in trees, you must plan your route well in advance. And it is likely you will very regularly experience novel branches. Additionally, computational models have found that the usefulness of planning

is directly tied to visual range. Visual range in water is so poor that computational models suggest planning in water is barely useful at all (Mugan and MacIver, 2020), whereas on land, such planning is highly adaptive.

Conclusion

Taken together, evidence across the in-group, out-group, and stem-group conditions is generally supportive of the hypothesis that VTE emerged in early mammals. First, VTE has been well observed even in mammals that diverged well before the first primates and some of the key neural substrates of VTE emerged in early mammals ("in-group condition"). Second, VTE behavior has not been reported in non-mammal vertebrates, outside of possible evidence in birds ("out-group condition"). And third, VTE has been shown to improve decision making, which would likely have been adaptive to the small arboreal and nocturnal mammals in the Mesozoic. However, the lack of published negative results in non-mammal vertebrates renders this hypothesis tentative. More studies will have to be done to further support or refute this proposal.

Hypothesis #7: Counterfactual Learning Emerged in Early Mammals

Counterfactual Learning in Mammals (In-Group Condition)

A hallmark of human intelligence is the ability to consider things that might have happened had we made a different past decision. Such counterfactual learning has been observed in rats, monkeys, and humans (Zhang et al., 2015). A test of counterfactual learning has called the restaurant row test. In such a test, rats can wait for a higher quality meal or take a smaller, lesser meal more quickly. Once a choice was made, they cannot reverse it. Two key results were found. First, rats that choose the worse deal tend to look back and glance at the room they skipped. Second, after experiencing a regret-inducing situation (where they took the smaller reward, but then saw had they waited, they would have gotten the better one), rats tend to wait longer at the high-cost option during the next trial (Lewis, 2014; Steiner and Redish, 2014). In a similar study it was shown that rats actively choose behavior to avoid these experiences of regret (Sweis et al., 2018). Counterfactual learning has also been shown in non-human primates, where if they are made to play rock paper scissors, selectively after losing, they show a strong bias toward playing the move that would have won the last round (Abe and Lee, 2011).

In the rat studies above, it was shown that neurons within the orbitofrontal cortex and VS represented the counterfactual action when rats looked back at their rejected option (Gilovich and Medvec, 1995). They further showed that the greater this representation of the counterfactual action in the orbitofrontal cortex and VS, the higher the likelihood rats would stay and wait for the counterfactual choice next time they were presented with the task. Note that, consistent with the idea that imagining an event or stimulus reactivates the same circuitry as experiencing the event or stimulus firsthand, across mammals it has been shown that neurons in the orbitofrontal cortex also encode the

reward values (Padoa-Schioppa and Assad, 2006; Sul et al., 2010; Abe and Lee, 2011). This implies that counterfactual learning includes re-simulating the reward of the alternative option, much the same way VTE simulates different paths. Humans also show orbitofrontal activity during regret (Coricelli et al., 2005, 2007). In fact, humans with damage to the orbitofrontal cortex seem to be unable to experience regret (Camille et al., 2004).

Counterfactual Learning in Non-mammal Vertebrates (Out-Group Condition)
To my knowledge, there have been no published evaluations of counterfactual learning in non-mammals.

Possible Adaptive Function(s) of Counterfactual Learning in Early Mammals (Stem-Group Condition)
Counterfactual learning has numerous benefits as demonstrated in various machine learning models that leverage hindsight experience replay (Andrychowicz et al., 2017). In a situation with three or more possible actions, when a mistake is made and the correct action is obvious after the mistake, an animal without counterfactual learning will only learn to inhibit the chosen action, but not to reinforce the observed correct one.

Conclusion
Taken together, evidence across the in-group, out-group, and stem-group conditions is generally supportive of the hypothesis that counterfactual learning emerged in early mammals. First, counterfactual learning has been observed in relatively early diverging mammals, key neural substrates of counterfactual learning are structures that emerged in early mammals (in-group condition). Second, counterfactual learning has not been reported in non-mammal vertebrates (out-group condition). And third, counterfactual learning offers many learning benefits, which would have been adaptive to early mammals (stem-group condition). However, the lack of published negative results in non-mammal vertebrates makes this hypothesis tentative.

Hypothesis #8: Episodic Memory Emerged in Early Mammals
Episodic Memory in Mammals (In-Group Condition)
A key test of episodic memory is whether an animal can answer an unexpected question about their own experience. The question must be unexpected because it requires the animal to inquire their own mind for the answer – if it is expected the behavior can simply be instrumentally associated with the past action via trace conditioning.

The ability to answer unexpected questions has been shown in mammals such as dogs (Fugazza et al., 2020), rats (Crystal, 2013), and non-human primates (Menzel, 1999). As an example, rats trained to get a reward differently depending on whether they recently experienced food or not, can be randomly asked this question throughout normal foraging and exploration. Rats can successfully report on whether they had just recently experienced food whenever they are unexpectedly asked this question (Crystal, 2013).

The neural mechanisms of episodic memory seem to be homologous across mammals. For example, in mammals,

episodic memory uses the same neural circuitry as simulating the future (Suddendorf and Corballis, 1997; Schacter et al., 2007; Martin et al., 2011; Allen and Fortin, 2013). The general view is that the frontal cortex asks a question to the hippocampus, inquiring about some past event and the hippocampus contains a pointer to the contents of that event, which then reactivates the whole episodic memory within the neocortex (just like with VTE) (McClelland and Goddard, 1996; Eichenbaum et al., 2007). The neocortical representation of the retrieved memory is then transmitted to the frontal cortex and striatum for evaluation and action selection (Goldman-Rakic, 1996; Fuster, 2001; Ninokura et al., 2003; Eichenbaum and Fortin, 2009).

Various studies are consistent with this view of episodic memory. For example, in the one study whereby rats answered unexpected questions about their past experiences, the experimenters temporarily inhibited the hippocampus specifically during the moment rats were asked questions. What they found was that when rats answered *expected* questions (hence not requiring internal inquiry), hippocampal inactivation had no effect on performance. However, when the questions were unexpected, rats completely lost the ability to successfully inquire on their own episodic memory. This is suggestive of the idea that the hippocampus is specifically crucial for the reactivation of episodic memories. The hippocampus has been proposed to play the same role in human episodic memory (Wixted et al., 2018). Further consistent with this model of episodic memory, studies have shown that remembering past events reactivates the exact same cortical representations of the experience of it itself (O'Craven and Kanwisher, 2000; Pearson et al., 2015).

Episodic Memory in Non-mammal Vertebrates (Out-Group Condition)
Episodic memory of this form, whereby animals answer unexpected questions about their past, has also been shown in pigeons (Zentall et al., 2001, 2008; Singer and Zentall, 2007) and cephalopods (Billard, 2020). However, the neural mechanisms underlying episodic memory in these species seem to be non-homologous with the neural mechanisms in mammals. Episodic memory in mammals is highly dependent on the neocortex, a structure with which at least cephalopods have no homologous region.

Admittedly, some evidence suggests that the dorsal ventricular ridge of birds (which contain the nidopallium and mesopallium) and the neocortex of mammals both derive from the pallium of their shared amniote ancestor (Karten, 1969, 1997; Reiner et al., 2004; Dugas-Ford et al., 2012). These studies demonstrate that the dorsal ventricular ridge and neocortex share many features, including the subcortical structures they interact with, and the molecular properties of their neurons. Hence if episodic memories in mammals is dependent on the neocortex, perhaps this function was derived from the pallium of the amniote or even vertebrate last common ancestor and is not dependent on the unique features of neocortex. However, this interpretation is unconvincing for two reasons.

First, birds are a poor model organism for the brain of the amniote last common ancestor. Some have suggested that the dorsal ventricular ridge is not homologous with the neocortex

and instead shares homology with the mammalian amygdaloid complex (Jarvis et al., 2005; Striedter, 2005). And even if the dorsal ventricular ridge does share homology with the neocortex, the dorsal ventricular ridge has completely distinct microcircuitry from the neocortex. The neocortex is organized into six layers, while the dorsal ventricular ridge is organized into clustered nuclei (Ulinski, 1983). The ontogeny of the dorsal ventricular ridge and the neocortex in mammals is also different (Jones and Levi-Montalcini, 1958; Striedter and Keefer, 2000; Dugas-Ford et al., 2012). Further, the pallial homolog of other extant amniotes such as non-bird reptiles, also have completely unique ontogeny and microcircuitry (Goffinet et al., 1986; Cheung et al., 2007). For example, turtles have a three layered cortex, instead of the clustered nuclei of the dorsal ventricular ridge, or the six layered neocortex. The turtle cortex is more like the three layered pallium of other non-amniote vertebrates, such as fish, than it is to the dorsal ventricular ridge of birds or the neocortex of mammals. This is thereby suggestive that many of the pallial homologs in birds and mammals have undergone substantial independent modification in the bird and mammal lineage since the amniote last common ancestor.

Second, episodic memory, of the type where animals answer unexpected questions, has not been reported in amniotes outside of birds and mammals. If episodic memory was in fact present in the amniote common ancestor, and relied on such older amniote structures, we would expect to see reports not only in birds and mammals, but also in other amniotes as well. However, given the fact that episodic memory has only been reported in specific amniotes including birds and mammals, the evidence is more consistent with the idea that the DVR and neocortex each independently implemented a mechanism of episodic memory.

It should also be noted that there are additional tests of episodic memory. For example, one such set of tests of episodic memory has been referred to as tests of "what-where-when" memory, where an animal must remember *what* happened, *where* it happened, and *when* it happened. This type of memory has been challenged as being a true assessment of episodic memory, and it is generally accepted that the "unexpected question" is a better test. "What-where-when" memory has been consistently shown across many phyla, including fish (Hamilton et al., 2016), rats (Bird et al., 2003; Babb and Crystal, 2005; Ergorul and Eichenbaum, 2007), pigeons (Skov-Rackette et al., 2006), apes (Schwartz et al., 2002, 2004, 2005; Mulcahy and Call, 2006), and birds (Zinkivskay et al., 2009) and seems likely to be an ability that emerged far before the first mammals.

Possible Adaptive Function(s) of Episodic Memory in Early Mammals (Stem-Group Condition)

It has been proposed that the adaptive benefit of episodic memory is to enable planning about the future (Eichenbaum and Fortin, 2009; Allen and Fortin, 2013; McGaugh, 2013). The ability to remember specific temporal, spatial, and semantic information about events that occurred in a previous situation is useful in anticipating what will happen in a future situation. If planning was a important feature of early mammals, as suggested above, then episodic memory may have been an additional feature that improved planning.

It has also been suggested that episodic memory is useful for remembering social information and hence enables stable social bonds and networks (Emery, 2004; Brennan and Kendrick, 2006; Davidson et al., 2012). The earliest fossil evidence of social behavior is that of early mammals in the Mesozoic (Weaver et al., 2020), suggestive that these early mammals may have been uniquely social compared to earlier amniotes.

Conclusion

Taken together, evidence across the in-group, out-group, and stem-group conditions is generally supportive of the hypothesis that episodic memory (of the type where animals answer unexpected questions) emerged within early mammals and evolved independently along the bird line. First, episodic memory has been reported across diverse species of mammals and key neural substrates of episodic memory seem to be structures that emerged within early mammals (in-group condition). Second, the only non-mammal species that have been reported to have such episodic memory are specifically those species known to have independently evolved many unique brain structures and intelligent abilities, such as birds and cephalopods (out-group condition). And third, episodic memory would have offered many adaptive benefits to early mammals, especially if planning head was part of their survival strategy. This hypothesis is consistent with proposals of others (Allen and Fortin, 2013). However, it should be noted that the lack of published evaluations of non-bird reptiles, as well as the lack of studies on the neural substrates of episodic memory in birds, makes this hypothesis tentative.

## BEHAVIORAL ABILITIES THAT EMERGED IN EARLY PRIMATES

See **Figure 6** for cladogram of primate divergence.

Hypothesis #9: The Ability to "Anticipate a Need in the Future" Emerged in Early Primates
Anticipating Future Needs in Primates (In-Group Condition)

The Bischof–Kohler hypothesis states that humans have a unique ability to make plans to alleviate a need that they will have in the future, but do not currently feel (example: buying food for the week at the grocery store even when not hungry), while other animals are only capable of making plans only to alleviate a need they currently feel (example: generating an optimal path through a maze to get to food when currently hungry as shown in VTE studies) (Bischof-Köhler, 1985). However, the view that only humans can do this has been challenged. Evidence now suggests that many different primates (including chimpanzees, squirrel monkeys, bonobos, and orangutans) are in fact capable of this anticipation of future needs (McKenzie et al., 2004; Mulcahy and Call, 2006; Naqshbandi and Roberts, 2006; Janmaat et al., 2014).

I am not aware of any studies that have examined neural activity during tests of the Bischof-Kohler hypothesis. However,

FIGURE 6 | Cladogram of primate divergence.
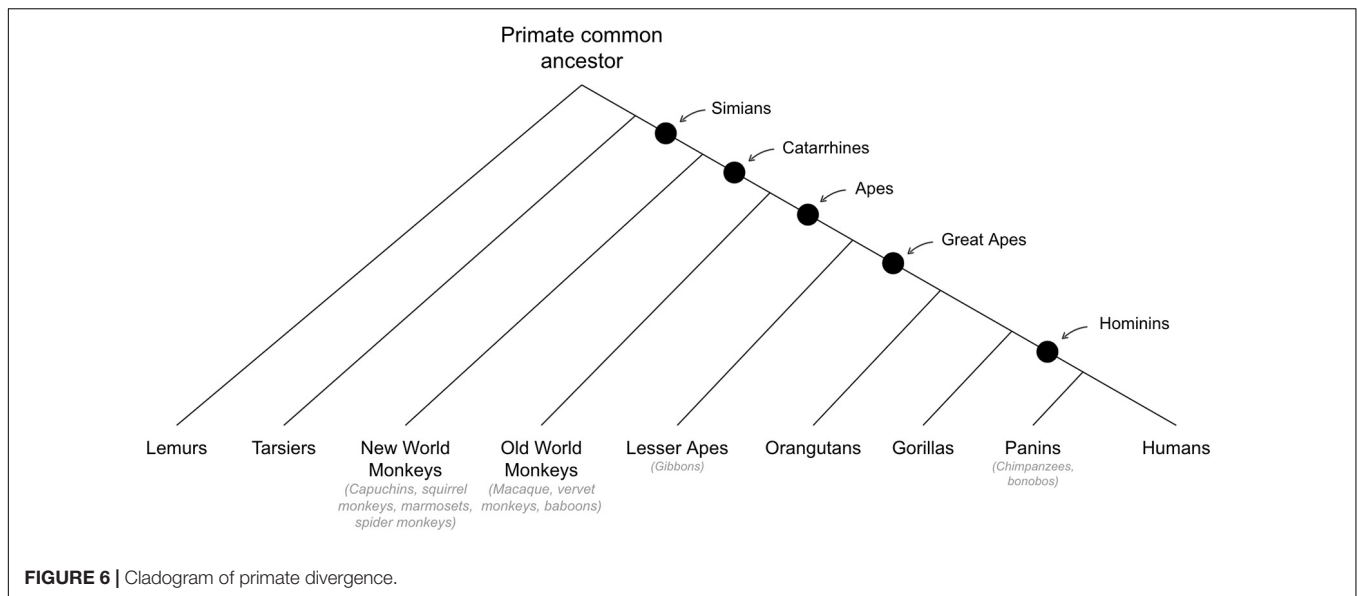
evidence is consistent with the idea that the dorsolateral prefrontal cortex is a substrate of this ability. And crucially, the dorsolateral prefrontal cortex is a structure that emerged in early primates (Semendeferi et al., 2001; Mansouri et al., 2017). The dorsolateral prefrontal cortex of non-human primates and humans seems to activate selectively during situations where you need to take an action to support a need that you will have in the future even if it doesn't support something you currently want. For example, the dorsolateral prefrontal cortex activates when considering future rewards, but not when considering present rewards (McClure, 2004; Tanaka et al., 2004; Kim et al., 2008). Dorsolateral prefrontal cortex activity is selectively activated when individuals choose delayed rewards over immediate rewards (McClure et al., 2004, 2007; Weber and Huettel, 2008) as well as when successfully avoiding temptation in self-control studies of dieters (Hare et al., 2009). Those with selective inactivation of dorsolateral prefrontal cortex are impaired in their ability to give up immediate rewards for futures ones (Figner et al., 2010). Selective inactivation of dorsolateral prefrontal cortex impairs people's ability to forgo the excitingly high reward (but high risk) gambles in favor of a lower reward but way lower risk option (i.e., it makes people risk seeking) (Knoch et al., 2006).

Anticipating Future Needs in Non-primate Mammals (Out-Group Condition)

I am only aware of one study where the Bischof–Kohler hypothesis was evaluated in non-primate mammals. In this study, non-thirsty squirrel monkeys and rats were tested in their ability to anticipate their future thirst and use this to change their current actions. While squirrel monkeys were shown to successfully anticipate future thirst, rats were incapable of doing so (Naqshbandi and Roberts, 2006). It should also be noted that there is some evidence that birds can solve this task (Roberts, 2007). But as noted above, birds are poor model organisms for the amniote last common ancestor.

Possible Adaptive Function(s) of Anticipating Future Needs in Early Primates (Stem-Group Condition)

Early primates are believed to have been foragers who lived in tightly knit social aggregations (Shultz et al., 2011). It has been proposed that living on forest fruits was much more difficult than other forms of foraging (Dunbar and Shultz, 2017). Forest fruits are variable, can face big shortages, are only available and ripe for short periods of time (sometimes only for 72 h, Milton, 1981), and are sought after by many animals (Milton, 1981, 1988; Chapman et al., 1999, 2004; Janmaat et al., 2014). As such, early arrival would be highly adaptive. This may then have required the motivation and ability to get food when it was available even in the absence of hunger (i.e., anticipate a future need).

Consistent with this, it has been shown that larger primate brains help buffer the risks associated with food scarcity during seasonality (van Woerden et al., 2011). Further, chimpanzees have been found to plan their entire foraging path at the beginning of the day to maximize the likelihood of getting food throughout the day (Janmaat et al., 2014). This requires anticipating hunger later in the day even when not yet hungry.

Conclusion

Taken together, evidence across the in-group, out-group, and stem-group conditions is generally supportive of the hypothesis that the ability to anticipate future needs is an ability that emerged in early primates. First, the ability to anticipate future needs has been observed in numerous non-human primates and the neural substrates of the ability to anticipate future needs seem to be structure that uniquely emerged in early primates (in-group condition). Second, negative results have been reported in non-primate mammals (out-group condition). And third, there are several plausible proposals as to the unique adaptive benefit of anticipating future needs in the ecological niche of early primates.

Hypothesis #10: Theory of Mind Emerged in Early Primates

Theory of Mind in Primates (In-Group Condition)

Theory of mind refers to the ability of an animal to take the perspective of someone else and understand that they can have different intentions, desires, and knowledge than you do. It continues to be controversial whether any animals other than humans have this ability. But there is compelling evidence that many primates do have theory of mind, even if it is not as robust as in humans. For example, many primates have passed the classic false belief test, whereby animals are tested in their ability to understand that another animal can hold a belief that the individual knows to be false. Macaques (Hayashi et al., 2020), Chimpanzees (Krupenye et al., 2016), and orangutans (Krupenye et al., 2016), have all passed this test.

In another test of theory of mind, non-human primates have been shown to be able to understand the intentions of others, as measured by their ability to distinguish between accidental and intentional actions as well as between someone unwilling to do something and someone unable to do something (Call and Tomasello, 1998; Tomasello et al., 2003, 2005; Call et al., 2004).

Another test of theory of mind is the goggle test, generally regarded to be more difficult than the false belief test. The goggle test includes showing an animal what it is like to look through opaque or transparent goggles, and then seeing if they treat humans wearing these different goggles differently, correctly inferring which human can see through their goggles. Apes have been shown to pass this test (Kano et al., 2019).

Some have proposed that there are degrees of theory of mind, and that while non-human primates have a form of theory of mind, it is far more limited than that of humans. One such interpretation is that non-human primates are aware that others have different beliefs, but only humans after the age of 4 can understand what those false beliefs are and use those beliefs to change their decisions (Kaminski et al., 2008). Although possible, this view is inconsistent with more modern studies that demonstrate the ability of non-human primates to indeed use false beliefs in decisions (Krupenye et al., 2016; Kano et al., 2019), which suggests that non-human primates have more sophisticated theory of mind than previously thought.

Consistent with the view that theory of mind emerged in early primates, the two structures most implicated in theory of mind are the superior temporal sulcus and temporoparietal junction, both of which are structures that emerged within early primates and are not present in non-primate mammals (Kaas, 2009). Theory of mind tasks in humans activates the superior temporal sulcus (Mars et al., 2011, 2012). Superior temporal sulcus activation correlates with subjective reports of considering the point of view of others (Dodell-Feder et al., 2011), when hearing stories (vs. nonsense speech), and observing faces (Beauchamp, 2015). Performance on theory of mind tasks has been shown to correlate with superior temporal sulcus activation (Otsuka et al., 2009). In monkeys, superior temporal sulcus activation is similarly sensitive to social information conveyed by faces, postures, and actions of others (Perrett et al., 1992; Dehaene et al., 2005).

The temporoparietal junction, sometimes considered part of the superior temporal sulcus (Beauchamp, 2015), is also highly implicated in theory of mind tasks. Temporoparietal junction activity in humans is correlated with how likely someone is to give a donation to someone else (Hare et al., 2010; Morishima et al., 2012), how risky a decision is that someone else makes in front of you (van den Bos et al., 2009), as well as altruism in a trust game (van den Bos et al., 2009). It also gets selectively activated when hearing false belief stories (versus false physical stories, Beauchamp, 2015). Although the temporoparietal junction is less studied in non-human primates, connectivity analysis has suggested a temporoparietal junction homolog in monkeys (Mars et al., 2013).

Theory of Mind in Non-primate Mammals (Out-Group Condition)

Most studies on non-primate mammals conclude that they do not have theory of mind (Byrne et al., 2001; Tomonaga et al., 2010; Bräuer, 2014; Aldhous, 2015). It should also be noted that there is some evidence of theory of mind in birds (Bugnyar et al., 2016), but the negative results in non-primate mammals, and the evidence of convergent brain evolution in birds, makes this unconvincing evidence for the presence of theory of mind in the amniote common ancestor of birds and primates.

Possible Adaptive Function(s) of Theory of Mind in Early Primates (Stem-Group Condition)

Early anthropoids likely lived in social societies with pair-living, had strong family bonds, and foraged in groups (Shultz et al., 2011). Like in the social groups of modern monkeys and apes, these ancestral social groups likely had hierarchies with coalitions and competition. The ability to infer intentions and beliefs of others would have enabled animals to more easily climb social hierarchies, build coalitions, collaborate, and deceive others when necessary (Dunbar and Shultz, 2017). Further, theory of mind may have also been a prerequisite for learning motor skills through observation, which would have been adaptive for group transmission of foraging skills and tool use (see next section).

Conclusion

Taken together, evidence across the in-group, out-group, and stem-group conditions is generally supportive of the hypothesis that theory of mind emerged, even in a simple form, with the first primates. First, theory of mind has been observed, even if in a primitive form, within many non-human primates, and the neural substrates of theory of mind in primates seem to be uniquely primate structures (in-group condition), Second, there are numerous negative results of theory of mind in non-primate mammals (out-group condition). And third, there are numerous plausible proposals for the unique adaptive benefit of theory of mind in the ecological niche of early primates.

Hypothesis #11: Learning Motor Skills Through Observation Emerged in Early Primates

Learning Motor Skills Through Observation in Primates (In-Group Condition)

There is evidence that non-human primates can learn motor skills such as tool use and manufacture simply by observing others perform such behaviors. This has been found in macaques (Ferrucci et al., 2019), rhesus monkeys (Meunier et al., 2007), and chimpanzees (Tomasello et al., 1987). For example, young chimpanzees who were able to observe adult chimpanzee use specific tools were able to figure out how to use the tools, while the young chimpanzees who did *not* observe adult chimpanzees use the tools failed to learn how to use them (Tomasello et al., 1987). Further evidence for this learning through observation was confirmed in diffusion experiments, whereby a new technique was taught to a few individuals in a group, and within a short period of time this technique had spread to other individuals in the group (Whiten et al., 2005; Dindo et al., 2009; Gunhold et al., 2014; van de Waal et al., 2015). Transmission of such skills has been shown across generations (Mercader et al., 2007; Haslam et al., 2016; Whiten, 2017).

There is even some emerging evidence that non-human primates will teach skills and tool use to peers and children, although this is still controversial. Such teaching would be a remarkable addition to the learning motor skills through observation ability, whereby one animal can anticipate the motor skills another must learn to achieve a goal. For example, macaques seem to exaggerate flossing when around their young, perhaps to help teach them (Choi, 2009). Chimpanzee's also show behaviors that some have interpreted as teaching (Everding, 2016; Musgrave et al., 2016). For example, skilled users will bring multiple fishing probes to a fishing activity and give one to a youngster. They will divide one in two if the child does not have one and will respond to begging if the child needs a tool. They can even identify that a youngster is struggling and will swap tools with them. But such teaching in non-human primates has been disputed (Premack, 2007; Hoppitt et al., 2008; Kline, 2015).

Further evidence for the idea that learning motor skills through observation emerged in early primates can be seen in the observation that many neural substrates of learning through observation are uniquely primate structures. The superior temporal sulcus in both humans and non-human primates activates in response to observing various forms of biological motion (Perrett et al., 1985; Puce and Perrett, 2003). Further, individual superior temporal sulcus neurons are highly selective for specific actions, as opposed to just being activated generically whenever biological movement is occurring (Perrett et al., 1985; Puce and Perrett, 2003). Superior temporal sulcus neurons are selective to the observation of others, and don't respond when the individual does the same action that they observed (Perrett et al., 1989; Jellema et al., 2000).

Humans and non-human primates have also been shown to also have mirror neurons in the premotor cortex and parietal lobes that activate both when an individual is taking a specific action as well as when they observe other conspecifics doing that same action (Gallese et al., 1996; Rizzolatti et al., 1996; Fogassi, 2005). Observed motor acts in others done by different body parts are somatotopically organized in the classic homunculus the same way as when the individual moves those same body parts themselves (Penfield and Rasmussen, 1950; Woolsey et al., 1952; Buccino et al., 2001). Interestingly, these mirror neurons respond robustly even if non-human primates can only infer the motion of others without explicitly observing it, such as when half of the motion is obscured (Umiltà et al., 2001; Kohler et al., 2002). The general view is that these mirror neurons encode the abstract goals of motor acts (Rizzolatti et al., 2001). Recording studies in non-human primates corroborate this view, as mirror neurons seem to discriminate different goals, even if the observed movements are the same (Fogassi, 2005).

As suggested by others (Tennie et al., 2009; Tomasello and Moll, 2009; Dean et al., 2013), it is likely that these abilities in non-human primate is not as sophisticated as they are in humans. One difference that has been proposed is that of shared intentionality, whereby although non-human primates understand intentions of others, they are much less motivated than are humans to share their mental states (Tomasello and Carpenter, 2007; Call, 2009). Another difference that has been suggested is that observational learning may be much more cumulative in humans than in non-human primates (Tennie et al., 2009; Dean et al., 2013). For example, although chimps can learn through observation, and can transmit these across transmission chains (Whiten et al., 2005), it has been suggested that such learning tends to be focused on the outcomes of movements, as opposed to also mimicking the entire nuanced movements themselves. This prevents a cumulative evolution of an action or cultural phenomena, since it always falls back to individuals finding the easiest way to achieve an outcome. Individuals will happily modify a learned skill to make it easier for themselves. It has been suggested that in humans however, there is a cumulative mimicry of actual actions, independent of outcome, which then creates a unique "ratcheting up the rachet" (Tennie et al., 2009) of culture, motor skills, and information. In contrast, in non-human primates such knowledge, while still transferable across generations, does not accumulate across generations as accurately.

Learning Motor Skills Through Observation in Non-Primate Mammals (Out-Group Condition)

Many animals, even fish, demonstrate socially coordinated movements. Coordinating movements with other conspecifics is likely a key anti-predator behavior evolved very early in evolution, perhaps even in early vertebrates. Further, there is also evidence of learning paths by observation in early diverging vertebrates such as fish (Lindeyer and Reader, 2010; Brown, 2015) and reptiles (Wilkinson et al., 2010), where they will learn to take navigational paths through simply observing conspecifics take the same path.

However, there are two differences between observational learning in non-human primates relative to that in non-primate vertebrates. First, learning through observation of the type in non-human primates includes an understanding of the intention of movement, not simply mirroring the movement. Coordination

and path following on the other hand can be achieved with much simpler mimicry. I am not aware of any studies demonstrating the ability to infer intentions in non-primate mammals. Further, goal-selective mirror neurons found in primates have not been found in other non-primate mammals. Although mirror neurons have indeed been found in songbirds (Prather et al., 2008), but as discussed previously, birds are a poor model organism for inferring abilities in our amniote common ancestor. Second, learning through observation in non-human primates has been shown to be incredibly transferable, and can be transmitted not only amongst peers but across many generations. Fish that learn paths through observation do not pass down this knowledge from generation to generation (Lindeyer and Reader, 2010).

Possible Adaptive Function(s) of Learning Motor Skills Through Observation in Early Primates (Stem-Group Condition)

As observed in many extant primates, it is likely that early primates (or at least anthropoids) used sophisticated motor skills and tools to obtain food from hard-to-get-places that other animals can't access. It is then reasonable to speculate that the ability for such skills to be transmitted throughout social groups would have improved a groups' ability to reliably obtain food.

Conclusion

Taken together, evidence across the in-group, out-group, and stem-group conditions is generally supportive of the hypothesis that learning motor skills through observation emerged in early primates. First, there are numerous reports of learning through observation throughout non-human primates, and the neural substrates of learning through observation seem to be structures that emerged first in early primates (in-group condition). Second, the observational learning in non-primate mammals seems to be much more limited (out-group condition). Specifically, non-primate mammals seem to lack the ability to understand the intention of observed motor movements as well as to consistently transfer such skills across generations – two features of observational learning consistently seen in non-human primates. And third, learning motor skills through observation would have offered adaptive benefits to the foraging ecological niche of early primates.

BEHAVIORAL ABILITIES THAT EMERGED IN EARLY HUMANS

Hypothesis #12: Language Emerged in Early Humans

Language in Humans (In-Group Condition)

The ability to name objects and organize words with grammar has been suggested to be what makes human language unique (Berwick and Chomsky, 2017; Terrace, 2019). The ontogeny of human language learning is revealing as to how these unique features of language emerge. Human infants even as young as a few months, far before they can speak, engage in affective imitation and rhythmic exchanges with their mother of various affective forms of communication – such as vocalizing, making

gestures and facial expressions (Meltzoff and Moore, 1977, 1989). Human infants and mothers will match the duration of each other's pauses, creating affective proto conversations (Beebe et al., 1988, 2016). By 6 months of age, infants begin to engage in *shared attention* of the same object as their mother. Infants have non-verbal mechanisms to confirm that they saw what you saw (Carpenter and Call, 2013). This enables parents to name things that have joint attention. As evidence for its role in language, the more joint attention expressed by a child, the larger the child's vocabulary 12 months later (Morales et al., 2000; Mundy et al., 2007). This is suggestive that two abilities required for the development of language are rhythmic exchanges of affective communication (proto conversations) and shared attention.

The two areas most implicated in language are Broca's area in on the left inferior prefrontal cortex and Wernicke's area in the left temporal lobe. Lesions to both these areas severely impair language abilities (Wernicke, 1995; DeWitt and Rauschecker, 2013). However, these areas have homologous regions in non-human primates, making it unlikely that these are completely new structures from which language emerged. It is more likely that language emerged via new connectivity that used existing structures in new ways.

The most notable connectivity difference is seen in the arcuate fasciculus, which connects Broca's area and Wernicke's area. In humans the arcuate fasciculus is massively expanded (Aboitiz and García, 1997; Aboitiz et al., 2006, 2010; Rilling et al., 2008, 2012; Aboitiz, 2012; Petrides, 2014; Rilling, 2014; Catani and Bambini, 2014; Stout and Hecht, 2017). The arcuate fasciculus also appears to contain unique connectivity in humans – whereby the frontal cortex of left hemisphere connects to left medial temporal gyrus and inferotemporal gyrus, close to the areas usually included in Wernicke's area – a connection not observed in non-human primates (Rilling et al., 2008). The proper functioning of arcuate fasciculus is also highly associated with various functions of language (Binder and Desai, 2011; Schomers et al., 2017). Damage to arcuate fasciculus impairs verbal working memory, fluency, and comprehension. Further, the development of AF correlates with language abilities in childhood (Friederici, 2011; Yeatman et al., 2011; Skeide et al., 2016; Goucha et al., 2017; Schomers et al., 2017), and the strength of connectivity of left arcuate fasciculus correlates with performance in word learning (Lopez-Barroso et al., 2013).

Language in Non-human Primates (Out-Group Condition)

Many animals communicate with one another – birds sing songs to attract mates and defend territories (Langmore, 1998), bees communicate location using dances (Aguilar et al., 2005), and vervet monkeys have alarm calls specific to different predators (Cheney and Seyfarth, 1996). But human language differs from these forms of communication in that humans have grammar and words. Human language can combine sounds (phonemes and words) using grammar to create an almost infinite number of novel meanings. The results from attempts to teach non-human primates such flexible language have mostly supported the conclusion that non-human primates are incapable of learning language with grammar and words. Initial potential positive

results in teaching the basics of sign language (Gardner and Gardner, 1969; Patterson, 1978) have been disconfirmed by later studies (Terrace et al., 1979), where it was shown that signing was always a sequence of prompted signs, had low diversity, and was non-grammatical. Further analysis of these earlier studies also demonstrated that the subjects only learned sequences of symbols to get rewards, and sequences were the result of simple non-grammatical rules (Thompson and Church, 1980). Some of these earlier studies may even have unintentionally cued the animals to make specific signs (Miles, 1983). Later work with bonobos showed impressive signing abilities, perhaps indicative of simple grammar and words (Savage-Rumbaugh et al., 1986), but this interpretation is still controversial (Pinker, 1994; Terrace, 2019). The most popular view seems to be that non-human primates can learn imperative functions of symbols (use them to attain rewards), but they fail to learn that objects have names, and that those names can be used conversationally (Terrace, 2019).

The underlying ontogenetic mechanisms whereby language emerges in infant humans, namely joint attention and proto-conversations, are not present in non-human primates. Attempts to show joint attention in chimps showed negative results (Warneken, 2006; Warneken et al., 2006).

There are no reports of these proto-conversations in non-human primates and others have suggested they do not occur (Terrace, 2019). It has been shown that collaborative activities that require joint attention become much more difficult for non-human primates than those that don't require joint attention (Tomasello and Carpenter, 2005; Tomasello et al., 2005; Melis and Tomasello, 2013).

Further evidence for the uniqueness of human language relative to animal communication can be seen in their completely different neural substrates. Primate calls, such as those in vervet monkeys, are driven by midbrain and limbic forebrain structures (Jürgens, 1988). Damage to monkey cortical areas homologous to human language areas do not produce call production deficits (Aitken, 1981; Jürgens et al., 1982). And voluntary speech in humans is driven by direct cortical projections. Circumventing the limbic and midbrain pathways wherein animal calls tend to be produced (Fitch, 2018; Jarvis, 2019). This is strong evidence that language is not an evolutionary elaboration of animal call production system, but a separate system all together.

Possible Adaptive Function(s) of Language in Early Humans (Stem-Group Condition)

There are numerous proposals for the original adaptive benefit that language provided early humans, including cooperative hunting (Washburn and Lancaster, 1968), promotion of pair bonding (Deacon, 1997), expediting of toolmaking (Greenfield, 1991), and the enhancement of teaching (Laland, 2017). One additional leading theory is the confrontational scavenging theory, which suggests that the ecological niche of early humans was to split up into groups to scout for dead animals. After finding a carcass they would recruit a large group to scare off other animals and work together to scavenge the remains. This would have required *words* to reference the identity and location of the discovered remains to other conspecifics (Terrace, 2019), this ability has been called displaced reference, whereby

an animal can refer to something that is currently not present (Blumenschine et al., 1994). Note that select invertebrates seem to have independently evolved certain abilities of displaced reference as well, although it seems reflexive and not flexible as in humans, including bees (Von Frisch, 1967) and ants (Wilson, 1962).

Conclusion

Taken together, evidence across the in-group, out-group, and stem-group conditions is generally supportive of the hypothesis that the ability to use language with words and grammar first emerged in early humans. First, language seems to be universal across human cultures, and key neural substrates of language in humans seem to use modifications that emerged uniquely in early humans (in-group condition). Second, most studies of language in non-human primates conclude that they are unable to learn language with words and grammar (out-group condition). And third, evidence suggests that the ecological niche of early humans was such that language would have offered a uniquely adaptive benefit (stem-group condition). Additionally, the ontogeny of language in humans seems to rely on proto-conversations and joint attention, two features that have been reported to not exist in non-human primates.

Hypothesis #13: Music Emerged in Early Humans

Music in Humans (In-Group Condition)

Music, like language, seems to be universal to humans – music is found across all human cultures, and across all these cultures music shares many features (Brown and Jordania, 2011). One feature of music across human cultures is beat-based timing. Evidence suggests that beat-based timing has a separate neural substrate than duration-based timing (Teki et al., 2011). Duration-based timing requires discriminating the absolute amount of time passed, whereas beat-based timing requires discriminating relative timing. Duration-based timing seems to be implemented in the inferior olivary nucleus and cerebellum (Teki et al., 2011). In contrast, beat-based timing seems to be implemented within a network of the striatum, thalamus, premotor cortex, and prefrontal cortex (Rao et al., 1997; Jäncke et al., 2000; Grahn and Brett, 2007; Wiener et al., 2010; Teki et al., 2011). The mere perception of a beat, even without movement, activates premotor areas (Grahn and Brett, 2007; Chen et al., 2008; Grahn and Rowe, 2009; Geiser et al., 2012; Teki et al., 2012; Kung et al., 2013) and creates enhanced coupling between auditory areas and motor areas (Kung et al., 2013). In humans this circuitry is involved in many observed beat-related behaviors such as dance, music, timing movements, and sequencing movements (Grafton et al., 1995; Grahn and Brett, 2007; Harrington et al., 2009; Wiener et al., 2010). Consistent with this, humans with cerebellar lesions are impaired in duration-based timing tasks, but not in beat-based timing tasks (Grube et al., 2010). And lesions of striatum or substantia nigra impair beat-based timing (Teki et al., 2011). Further, Parkinson's patients, who have impairment in basal ganglia function, also seem to struggle with beat-based timing tasks (Teki et al., 2011).

Music in Non-human Primates (Out-Group Condition)

All mammals, especially non-human primates, have very impressive auditory perception. They can perceive pitches (Song et al., 2015) and they can learn very subtle differences in sound structures (Selezneva et al., 2006; Tsunada et al., 2011). Many mammals and birds can perceive the fundamental frequency of a chord even if it has been removed (showing complex pitch perception) (Heffner and Whitfield, 1976; Cynx and Shapiro, 1986; Tomlinson and Schwarz, 1988). Non-human primates can also readily observe the interval timing between two events, such as two clicks, and repeat back almost the same interval (Zarco et al., 2009).

However, there are differences in music-related cognitive abilities between humans and non-human primates. The most obvious difference is with beat perception. Monkeys cannot learn to synchronize taps with an auditory or visual metronome, even after a year of training (Zarco et al., 2009; Honing et al., 2012). Electro-encephalogram studies in non-human primates have confirmed that they do not seem to perceive beats in auditory signals (Honing et al., 2012). Even birds, bats, and dolphins with exceptional natural song abilities, have weak beat perception (Patel et al., 2009; Schachner et al., 2009; Merchant and Honing, 2014). Moreover, this very basic skill of tapping to a beat is remarkably universal and accurate amongst human (Wallin et al., 2000; Iversen and Patel, 2008; Rankin et al., 2009; Repp and Su, 2013; van der Steen and Keller, 2013). Human beat perception is also flexible: humans can adjust the same melody to changing beats across a very wide range of tempos (Large and Jones, 1999; Honing, 2013). Even the few animals that show some form of rhythmic entrainment (such as birds), do not show the ability to synchronize across a wide range of tempos as humans do (Large, 2000; Fitch, 2009; Patel et al., 2009).

The cortico-basal-ganglia-thalamic circuit important for human musical ability also exists in monkeys and is also related to timing and sequencing movements (Tanji, 2001; Merchant et al., 2013; Perez et al., 2013). However, there are differences in the circuitry between humans and non-human primates that may explain the lack of beat-perception in non-human primates. First, the arcuate fasciculus in humans connects auditory superior temporal areas and premotor areas much more strongly than in non-human primates. Broca's area (as well as homolog on right side), known to be less developed in non-human primates (Petrides and Pandya, 2009), has been implicated in concatenating and organizing hierarchical temporal sequences of simple motor actions (Koechlin and Jubault, 2006). Second, superior temporal areas in humans project much more extensively to the basal ganglia than they do in macaques (Yeterian and Pandya, 1998; Borgmann and Jürgens, 1999; Rilling et al., 2008; Thiebaut de Schotten et al., 2012). Consistent with this, the cortico-basal-ganglia-thalamic circuit is activated less by audiomotor tasks in Rhesus monkeys than in humans (Grahn and Brett, 2007).

Non-human primates also struggle with *relative* pitch perception more than humans do. Non-human animals that learn melodies often struggle to generalize these to transpositions (Hulse et al., 1984; D'Amato, 1988), a task that is mostly automatic for humans. The reports that demonstrate basic forms of relative pitch perception (Ferrets: Yin et al., 2010; Rhesus monkeys: Wright et al., 2000; Dolphins: Ralston and Herman, 1995) use very simple paradigms, where a subject merely as to perceive an octave generalization or to simply discriminate increasing pitch versus decreasing pitch– and even learning these simple tasks often requires a lot of training. Further, evidence suggests non-human primates show no preference for consonance, a feature of certain pitch intervals that humans across cultures tend to report finding pleasant (McDermott and Hauser, 2004; Vassilakis, 2005). Although some birds do show this preference (Chiandetti and Vallortigara, 2011).

Possible Adaptive Function(s) of Music in Early Humans (Stem-Group Condition)

Numerous adaptive benefits have been proposed for music, most of which hinge on the social interactions of early humans (reviewed in Cross and Morley, 2008). For example, it has been proposed that music was useful for coordinating cooperative social behaviors (Brown, 2006; Cross, 2009), for more nuanced forms of communication that language alone did not sufficiently satisfy (Cross, 2009), or for promoting group cohesion (Roederer, 1984). Musical instruments have been found over 50,000 years ago, suggesting that music emerged around a similar period as language (Cross, 1999; Morley, 2003).

Conclusion

Taken together, evidence across the in-group, out-group, and stem-group conditions is supportive of the hypothesis that the totality of human-level music perception, inclusive of beat perception, relative pitch, and consonance preference, emerged in early humans and did not exist in our last common ancestor with chimpanzees. First, music has been observed across human cultures, suggesting that it is universal to humans and key neural substrates of music include neural modifications that are unique to humans (in-group condition). Second, negative results of various features of music, such as beat perception, have been reported in non-human primates (out-group condition). And third, the ecological niche of early humans was such that music may have offered unique adaptive benefits (stem-group condition).

DISCUSSION

I have presented 13 hypotheses of specific behavioral abilities that emerged at major milestones in the evolution of the human brain. I hope they provide a guide for further comparative analysis so they may be refined as more evidence emerges. Deciphering the phylogenetic history of human intelligence offers many opportunities to better understand both our own history and how the human brain works.

Comparison to Other Work

Ginsburg and Jablonka (2019, 2021) have done extensive work in chronicling the evolutionary transitions by which various learning abilities emerged. The hypotheses presented

here were informed by their work. I have attempted to add to their framework by adding specific behavioral capabilities that likely emerged at each transition, in addition to various features of learning. They define three levels of learning: non-associative learning (sensitization and habituation), limited associative learning (LAL), and unlimited associative learning (UAL). They describe LAL as the ability to reinforce the association between "simple" stimuli and actions. This is what I have called "associative learning" in hypothesis #2. They define UAL as a form of learning that enabled *unlimited* numbers of associations between complex stimuli and actions. They argue that UAL enabled four key advantages: (1) animals could perceive compound stimuli and actions, (2) animals could engage in pattern completion, (3) animals could make guesses as to the right behavior based on past experiences, and (4) animals could make cumulative improvements to cope with new situations. They argue that UAL independently evolved in three lineages. The hypotheses here parse out some of the specific features of UAL that emerged within the vertebrate lineage – such as omission learning, map-based navigation, VTE, and episodic memory. Further, some abilities presented here are not "learning" abilities *per se*, and hence are additive to their analysis. For example, the ability to anticipate a need in the future or theory of mind are not necessarily only unique in the ability to *learn* something new, but also the ability to *apply* something previously learned in a new way.

Cisek's (2007, 2019) work on phylogenetic refinement also hypothesizes specific initial behaviors that were present in ancestral species. Some of the hypotheses here are consistent with his. For example, his proposal that early bilaterians had mechanisms for exploitation and exploration, modulated by neuromodulators and peptides that signaled internal states, is consistent with the hypotheses here that taxis navigation with valence and cross-modal integration first emerged in early bilaterians. Cisek also presents various interesting ideas on the sensorimotor and other functions of the neocortex, which are not explored in this paper, such as his theory of affordance competition, whereby the neocortex provided novel mechanisms for flexible action specification and selection (Cisek, 2007; Cisek and Kalaska, 2010; Pezzulo and Cisek, 2016).

There has also been work done chronicling the evolution of memory systems (Murray et al., 2017). Relevant hypotheses here are largely consistent with their arguments. For example, they similarly propose that a spatial navigation system emerged in early vertebrates. Further, they similarly propose that granular prefrontal areas that emerged in early primates enabled more complex goal representations, consistent with the idea that granular prefrontal areas were supportive of anticipating future needs. They also propose unique features that emerged in the primate lineage (not discussed in the current paper), such as grasping and manual foraging mechanisms. The hypotheses here provide additional proposals for behavioral abilities that emerged.

Testable Predictions

Testable predictions can be formulated for each of the hypotheses presented here. The three conditions (in-group, out-group, and stem-group conditions) provides a framework for formulating such predictions; each condition has a corresponding set of predictions. Consider the hypothesis that omission learning emerged in early vertebrates. Testable predictions of the in-group condition are that additional early diverging vertebrates will succeed in omission learning studies. Testable predictions of the out-group condition are that additional tests of protostomes will demonstrate failures in omission learning, or if they succeed, it will be implemented in neural substrates that independently evolved. Testable predictions of the stem-group condition are less specific but can be formulated as the prediction that the ecological niche of early vertebrates will prove to have been consistent with an environment where omission learning would have been adaptive.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

MB conceived the overall theory and wrote the entire manuscript.

REFERENCES

Abe, H., and Lee, D. (2011). Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron* 70, 731–741. doi: 10.1016/j.neuron.2011.03.026

Abellan, A., Desfilis, E., and Medina, L. (2014). Combinatorial expression of Lef1, Lhx2, Lhx5, Lhx9, Lmo3, Lmo4, and Prox1 helps to identify comparable subdivisions in the developing hippocampal formation of mouse and chicken. *Front. Neuroanat.* 8:59. doi: 10.3389/fnana.2014.00059

Aboitiz, F. (2012). Gestures, vocalizations, and memory in language origins. *Front. Evol. Neurosci.* 4:2. doi: 10.3389/fnevo.2012.00002

Aboitiz, F., Aboitiz, S., and García, R. R. (2010). The phonological loop. *Curr. Anthropol.* 51, S55–S65. doi: 10.1086/650525

Aboitiz, F., and García, V. R. (1997). The evolutionary origin of the language areas in the human brain. A neuroanatomical perspective. *Brain Res. Rev.* 25, 381–396. doi: 10.1016/s0165-0173(97)00053-2

Aboitiz, F., García, R. R., Bosman, C., and Brunetti, E. (2006). Cortical memory mechanisms and language origins. *Brain Lang.* 98, 40–56. doi: 10.1016/j.bandl.2006.01.006

Abramson, C. I., and Bitterman, M. E. (1986). Latent inhibition in honeybees. *Anim. Learn. Behav.* 14, 184–189. doi: 10.3758/bf03200054

Abramson, C. I., and Boyd, B. J. (2001). An automated apparatus for conditioning proboscis extension in honey bees, *Apis mellifera* L. *J. Entomol. Sci.* 36, 78–92. doi: 10.18474/0749-8004-36.1.78

Abramson, C. I., and Feinman, R. D. (1990). Lever-press conditioning in the crab. *Physiol. Behav.* 48, 267–272. doi: 10.1016/0031-9384(90)90311-q

Abramson, C. I., and Wells, H. (2018). An inconvenient truth: some neglected issues in invertebrate learning. *Perspect. Behav. Sci.* 41, 395–416. doi: 10.1007/s40614-018-00178-8

Abramson, C. I., Armstrong, P. M., Feinman, R. A., and Feinman, R. D. (1988). Signaled avoidance in the eye withdrawal reflex of the green crab. *J. Exper. Analys. Behav.* 50, 483–492. doi: 10.1901/jeab.1988.50-483

Abramson, C. I., Dinges, C. W., and Wells, H. (2016). Operant conditioning in honey bees (*Apis mellifera* L.): the cap pushing response. *PLoS One* 11:e0162347. doi: 10.1371/journal.pone.0162347

Acebes, F., Solar, P., Carnero, S., and Loy, I. (2009). Blocking of conditioning of tentacle lowering in the snail (*Helix aspersa*). *Q. J. Exper. Psychol.* 62, 1315–1327. doi: 10.1080/17470210802483545

Ackil, J. E., Mellgren, R. L., Halgren, C., and Frommer, G. P. (1969). Effects of CS preexposures on avoidance learning in rats with hippocampal lesions. *J. Compar. Physiol. Psychol.* 69(4 Pt.1), 739–747. doi: 10.1037/h0028159

Agetsuma, M., Aizawa, H., Aoki, T., Nakayama, R., Takahoko, M., Goto, M., et al. (2010). The habenula is crucial for experience-dependent modification of fear responses in zebrafish. *Nat. Neurosci.* 13, 1354–1356. doi: 10.1038/nn.2654

Aguilar, I., Fonseca, A., and Biesmeijer, J. C. (2005). Recruitment and communication of food source location in three species of stingless bees (Hymenoptera, Apidae, Meliponini). *Apidologie* 36, 313–324. doi: 10.1051/apido:2005005

Aitken, P. G. (1981). Cortical control of conditioned and spontaneous vocal behavior in rhesus monkeys. *Brain Lang.* 13, 171–184. doi: 10.1016/0093-934x(81)90137-1

Aldhous, P. (2015). *Wellcome. The Smartest Animal you've Never Heard of.* Available online at: https://mosaicscience.com/extra/smartest-animal-youve-never-heard/ (accessed March 23, 2021).

Allen, T. A., and Fortin, N. J. (2013). The evolution of episodic memory. *Proc. Natl. Acad. Sci. U.S.A.* 110(Suppl. 2), 10379–10386. doi: 10.1073/pnas.1301199110

Andrychowicz, M., Wolski, F., and Roy, A. (2017). "Hindsight experience replay," in *Proceedings of the Part of Advances in Neural Information Processing Systems 30*, (California: Long Beach), 5048–5058.

Arcediano, F., Matute, H., and Miller, R. R. (1997). Blocking of pavlovian conditioning in humans. *Learn. Motiv.* 28, 188–199. doi: 10.1006/lmot.1996.0957

Ardiel, E. L., and Rankin, C. H. (2010). An elegant mind: learning and memory in *Caenorhabditis elegans*. *Learn. Mem.* 17, 191–201. doi: 10.1101/lm.960510

Avcu, P., Jiao, X., Myers, C. E., Beck, K. D., Pang, K. C., and Servatius, R. J. (2014). Avoidance as expectancy in rats: sex and strain differences in acquisition. *Front. Behav. Neurosci.* 8:334. doi: 10.3389/fnbeh.2014.00334

Babb, S. J., and Crystal, J. D. (2005). Discrimination of what, when, and where: implications for episodic-like memory in rats. *Learn. Motiv.* 36, 177–189. doi: 10.1016/j.lmot.2005.02.009

Balci, F. (2015). Interval timing behavior: comparative and integrative approaches. *Intern. J. Compar. Psychol.* 28:510.

Balci, F., Meck, W. H., Moore, H., and Brunner, D. (2009). "Timing deficits in aging and neuropathology," in *Animal Models of Human Cognitive Aging*, eds J. Bizon and A. Woods (Totowa, NJ: Humana Press), 1–41. doi: 10.1007/978-1-59745-422-3_8

Balleine, B. W. (2001). "Incentive processes in instrumental conditioning," in *Handbook of Contemporary Learning Theories*, eds R. R. Mowrer and S. B. Klein (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 307–366. doi: 10.4324/9781410600691-14

Balleine, B. W. (2018). The motivation of action and the origins of reward. *Goal Direct. Dec. Mak.* 2018, 429–455. doi: 10.1016/b978-0-12-812098-9.00019-x

Bateson, M., and Kacelnik, A. (1997). Starlings' preferences for predictable and unpredictable delays to food. *Anim. Behav.* 53, 1129–1142. doi: 10.1006/anbe.1996.0388

Batham, E. J., and Pantin, C. F. A. (1950a). Inherent activity in the sea-anemone, *Metridium senile* (L.). *J. Exper. Biol.* 27, 290–301.

Batham, E. J., and Pantin, C. F. A. (1950b). Phases of activity in the sea-anemone, metridium senile (L.), and their relation to external stimuli. *J. Exp. Biol.* 27, 377–399.

Beauchamp, M. S. (2015). The social mysteries of the superior temporal sulcus. *Trends Cogn. Sci.* 19, 489–490. doi: 10.1016/j.tics.2015.07.002

Beebe, B., Alson, D., Jaffe, J., Feldstein, S., and Crown, C. (1988). Vocal congruence in mother-infant play. *J. Psycholing. Res.* 17, 245–259. doi: 10.1007/BF01686358

Beebe, B., Cohen, P., and Lachmann, F. M. (2016). *The Mother-Infant Interaction Picture Book: Origins of Attachment*, 1st Edn, New York, NY: W. Norton.

Benchenane, K., Peyrache, A., Khamassi, M., Tierney, P. L., Gioanni, Y., Battaglia, F. P., et al. (2010). Coherent theta oscillations and reorganization of spike timing in the hippocampal- prefrontal network upon learning. *Neuron* 66, 921–936. doi: 10.1016/j.neuron.2010.05.013

Ben-Hamo, M., Muñoz-Garcia, A., Larrain, P., Pinshow, B., Korine, C., and Williams, J. B. (2016). The cutaneous lipid composition of bat wing and tail membranes: a case of convergent evolution with birds. *Proc. R. Soc. B Biol. Sci.* 283:20160636. doi: 10.1098/rspb.2016.0636

Benhamou, S., Sauve, J. P., and Bovet, P. (1990). Spatial memory in large scale movements: efficiency and limitation of the egocentric coding process. *J. Theoret. Biol.* 145, 1–12.

Ben-Yishay, E., Krivoruchko, K., Ron, S., Ulanovsky, N., Derdikman, D., and Gutfreund, Y. (2020). Head-direction coding in the hippocampal formation of birds. *bioRxiv* [Preprint]. doi: 10.1101/2020.08.31.274928

Berridge, K. C. (2019). Affective valence in the brain: modules or modes? *Nat. Rev. Neurosci.* 20, 225–234. doi: 10.1038/s41583-019-0122-8

Berwick, R. C., and Chomsky, N. (2017). *Why Only Us: Language and Evolution*. Cambridge, MA: MIT Press.

Bett, D., Murdoch, L. H., Wood, E. R., and Dudchenko, P. A. (2015). Hippocampus, delay discounting, and vicarious trial-and-error. *Hippocampus* 25, 643–654. doi: 10.1002/hipo.22400

Billard, P. (2020). Comparative Study of Episodic Memory in Common Cuttlefish (Sepia Officinalis) and Eurasian jay (Garrulus glandarius): Request PDF. ResearchGate. Available online at: https://www.researchgate.net/publication/349622632_Comparative_study_of_episodic_memory_in_common_cuttlefish_Sepia_officinalis_and_Eurasian_jay_Garrulus_glandarius (accessed May 15, 2021).

Binder, J. R., and Desai, R. H. (2011). The neurobiology of semantic memory. *Trends Cogn. Sci.* 15, 527–536. doi: 10.1016/j.tics.2011.10.001

Bird, L. R., Roberts, W. A., Abroms, B., Kit, K. K., and Crupi, C. (2003). Spatial memory for food hidden by rats (*Rattus norvegicus*) on the radial maze: studies of memory for where, what, and when. *J. Comp. Psychol.* 117, 176–187. doi: 10.1037/0735-7036.117.2.176

Bischof-Köhler, D. (1985). "Zur phylogenese menschticher motivation [On the phylogeny of human motivation]," in *Emotion und Reflexivitat*, eds L. H. Eckensberger and E. D. Lantermann (Vienna: Urban & Schwarzenberg), 3–47.

Blumenschine, R. J., Cavallo, J. A., and Capaldo, S. D. (1994). Competition for carcasses and early hominid behavioral ecology: a case study and conceptual framework. *J. Hum. Evol.* 27, 197–213.

Boisvert, M. J., and Sherry, D. F. (2006). Interval timing by an invertebrate, the bumble bee bombus impatiens. *Curr. Biol.* 16, 1636–1640. doi: 10.1016/j.cub.2006.06.064

Boles, L. C., and Lohmann, K. J. (2003). True navigation and magnetic maps in spiny lobsters. *Nature* 421, 60–63. doi: 10.1038/nature01226

Borgmann, S., and Jürgens, U. (1999). Lack of cortico-striatal projections from the primary auditory cortex in the squirrel monkey. *Brain Res.* 836, 225–228. doi: 10.1016/s0006-8993(99)01704-7

Bosch, T., Klimovich, A., Domazet-Lošo, T., Gründer, S., Holstein, T. W., Jékely, G., et al. (2017). Back to the basics: cnidarians start to fire. *Trends Neurosci.* 40, 92–105. doi: 10.1016/j.tins.2016.11.005

Boughner, R. L., and Papini, M. R. (2006). Appetitive latent inhibition in rats: preexposure performance does not predict conditioned performance. *Behav. Process.* 72, 42–51. doi: 10.1016/j.beproc.2005.11.015

Bräuer, J. (2014). "Chapter 10 - What dogs understand about humans," in *The Social Dog: Behaviour and Cognition*, eds J. Kaminski and S. Marshall-Pescini (Cambridge, MA: Academic Press), 295–317.

Brennan, P. A., and Kendrick, K. M. (2006). Mammalian social odours: attraction and individual recognition. *Philos. Trans. R. Soc. B Biol. Sci.* 361, 2061–2078. doi: 10.1098/rstb.2006.1931

Broadbent, N. J., Squire, L. R., and Clark, R. E. (2004). Spatial memory, recognition memory, and the hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* 101, 14515–14520. doi: 10.1073/pnas.0406344101

Broglio, C., Martín-Monzón, I., Ocaña, F. M., Gómez, A., Durán, E., Salas, C., et al. (2015). Hippocampal pallium and Map-Like memories through vertebrate evolution. *J. Behav. Brain Sci.* 05, 109–120. doi: 10.4236/jbbs.2015.53011

Brown, C. (2015). Fish intelligence, sentience and ethics. *Anim. Cogn.* 18, 1–17. doi: 10.1007/s10071-014-0761-0

Brown, J., Bullock, D., and Grossberg, S. (1999). How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. *J. Neurosci.* 19, 10502–10511. doi: 10.1523/jneurosci.19-23-10502.1999

Brown, S. (2006). "How does music work? Towards a pragmatics of musical communication," in *Music and Manipulation: On the Social Used and Social Control of Music*, eds S. Brown and U. Volgsten (New York, NY: Berghahn Books), 1–27.

Brown, S., and Jordania, J. (2011). Universals in the world's musics. *Psychol. Music* 41, 229–248. doi: 10.1177/0305735611425896

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* 13, 400–404.

Bugnyar, T., Reber, S. A., and Buckner, C. (2016). Ravens attribute visual access to unseen competitors. *Nat. Commun.* 7:10506. doi: 10.1038/ncomms10506

Buhusi, C. V., Perera, D., and Meck, W. H. (2005). Memory for timing visual and auditory signals in albino and pigmented rats. *J. Exper. Psychol. Anim. Behav. Process.* 31, 18–30. doi: 10.1037/0097-7403.31.1.18

Buhusi, C. V., Sasaki, A., and Meck, W. H. (2002). Temporal integration as a function of signal and gap intensity in rats (*Rattus norvegicus*) and pigeons (*Columba livia*). *J. Compar. Psychol.* 116, 381–390. doi: 10.1037/0735-7036.116.4.381

Burke, C. J., Huetteroth, W., Owald, D., Perisse, E., Krashes, M. J., Das, G., et al. (2012). Layered reward signalling through octopamine and dopamine in *Drosophila*. *Nature* 492, 433–437. doi: 10.1038/nature11614

Burt de Perera, T., Holbrook, R. I., and Davis, V. (2016). The representation of three- dimensional space in fish. *Front. Behav. Neurosci.* 10:40. doi: 10.3389/fnbeh.2016.00040

Byrne, R., Mendl, M., Devereux, C., and Held, S. (2001). Behaviour of domestic pigs in a visual perspective taking task. *Behaviour* 138, 1337–1354. doi: 10.1163/156853901317367627

Call, J. (2009). Contrasting the social cognition of humans and nonhuman apes: the shared intentionality hypothesis. *Top. Cogn. Sci.* 1, 368–379. doi: 10.1111/j.1756-8765.2009.01025.x

Call, J., and Tomasello, M. (1998). Distinguishing intentional from accidental actions in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*), and human children (*Homo sapiens*). *J. Compar. Psychol.* 112, 192–206. doi: 10.1037/0735-7036.112.2.192

Call, J., Hare, B., Carpenter, M., and Tomasello, M. (2004). "Unwilling" versus "unable": Chimpanzees' understanding of human intentional action. *Dev. Sci.* 7, 488–498. doi: 10.1111/j.1467-7687.2004.00368.x

Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J. R., and Sirigu, A. (2004). The involvement of the orbitofrontal cortex in the experience of regret. *Science* 304, 1167–1170. doi: 10.1126/science.1094550

Carpenter, M., and Call, J. (2013). How joint is the joint attention of apes and human infants? *Agency Joint Attent.* 25, 49–61. doi: 10.1093/acprof:oso/9780199988341.003.0003

Cassenaer, S., and Laurent, G. (2012). Conditional modulation of spike-timing-dependent plasticity for olfactory learning. *Nature* 482, 47–52.

Catani, M., and Bambini, V. (2014). A model for social communication and language evolution and development (SCALED). *Curr. Opin. Neurobiol.* 28, 165–171. doi: 10.1016/j.conb.2014.07.018

Chalasani, S. H., Chronis, N., Tsunozaki, M., Gray, J. M., Ramot, D., Goodman, M. B., et al. (2007). Dissecting a circuit for olfactory behaviour in *Caenorhabditis elegans*. *Nature* 450, 63–70. doi: 10.1038/nature06292

Chalasani, S. H., Kato, S., Albrecht, D. R., Nakagawa, T., Abbott, L. F., and Bargmann, C. I. (2010). Neuropeptide feedback modifies odor-evoked dynamics in *Caenorhabditis elegans* olfactory neurons. *Nat. Neurosci.* 13, 615–621. doi: 10.1038/nn.2526

Chandra, S. B., Hosler, J. S., and Smith, B. H. (2000). Heritable variation for latent inhibition and its correlation with reversal learning in honeybees

(*Apis mellifera*). *J. Compar. Psychol.* 114, 86–97. doi: 10.1037/0735-7036.114.1.86

Chandra, S. B., Wright, G. A., and Smith, B. H. (2010). Latent inhibition in the honey bee, *Apis mellifera*: Is it a unitary phenomenon? *Anim. Cogn.* 13, 805–815. doi: 10.1007/s10071-010-0329-6

Chao, M. Y., Komatsu, H., Fukuto, H. S., Dionne, H. M., and Hart, A. C. (2004). Feeding status and serotonin rapidly and reversibly modulate a *Caenorhabditis elegans* chemosensory circuit. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15512–15517. doi: 10.1073/pnas.0403369101

Chapman, C. A., Chapman, L. J., Struhsaker, T. T., Zanne, A. E., Clark, C. J., and Poulsen, J. R. (2004). A long-term evaluation of fruiting phenology: importance of climate change. *J. Trop. Ecol.* 21, 1–14. doi: 10.1017/s026646740400 1993

Chapman, C. A., Wrangham, R. W., Chapman, L. J., Kennard, D. K., and Zanne, A. E. (1999). Fruit and flower phenology at two sites in Kibale National Park, Uganda. *J. Trop. Ecol.* 15, 189–211. doi: 10.1017/s0266467499000759

Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cereb. Cortex* 18, 2844–2854. doi: 10.1093/cercor/bhn042

Chen, Z., Zhou, C., Meyer, M., Xiang, K., Schiffbauer, J. D., Yuan, X., et al. (2013). Trace fossil evidence for *Ediacaran bilaterian* animals with complex behaviors. *Precamb. Res.* 224, 690–701. doi: 10.1016/j.precamres.2012.11.004

Cheney, D. L., and Seyfarth, R. M. (1996). *How Monkeys See The World: Inside the Mind of Another Species*. Chicago: University of Chicago Press.

Cheng, K. (2012). How to navigate without maps: the power of taxon-like navigation in ants. *Compar. Cogn. Behav. Rev.* 7, 1–22.

Cheng, R.-K., Jesuthasan, S. J., and Penney, T. B. (2014). Zebrafish forebrain and temporal conditioning. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20120462. doi: 10.1098/rstb.2012.0462

Cheung, A. F., Pollen, A. A., Tavare, A., DeProto, J., and Molnár, Z. (2007). Comparative aspects of cortical neurogenesis in vertebrates. *J. Anat.* 211, 164–176. doi: 10.1111/j.1469-7580.2007.00769.x

Cheung, T. H. C., and Cardinal, R. N. (2005). Hippocampal lesions facilitate instrumental learning with delayed reinforcement but induce impulsive choice in rats. *BMC Neurosci.* 6:36. doi: 10.1186/1471-2202-6-36

Chiandetti, C., and Vallortigara, G. (2011). Chicks like consonant music. *Psychol. Sci.* 22, 1270–1273. doi: 10.1177/0956797611418244

Choi, C. Q. (2009). *How Monkeys Teach Tool Use. Scientific American*. Available online at: https://www.scientificamerican.com/article/monkey-education/ (accessed March 23, 2021).

Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos. Trans. R. Soc. B Biol. Sci.* 362, 1585–1599. doi: 10.1098/rstb.2007.2054

Cisek, P. (2019). Resynthesizing behavior through phylogenetic refinement. *Attent. Percept. Psychophys.* 81, 2265–2287. doi: 10.3758/s13414-019-01760-1

Cisek, P., and Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annu. Rev. Neurosci.* 33, 269–298. doi: 10.1146/annurev.neuro.051508.135409

Clark, R. E., Broadbent, N. J., and Squire, L. R. (2005). Impaired remote spatial memory after hippocampal lesions despite extensive training beginning early in life. *Hippocampus* 15, 340–346. doi: 10.1002/hipo.20076

Cole, M., and Wahlsten, D. (1968). Response-contingent CS termination as a factor in avoidance conditioning. *Psychonom. Sci.* 12, 15–16. doi: 10.3758/bf033 31165

Colwill, R. M., Absher, R. A., and Roberts, M. L. (1988). Context-US learning in Aplysia californica. *J. Neurosci.* 8, 4434–4439. doi: 10.1523/jneurosci.08-12-04434.1988

Cook, M., and Mineka, S. (1987). Second-order conditioning and overshadowing in the observational conditioning of fear in monkeys. *Behav. Res. Ther.* 25, 349–364. doi: 10.1016/0005-7967(87)90013-1

Cope, A. J., Vasilaki, E., Minors, D., Sabo, C., Marshall, J. A. R., and Barron, A. B. (2018). Abstract concept learning in a simple neural network inspired by the insect brain. *PLoS Comput. Biol.* 14:e1006435. doi: 10.1371/journal.pcbi.1006435

Coricelli, G., Critchley, H. D., Joffily, M., O'Doherty, J. P., Sirigu, A., and Dolan, R. J. (2005). Regret and its avoidance: a neuroimaging study of choice behavior. *Nat. Neurosci.* 8, 1255–1262. doi: 10.1038/nn1514

Coricelli, G., Dolan, R. J., and Sirigu, A. (2007). Brain, emotion and decision making: the paradigmatic example of regret. *Trends Cogn. Sci.* 11, 258–265. doi: 10.1016/j.tics.2007.04.003

Couvillon, P. A., and Bitterman, M. E. (1989). Reciprocal overshadowing in the discrimination of color-odor compounds by honeybees: further tests of a continuity model. *Anim. Learn. Behav.* 17, 213–222. doi: 10.3758/bf03207637

Couvillon, P. A., Arakaki, L., and Bitterman, M. E. (1997). Intramodal blocking in honeybees. *Anim. Learn. Behav.* 25, 277–282. doi: 10.3758/bf03199085

Craig, D. P., Varnon, C. A., Sokolowski, M. B., Wells, H., and Abramson, C. I. (2014). An assessment of fixed interval timing in free-flying honey bees (*Apis mellifera* ligustica): an analysis of individual performance. *PLoS One* 9:e0101262. doi: 10.1371/journal.pone.0101262

Cross, I. (1999). "Is music the most important thing we ever did?," in *Music, Mind and Science*, ed. S. W. Yi (Seoul: Seoul National University Press), 10–39.

Cross, I. (2009). "Music as a communicative medium," in *Prehistory of Language*, Vol. 1, eds R. Botha and C. Knight (Oxford: Oxford University Press), 113–144.

Cross, I., and Morley, I. (2008). "The evolution of music: theories, definitions and nature of the evidence," in *Communicative Musicality*, eds S. Malloch and C. Trevarthen (Oxford: Oxford University Press), 61–82.

Crystal, J. D. (2013). Remembering the past and planning for the future in rats. *Behav. Process.* 93, 39–49. doi: 10.1016/j.beproc.2012.11.014

Cynx, J., and Shapiro, M. (1986). Perception of missing fundamental by a species of songbird (*Sturnus vulgaris*). *J. Compar. Psychol.* 100, 356–360.

D'Amato, M. R. (1988). A search for tonal pattern perception in Cebus monkeys: why monkeys can't hum a tune. *Music Percept.* 5, 453–480. doi: 10.2307/40285410

Davidson, P. S., Drouin, H., Kwan, D., Moscovitch, M., and Rosenbaum, R. S. (2012). Memory as social glue: close interpersonal relationships in amnesic patients. *Front. Psychol.* 3:531. doi: 10.3389/fpsyg.2012.00531

Davis, K., Cheong, M. C., Park, J. S., and You, Y.-J. (2017). "Appetite control in *C. elegans*," in *Appetite and Food Intake: Central Control*, 2nd Edn, (Boca Raton, FL: CRC Press/Taylor & Francis), 1–16. Available online at: https://www.ncbi.nlm.nih.gov/books/NBK453151/

Deacon, T. W. (1997). *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York, NY: Norton.

Dean, L. G., Vale, G. L., Laland, K. N., Flynn, E., and Kendal, R. L. (2013). Human cumulative culture: a comparative perspective. *Biol. Rev.* 89, 284–301. doi: 10.1111/brv.12053

Dehaene, S., Duhamel, J. R., Hauser, M. D., Rizzolatti, G., Lorincz, E. N., et al. (2005). "Do monkeys understand the actions and minds of others? Studies of single cells and eye movements," in *From Monkey Brain to Human Brain*, eds S. Dehaene, J. R. Duhamel, M. D. Hauser, and G. Rizzolatti (Cambridge: MIT Press), 190–210.

DeWitt, I., and Rauschecker, J. P. (2013). Wernicke's area revisited: parallel streams and word processing. *Brain Lang.* 127, 181–191. doi: 10.1016/j.bandl.2013.09.014

Dindo, M., Whiten, A., and de Waal, F. B. (2009). In-group conformity sustains different foraging traditions in capuchin monkeys (*Cebus apella*). *PLoS One* 4:e7858. doi: 10.1371/journal.pone.0007858

Dodell-Feder, D., Koster-Hale, J., Bedny, M., and Saxe, R. (2011). fMRI item analysis in a theory of mind task. *Neuroimage* 55, 705–712. doi: 10.1016/j.neuroimage.2010.12.040

Drew, M. R., Zupan, B., Cooke, A., Couvillon, P. A., and Balsam, P. D. (2005). Temporal control of conditioned responding in goldfish. *J. Exper. Psychol. Anim. Behav. Process.* 31, 31–39. doi: 10.1037/0097-7403.31.1.31

Dubnau, J., Chiang, A.-S., and Tully, T. (2002). Neural substrates of memory: from synapse to system. *J. Neurobiol.* 54, 238–253. doi: 10.1002/neu.10170

Dugas-Ford, J., Rowell, J. J., and Ragsdale, C. W. (2012). Cell-type homologies and the origins of the neocortex. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16974–16979. doi: 10.1073/pnas.1204773109

Dunbar, R. I., and Shultz, S. (2017). Why are there so many explanations for primate brain evolution? *Philos. Trans. R. Soc. B Biol. Sci.* 372:20160244. doi: 10.1098/rstb.2016.0244

Durán, E., Ocaña, F. M., Broglio, C., Rodríguez, F., and Salas, C. (2010). Lateral but not medial *Telencephalic pallium* ablation impairs the use of goldfish spatial allocentric strategies in a "hole-board" task. *Behav. Brain Res.* 214, 480–487. doi: 10.1016/j.bbr.2010.06.010

Dylla, K. V., Galili, D. S., Szyszka, P., and Lüdke, A. (2013). Trace conditioning in insects— keep the trace! *Front. Physiol.* 4:67. doi: 10.3389/fphys.2013.00067

Eichenbaum, H., and Fortin, N. J. (2009). The neurobiology of memory based predictions. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 364, 1183–1191. doi: 10.1098/rstb.2008.0306

Eichenbaum, H., Yonelinas, A. P., and Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci.* 30, 123–152. doi: 10.1146/annurev.neuro.30.051606.094328

Emery, N. J. (2004). The mentality of crows: convergent evolution of intelligence in Corvids and Apes. *Science* 306, 1903–1907. doi: 10.1126/science.1098410

Ergorul, C., and Eichenbaum, H. (2007). The hippocampus and memory for "what," "where," and "when". *Learn. Mem.* 11, 397–405. doi: 10.1101/lm.73304

Everding, G. (2016). *Wild Chimpanzee Mothers Teach Young to Use Tools, Video Study Confirms: The Source: Washington University in St. Louis. The Source*. Available online at: https://source.wustl.edu/2016/10/wild-chimpanzee-mothers-teach-young-use-tools-video-study-confirms/ (accessed March 23, 2021).

Ewer, R. F., and Fox, H. M. (2009). On the functions and mode of action of the nematocysts of hydra. *Proc. Zool. Soc. Lond.* 117, 365–376. doi: 10.1111/j.1096-3642.1947.tb00524.x

Farooqui, T., Robinson, K., Vaessin, H., and Smith, B. H. (2003). Modulation of Early olfactory processing by an octopaminergic reinforcement pathway in the honeybee. *J. Neurosci.* 23, 5370–5380. doi: 10.1523/jneurosci.23-12-05370.2003

Farris, S. M. (2008). Evolutionary convergence of higher brain centers spanning the protostome-deuterostome boundary. *Brain Behav. Evol.* 72, 106–122. doi: 10.1159/000151471

Ferguson, H. J., Cobey, S., and Smith, B. H. (2001). Sensitivity to a change in reward is heritable in the honeybee, *Apis mellifera*. *Anim. Behav.* 61, 527–534. doi: 10.1006/anbe.2000.1635

Fernández, V. M., Arenas, A., and Farina, W. M. (2009). Volatile exposure within the honeybee hive and its effect on olfactory discrimination. *J. Compar. Physiol. A* 195, 759–768. doi: 10.1007/s00359-009-0453-4

Fernandez, V. M., Giurfa, M., Devaud, J.-M., and Farina, W. M. (2012). Latent inhibition in an insect: the role of aminergic signaling. *Learn. Mem.* 19, 593–597. doi: 10.1101/lm.028167.112

Ferrucci, L., Nougaret, S., and Genovesio, A. (2019). Macaque monkeys learn by observation in the ghost display condition in the object-in-place task with differential reward to the observer. *Sci. Rep.* 9:401. doi: 10.1038/s41598-018-36803-4

Figner, B., Knoch, D., Johnson, E. J., Krosch, A. R., Lisanby, S. H., Fehr, E., et al. (2010). Lateral prefrontal cortex and self-control in intertemporal choice. *Nat. Neurosci.* 13, 538–539. doi: 10.1038/nn.2516

Fitch, W. M. (1971). Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Zool.* 20:406. doi: 10.2307/2412116

Fitch, W. T. (2009). Biology of music: another one bites the dust. *Curr. Biol.* 19, R403–R404. doi: 10.1016/j.cub.2009.04.004

Fitch, W. T. (2018). The biology and evolution of speech: a comparative analysis. *Annu. Rev. Linguist.* 4, 255–279. doi: 10.1146/annurev-linguistics-011817-045748

Fogassi, L. (2005). Parietal lobe: from action organization to intention understanding. *Science* 308, 662–667. doi: 10.1126/science.1106138

Foo, S. A., Liddell, L., Grossman, A., and Caldeira, K. (2019). Photo-movement in the sea anemone Aiptasia influenced by light quality and symbiotic association. *Coral Reefs* 39, 47–54. doi: 10.1007/s00338-019-01866-w

Fossette, S., Gleiss, A. C., Chalumeau, J., Bastian, T., Armstrong, C. D., Vandenabeele, S., et al. (2015). Current-oriented swimming by jellyfish and its role in bloom maintenance. *Curr. Biol.* 25, 342–347. doi: 10.1016/j.cub.2014.11.050

Fotowat, H., Lee, C., Jun, J. J., and Maler, L. (2019). Neural activity in a hippocampus-like region of the teleost pallium is associated with active sensing and navigation. *eLife* 8:e044119. doi: 10.7554/elife.44119

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiol. Rev.* 91, 1357–1392. doi: 10.1152/physrev.00006.2011

Fröbisch, J., and Reisz, R. R. (2009). The late permian herbivore suminia and the early evolution of arboreality in terrestrial vertebrate ecosystems. *Proc. R. Soc. B Biol. Sci.* 276, 3611–3618. doi: 10.1098/rspb.2009.0911

Fugazza, C., Pongrácz, P., Pogány, Á, Lenkei, R., and Miklósi, Á (2020). Mental representation and episodic-like memory of own actions in dogs. *Sci. Rep.* 10:10449. doi: 10.1038/s41598-020-67302-0

Fuster, J. M. (2001). The prefrontal cortex–an update: time is of the essence. *Neuron* 30, 319–333. doi: 10.1016/s0896-6273(01)00285-9

Fyhn, M., Molden, S., Witter, M. P., Moser, E. I., and Moser, M.-B. (2004). Spatial representation in the entorhinal cortex. *Science* 305, 1258–1264.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119(Pt 2), 593–609. doi: 10.1093/brain/119.2.593

Gallistel, C. R., King, A., and McDonald, R. (2004). Sources of variability and systematic error in mouse timing behavior. *J. Exper. Psychol. Anim. Behav. Process.* 30, 3–16. doi: 10.1037/0097-7403.30.1.3

Gardner, R. A., and Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science* 165, 664–672. doi: 10.1126/science.165.3894.664

Gardner, R. S., Uttaro, M. R., Fleming, S. E., Suarez, D. F., Ascoli, G. A., and Dumas, T. C. (2013). A secondary working memory challenge preserves primary place strategies despite overtraining. *Learn. Mem.* 20, 648–656. doi: 10.1101/lm.031336.113

Garm, A., O'Connor, M., Parkefelt, L., and Nilsson, D.-E. (2007). Visually guided obstacle avoidance in the box jellyfish *Tripedalia cystophora* and *Chiropsella bronzie*. *J. Exper. Biol.* 210, 3616–3623. doi: 10.1242/jeb.004044

Garrity, P. A., Goodman, M. B., Samuel, A. D., and Sengupta, P. (2010). Running hot and cold: behavioral strategies, neural circuits, and the molecular machinery for thermotaxis in *C. elegans* and *Drosophila*. *Genes Dev.* 24, 2365–2382. doi: 10.1101/gad.1953710

Geiser, E., Notter, M., and Gabrieli, J. D. (2012). A corticostriatal neural system enhances auditory perception through temporal context processing. *J. Neurosci.* 32, 6177–6182. doi: 10.1523/jneurosci.5153-11.2012

Gerkema, M. P., Davies, W. I., Foster, R. G., Menaker, M., and Hut, R. A. (2013). The nocturnal bottleneck and the evolution of activity patterns in mammals. *Proc. R. Soc. B Biol. Sci.* 280:20130508. doi: 10.1098/rspb.2013.0508

Giassi, A. C. C., Harvey-Girard, E., Valsamis, B., and Maler, L. (2012). Organization of the gymnotiform fish pallium in relation to learning and memory: I. Cytoarchitectonics and cellular morphology. *J. Compar. Neurol.* 520, 3314–3337. doi: 10.1002/cne.23097

Gilovich, T., and Medvec, V. H. (1995). The experience of regret: what, when, and why. *Psychol. Rev.* 102, 379–395. doi: 10.1037/0033-295x.102.2.379

Ginsburg, S., and Jablonka, E. (2019). *The Evolution of the Sensitive Soul: Learning and the Origins of Consciousness*. Cambridge, MA: The MIT Press.

Ginsburg, S., and Jablonka, E. (2021). Evolutionary transitions in learning and cognition. *Philos. Trans. R. Soc. B Biol. Sci.* 376:20190766. doi: 10.1098/rstb.2019.0766

Glanzman, D. L. (2010). Common mechanisms of synaptic plasticity in vertebrates and invertebrates. *Curr. Biol.* 20:23. doi: 10.1016/j.cub.2009.10.023

Goffinet, A. M., Daumerie, C., Langerwerf, B., and Pieau, C. (1986). Neurogenesis in reptilian cortical structures:3H-thymidine autoradiographic analysis. *J. Compar. Neurol.* 243, 106–116. doi: 10.1002/cne.902430109

Goldman-Rakic, P. S. (1996). Regional and cellular fractionation of working memory. *Proc. Natl. Acad. Sci. U.S.A.* 93, 13473–13480. doi: 10.1073/pnas.93.24.13473

Gómez-Laplaza, L. M., and Gerlai, R. (2010). Latent learning in zebrafish (*Danio rerio*). *Behav. Brain Res.* 208, 509–515. doi: 10.1016/j.bbr.2009.12.031

Goodroe, S. C., Starnes, J., and Brown, T. I. (2018). The complex nature of hippocampal-striatal interactions in spatial navigation. *Front. Hum. Neurosci.* 12:250. doi: 10.3389/fnhum.2018.00250

Goucha, T., Zaccarella, E., and Friederici, A. D. (2017). A revival of homo loquens as a builder of labeled structures: neurocognitive considerations. *Neurosci. Biobehav. Rev.* 81, 213–224. doi: 10.1016/j.neubiorev.2017.01.036

Grafton, S. T., Hazeltine, E., and Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *J. Cogn. Neurosci.* 7, 497–510. doi: 10.1162/jocn.1995.7.4.497

Grahn, J. A., and Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *J. Cogn. Neurosci.* 19, 893–906. doi: 10.1162/jocn.2007.19.5.893

Grahn, J. A., and Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions inmusicians and nonmusicians during beat perception. *J. Neurosci.* 29, 7540–7548. doi: 10.1523/jneurosci.2018-08.2009

Granon, S., and Poucet, B. (1995). Medial prefrontal lesions in the rat and spatial navigation: Evidence for impaired planning. *Behav. Neurosci.* 109, 474–484. doi: 10.1037/0735-7044.109.3.474

Greenfield, P. M. (1991). Language, tools and brain: the ontogeny and phylogeny of hierarchically organized sequential behavior. *Behav. Brain Sci.* 14, 577–595.

Gribova, A., Donchin, O., Bergman, H., Vaadia, E., and de Oliveira, S. C. (2002). Timing of bimanual movements in human and nonhuman primates in relation to neuronal activity in primary motor cortex and supplementary motor area. *Expe. Brain Res.* 146, 322–335. doi: 10.1007/s00221-002-1174-x

Griffith, R. W. (1994). The life of the first vertebrates. *Bioscience* 44, 408–417. doi: 10.2307/1312364

Grillner, S., and Robertson, B. (2016). The basal ganglia over 500 Million years. *Curr. Biol.* 26, R1088–R1100. doi: 10.1016/j.cub.2016.06.041

Grossmann, K. E. (1973). Continuous, fixed-ratio, and fixed-interval reinforcement in honey bees. *J. Exper. Analys. Behav.* 20, 105–109. doi: 10.1901/jeab.1973.20-105

Grube, M., Cooper, F. E., Chinnery, P. F., and Griffiths, T. D. (2010). Dissociation of duration-based and beat-based auditory timing in cerebellar degeneration. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11597–11601. doi: 10.1073/pnas.0910473107

Guillermin, M. L., Carrillo, M. A., and Hallem, E. A. (2017). A single set of interneurons drives opposite behaviors in *C. elegans*. *Curr. Biol.* 27, 2630–2639. doi: 10.1016/j.cub.2017.07.023

Gunhold, T., Whiten, A., and Bugnyar, T. (2014). Video demonstrations seed alternative problem- solving techniques in wild common marmosets. *Biol. Lette.* 10:20140439. doi: 10.1098/rsbl.2014.0439

Gupta, A. S., van der Meer, M. A., Touretzky, D. S., and Redish, A. D. (2012). Segmentation of spatial experience by hippocampal θ sequences. *Nat. Neurosci.* 15, 1032–1039. doi: 10.1038/nn.3138

Gupta, S., Maurya, R., Saxena, M., and Sen, J. (2012). Defining structural homology between the mammalian and avian hippocampus through conserved gene expression patterns observed in the chick embryo. *Dev. Biol.* 366, 125–141. doi: 10.1016/j.ydbio.2012.03.027

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., and Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature* 436, 801–806.

Hamilton, T. J., Myggland, A., Duperreault, E., May, Z., Gallup, J., Powell, R. A., et al. (2016). Episodic-like memory in zebrafish. *Anim. Cogn.* 19, 1071–1079. doi: 10.1007/s10071-016-1014-1

Hammer, M., and Menzel, R. (1998). Multiple sites of associative odor learning as revealed by local brain microinjections of octopamine in honeybees. *Learn. Mem.* 5, 146–156.

Haralson, J. (1975). Classical conditioning in the sea anemone, *Cribrina xanthogrammica*. *Physiol. Behav.* 15, 455–460. doi: 10.1016/0031-9384(75)90214-0

Hare, T. A., Camerer, C. F., and Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 324, 646–648. doi: 10.1126/science.1168450

Hare, T. A., Camerer, C. F., Knoepfle, D. T., O'Doherty, J. P., and Rangel, A. (2010). Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *J. Neurosci.* 30, 583–590. doi: 10.1523/jneurosci.4089-09.2010

Harrington, D. L., Zimbelman, J. L., Hinton, S. C., and Rao, S. M. (2009). Neural modulation of temporal encoding, maintenance, and decision processes. *Cereb. Cortex* 20, 1274–1285. doi: 10.1093/cercor/bhp194

Harris, V. A. (1990). "Hydroids," in *Sessile Animals of the Sea Shore*, ed. V. A. Harris (London: Chapman and Hall), 223.

Haslam, M., Luncz, L. V., Staff, R. A., Bradshaw, F., Ottoni, E. B., and Falótico, T. (2016). Pre Columbian monkey tools. *Curr. Biol.* 26, R521–R522. doi: 10.1016/j.cub.2016.05.046

Haug, G. (1933). Die Lichtreaktionen der Hydren (Chlorohydra viridissima und Pelmatohydra oligactis). *Z. Vergleich. Physiol.* 19, 246–303.

Hawkins, R. D., and Byrne, J. H. (2015). Associative learning in invertebrates. *Cold Spring Harb. Perspect. Biol.* 7:a021709. doi: 10.1101/cshperspect.a021709

Hawkins, R. D., and Kandel, E. R. (1984). Is there a cell-biological alphabet for simple forms of learning? *Psychol. Rev.* 91, 375–391. doi: 10.1037/0033-295x.91.3.375

Hawkins, R. D., Greene, W., and Kandel, E. R. (1998). Classical conditioning, differential conditioning, and second-order conditioning of the Aplysia gill-withdrawal reflex in a simplified mantle organ preparation. *Behav. Neurosci.* 112, 636–645. doi: 10.1037/0735-7044.112.3.636

Hawkins, R. D., Lalevic, N., Clark, G. A., and Kandel, E. R. (1989). Classical conditioning of the Aplysia siphon-withdrawal reflex exhibits response specificity. *Proc. Natl. Acad. Sci. U.S.A.* 86, 7620–7624. doi: 10.1073/pnas.86.19.7620

Hayashi, T., Akikawa, R., Kawasaki, K., Egawa, J., Minamimoto, T., Kobayashi, K., et al. (2020). macaques exhibit implicit gaze bias anticipating others' False-belief-driven actions via medial prefrontal cortex. *Cell Rep.* 30, 4433–4444. doi: 10.1016/j.celrep.2020.03.013

Hays, G. C., Bastian, T., Doyle, T. K., Fossette, S., Gleiss, A. C., Gravenor, M. B., et al. (2011). High activity and Lévy searches: jellyfish can search the water column like fish. *Proc. R. Soc. B Biologic. Sci.* 279, 465–473. doi: 10.1098/rspb.2011.0978

Heffner, H., and Whitfield, I. C. (1976). Perception of the missing fundamental by cats. *J. Acoustic. Soc. Am.* 59, 915–919. doi: 10.1121/1.380951

Hinde, R. T. (1998). "The Cnidaria and Ctenophora," in *Invertebrate Zoology*, ed. D. T. Anderson (Oxford: Oxford University Press), 28–57.

Hobert, O. (2002). Behavioral plasticity in *C. elegans*: paradigms, circuits, genes. *J. Neurobiol.* 54, 203–223. doi: 10.1002/neu.10168

Honing, H. (2013). "Structure and interpretation of rhythm in music," in *The Psychology of Music*, ed. D. Deutsch (Amsterdam: Elsevier Academic Press), 369–404. doi: 10.1016/b978-0-12-381460-9.00009-2

Honing, H., Merchant, H., Háden, G. P., Prado, L., and Bartolo, R. (2012). Rhesus monkeys (*Macaca mulatta*) detect rhythmic groups in music, but not the beat. *PLoS One* 7:e051369. doi: 10.1371/journal.pone.0051369

Hoppitt, W. J. E., Brown, G. R., Kendal, R., Rendell, L., Thornton, A., Webster, M. M., et al. (2008). Lessons from animal teaching. *Trends Ecol. Evol.* 23, 486–493.

Hough, G. E., and Bingman, V. P. (2004). Spatial response properties of homing pigeon hippocampal neurons: correlations with goal locations, movement between goals, and environmental context in a radial-arm arena. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 190, 1047–1062.

Hulse, S. H., Cynx, J., and Humpal, J. (1984). Absolute and relative pitch discrimination in serial pitch perception by birds. *J. Exper. Psychol. Gen.* 113, 38–54. doi: 10.1037/0096-3445.113.1.38

Illich, P. A., Salinas, J. A., and Grau, J. W. (1994). Latent inhibition, overshadowing, and blocking of a conditioned antinociceptive response in spinalized rats. *Behav. Neural Biol.* 62, 140–150. doi: 10.1016/s0163-1047(05)80035-4

Inoue, T., Hoshino, H., Yamashita, T., Shimoyama, S., and Agata, K. (2015). Planarian shows decision-making behavior in response to multiple stimuli by integrative brain function. *Zool. Lett.* 1:7. doi: 10.1186/s40851-014-0010-z

Ishihara, T., Iino, Y., Mohri, A., Mori, I., Gengyo-Ando, K., Mitani, S., et al. (2002). HEN-1, a secretory protein with an LDL receptor motif, regulates sensory integration and learning in *Caenorhabditis elegans*. *Cell* 109, 639–649. doi: 10.1016/s0092-8674(02)00748-1

Ito, H. T., Zhang, S. J., Witter, M. P., Moser, E. I., and Moser, M. B. (2015). A prefrontal-thalamo-hippocampal circuit for goal-directed spatial navigation. *Nature* 522, 50–55. doi: 10.1038/nature14396

Ito, R., Robbins, T. W., Pennartz, C. M., and Everitt, B. J. (2008). Functional interaction between the hippocampus and nucleus Accumbens shell is necessary for the acquisition of appetitive spatial context conditioning. *J. Neurosci.* 28, 6950–6959. doi: 10.1523/jneurosci.1615-08.2008

Iversen, J. R., and Patel, A. D. (2008). "The Beat Alignment Test (BAT): surveying beat processing abilities in the general population," in *Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC10)*, (New York, NY), 465–468.

Ivry, R. B., and Spencer, R. M. C. (2004). The neural representation of time. *Curr. Opin. Neurobiol.* 14, 225–232. doi: 10.1016/j.conb.2004.03.013

Jäncke, L., Loose, R., Lutz, K., Specht, K., and Shah, N. J. (2000). Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. *Cogn. Brain Res.* 10, 51–66. doi: 10.1016/s0926-6410(00)00022-7

Janmaat, K. R., Polansky, L., Ban, S. D., and Boesch, C. (2014). Wild chimpanzees plan their breakfast time, type, and location. *Proc. Natl. Acad. Sci. U.S.A.* 111, 16343–16348. doi: 10.1073/pnas.1407524111

Jarvis, E. D. (2019). Evolution of vocal learning and spoken language. *Science* 366, 50–54. doi: 10.1126/science.aax0287

Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nat. Rev. Neurosci.* 6, 151–159. doi: 10.1038/nrn1606

Jellema, T., Baker, C. I., Wicker, B., and Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain Cogn.* 44, 280–302. doi: 10.1006/brcg.2000.1231

Jerison, H. J. (1973). *Evolution of the Brain and Intelligence*. New York, NY: Academic Press.

Johnson, A., and Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J. Neurosci.* 27, 12176–12189. doi: 10.1523/JNEUROSCI.3761-07.2007

Jones, A. W., and Levi-Montalcini, R. (1958). Patterns of differentiation of the nerve centers and fiber tracts of the avian cerebral hemispheres. *Arch. Ital Biol.* 96, 231–284.

Jürgens, U. (1988). "Central control of monkey calls," in *Primate Vocal Communication*, eds D. Todt, P. Goedeking, and D. Symmes (Berlin: Springer), 162–167. doi: 10.1007/978-3-642-73769-5_12

Jürgens, U., Kirzinger, A., and von Cramon, D. (1982). The effects of deep-reaching lesions in the cortical face area on phonation a combined case report and experimental monkey study. *Cortex* 18, 125–139. doi: 10.1016/s0010-9452(82)80024-5

Kaas, J. H. (ed.) (2009). *Evolutionary Neuroscience*. Cambridge: Academic Press, 815–817.

Kahn, M. C., Siegel, J. J., Jechura, T. J., and Bingman, V. P. (2008). Response properties of avian hippocampal formation cells in an environment with unstable goal locations. *Behav. Brain Res.* 191, 153–163.

Kamin, L. J. (1957). The effects of termination of the CS and avoidance of the US on avoidance learning: an extension. *Can. J. Psychol.* 11, 48–56. doi: 10.1037/h0083690

Kamin, L. J. (1968). "Attention-like" processes in classical conditioning," in *Proceedings of the Miami Symposium on the Prediction of Behavior, 1967: Aversive Stimulation*, ed. M. R. Jones (Coral Gables, FL: University of Miami Press), 9–31.

Kaminski, J., Call, J., and Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition* 109, 224–234. doi: 10.1016/j.cognition.2008.08.010

Kandel, E. R. (1979). *Behavioural Bio of Aplysia: Origin & Evolution*. New York, NY: W.H. Freeman & Company.

Kandel, E. R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science* 294, 1030–1038. doi: 10.1126/science.1067020

Kandel, E. R. (2006). *Search of Memory: the Emergence of a New Science of Mind*. New York, NY: W.W. Norton and Company.

Kano, F., Krupenye, C., Hirata, S., Tomonaga, M., and Call, J. (2019). Great apes use self-experience to anticipate an agent's action in a false-belief test. *Proc. Natl. Acad. Sci. U.S.A.* 116, 20904–20909. doi: 10.1073/pnas.1910095116

Karnik, I., and Gerlai, R. (2012). Can zebrafish learn spatial tasks? An empirical analysis of place and single CS-US associative learning. *Behav. Brain Res.* 233, 415–421. doi: 10.1016/j.bbr.2012.05.024

Karten, H. J. (1969). The organization of the avian telencephalon and some speculations on the phylogeny of the amniote telencephalon*. *Ann. N. Y. Acad. Sci.* 167, 164–179. doi: 10.1111/j.1749-6632.1969.tb20442.x

Karten, H. J. (1997). Evolutionary developmental biology meets the brain: the origins of mammalian cortex. *Proc. Natl. Acad. Sci. U.S.A.* 94, 2800–2804. doi: 10.1073/pnas.94.7.2800

Katie (2015). *Goldfish as a Model for Understanding Learning and Memory: More Complex Than You Think: AMNH*. New York, NY: American Museum of Natural History.

Kielan-Jaworowska, Z., Bown, T. M., and Lillegraven, J. A. (1979). "Eutheria," in *Mesozoic Mammals: The First Two-Thirds of Mammalian History*, eds J. A. Lillegraven and Z. Kielan-Jaworowska (Berkeley, CA: University of California Press), 221–228.

Killcross, S., and Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* 13, 400–408. doi: 10.1093/cercor/13.4.400

Kim, S., Hwang, J., and Lee, D. (2008). Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron* 59, 161–172. doi: 10.1016/j.neuron.2008.05.010

Kimata, T., Sasakura, H., Ohnishi, N., Nishio, N., and Mori, I. (2012). Thermotaxis of *C. elegans* as a model for temperature perception, neural information processing and neural plasticity. *Worm* 1, 31–41. doi: 10.4161/worm.19504

Kleniginna, P. R., and Currie, J. A. (1979). Effects of intermittent reinforcement in the Florida kingsnake. *Biol. Psychol.* 21, 14–16.

Kline, M. A. (2015). How to learn about teaching: an evolutionary framework for the study of teaching behavior in humans and other animals. *Beahv. Brain Sci.* 2015:e31.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., and Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314, 829–832. doi: 10.1126/science.1129156

Koechlin, E., and Jubault, T. (2006). Broca's Area and the hierarchical organization of human behavior. *Neuron* 50, 963–974. doi: 10.1016/j.neuron.2006.05.017

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297, 846–848. doi: 10.1126/science.1070311

Krupenye, C., Kano, F., Hirata, S., Call, J., and Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science* 354, 110–114. doi: 10.1126/science.aaf8110

Kung, S.-J., Chen, J. L., Zatorre, R. J., and Penhune, V. B. (2013). Interacting cortical and basal ganglia networks underlying finding and tapping to the musical beat. *J. Cogn. Neurosci.* 25, 401–420. doi: 10.1162/jocn_a_00325

Kusayama, T., and Watanabe, S. (2000). Reinforcing effects of methamphetamine in planarians. *Neuroreport* 11, 2511–2513. doi: 10.1097/00001756-200008030-00033

Laland, K. N. (2017). *Darwin's Unfinished Symphony: How Culture Made the Human Mind*. Princeton, NH: Princeton University Press.

Langmore, N. E. (1998). Functions of duet and solo songs of female birds. *Trends Ecol. Evol.* 13, 136–140. doi: 10.1016/s0169-5347(97)01241-x

Large, E. W. (2000). On synchronizing movements to music. *Hum. Mov. Sci.* 19, 527–566. doi: 10.1016/s0167-9457(00)00026-9

Large, E. W., and Jones, M. R. (1999). The dynamics of attending: how people track time- varying events. *Psychol. Rev.* 106, 119–159. doi: 10.1037/0033-295x.106.1.119

Larsch, J., Flavell, S. W., Liu, Q., Gordus, A., Albrecht, D. R., and Bargmann, C. I. (2015). A circuit for gradient climbing in *C. elegans* chemotaxis. *Cell Rep.* 12, 1748–1760. doi: 10.1016/j.celrep.2015.08.032

Laurent, E., and Lejeune, H. (1985). Temporal regulation of behavior in a fresh water turtle, *Pseudemys scripta elegans* (Wied). *Behav. Process.* 10, 159–160.

Lay, B. P. P., Westbrook, R. F., Glanzman, D. L., and Holmes, N. M. (2018). Commonalities and differences in the substrates underlying consolidation of first- and second-order conditioned fear. *J. Neurosci.* 38, 1926–1941. doi: 10.1523/jneurosci.2966-17.2018

Lee, A., Mathuru, A. S., The, C., Kibat, C., Korzh, V., Penney, T. B., et al. (2010). The habenula prevents helpless behavior in larval zebrafish. *Curr. Biol.* 20, 2211–2216.

Lehrer, M. (1997a). "Honeybees' visual spatial orientation at the feeding site," in *Orientation and Communication in Arthropods*, ed. M. Lehrer (Zürich: University of Zurich), 115–144.

Lehrer, M. (1997b). *Orientation and Communication in Arthropods*. Basel: Birkhäuser.

Lejeune, H., and Wearden, J. H. (1991). The comparative psychology of fixed-interval responding: Some quantitative analyses. *Learn. Motiv.* 22, 84–111. doi: 10.1016/0023-9690(91)90018-4

Lewis, T. (2014). *Like Humans, Rats Experience Regret, Study Suggests. LiveScience.* Available online at: https://www.livescience.com/46184-rats-experience-regret.html (accessed March 22, 2021).

Leys, S. P., and Degnan, B. M. (2001). Cytological basis of photoresponsive behavior in a sponge larva. *Biol. Bull.* 201, 323–338. doi: 10.2307/1543611

Leys, S., Cronin, T., Degnan, B., and Marshall, J. (2002). Spectral sensitivity in a sponge larva. *J. Compar. Physiol. A Sens. Neural Behav. Physiol.* 188, 199–202. doi: 10.1007/s00359-002-0293-y

Li, M. J. (2012). *Identification of an Operant Learning Circuit by Whole Brain Functional Imaging in Larval Zebrafish*. Doctoral thesis. Cambridge, MA: Harvard University.

Li, Z., Liu, J., Zheng, M., and Xu, X. Z. (2014). Encoding of both analog- and digital-like behavioral outputs by one *C. elegans* interneuron. *Cell* 159, 751–765. doi: 10.1016/j.cell.2014.09.056

Liebeskind, B. J., Hofmann, H. A., Hillis, D. M., and Zakon, H. H. (2017). Evolution of animal neural systems. *Annu. Rev. Ecol. Evol. Syst.* 48, 377–398. doi: 10.1146/annurev-ecolsys-110316-023048

Lindeyer, C. M., and Reader, S. M. (2010). Social learning of escape routes in zebrafish and the stability of behavioural traditions. *Anim. Behav.* 79, 827–834. doi: 10.1016/j.anbehav.2009.12.024

Liu, C., Plaçais, P.-Y., Yamagata, N., Pfeiffer, B. D., Aso, Y., Friedrich, A. B., et al. (2012). A subset of dopamine neurons signals reward for odour memory in *Drosophila*. *Nature* 488:512. doi: 10.1038/nature11304

Lofgren, D. L., Lillegraven, J. A., Clemens, W. A., Gingerich, P. D., and Williamson, T. E. (2004). "Paleocene Biochronology: the Puercan through Clarkforkian land mammal ages," in *Late Cretaceous and Cenozoic Mammals of North America*, ed. M. O. Woodburne (New York, NY: Columbia University Press), 43–105. doi: 10.7312/wood13040-005

López, J. C., Gómez, Y., Rodríguez, F., Broglio, C., Vargas, J. P., and Salas, C. (2001). Spatial learning in turtles. *Anim. Cogn.* 4, 49–59. doi: 10.1007/s100710100091

López, J. C., Vargas, J. P., Gómez, Y., and Salas, C. (2003). Spatial and non-spatial learning in turtles: the role of medial cortex. *Behav. Brain Res.* 143, 109–120. doi: 10.1016/S0166-4328(03)00030-5

Lopez-Barroso, D., Catani, M., Ripolles, P., Dell'Acqua, F., Rodriguez-Fornells, A., and de Diego-Balaguer, R. (2013). Word learning is mediated by the left *Arcuate fasciculus*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13168–13173. doi: 10.1073/pnas.1301696110

Loy, I., Fernández, V., and Acebes, F. (2006). Conditioning of tentacle lowering in the snail (*Helix aspersa*): acquisition, latent inhibition, overshadowing, second-order conditioning, and sensory preconditioning. *Learn. Behav.* 34, 305–314. doi: 10.3758/bf03192885

Lubow, R. E., and Moore, A. U. (1959). Latent inhibition: the effect of nonreinforced pre- exposure to the conditional stimulus. *J. Compar. Physiol. Psychol.* 52, 415–419. doi: 10.1037/h0046700

Lucon-Xiccato, T., and Bisazza, A. (2017). Complex maze learning by fish. *Anim. Behav.* 125, 69–75. doi: 10.1016/j.anbehav.2016.12.022

Luersen, K., Faust, U., Gottschling, D.-C., and Doring, F. (2014). Gait-specific adaptation of locomotor activity in response to dietary restriction in *Caenorhabditis elegans*. *J. Exper. Biol.* 217, 2480–2488. doi: 10.1242/jeb.099382

Luo, Z.-X., Meng, Q.-J., Ji, Q., Liu, D., Zhang, Y.-G., and Neander, A. I. (2015). Evolutionary development in basal mammaliaforms as revealed by a docodontan. *Science* 347, 760–764. doi: 10.1126/science.1260880

MacDonald, C. J., Lepage, K. Q., Eden, U. T., and Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron* 71, 737–749. doi: 10.1016/j.neuron.2011.07.012

Mackie, G. O. (2004). Central neural circuitry in the jellyfish *Aglantha*. *Neurosignals* 13, 5–19. doi: 10.1159/000076155

Malapani, C., Deweer, B., and Gibbon, J. (2002). Separating storage from retrieval dysfunction of temporal memory in Parkinson's disease. *J. Cogn. Neurosci.* 14, 311–322. doi: 10.1162/089892902317236920

Malapani, C., Rakitin, B., Levy, R., Meck, W. H., Deweer, B., Dubois, B., et al. (1998). Coupled temporal memories in Parkinson's Disease: a dopamine-related dysfunction. *J. Cogn. Neurosci.* 10, 316–331. doi: 10.1162/089892998562762

Mansouri, F. A., Koechlin, E., Rosa, M. G., and Buckley, M. J. (2017). Managing competing goals — a key role for the frontopolar cortex. *Nat. Rev. Neurosci.* 18, 645–657. doi: 10.1038/nrn.2017.111

Mars, R. B., Neubert, F. X., Noonan, M. P., Sallet, J., Toni, I., and Rushworth, M. F. (2012). On the relationship between the "default mode network" and the "social brain". *Front. Hum. Neurosci.* 6:189. doi: 10.3389/fnhum.2012.00189

Mars, R. B., Sallet, J., Neubert, F.-X., and Rushworth, M. F. (2013). Connectivity profiles reveal the relationship between brain areas for social cognition in human and monkey temporoparietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 110, 10806–10811. doi: 10.1073/pnas.1302956110

Mars, R. B., Sallet, J., Schuffelgen, U., Jbabdi, S., Toni, I., and Rushworth, M. F. (2011). Connectivity-Based subdivisions of the human right "Temporoparietal

Junction Area": evidence for different areas participating in different cortical networks. *Cereb. Cortex* 22, 1894–1903. doi: 10.1093/cercor/bhr268

Martin, V. C., Schacter, D. L., Corballis, M. C., and Addis, D. R. (2011). A role for the hippocampus in encoding simulations of future events. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13858–13863. doi: 10.1073/pnas.1105816108

Matanoski, J., Hood, R., and Purcell, J. (2001). Characterizing the effect of prey on swimming and feeding efficiency of the scyphomedusa *Chrysaora quinquecirrha*. *Mar. Biol.* 139, 191–200. doi: 10.1007/s002270100558

Matell, M. S., and Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *Bioessays* 22, 94–103.

Matell, M. S., and Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cogn. Brain Res.* 21, 139–170. doi: 10.1016/j.cogbrainres.2004.06.012

McClelland, J. L., and Goddard, N. H. (1996). Considerations arising from a complementary learning systems perspective on hippocampus and neocortex. *Hippocampus* 6, 654–665. doi: 10.1002/(sici)1098-106319966:6<654::aid-hipo8<3.0.co;2-g

McClure, S. M. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507.

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2007). Time discounting for primary rewards. *J. Neurosci.* 27, 5796–5804. doi: 10.1523/jneurosci.4246-06.2007

McClure, S. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507. doi: 10.1126/science.1100907

McDermott, J., and Hauser, M. (2004). Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. *Cognition* 94, B11–B21. doi: 10.1016/j.cognition.2004.04.004

McGaugh, J. L. (2013). Making lasting memories: remembering the significant. *Proc. Natl. Acad. Sci. U.S.A.* 110, 10402–10407. doi: 10.1073/pnas.1301209110

McKenzie, T., Cherman, T., Bird, L. R., Naqshbandi, M., and Roberts, W. A. (2004). Can squirrel monkeys (*Saimiri sciureus*) plan for the future? Studies of temporal myopia in food choice. *Learn. Behav.* 32, 377–390.

Meck, W. H., and Benson, A. M. (2002). Dissecting the brain's internal clock: how frontal-striatal circuitry keeps time and shifts attention. *Brain Cogn.* 48, 195–211.

Meck, W. H., Church, R. M., and Olton, D. S. (1984). Hippocampus, time, and memory. *Behav. Neurosci.* 98, 3–22. doi: 10.1037/0735-7044.98.1.3

Meck, W. H., Church, R. M., Wenk, G. L., and Olton, D. S. (1987). Nucleus basalis magnocellularis and medial septal area lesions differentially impair temporal memory. *J. Neurosci.* 7, 3505–3511. doi: 10.1523/jneurosci.07-11-03505.1987

Meck, W. H., Doyère, V., and Gruart, A. (2012). *Interval Timing and Time-Based Decision Making. Google Books*. Available online at: https://books.google.com/books?id=UqF8DwAAQBAJ (accessed March 21, 2021).

Melgire, M., Ragot, R., Samson, S., Penney, T. B., Meck, W. H., and Pouthas, V. (2005). Auditory/visual duration bisection in patients with left or right medial-temporal lobe resection. *Brain Cogn.* 58, 119–124. doi: 10.1016/j.bandc.2004.09.013

Melis, A. P., and Tomasello, M. (2013). Chimpanzees' (Pan troglodytes) strategic helping in a collaborative task. *Biol. Lett.* 9:20130009. doi: 10.1098/rsbl.2013.0009

Meltzoff, A. N., and Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science* 198, 74–78. doi: 10.1126/science.897687

Meltzoff, A. N., and Moore, M. K. (1989). Imitation in newborn infants: exploring the range of gestures imitated and the underlying mechanisms. *Dev. Psychol.* 25, 954–962. doi: 10.1037/0012-1649.25.6.954

Meng, Q.-J., Ji, Q., Zhang, Y.-G., Liu, D., Grossnickle, D. M., and Luo, Z.-X. (2015). An arboreal docodont from the Jurassic and mammaliaform ecological diversification. *Science* 347, 764–768. doi: 10.1126/science.1260879

Menzel, C. R. (1999). Unprompted recall and reporting of hidden objects by a chimpanzee (Pantroglodytes) after extended delays. *J. Compar. Psychol.* 113, 426–434. doi: 10.1037/0735-7036.113.4.426

Menzel, R., Greggers, U., Smith, A., Berger, S., Brandt, R., Brunke, S., et al. (2005). Honey bees navigate according to a map-like spatial memory. *Proc. Natl. Acad. Sci. U.S.A.* 102, 3040–3045. doi: 10.1073/pnas.0408550102

Menzel, R., Kirbach, A., Haass, W. D., Fischer, B., Fuchs, J., Koblofsky, M., et al. (2011). A common frame of reference for learned and communicated vectors in honeybee navigation. *Curr. Biol.* 21, 645–650. doi: 10.1016/j.cub.2011.02.039

Mercader, J., Barton, H., Gillespie, J., Harris, J., Kuhn, S., Tyler, R., et al. (2007). 4,300-year-old chimpanzee sites and the origins of percussive stone technology. *Proc. Natl. Acad. Sci. U.S.A.* 104, 3043–3048. doi: 10.1073/pnas.0607909104

Merchant, H. G., and Moore, J. W. (1973). Blocking of the rabbit's conditioned nictitating membrane response in Kamin's two-stage paradigm. *J. Exper. Psychol.* 101, 155–158. doi: 10.1037/h0035784

Merchant, H., and Honing, H. (2014). Are nonhuman primates capable of rhythmicentrainment? Evidence for the gradual audiomotor evolution hypothesis. *Front. Neurosci.* 7:274. doi: 10.3389/fnins.2013.00274

Merchant, H., Harrington, D., and Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Ann. Rev. Neurosci.* 36, 313–336. doi: 10.1146/annurev-neuro-062012-170349

Meunier, M., Monfardini, E., and Boussaoud, D. (2007). Learning by observation in rhesus monkeys. *Neurobiol. Learn. Mem.* 88, 243–248. doi: 10.1016/j.nlm.2007.04.015

Miles, H. L. (1983). "Apes and language: the search for communicative competence," in *Language in Primates* (New York, NY: Springer), 43–61. Available online at: https://link.springer.com/chapter/10.1007%2F978-1-4612-5496-6_3

Milton, K. (1981). Distribution Patterns of tropical plant foods as an evolutionary stimulus to primate mental development. *Am. Anthropol.* 83, 534–548. doi: 10.1525/aa.1981.83.3.02a00020

Milton, K. (1988). "Foraging behaviour and the evolution of primate intelligence," in *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*, eds R. W. Byrne and A. Whiten (Oxford: Clarendon Press), 285–305.

Mitchell, M., McCormick, M., Ferrari, M., and Chivers, D. (2011). Friend or foe? The role of latent inhibition in predator and non-predator labelling by coral reef fishes. *Anim. Cogn.* 14, 707–714. doi: 10.1007/s10071-011-0405-6

Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R., et al. (2000). Responding to joint attention across the 6- Through 24-month age period and early language acquisition. *J. Appl. Dev. Psychol.* 21, 283–298. doi: 10.1016/s0193-3973(99)00040-4

Morishima, Y., Schunk, D., Bruhin, A., Ruff, C. C., and Fehr, E. (2012). Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. *Neuron* 75, 73–79. doi: 10.1016/j.neuron.2012.05.021

Morley, I. (2003). *The Evolutionary Origins and Archeology of Music: and Investigation into the Prehistory of Human Musical Capacities and Behaviors*. Ph.D thesis. Cambridge, MA: University of Cambridge.

Moroz, L. L. (2015). Convergent evolution of neural systems in ctenophores. *J. Exper. Biol.* 218, 598–611. doi: 10.1242/jeb.110692

Moroz, L. L., and Kohn, A. B. (2016). Independent origins of neurons and synapses: insights from ctenophores. *Philos. Trans. R. Soc. B Biol. Sci.* 371:20150041. doi: 10.1098/rstb.2015.0041

Muenzinger, K. F., and Gentry, E. (1931). Tone discrimination in white rats. *J. Comp. Psychol.* 12, 195–206.

Mugan, U., and MacIver, M. A. (2020). Spatial planning with long visual range benefits escape from visual predators in complex naturalistic environments. *Nat. Commun.* 11:3057. doi: 10.1038/s41467-020-16102-1

Mulcahy, N. J., and Call, J. (2006). Apes save tools for future use. *Science* 312, 1038–1040. doi: 10.1126/science.1125456

Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., and Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Dev.* 78, 938–954. doi: 10.1111/j.1467-8624.2007.01042.x

Murray, E. A., Wise, S. P., and Graham, K. S. (2017). *The Evolution of Memory Systems: Ancestors, Anatomy, and Adaptations*. Oxford: Oxford University Press.

Musgrave, S., Morgan, D., Lonsdorf, E., Mundry, R., and Sanz, C. (2016). Tool transfers are a form of teaching among chimpanzees. *Sci. Rep.* 6:34783. doi: 10.1038/srep34783

Naqshbandi, M., and Roberts, W. A. (2006). Anticipation of future events in squirrel monkeys (*Saimiri sciureus*) and rats (*Rattus norvegicus*): Tests of the Bischof-Kohler hypothesis. *J. Compar. Psychol.* 120, 345–357. doi: 10.1037/0735-7036.120.4.345

Nicola, S. M., Yun, I. A., Wakabayashi, K. T., and Fields, H. L. (2004). Cue-evoked firing of nucleus accumbens neurons encodes motivational significance during a discriminative stimulus task. *J. Neurophysiol.* 91, 1840–1865. doi: 10.1152/jn.00657.2003

Nilsson, D. E. (2013). Eye evolution and its functional basis. *Vis. Neurosci.* 30, 5–20. doi: 10.1017/S0952523813000035

Ninokura, Y., Mushiake, H., and Tanji, J. (2003). Representation of the temporal order of visual objects in the primate lateral prefrontal cortex. *J. Neurophysiol.* 89, 2868–2873. doi: 10.1152/jn.00647.2002

O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Exp. Neurol.* 51, 78–109.

O'Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* 34, 171–175.

O'Craven, K. M., and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stiimulus-specific brain regions. *J. Cogn. Neurosci.* 12, 1013–1023. doi: 10.1162/08989290051137549

Ogura, A. (2004). Comparative analysis of gene expression for convergent evolution of camera eye between octopus and human. *Genome Res.* 14, 1555–1561. doi: 10.1101/gr.2268104

Ohyama, T., Gibbon, J., Deich, J. D., and Balsam, E. D. (1999). Temporal control during maintenance and extinction of conditioned keypecking in ring doves. *Anim. Learn. Behav.* 27, 89–98. doi: 10.3758/bf03199434

O'Neill, P. K., Gordon, J. A., and Sigurdsson, T. (2013). Theta oscillations in the medial prefrontal cortex are modulated by spatial working memory and synchronize with the hippocampus through its ventral subregion. *J. Neurosci.* 33, 14211–14224. doi: 10.1523/JNEUROSCI.2378-13.2013

Oprisan, S. A., Aft, T., Buhusi, M., and Buhusi, C. V. (2018). Scalar timing in memory: a temporal map in the hippocampus. *J. Theoret. Biol.* 438, 133–142. doi: 10.1016/j.jtbi.2017.11.012

Oprisan, S. A., and Buhusi, C. V. (2013). How noise contributes to time-scale invariance of interval timing. *Phys. Rev. E* 87:052717. doi: 10.1103/physreve.87.052717

Otsuka, Y., Osaka, N., Ikeda, T., and Osaka, M. (2009). Individual differences in the theory of mind and superior temporal sulcus. *Neurosci. Lett.* 463, 150–153. doi: 10.1016/j.neulet.2009.07.064

Padoa-Schioppa, C., and Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226. doi: 10.1038/nature04676

Pantin, C. F. A. (1935). The nerve net of the Actinozoa: II. Plan of the nerve net. *J. Exper. Biol.* 2, 139–155.

Parker, G. H. (1916). Locomotion of sea-anemones. *Proc. Natl. Acad. Sci. U.S.A.* 2, 449–450. doi: 10.1073/pnas.2.8.449

Patel, A. D., Iversen, J. R., Bregman, M. R., and Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Curr. Biol.* 19, 827–830. doi: 10.1016/j.cub.2009.03.038

Patterson, F. G. (1978). The gestures of a gorilla: language acquisition in another pongid. *Brain Lang.* 5, 72–97. doi: 10.1016/0093-934x(78)90008-1

Payne, H. L., Lynch, G. F., and Aronov, D. (2020). Precise spatial representations in the hippocampus of a food-caching bird. *bioRxiv* [Preprint]. doi: 10.1101/2020.11.27.399444

Pearl, R. (1903). The movements and reactions of fresh-water planarians: a study in animal behaviour / by Raymond Pearl. *J. Cell Sci.* s2–s46, 509–714. doi: 10.5962/bhl.title.11145

Pearse, V. B. (1974). Modification of sea anemone behavior by symbiotic zooxanthellae: expansion and contraction. *Biol. Bull.* 147, 641–651. doi: 10.2307/1540747

Pearson, J., Naselaris, T., Holmes, E. A., and Kosslyn, S. M. (2015). Mental imagery: functional mechanisms and clinical applications. *Trends Cogn. Sci.* 19, 590–602. doi: 10.1016/j.tics.2015.08.003

Penfield, W., and Rasmussen, T. (1950). *The Cerebral Cortex of Man; a Clinical Study of Localization of Function.* Basingstoke: Macmillan.

Perez, O., Kass, R. E., and Merchant, H. (2013). Trial time warping to discriminate stimulus-related from movement-related neural activity. *J. Neurosci. Methods* 212, 203–210. doi: 10.1016/j.jneumeth.2012.10.019

Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., et al. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *J. Exper. Biol.* 146, 87–113.

Perrett, D. I., Hietanen, J. K., Oram, M. W., and Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 335, 23–30. doi: 10.1098/rstb.1992.0003

Perrett, D. I., Smith, P. A. J., Mistlin, A. J., Chitty, A. J., Head, A. S., Potter, D. D., et al. (1985). Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: a preliminary report. *Behav. Brain Res.* 16, 153–170. doi: 10.1016/0166-4328(85)90089-0

Perry, C. J., Barron, A. B., and Cheng, K. (2013). Invertebrate learning and cognition: relating phenomena to neural substrate. *Wiley Interdiscipl. Rev. Cogn. Sci.* 4, 561–582. doi: 10.1002/wcs.1248

Petrides, M. (2014). *Neuroanatomy of Language Regions of the Human Brain.* London: Academic Press.

Petrides, M., and Pandya, D. N. (2009). Distinct parietal and temporal pathways to the homologues of Broca's Area in the Monkey. *PLoS Biol.* 7:e00170. doi: 10.1371/journal.pbio.1000170

Pezzulo, G., and Cisek, P. (2016). Navigating the affordance landscape: Feedback control as a process model of behavior and cognition. *Trends Cogn. Sci.* 20, 414–424.

Phillips, J. B., Adler, K., and Borland, S. C. (1995). True navigation by an amphibian. *Anim. Behav.* 50, 855–858. doi: 10.1016/0003-3472(95)80146-4

Pinker, S. (1994). *The Language Instinct: How the Mind Creates Language.* New York, NY: Perennial Classics.

Polyak, S. (1957). *The Vertebrate Visual System: its Origin, Structure, and Function and its Manifestations in Disease with an Analysis of its Role in the Life of Animals and in the Origin of Man.* Chicago & London: The University of Chicago Press.

Portavella, M. (2004). Avoidance response in goldfish: emotional and temporal involvement of medial and lateral *Telencephalic pallium. J. Neurosci.* 24, 2335–2342. doi: 10.1523/jneurosci.4930-03.2004

Poucet, B. (1989). Object exploration, habituation and response to a spatial change in rats following septal or medial frontal cortical damage. *Behav. Brain Res.* 41, 229–237.

Poucet, B. (1990). A further characterization of the spatial problem-solving deficit induced by lesions of the medial frontal cortex in the rat. *Behav. Brain Res.* 41, 229–237.

Prados, J., Alvarez, B., Acebes, F., Loy, I., Sansa, J., and Moreno-Fernández, M. M. (2013). Blocking in rats, humans and snails using a within-subjects design. *Behav. Process.* 100, 23–31. doi: 10.1016/j.beproc.2013.07.014

Prados, J., Alvarez, B., Howarth, J., Stewart, K., Gibson, C. L., Hutchinson, C. V., et al. (2012). Cue competition effects in the planarian. *Anim. Cogn.* 16, 177–186. doi: 10.1007/s10071-012-0561-3

Prather, J. F., Peters, S., Nowicki, S., and Mooney, R. (2008). Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature* 451, 305–310. doi: 10.1038/nature06492

Premack, D. (2007). Human and animal cognition: continuity and discontinuity. *Proc. Natl. Acad. Sci. U.S.A.* 104:13861.

Puce, A., and Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 358, 435–445. doi: 10.1098/rstb.2002.1221

Qin, J., and Wheeler, A. R. (2007). Maze exploration and learning in *C. elegans. Lab. Chip* 7, 186–192. doi: 10.1039/b613414a

Ragozzino, M. E., Detrick, S., and Kesner, R. P. (1999). Involvement of the prelimbic-infralimbic areas of the rodent prefrontal cortex in behavioral flexibility for place and response learning. *J. Neurosci.* 19, 4585–4594. doi: 10.1523/JNEUROSCI.19-11-04585.1999

Ralston, J. V., and Herman, L. M. (1995). Perception and generalization of frequency contours by a bottlenose dolphin (*Tursiops truncatus*). *J. Compar. Psychol.* 109, 268–277. doi: 10.1037/0735-7036.109.3.268

Ranck, J. B. Jr. (1973). Studies on single neurons in dorsal hippocampal formation and septum in unrestrained rats: I. Behavioral correlates and firing repertoires. *Exp. Neurol.* 41, 461–531.

Rankin, S. K., Large, E. W., and Fink, P. W. (2009). Fractal tempo fluctuation and pulse prediction. *Music Percept.* 26, 401–413. doi: 10.1525/mp.2009.26.5.401

Rao, S. M., Harrington, D. L., Haaland, K. Y., Bobholz, J. A., Cox, R. W., and Binder, J. R. (1997). Distributed neural systems underlying the timing of movements. *J. Neurosci.* 17, 5528–5535. doi: 10.1523/jneurosci.17-14-05528.1997

Rattenborg, N. C., Martinez-Gonzalez, D., Roth, T. C., and Pravosudov, V. V. (2011). Hippocampal memory consolidation during sleep: a comparison of mammals and birds. *Biol. Rev.* 86, 658–691.

Redish, A. D. (2016). Vicarious trial and error. *Natu. Rev. Neurosci.* 17, 147–159. doi: 10.1038/nrn.2015.30

Regier, P. S., Amemiya, S., and Redish, A. D. (2015). Hippocampus and subregions of the dorsal striatum respond differently to a behavioral strategy change on a spatial navigation task. *J. Neurophysiol.* 114, 1399–1416. doi: 10.1152/jn.00189.2015

Reiner, A., Perkel, D. J., Bruce, L. L., Butler, A. B., Csillag, A., Kuenzel, W., et al. (2004). Revised nomenclature for avian telencephalon and some related brainstem nuclei. *J. Compar. Neurol.* 473, 377–414. doi: 10.1002/cne.20118

Rengarajan, S., Yankura, K. A., Guillermin, M. L., Fung, W., and Hallem, E. A. (2019). Feeding state sculpts a circuit for sensory valence in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* 116, 1776–1781. doi: 10.1073/pnas.1807454116

Repp, B. H., and Su, Y.-H. (2013). Sensorimotor synchronization: a review of recent research (2006-2012). *Psychon. Bull. Rev.* 20, 403–452. doi: 10.3758/s13423-012-0371-2

Rescorla, R. A., and Wagner, A. R. (1972). "A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement," in *Classical Conditioning II*, eds A. H. Black and W. F. Prokasy (New York, NY: Appleton-Century-Crofts), 64–99.

Rilling, J. K. (2014). Comparative primate neurobiology and the evolution of brain language systems. *Curr. Opin. Neurobiol.* 28, 10–14. doi: 10.1016/j.conb.2014.04.002

Rilling, J. K., Glasser, M. F., Jbabdi, S., Andersson, J., and Preuss, T. M. (2012). Continuity, divergence, and the evolution of brain language pathways. *Front. Evol. Neurosci.* 3:11. doi: 10.3389/fnevo.2011.00011

Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., et al. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat. Neurosci.* 11, 426–428. doi: 10.1038/nn2072

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Res.* 3, 131–141. doi: 10.1016/0926-6410(95)00038-0

Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2, 661–670. doi: 10.1038/35090060

Robbe, D., Montgomery, S. M., Thome, A., Rueda-Orozco, P. E., McNaughton, B. L., and Buzsaki, G. (2006). Cannabinoids reveal importance of spike timing coordination in hippocampal function. *Nat. Neurosci.* 9, 1526–1533. doi: 10.1038/nn1801

Roberts, S., and Church, R. M. (1978). Control of an internal clock. *J. Exper. Psychol. Anim. Behav. Process.* 4, 318–337. doi: 10.1037/0097-7403.4.4.318

Roberts, W. A. (2007). Mental time travel: animals anticipate the future. *Curr. Biol.* 17:10. doi: 10.1016/j.cub.2007.04.010

Rodríguez, F., López, J. C., Vargas, J. P., Broglio, C., Gómez, Y., and Salas, C. (2002a). Spatial memory and hippocampal pallium through vertebrate evolution: insights from reptiles and teleostfish. *Brain Res. Bull.* 57, 499–503. doi: 10.1016/s0361-9230(01)00682-7

Rodríguez, F., López, J. C., Vargas, J. P., Gómez, Y., Broglio, C., and Salas, C. (2002b). Conservation of spatial memory function in the pallial forebrain of reptiles and ray-finned fishes. *J. Neurosci.* 22, 2894–2903. doi: 10.1523/JNEUROSCI.22-07-02894.2002

Roederer, J. G. (1984). The search for a survival value of music. *Music Percept.* 1, 350–356. doi: 10.2307/40285265

Roitman, M. F., Wheeler, R. A., and Carelli, R. M. (2005). Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron* 45, 587–597. doi: 10.1016/j.neuron.2004.12.05

Rolls, E. T., and Mills, P. (2019). The generation of time in the hippocampal memory system. *Cell Rep.* 28, 1649.e–1658.e. doi: 10.1016/j.celrep.2019.07.042

Ruppert, E. E., Fox, S. R., and Barnes, D. R. (2004). *Ivertebrate Zoology*, 7th Edn. Boston, MA: Cengage Learning, 130–132.

Rushforth, N. B. (1973). "Behavioral modifications in coelenterates," in *Invertebrate Learning*, eds W. C. Corning, J. A. Dyal, and A. O. D. Willows (Boston, MA: Springer). doi: 10.1007/978-1-4684-3006-6_3

Ryan, J. F. (2014). Did the ctenophore nervous system evolve independently? *Zoology* 117, 225–226. doi: 10.1016/j.zool.2014.06.001

Sahley, C., Rudy, J. W., and Gelperin, A. (1981). An analysis of associative learning in a terrestrial mollusc. *J. Compar. Physiol. A* 144, 1–8. doi: 10.1007/bf00612791

Sallan, L., Friedman, M., Sansom, R. S., Bird, C. M., and Sansom, I. J. (2018). The nearshore cradle of early vertebrate diversification. *Science* 362, 460–464. doi: 10.1126/science.aar3689

Sanderson, C. E., Cook, P., Hill, P. S. M., Orozco, B. S., Abramson, C. I., and Wells, H. (2013). Nectar quality perception by honey bees (*Apis mellifera* ligustica). *J. Compar. Psychol.* 127, 341–351. doi: 10.1037/a0032613

Sandoz, J. C., Laloi, D., Odoux, J. F., and Pham-Delègue, M. H. (2000). Olfactory information transfer in the honeybee: compared efficiency of classical conditioning and early exposure. *Anim. Behav.* 59, 1025–1034. doi: 10.1006/anbe.2000.1395

Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B. L., Witter, M. P., Moser, M.-B., et al. (2006). Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science* 312, 758–762. doi: 10.1126/science.1125572

Savage-Rumbaugh, S., McDonald, K., Sevcik, R. A., Hopkins, W. D., and Rubert, E. (1986). Spon-taneous symbol acquisition and communicative use by pygmychimpanzees (*Pan paniscus*). *J. Exp. Psychol. Gen.* 115, 211–235.

Schachner, A., Brady, T. F., Pepperberg, I. M., and Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Curr. Biol.* 19, 831–836. doi: 10.1016/j.cub.2009.03.061

Schacter, D. L., Addis, D. R., and Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nat. Rev. Neurosci.* 8, 657–661. doi: 10.1038/nrn2213

Schmidt, B., Papale, A., Redish, A. D., and Markus, E. J. (2013). Conflict between place and response navigation strategies: effects on vicarious trial and error (VTE) behaviors. *Learn. Mem.* 20, 130–138. doi: 10.1101/lm.028753.112

Schomers, M. R., Garagnani, M., and Pulvermüller, F. (2017). Neurocomputational consequences of evolutionary connectivity changes in perisylvian language cortex. *J. Neurosci.* 37, 3045–3055. doi: 10.1523/jneurosci.2693-16.2017

Schwartz, B. L., Colon, M. R., Sanchez, I. C., Rodriguez, I. A., and Evans, S. (2002). Single-trial learning of "what" and "who" information in a gorilla (*Gorilla gorilla gorilla*): implications for episodic memory. *Anim. Cogn.* 5, 85–90. doi: 10.1007/s10071-002-0132-0

Schwartz, B. L., Hoffman, M. L., and Evans, S. (2005). Episodic-like memory in a gorilla: a review and new findings. *Learn. Motiv.* 36, 226–244. doi: 10.1016/j.lmot.2005.02.012

Schwartz, B. L., Meissner, C. M., Hoffman, M., Evans, S., and Frazier, L. D. (2004). Event memory and misinformation effects in a gorilla (*Gorilla gorilla gorilla*). *Anim. Cogn.* 7, 93–100. doi: 10.1007/s10071-003-0194-7

Selezneva, E., Scheich, H., and Brosch, M. (2006). Dual time scales for categorical decision making in auditory cortex. *Curr. Biol.* 16, 2428–2433. doi: 10.1016/j.cub.2006.10.027

Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K., and Van Hoesen, G. W. (2001). Prefrontal cortex in humans and apes: a comparative study of area 10. *Am. J. Phys. Anthropol.*, 114, 224-241.

Sengupta, P. (2013). The belly rules the nose: feeding state-dependent modulation of peripheral chemosensory responses. *Curr. Opin. Neurobiol.* 23, 68–75. doi: 10.1016/j.conb.2012.08.001

Sharpe, M. J., and Killcross, S. (2015). The prelimbic cortex directs attention toward predictive cues during fear learning. *Learn. Mem.* 22, 289–293. doi: 10.1101/lm.038273.115

Shimbo, A., Izawa, E.-I., and Fujisawa, S. (2021). Scalable representation of time in the hippocampus. *Sci. Adv.* 7:eabd7013. doi: 10.1126/sciadv.abd7013

Shu, D. (2003). A paleontological perspective of vertebrate origin. *Chin. Sci. Bull.* 48, 725–735. doi: 10.1007/bf03187041

Shu, D.-G., Conway Morris, S., Zhang, Z.-F., and Han, J. (2009). The earliest history of the deuterostomes: the importance of the Chengjiang Fossil-Lagerstätte. *Proc. R. Soc. B Biol. Sci.* 277, 165–174. doi: 10.1098/rspb.2009.0646

Shu, D.-G., Morris, S. C., Han, J., Zhang, Z.-F., Yasui, K., Janvier, P., et al. (2003). Head and backbone of the early cambrian vertebrate haikouichthys. *Nature* 421, 526–529. doi: 10.1038/nature01264

Shultz, S., Opie, C., and Atkinson, Q. D. (2011). Stepwise evolution of stable sociality in primates. *Nature* 479, 219–222. doi: 10.1038/nature10601

Singer, R. A., and Zentall, T. R. (2007). Pigeons learn to answer the question "where did you just peck?" and can report peck location when unexpectedly asked. *Learn. Behav.* 35, 184–189. doi: 10.3758/bf03193054

Skeide, M. A., Brauer, J., and Friederici, A. D. (2016). Brain functional and structural predictors of language performance. *Cereb. Cortex* 26, 2127–2139. doi: 10.1093/cercor/bhv042

Skov-Rackette, S. I., Miller, N. Y., and Shettleworth, S. J. (2006). What-where-when memory in pigeons. *J. Exp. Psychol. Anim. Behav. Process.* 32, 345–358. doi: 10.1037/0097-7403.32.4.345

Smith, B. H. (1997). An analysis of blocking in odorant mixtures: an increase but not a decrease in intensity of reinforcement produces unblocking. *Behav. Neurosci.* 111, 57–69. doi: 10.1037/0735-7044.111.1.57

Smith, B. H., and Cobey, S. (1994). The olfactory memory of the honeybee *Apis mellifera*. II. Blocking between odorants in binary mixtures. *J. Exper. Biol.* 195, 91–108.

Song, X., Osmanski, M. S., Guo, Y., and Wang, X. (2015). Complex pitch perception mechanisms are shared by humans and a New World monkey. *Proc. Natl. Acad. Sci. U.S.A.* 113, 781–786. doi: 10.1073/pnas.1516120113

Spellman, T., Rigotti, M., Ahmari, S. E., Fusi, S., Gogos, J. A., and Gordon, J. A. (2015). Hippocampal-prefrontal input supports spatial encoding in working memory. *Nature* 522, 309–314. doi: 10.1038/nature14445

Steiner, A. P., and Redish, A. D. (2012). The road not taken: neural correlates of decision making in orbitofrontal cortex. *Front. Neurosci.* 6:131. doi: 10.3389/fnins.2012.00131

Steiner, A. P., and Redish, A. D. (2014). Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nat. Neurosci.* 17, 995–1002. doi: 10.1038/nn.3740

Stott, J. J., and Redish, A. D. (2014). A functional difference in information processing between orbitofrontal cortex and ventral striatum during decision-making behaviour. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 369:20130472.

Stout, D., and Hecht, E. E. (2017). Evolutionary neuroscience of cumulative culture. *Proc. Natl. Acad. Sci. U.S.A.* 114, 7861–7868. doi: 10.1073/pnas.1620738114

Strausfeld, N. J., and Hirth, F. (2013). Deep homology of arthropod central complex and vertebrate basal ganglia. *Science* 340, 157–161. doi: 10.1126/science.1231828

Striedter, G. F. (2005). *Principles of Brain Evolution*. Sunderland, MA: Sinauer Associates, Inc. Publishers.

Striedter, G. F., and Keefer, B. P. (2000). Cell migration and aggregation in the developing telencephalon: pulse-labeling chick embryos with Bromodeoxyuridine. *J. Neurosci.* 20, 8021–8030. doi: 10.1523/jneurosci.20-21-08021.2000

Suddendorf, T., and Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genet. Soc. Gen. Psychol. Monogr.* 123, 133–167.

Sul, J. H., Kim, H., Huh, N., Lee, D., and Jung, M. W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron* 66, 449–460. doi: 10.1016/j.neuron.2010.03.033

Sulikowski, D., and Burke, D. (2015). Noisy miners plan ahead: cryptic signalling of reward location impairs search for nectar, but not for invertebrates. *Anim. Behav.* 102, 149–155. doi: 10.1016/j.anbehav.2015.01.005

Sumbre, G., Muto, A., Baier, H., and Poo, M. M. (2008). Entrained rhythmic activities of neuronal ensembles as perceptual memory of time interval. *Nature* 456, 102–107.

Sund, P. N. (1958). A study of the muscular anatomy and swimming behaviour of the sea anemone, *Stomphia coccinea*. *J. Cell Sci.* 99, 401–420.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Mach. Learn.* 3, 9–44. doi: 10.1007/bf00115009

Sweis, B. M., Thomas, M. J., and Redish, A. D. (2018). Mice learn to avoid regret. *PLoS Biol.* 16:e2005853. doi: 10.1371/journal.pbio.2005853

Tanaka, S. C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., and Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* 7, 887–893. doi: 10.1038/nn1279

Tanji, J. (2001). Sequential organization of multiple movements: involvement of cortical motor areas. *Annu. Rev. Neurosci.* 24, 631–651. doi: 10.1146/annurev.neuro.24.1.631

Teki, S., Grube, M., and Griffiths, T. D. (2012). A unified model of time perception accounts for duration-based and beat-based timing mechanisms. *Front. Integr. Neurosci.* 5:90. doi: 10.3389/fnint.2011.00090

Teki, S., Grube, M., Kumar, S., and Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *J. Neurosci.* 31, 3805–3812. doi: 10.1523/jneurosci.5561-10.2011

Tennie, C., Call, J., and Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 2405–2415. doi: 10.1098/rstb.2009.0052

Terrace, H. S. (2019). *Why Chimpanzees can't Learn Language and Only Humans Can*. New York, NY: Columbia University Press.

Terrace, H. S., Petitto, L. A., Sanders, R. J., and Bever, T. G. (1979). Can an ape create a sentence? *Science* 206, 891–902. doi: 10.1126/science.504995

Thiebaut de Schotten, M., Dell'Acqua, F., Valabregue, R., and Catani, M. (2012). Monkey to human comparative anatomy of the frontal lobe association tracts. *Cortex* 48, 82–96. doi: 10.1016/j.cortex.2011.10.001

Thompson, C., and Church, R. (1980). An explanation of the language of a chimpanzee. *Science* 208, 313–314. doi: 10.1126/science.7367862

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift Tierpsychol.* 20, 410–433. doi: 10.1111/j.1439-0310.1963.tb01161.x

Tolman, E. C. (1939). Prediction of vicarious trial and error by means of the schematic sowbug. *Psychol. Rev* 46, 318–336.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208.

Tolman, E. C., and Minium, E. (1942). VTE in rats: overlearning and difficulty of discrimination. *J. Compar. Psychol.* 34, 301–306. doi: 10.1037/h0055388

Tomasello, M., and Carpenter, M. (2005). Intention reading and imitative learning. *Perspect. Imitat.* 2, 133–148. doi: 10.7551/mitpress/5331.003.0007

Tomasello, M., and Carpenter, M. (2007). Shared intentionality. *Dev. Sci.* 10, 121–125. doi: 10.1111/j.1467-7687.2007.00573.x

Tomasello, M., and Moll, H. (2009). "The gap is social: human shared intentionality and culture," in *Mind the Gap*, eds P. Kappeler, and J. Silk (Berlin, Heidelberg: Springer), 331–349. doi: 10.1007/978-3-642-02725-3_16

Tomasello, M., Call, J., and Hare, B. (2003). Chimpanzees understand psychological states - the question is which ones and to what extent. *Trends Cogn. Sci.* 7, 153–156. doi: 10.1016/s1364-6613(03)00035-4

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691. doi: 10.1017/s0140525x05000129

Tomasello, M., Davis-Dasilva, M., Camak, L., and Bard, K. (1987). Observational learning of tool-use by young chimpanzees. *Hum. Evol.* 2, 175–183. doi: 10.1007/bf02436405

Tomlinson, R. W., and Schwarz, D. W. (1988). Perception of the missing fundamental in nonhuman primates. *J. Acoust. Soc. Am.* 84, 560–565. doi: 10.1121/1.396833

Tomonaga, M., Uwano, Y., Ogura, S., and Saito, T. (2010). Bottlenose Dolphins' (*Tursiops truncatus*) Theory of mind as demonstrated by responses to their trainers' Attentional states. *Intern. Soc. Compar. Psychol.* 23, 386–400.

Torley, V. J. (2009). *The Anatomy of a Minimal Mind*. Ph.D thesis. Melbourne, VIC: University of Melbourne.

Tosches, M. A., Yamawaki, T. M., Naumann, R. K., Jacobi, A. A., Tushev, G., and Laurent, G. (2018). Evolution of pallium, hippocampus, and cortical cell types revealed by single-cell transcriptomics in reptiles. *Science* 360, 881–888. doi: 10.1126/science.aar4237

Tremblay, L., and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature* 398, 704–708. doi: 10.1038/19525

Tsunada, J., Lee, J. H., and Cohen, Y. E. (2011). Representation of speech categories in the primate auditory cortex. *J. Neurophysiol.* 105, 2634–2646. doi: 10.1152/jn.00037.2011

Ulinski, P. S. (1983). *Dorsal Ventricular Ridge: a Treatise on Forebrain Organization in Reptiles and Birds*. New York, NY: Wiley.

Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al. (2001). I Know what you are doing. *Neuron* 31, 155–165. doi: 10.1016/s0896-6273(01)00337-3

Valente, A., Huang, K. H., Portugues, R., and Engert, F. (2012). Ontogeny of classical and operant learning behaviors in zebrafish. *Learn. Mem.* 19, 170–177. doi: 10.1101/lm.025668.112

van de Waal, E., Claidière, N., and Whiten, A. (2015). Wild vervet monkeys copy alternative methods for opening an artificial fruit. *Anim. Cogn.* 18, 617–627. doi: 10.1007/s10071-014-0830-4

van den Bos, W., van Dijk, E., Westenberg, M., Rombouts, S. A., and Crone, E. A. (2009). What motivates repayment? Neural correlates of reciprocity in the Trust Game. *Soc. Cogn. Affect. Neurosci.* 4, 294–304. doi: 10.1093/scan/nsp009

van der Meer, M. A. A., and Redish, A. D. (2009). Covert expectation-of-reward in rat ventral striatum at decision points. *Front. Integr. Neurosci.* 3:1. doi: 10.3389/neuro.07.001.2009

van der Meer, M. A. A., Johnson, A., Schmitzer-Torbert, N. C., and Redish, A. D. (2010). Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron* 67, 25–32. doi: 10.1016/j.neuron.2010.06.023

van der Steen, M. C., and Keller, P. E. (2013). The adaptation and anticipation model (ADAM) of sensorimotor synchronization. *Front. Hum. Neurosci.* 7:253. doi: 10.3389/fnhum.2013.00253

Van Twyver, H., and Allison, T. (1972). A polygraphic and behavioral study of sleep in the pigeon (*Columba livia*). *Exp. Neurol.* 35, 138–153.

van Woerden, J. T., Willems, E. P., van Schaik, C. P., and Isler, K. (2011). Large brains buffer energetic effects of seasonal habitats in catarrhine primates. *Evolution* 66, 191–199. doi: 10.1111/j.1558-5646.2011.01434.x

Vassilakis, P. (2005). Auditory roughness as a means of musical expression. *Select. Rep. Ethnomusicol.* 12, 119–144.

Vergoz, V., Roussel, E., Sandoz, J.-C., and Giurfa, M. (2007). Aversive learning in honeybees revealed by the olfactory conditioning of the sting extension reflex. *PLoS One* 2:288. doi: 10.1371/journal.pone.0000288

Vindas, M. A., Folkedal, O., Kristiansen, T. S., Stien, L. H., Braastad, B. O., Mayer, I., et al. (2012). Omission of expected reward agitates Atlantic salmon (*Salmo salar*). *Anim. Cogn.* 15, 903–911. doi: 10.1007/s10071-012-0517-7

Vindas, M. A., Johansen, I. B., Vela-Avitua, S., Nørstrud, K. S., Aalgaard, M., Braastad, B. O., et al. (2014). Frustrative reward omission increases aggressive behaviour of inferior fighters. *Proc. R. Soc. B Biol. Sci.* 281, 20140300. doi: 10.1098/rspb.2014.0300

Vinepinsky, E., Cohen, L., Perchik, S., Ben-Shahar, O., Donchin, O., and Segev, R. (2018). Representation of borders and swimming kinematics in the brain of freely-navigating fish. *Sci. Rep.* doi: 10.1101/291013

Vinepinsky, E., Cohen, L., Perchik, S., Ben-Shahar, O., Donchin, O., and Segev, R. (2020). Representation of edges, head direction, and swimming kinematics in the brain of freely- navigating fish. *Sci Rep* 10:14762. doi: 10.1038/s41598-020-71217-1

Von Frisch, K. (1967). Honeybees: Do they use direction and distance information provided by their dancers? *Science* 158, 1072–1077.

Wakabayashi, T., Kitagawa, I., and Shingai, R. (2004). Neurons regulating the duration of forward locomotion in *Caenorhabditis elegans*. *Neurosci. Res.* 50, 103–111. doi: 10.1016/j.neures.2004.06.005

Walker, M. M. (1997). "Magnetic orientation and the magnetic sense in arthropods," in *Orientation and Communication in Arthropods*, ed. M. Lehrer (Basel: Birkhäuser), 187–213. doi: 10.1007/978-3-0348-8878-3_7

Wallach, A., Harvey-Girard, E., Jun, J. J., Longtin, A., and Maler, L. (2018). A time-stamp mechanism may provide temporal information necessary for egocentric to allocentric spatial transformations. *eLife* 7:e036769. doi: 10.7554/elife.36769

Wallin, N. L., Merker, B., and Brown, S. (eds) (2000). *The Origins of Music*. Cambridge, MA: The MIT Press.

Wapstra, M., and van Soest, R. W. (1987). Sexual reproduction, larval morphology and behaviour in demosponges from the southwest of the Netherlands. *Taxon. Porifera* 13, 281–307. doi: 10.1007/978-3-642-70892-3_15

Warneken, F. (2006). Altruistic helping in human infants and young Chimpanzees. *Science* 311, 1301–1303. doi: 10.1126/science.1121448

Warneken, F., Chen, F., and Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Dev.* 77, 640–663. doi: 10.1111/j.1467-8624.2006.00895.x

Washburn, S. L., and Lancaster, C. (1968). "The evolution of hunting," in *Man the Hunter*, eds R. B. Lee and I. DeVore (Chicago: Aldine), 293–303.

Weaver, L. N., Varricchio, D. J., Sargis, E. J., Chen, M., Freimuth, W. J., and Wilson Mantilla, G. P. (2020). Early mammalian social behaviour revealed by multituberculates from a dinosaur nesting site. *Nat. Ecol. Evol.* 5, 32–37. doi: 10.1038/s41559-020-01325-8

Weber, B. J., and Huettel, S. A. (2008). The neural substrates of probabilistic and intertemporal decision making. *Brain Res.* 1234, 104–115. doi: 10.1016/j.brainres.2008.07.105

Wehner, R., Boyer, M., Loertscher, F., Sommer, S., and Menzi, U. (2006). Ant navigation: one-way routes rather than maps. *Curr. Biol.* 16, 75–79. doi: 10.1016/j.cub.2005.11.035

Wehner, R., Michael, B., and Antonsen, P. (1996). Visual navigation in insects: coupling of egocentric and geocentric information. *J. Exper. Biol.* 199, 129–140.

Wenner, A. M., and Wells, P. H. (1990). *An Anatomy of a Controversy*. New York, NY: Columbia University Press.

Wernicke, K. (1995). "The aphasia symptom-complex: a psychological study on an anatomical basis (1875)," in *Reader in the History of Aphasia: From sasi(Franz Gall to). 4*, ed. P. Eling (Amsterdam: John Benjamins Pub Co), 69–89.

Whiten, A. (2017). A second inheritance system: the extension of biology through culture. *Interf. Focus* 7:20160142. doi: 10.1098/rsfs.2016.0142

Whiten, A., Horner, V., and de Waal, F. B. (2005). Conformity to cultural norms of tool use in chimpanzees. *Nature* 437, 737–740. doi: 10.1038/nature04047

Wiener, M., Turkeltaub, P., and Coslett, H. B. (2010). The image of time: a voxel-wise meta-analysis. *Neuroimage* 49, 1728–1740. doi: 10.1016/j.neuroimage.2009.09.064

Wikenheiser, A. M., and Redish, A. D. (2015). Hippocampal theta sequences reflect current goals. *Nat. Neurosci.* 18, 289–294. doi: 10.1038/nn.3909

Wilkinson, A., and Huber, L. (2012). "Cold-blooded cognition: reptilian cognitive abilities," in *The Oxford Handbook of Comparative Evolutionary Psychology*, eds J. Vonk and T. K. Shackelford (Oxford: Oxford University Press), doi: 10.1093/oxfordhb/9780199738182.013.0008

Wilkinson, A., Chan, H.-M., and Hall, G. (2007). Spatial learning and memory in the tortoise (*Geochelone carbonaria*). *J. Compar. Psychol.* 121, 412–418.

Wilkinson, A., Kuenstner, K., Mueller, J., and Huber, L. (2010). Social learning in a non-social reptile (*Geochelone carbonaria*). *Biol. Lett.* 6, 614–616. doi: 10.1098/rsbl.2010.0092

Wilson, E. B. (1891). The heliotropism of hydra. *Am. Natur.* 25, 413–433. doi: 10.1086/275328

Wilson, E. O. (1962). Chemical communication in the fire ant *Solenopsis saevissima*. *Anim. Behav.* 10, 134–138.

Wixted, J. T., Goldinger, S. D., Squire, L. R., Kuhn, J. R., Papesh, M. H., Smith, K. A., et al. (2018). Coding of episodic memory in the human hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* 115, 1093–1098. doi: 10.1073/pnas.1716443115

Woollacott, R. M. (1993). Structure and swimming behavior of the larva of *Haliclona tubifera* (Porifera: Demospongiae). *J. Morphol.* 218, 301–321. doi: 10.1002/jmor.1052180306

Woolsey, C. N., Settlage, P. H., Meyer, D. R., Sencer, W., Pinto Hamuy, T., and Travis, A. M. (1952). Patterns of localization in precentral and "supplementary" motor areas and their relation to the concept of a premotor area. *Res. Public. Assoc. Res. Nervo. Ment. Dis.* 30, 238–264.

Wright, A. A., Rivera, J. J., Hulse, S. H., Shyan, M., and Neiworth, J. J. (2000). Music perception and octave generalization in rhesus monkeys. *J. Exp. Psychol. Gen.* 129, 291–307. doi: 10.1037//0096-3445.129.3.291

Wu, Y., Wang, H., and Hadly, E. A. (2017). Invasion of ancestral mammals into dim-light environments inferred from adaptive evolution of the *Phototransduction* genes. *Sci. Rep.* 7:46542. doi: 10.1038/srep46542

Yeatman, J. D., Dougherty, R. F., Rykhlevskaia, E., Sherbondy, A. J., Deutsch, G. K., Wandell, B. A., et al. (2011). Anatomical properties of the *Arcuate fasciculus* predict phonological and reading skills in children. *J. Cogn. Neurosci.* 23, 3304–3317. doi: 10.1162/jocn_a_00061

Yeterian, E. H., and Pandya, D. N. (1998). Corticostriatal connections of the superior temporal region in rhesus monkeys. *J. Compar. Neurol.* 399, 384–402. doi: 10.1002/(sici)1096-9861(19980928)399:3<384::aid-cne7>3.0.co;2-x

Yin, B., and Meck, W. H. (2014). Comparison of interval timing behaviour in mice following dorsal or ventral hippocampal lesions with mice having δ-opioid receptor gene deletion. *Philos. Trans. Roy. Soc. Lond. Ser. B Biol. Sci.* 369, 20120466. doi: 10.1098/rstb.2012.0466

Yin, P., Fritz, J. B., and Shamma, S. A. (2010). Do ferrets perceive relative pitch? *J. Acoust. Soc. Am.erica* 127, 1673–1680. doi: 10.1121/1.3290988

Yuan, X.-L., Chen, Z., Xiao, S.-H., Zhou, C.-M., and Hua, H. (2011). An early Ediacaran assemblage of macroscopic and morphologically differentiated eukaryotes. *Nature* 470, 390–393.

Zarco, W., Merchant, H., Prado, L., and Mendez, J. C. (2009). Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. *J. Neurophysiol.* 102, 3191–3202. doi: 10.1152/jn.00066.2009

Zentall, T. R., Clement, T. S., Bhatt, R. S., and Allen, J. (2001). Episodic-like memory in pigeons. *Psychon. Bull. Rev.* 8, 685–690. doi: 10.3758/bf03196204

Zentall, T. R., Singer, R. A., and Stagner, J. P. (2008). Episodic-like memory: pigeons can report location pecked when unexpectedly asked. *Behav. Process.* 79, 93–98. doi: 10.1016/j.beproc.2008.05.003

Zhang, S., Mizutani, A., and Srinivasan, M. V. (2000). Maze navigation by honeybees: learning path regularity. *Learn. Mem.* 7, 363–374. doi: 10.1101/lm.32900

Zhang, X., Shu, D., Li, Y., and Han, J. (2001). New sites of Chengjiang fossils: crucial windows on the Cambrian explosion. *J. Geol. Soc.* 158, 211–218. doi: 10.1144/jgs.158.2.211

Zhang, Y., Paik, J., and Pirolli, P. (2015). Reinforcement learning and counterfactual reasoning explain adaptive behavior in a changing environment. *Top. Cogn. Sci.* 7, 368–381. doi: 10.1111/tops.12143

Zinkivskay, A., Nazir, F., and Smulders, T. V. (2009). What-Where-When memory in magpies (Pica pica). *Anim. Cogn.* 12, 119–125. doi: 10.1007/s10071-008-0176-x