Appears in B. Huebner (ed.), The Philosophy of Daniel Dennett.
Oxford: Oxford University Press, 2018, pp. 225-249.
This version corrects a typo on page 234.

8.1 TOWERS AND TREES IN COGNITIVE EVOLUTION

Peter Godfrey-Smith

What are the main stages in the evolution of cognition, the crucial transitions between simpler and more complex forms? When addressing this question, it is tempting to adopt a kind of linear thinking—*X* leads to *Y* which leads to *Z*—and common also to push back against these temptations. Evolution is not a matter of ladders and scales, we're reminded, but a process of branching and divergence. Discrete stages, as opposed to seamless gradations, might also be seen as unlikely. But amid all the radiations and gradual shifts, and all the different niches and lifestyles, might there be some especially important inventions that fall into a natural order? Perhaps such a sequence might be important, also, as a manifestation of some general principle about how organisms control behavior and deal with the world.

One of my favorite papers by Dan Dennett is "Why the Law of Effect Will Not Go Away" (1975; reprinted in *Brainstorms*, 1978). Here Dennett made some early moves (following up the even earlier *Content and Consciousness*, 1969) in what has become a lengthy defense of the idea that there is a single pattern essential to all processes of adaptation, cognitive improvement, and "R & D" in a broad sense of the term. That pattern is the Darwinian one: trial and error, generate and test, variation and selection. Dennett argues that whenever adaptive improvement occurs, a generate-and-test process must be at the bottom of it. Such a process can be realized on many timescales; Thorndike's "law of effect" of 1905 is a manifestation of the pattern within an individual learner's lifetime. Successful behaviors are repeated, unsuccessful ones abandoned.[1]

1. The idea that variation and selection are omnipresent in adaptive processes was also defended around the same time by Donald Campbell (1974), and one kind of "evolutionary epistemology" takes off from this idea.

In his books *Darwin's Dangerous Idea* (1995) and *Kinds of Minds* (1996), Dennett developed these sketches into a "Tower of Generate and Test," with a series of transitions between ways of realizing the Darwinian pattern on different scales and with different degrees of sophistication. At the bottom are *Darwinian creatures*, which can only adapt through genetic mutation and natural selection. Above them in the tower, *Skinnerian creatures* also adapt during their lifetimes by trial and error learning. Above those are *Popperian creatures*, which don't have to actually perform the behavioral options being tested. Instead, they can run internal experiments to assess the consequences of a behavior "offline," discarding bad options and only exposing the better ones. *Gregorian creatures* use social tools, especially language, to make use of innovations discovered by other individuals. Rather than test all the options themselves, even internally, they draw on the collective experience of many.

In 2015, at a conference at Macquarie University, the psychologists Russell Gray and Alex Taylor discussed a more elaborate tower, due to Taylor, that added a number of floors to Dennett's.[2] But both Gray and Taylor worried that a "tower" is not the right framework to be using at all; it has too much of the *scala natura* about it. Darwinians don't expect scales but trees. That concern is reasonable, but there can still be cumulative and directional processes in a Darwinian context, processes in which one move must occur before another. So given all that we now know in comparative psychology and evolutionary biology, what should we make of Dennett's tower and other attempts to mark out the stages in an advance of cognitive complexity?

I'll investigate this question by offering a new taxonomy of creatures, drawing on Dennett, Taylor, and other work. Dennett's own tower had a dual role. First, it was intended to mark out genuine historical stages, though it was not intended as an overall map of the evolution of the mind, which involves a lot more than the proliferation of generate-and-test mechanisms. (Perhaps Dennett is not so convinced there's a *lot* more.) Second, Dennett sees the sequence of realizations of the Darwinian pattern as part of a general account of how adaptation, design, and meaning are possible in a wholly physical world. My project here lies more in the history of cognition, and I will set up a sequence in which generate-and-test mechanisms do have an important role but are mixed in with others. My interest is partly in the question of where there *is* a tower-like structure, and where there is not, both empirically and in principle. By "tower-like" structure I mean a situation in which the evolution of one cognitive mechanism must, or at least always does, precede the

---

2. The conference was *Understanding Complex Animal Cognition: An Interdisciplinary Workshop*, organized by Rachael Brown, at Macquarie University, Sydney, Australia.

evolution of another, especially in situations where both are retained. If this situation holds, there will be a nesting: every organism with *Y* will also have *X*, and so on. For which cognitive kinds is there such a sequence, and for which is there a different pattern?[3]

If there is nesting in some cases, this will still be laid out as a tree in the basic sense relevant to Darwinism. You might have a situation where *X* appears early and *Y* later appears independently in a subset of the branches with *X*. Then not all the branches will have the same things on them, but if we set aside losses, there will be no cases where *Y* appears without *X*. When that is true, there is something of a tower present within the Darwinian tree.

Before going on, I'll note a terminological point. A categorization like the one seen here can be set up either with nested or exclusive categories. (Are Skinnerian creatures *also* Darwinian creatures, or does becoming Skinnerian take an animal out of the Darwinian category?) Dennett has mostly (perhaps not always) written about his categories as if they are exclusive. Given the way my story will turn out, it's best to use nested and non-exclusive categories (so a Skinnerian can also be Darwinian). Sometimes I'll say "*merely* Darwinian," or something similar, to indicate contrasts, and occasionally this role will be played by the context, in a way that I hope remains clear.

1. Humean Creatures

In Dennett's tower. there are Darwinian creatures, which evolve by natural selection over generations, and then Skinnerian creatures, which adapt during their lifetime by trial and error learning. I'll discuss a stage between these—though whether it really lies *between* them will be discussed in detail. These are *Humean creatures*, creatures capable of associative learning that does not have a trial-and-error character, also known as classical conditioning.

I take the essential feature of classical conditioning to be the learning of correlations between events perceived. Pavlov's dogs regularly heard the bell ring before the arrival of food. The usefulness of classical conditioning is predictive; you can learn that one event predicts another. There's an *unconditioned stimulus* (US), to which you have a preexisting behavior, an *unconditioned response* (UR), that has been established through evolution or earlier learning. You learn

3. There might be a directional process in which replacement rather than retention was the prevailing pattern—*X* is always *replaced* by *Y*, and so on. I'll assume in this chapter that capacities are augmented rather than replaced, and will ignore the possibility of losses except where this is explicitly discussed.

to produce this behavior in response to a new *conditioned stimulus* (CS) which is found to be associated with the US.

I used Pavlov's example but named the creatures after Hume. I do see Hume as the first to describe a mechanism of this general kind. As Hume had it, ideas that occur in the mind together tend to prompt or stimulate each other. A kind of "attraction" operates between them. There is no essential link to behavior or reward in Hume's story. Though he was attuned to adaptive considerations, in a loose sort of way, Hume conceived of his mechanism more along the lines of a quasi-physics, a "power of attraction" in a sense analogous to Newtonian gravity. After Hume, this picture evolved into a psychological theory, via James Mill and his son J. S., also Bain and Spencer, and then the experimental tradition of Pavlov, Watson, and others.[4] People sometimes talk of "associative learning" as a single thing, or a single package with minor variants, but I am emphasizing what I take to be some deep differences between its forms.

Who is a Humean creature and how did this capacity evolve? Classical conditioning is *very* widespread. To put the distribution of Humean capacities into context, I will give a quick sketch of the evolutionary history of animals. Animals evolved from single-celled protists somewhere between 700 million to a billion years ago.[5] The early branchings are currently unclear. It used to be thought that sponges are a "sister group" to all other animals—they are the present-day animals whose lineage branched off first. Some data presently suggest that *ctenophores* (comb jellies) branched off earlier. If so, there may have been two independent origins of nervous systems, as ctenophores have them and sponges do not. Either way, later than the branching between sponges and others was a split between cnidarians and a large group of *bilaterian* animals, those with a left and right side as well as a distinction between top and bottom. Nearly all the familiar animals are in this group. Cnidarians, which are radially rather than bilaterally symmetrical, include jellyfish, anemones, and corals. Within the bilaterians is a further split into *protostomes* and *deuterostomes*. Protostomes include arthropods (like insects and crabs), molluscs (like clams and octopuses), annelids (like earthworms) and some others. In deuterostomes, the main groups include vertebrates, like ourselves, and echinoderms, including starfish.

I will draw on a 2013 review of invertebrate learning by Clint Perry, Andrew Barron, and Ken Cheng (2013), which includes a large chart summarizing which animals have been shown to do various kinds of learning. Classical conditioning is extremely widespread among bilaterians, and it is seen in some very simple

4. For this history, see Greenwood (2009).

5. See Budd and Jensen (2015); Peterson, Cotton, Gehling, and Pisani (2008); Ryan et al. (2013).

animals.[6] The nematode worm *Caenorhabditis elegans*, which has only 959 cells and 302 neurons (in the hermaphrodite) has shown classical conditioning. In the summary given in the Perry et al. (2013) paper, there are only a few bilaterian groups in which classical conditioning has *not* been shown, including millipedes, barnacles, rotifers, and also sea squirts, which are quite close to us on the tree but lead simple lives.

Classical conditioning might, then, be an early bilaterian invention, evolving once and ramifying through the many groups, perhaps lost in some, or just hard to find. If this is true, classical conditioning is very old—600 million years or so. Alternatively, it might have been invented several times, being useful and easy to build. I've not been able to find much information, or even confident claims, on that last point. Classical conditioning has also been reported in one animal (in a study accepted by Perry et al. [2013]) outside the bilaterians: in an anemone.[7] So if classical conditioning has a single origin, then it is even older than I just said it is, or perhaps there's one origin here and another (or several) in the bilaterians. Some old reports suggest classical conditioning in paramecia, which are single-celled protists, and not animals at all. But these reports seem to have been discredited, as far as I can tell, and I will set those aside.

The Perry, Barron, and Cheng (2013) paper acknowledges the difficulty in showing that an animal *can't* do some form of learning. If the experiment was done at all, it might have used the wrong stimuli; or, it might not have been reported because negative results are seen as less interesting and harder to publish. Their paper is a list of what has been shown to be present, not what has shown to be absent. In some cases here, however, I will take—very provisionally—the absence of evidence to be evidence of absence. In well-studied organisms, the non-report of an interesting trait is informative. At other places I will treat absences more hypothetically: *if* we assume the distribution is a certain way (including absences), then a certain picture results. Anyway, a very large range of animals are Humean creatures, and there even seems to be a radially symmetrical Humean.[8]

The evolutionary history of classical conditioning also bears on other questions. In the early evolution of animals there is an enigmatic period before the Cambrian explosion, called the *Ediacaran*, between about 635 and 540 million years ago. Before this time there is no fossil trace of animal life at all. In the Ediacaran there are fossils of many sizes and kinds. It's hard to work out how these animals lived and, in many cases, whether they were animals at all. Some have

6. The main figure in the paper does have one typographical error: the molluscs are split in a way that is not accurate.

7. See J. Haralson, Groff, and S. Haralson. (1975).

8. Cubozoa are usually seen as the most behaviorally sophisticated non-bilaterians. I don't know if anyone has looked there.

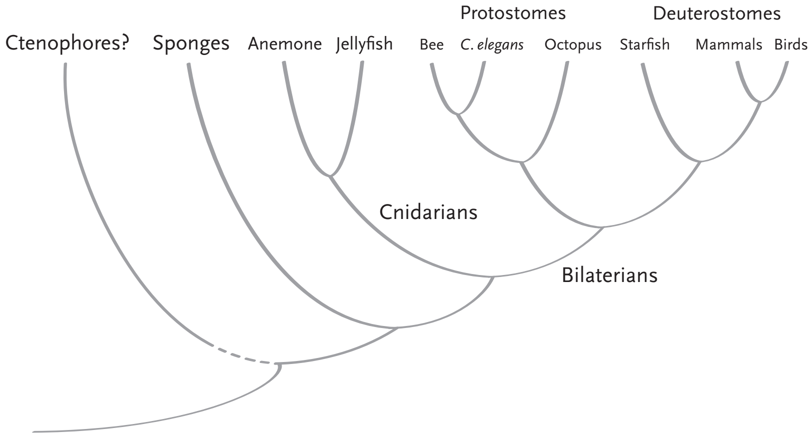

FIGURE 1. Evolutionary relationships between some major animal groups. The figure is not to scale with respect to time, and it mixes taxonomic ranks. The order of the early branchings is controversial.

three-sided forms that have never been seen since. Genetic evidence strongly suggests that animals were around, and that they had evolved nervous systems, even before this time. The branching between cnidarians and bilaterians in Figure 1 may date to 650 or 700 million years ago, and the animals on both sides have nervous systems.[9] Behavior in the Ediacaran seems to have been very simple, though; there are no known bodies with legs, claws, sophisticated eyes, spines, or shells. These animals seem to have lived lives much more self-contained than a typical animal life now. This limited traffic with the world makes Ediacaran animals appear to be—in a phrase suggested by Hans Pols—*Leibnizian creatures.* They were not truly windowless—no life is windowless, even the simplest known bacteria have sensory "windows" (signal transduction mechanisms) with which they track what is outside.[10] But the evolution of animals might have included a somewhat less windowy time. This picture becomes less likely, however, if classical conditioning was around. Classical conditioning is essentially a tool for dealing with external patterns and events, in situations where timing matters. So *if* it were to be shown that classical conditioning is of Ediacaran age, this would suggest that this time was not so Leibnizian—unless the only living remains of that time are descended from the animals most interested in looking out the window.[11]

9. See Peterson et al. (2008).

10. See Lyon (2015).

11. In standard taxonomies of learning, there is also "non-associative" learning—habituation and sensitization. In the Perry et al. (2013) review, habituation is represented as even more

2. Skinnerian Creatures

Skinnerian creatures, in Dennett's hierarchy, learn by trial and error, tracking the good and bad consequences of their actions. They retain and reproduce actions that lead to good outcomes and discard ones that don't. I'll refer to this broad category as *instrumental* learning or conditioning.[12]

Instrumental learning features an essential role for reward and/or punishment. It's not merely reward-driven behavior in an immediate, here-and-now sense (pull back when shocked, keep coming when rewarded). It has to involve longer-term change and the shaping of dispositions, so the animal responds adaptively to the *next* situation of a certain kind, as opposed to maintaining or ceasing a behavior in response to what it experiences in the present.

Drawing again on the review by Perry, Barron, and Cheng (2013), who can learn in this way? First, it is found only in bilaterian animals. It has also not been found in nearly as many bilaterians as classical conditioning; there are many gaps. It is seen in some arthropods (bees, crabs, crickets, and some others), but there are others in which it has not been reported (spiders, wasps, millipedes). Of those, spiders and wasps can be classically conditioned, though it has not been reported in millipedes. In molluscs, some can, and some (apparently) can't. No worms have been reported to do it, though annelid worms can be classically conditioned.

If we take this survey at face value, animals who can learn by instrumental conditioning are a strict subset of those who can learn by classical conditioning. That is: *all Skinnerian creatures are Humean creatures, but not vice versa*. If so, we have a cumulative, tower-like structure so far.

The distribution of this trait also makes it seem likely that instrumental conditioning did not evolve once and get passed down many branches of the tree. It's likely to have evolved a number of times. Otherwise, instrumental learning has very often been lost, and/or many non-reports of the trait are misleading.

I don't discount that last possibility, and there are interesting grey areas, too. Nematode worms are an important case in which classical conditioning has been shown, but, in the surveys I'm using, instrumental learning has not. These are

---

widespread—seen in jellyfish and anemones, for example, as well as all the bilaterian groups they include. Sensitization is not quite as common. These forms of plasticity are so minimal I leave them out of further discussion, though. It is true that there are interesting possibilities concerning the evolution of associative from non-associative learning (Wells, 1968), and there is also a detailed and interesting report of habituation in plants (Gagliano, Renton, Depczynski, & Mancuso, 2014).

12. It is sometimes called *operant* conditioning, but I'll use that term in a narrower sense (as some others do). In this sense, operant conditioning involves the addition of new behaviors to a repertoire, rather than merely adjusting the frequency or setting in which pre-existing and stereotypical behaviors are produced. The Perry et al. (2013) review uses a broader sense of "operant."

very well-studied animals. A closer look at the literature shows some borderline phenomena, too. Although a lot of what nematodes can do fits classical conditioning, they have been reported to show rapid aversion learning by taste. Food that makes them sick is avoided when it appears again, in a way that has been compared to the "Garcia effect" in mammals.[13] Might this, also, be a Humean phenomenon? Perhaps the taste of the (bacterial) food is associated with an experience of unease that has an unconditioned response (UR) of avoidance. Then might the taste come to be used as a predictor of that experience, leading to a pairing of the avoidance with the taste? Perhaps, and in a moment I'll look more closely at attempts to Humeanize behaviors of this general kind.

Another borderline case is further away on the tree: anemones again. They have not been reported to engage in instrumental learning, but in a recent study, Mark Briffa found the following. In an anemone species in which individuals sometimes fight each other, there are meeker and bolder anemones. Briffa found that anemones become less bold after they lose a fight, and the change in boldness is seen across different contexts, not merely in later fights. Is this a case of instrumental learning? Do they learn *that* they're no good as fighters, or that fighting is a bad idea for them?[14]

Pressing on these borderline cases, especially the one with nematodes, leads to a general problem. What kind of separation is there, in principle, between classical and instrumental conditioning, between Humean and Skinnerian capacities? Many papers assume the distinction and treat it as an important classificatory tool. But quite a few discussions, especially by specialists, either reject the standard distinction or see it as very elusive in contexts of testing.[15] Björn Brembs, for example, has suggested that the distinction "needs to be reconsidered," as most situations in which associative learning occurs can be described in either way, or perhaps contain both processes. When an animal performs some behavior, gets a reward, and consequently performs the behavior more often, this can be

13. See Ardiel and Rankin (2010); and Nuttley, Atkinson-Leadbeater, and van der Kooy (2002). For the "Garcia effect," see Garcia and Koelling (1966).

14. Rudin and Briffa (2012). Limpets, a kind of gastropod not on the Perry et al. (2013) list for instrumental conditioning, have been reported to do something similar: see Shanks (2002). Spiders have also been reported to do something like this—fight winners (and losers) behave differently the next time they fight—but the decay of memory is quick, with a complete reset by 24 hours (Kasumovic, Elias, Sivalinghem, Mason, & Andrade, 2010).

15. The nematode case illustrates the problem. Andrew Barron, one of the authors of the Perry et al. (2013) review I use, is very cautious about these categorizations (personal communication). For the Brembs comment, see Bjorn Brembs, (2013, July 11), Brains as output/input systems [Web blog post]. Retrieved from http://bjoern.brembs.net/2013/07/brains-as-outputinput-systems/.

described in terms of an association between behavior and reward, *or* in terms of a classically conditioned association between the experience of performing the behavior and the appearance of (say) food. The animal learns *that* the experience of doing *X* predicts the arrival of food, much as it might learn that a bell predicts food.[16] Brembs thinks that with care it is possible to remove the classical element from some cases of instrumental learning, but it's not easy, and the results should then be seen as akin to the learning of skills.

In this way and others, the simple and neat Humean/Skinnerian relationship could be made more complicated. I'm reluctant to admit that the distinction is problematic in principle, though. Consider some extreme cases. An animal could be completely blind to reward—completely unable to distinguish between good and bad outcomes—and yet be able to learn by classical conditioning; it could still learn that when *X* happens, *Y* happens soon after, and it could come to produce its *Y*-adapted behavior in response to *X*, as well. It might not register the good consequences of that production of a *Y*-adapted behavior; only blind Darwinian processes might assign the appropriate behavior to stimulus *Y*. But it can usefully modify when the behavior is produced. An animal might be sensitive to facts but not to the consequences of its actions and still usefully learn.

On the other side, can an animal be sensitive to the consequences of its actions but not to the facts? In a sense, no, because the consequences of actions *are* facts. An animal sensitive to consequences can't be insensitive to all the facts. (This is part of what Brembs was getting at.) However, whereas a sensitivity to facts, not consequences, can be sufficient to change behavior using classical conditioning, the kind of "facts" discussed just above (facts about the consequences of actions) can't be used to change behavior in that way. It might be a fact that lever-pressing predicts the arrival of food, but lever-pressing only occurs when the animal decides to *do* it. It's not an exogenous event that can be sensed and then associated with other events. The animal might learn the fact that next time it does a lever press, food will arrive, but in a classical conditioning scenario, that fact cannot induce the animal to do the lever press. The animal would have to wait until presses occurred, at which point it could expect food and act in a way adapted to food's arrival. But "waiting" is not a way for the lever to become pressed, at least in normal circumstances. Something must induce the animal to *make* the press.

In instrumental conditioning, the initial presses might be entirely random, and the later ones are guided by reward. Pressing occurs *because* it has had good effects. If an animal is only classically conditionable, there is no reason why it should go from an initial random lever press to additional ones. All it can do is

16. Dickinson and Balleine (1994).

note that *if* the lever happens to be pressed, there are other things it can expect to happen.

If sensitivity to correlations between perceived events and sensitivity to reward are two distinct capacities, then we can ask straightforward questions about which comes first, if either does, and which animals have one without the other. The minimally sufficient mechanisms for each kind of learning are different, too. Simple arrangements of Hebbian synapses are sufficient for classical conditioning: suppose neurons *A* and *B* both excite *C*, with *A* but not *B* being initially sufficient to make *C* fire. Then if *A* and *B* register distinct but correlated events, *B* will fire when *C* does (as *C* has been made to fire by *A*), and the link between *B* and *C* can strengthen according to Hebb's rule (neurons that fire together, wire B together), until ~~*X*~~ is also sufficient to make *C* fire. This is a local and undemanding process. Nothing so simple suffices for instrumental conditioning. The animal must have a way of registering the consequences of its actions in a way that feeds back and affects how it will behave in the future. The apparent distribution of Humean and Skinnerian capacities on the phylogenetic tree appears to reflect this difference in how demanding the processes are. From what I've said, there's no reason why an animal *couldn't* have instrumental without classical conditioning. But perhaps classical conditioning is easy enough and useful enough for this never to happen, whereas instrumental conditioning—which is even more useful, one might think—is not so easy.

It might reasonably be objected that I am assuming here a very simple form of classical conditioning—a form that is simpler than what many animals seem to use. Animals don't just track where two stimuli occur near to each other in time; they track whether the conditioned stimulus is genuinely *predictive* of the unconditioned stimulus. This is a much harder question for the learning brain to answer, but classical conditioning seems to be sensitive to it.[17] And this is the tip of the iceberg. Here is a result discussed in this connection by Anthony Dickinson, from an experiment by Balleine, Espinet, and Gonzalez (2005). Rats were fed initially on a sweetened drink that had two flavors, orange and lime, combined. Then they let the rats drink an unsweetened lime-only drink. Later, those rats showed a stronger preference for an orange-only drink than did rats for whom the second lime-only drink *had* been sweet. Put in folk-psychological terms, the rats who drank the unsweetened lime-only drink worked out that it was probably the orange that was the source of the sugar when they'd encountered the two together. Tasting the unsweetened lime-only drink made them *retrospectively re-evaluate* the orange flavor.

17. Rescorla (1967); Dickinson (2012).

Contemporary work on associationism posits mechanisms that have much more complexity than Hume or Pavlov envisaged.[18] There also seems to be a tendency, though, to try to give quite *unified* models of how classical conditioning and instrumental conditioning work.[19] But if classical conditioning arose independently more than once, there's no reason to expect it to work the same way all the time. It might appear in simpler forms in some animals and complicated forms in others. Even if it arose just once, it presumably arose first in a simple form, and there would then be no reason to expect it to become complex in the same way in all the independently evolving lineages on which it's now found. Or perhaps there *is* reason to expect this, if there's a single way of doing it that is both more complex than the initial form and the best possible way (or a very good and easily engineered way) of doing it. Dickinson (2012) does not consider multiple evolutionary lineages, but he does consider such a sequence: "The ancestral form of associative learning may have been based on simple temporal contiguity between events. However, this simple system was prone to developing superstitious 'beliefs' based on fortuitous event pairings" (p. 2737). From there, a more complex process arose that made the organism sensitive to the predictive relationships, not merely a pairing, between events.

As Dickinson sketches it, the simple version evolves, and then the complex version follows. But given the shape of the evolutionary tree, there are many possibilities. Perhaps the simple version evolves once, and the same complex version appears in various places later on because it's so useful. Perhaps the same path from nothing to simple to complex is trodden many times. Or perhaps both the simple and the complex forms were one-time inventions, much deeper in the past than a paleontological behavioral ecologist would ever have suspected. It's interesting, given this sketch of possibilties, that the old anemone study I mentioned earlier

---

18. Dickinson (2012):

> Within contemporary associationism, knowledge takes the form of what is often called a representation of the relationship between events, be they stimuli or responses, by a connection (or association) between representations of these events. The process by which this knowledge is deployed is the transmission of excitation or activation (and inhibition) from one event representation to the other via the connection with stronger connections producing greater transmission. Finally, this associative knowledge is acquired by the progressive strengthening of the connection with each effective experience of a relationship between the events. (p. 2733)

Associationism is no longer a quasi-physics, or does some of that character remain?

19. See, for example, the way the cricket work in Terao, Matsumoto, and Mizunami (2015) is set up.

claims to have shown a complex form of classical conditioning, one attuned to predictive relationships. And a recent study of crickets reported that they, too, exhibited a quite complex form of classical conditioning.[20]

I finish this section with a point of purely historical interest about instrumental learning and its relatives. Dennett used Skinner as namesake for this kind of creature. Why not Thorndike, as in Dennett's old paper? I associated Hume, an 18th-century name, with classical conditioning, so we might wonder whether an earlier figure deserves the credit here as well. In intellectual history, as well as animal evolution, instrumental learning seems to have come later. One of the first to describe it was Alexander Bain, in his 1859(!) book *Emotions and the Will.*[21] Bain's work, a mix of psychology and philosophy, is quite neglected, though the American pragmatists recognized its importance. I suspect that Dennett gave Skinner the credit because Skinner saw and emphasized the analogy between instrumental learning and Darwinian evolution (1938). Bain did not, and I don't know whether Thorndike did. William James saw it, and James taught Thorndike, but James tended toward saltationism about both kinds of evolution in a way that alters the explanatory role of selection quite substantially.[22] Hume (1739/2000), incidentally, did hit on a momentary anticipation of adaptation by trial and error, in a passage about social behavior:

> Two men who pull the oars of a boat, do it by an agreement or convention, although they have never given promises to each other. Nor is the rule concerning the stability of possessions the less derived from human conventions, that it arises gradually, and acquires force by a slow progression *and by our repeated experience of the inconveniences of transgressing it* [italics added]. (bk. 3, pt. 2, section 2)[23]

20. The paper by Haralson et al. (1975) reports that the work did include a Rescorla-type "truly random" control. The cricket study is Terao et al. (2015).

21. Bain (1859): "In the primitive aspect of volition, which also continues to be exemplified through the whole of life, an action once begun by spontaneous accident is maintained, when it sensibly alleviates a pain, or nurses a pleasure." I had thought Bain might deserve credit for being *the* first to state a principle of this kind, but according to John Greenwood (2009), he was not the first and he might have picked up this idea from the German psychologist Johannes Müller (1801–1858).

22. See Godfrey-Smith (1996, chapter 3).

23. Another near-miss is monumentally ironic. Lamarck, in his 1809 defense of his doctrine of the inheritance of acquired characteristics, notes that a version of his view exists as a proverb, "*Habits form a second nature.*" Then, "if the habits and nature of each animal could never vary, the proverb would have been false and would not have come into existence, nor been preserved in the event of anyone suggesting it" (1809/2011, p. 114).

It is interesting that Hume's theory of psychological dynamics did not include anything like this, another aspect of the blind spot about selectionist mechanisms that seemed firmly in place before Darwin and Bain.

3. Carnapian, Pearlian, Mimetic, Popperian, and Tolmanian Creatures

Dennett's tower had *Popperian creatures* next. Rather than trying out behaviors in potentially risky actions, they internalize the generate-and-test mechanism entirely. I'll work my way to these creatures but will do that by thinking more generally about steps that might be made onward from the Skinnerian state.

At this point it's surely likely that the tower-like structure breaks down. I'll discuss four separate sophistications that might appear next, without offering an ordering of them. I'll also suggest that the Popperian grade of cognition builds on elements from this set—perhaps one, perhaps several.

The first thing to mention is not a step to a new algorithm or mechanism but a road forward that can be trod to different degrees. In current work on learning in some invertebrates, especially bees, there is a lot of interest in the learning of abstract concepts and logical relations—conjunctions and disjunctions, relational concepts such as *larger than*, and others. In the review by Perry, Barron, and Cheng (2013) used in earlier sections of this paper, this capacity to deal with abstraction is the main path away from ordinary instrumental conditioning that they consider (along with navigation, which I'll discuss below). Bees seem to be the most notable masters of abstraction among the invertebrates—or at least, the animals in which abstraction has most clearly been demonstrated so far. For example, after being conditioned to choose the *larger* of two stimuli, bees can extrapolate this rule to a situation in which two objects of different sizes from either of those they saw during training are presented to them, and in which the object most similar to the object they learned to choose during the training period is *not* the one they need to pick in the new situation, because it's not the larger of the new pair.[24] That is not bad for a brain one cubic millimeter in size. A likely empirical picture, in any case, is that one road away from simple instrumental learning is a road on which the basic rules are the same, but the concepts and discriminations employed are more abstract and complex. I'll call animals like bees *Carnapian creatures.*

24. Avarguès-Weber, Dyer, Combe, and Giurfa (2012); Avarguès-Weber, d'Amaro, Metzler, and Dyer (2014).

Here is a second kind of sophistication, which might be reached by traveling the Carnapian path, but might not be. Taking up a concept from Taylor's tower, I'll recognize *Pearlian creatures.* These animals can make use of causal reasoning based on a notion of intervention, the kind of causal thinking described in detail by the computer scientist Judea Pearl and also by a number of philosophers, including Jim Woodward, Clark Glymour, Richard Scheines, and Peter Spirtes. The psychological application of these ideas has been developed by Alison Gopnik.[25] The controversial idea to look at here is that thinking based on a rich notion of cause is different from thinking based on mere extraction and extrapolation of patterns. That is, Carnapian and Pearlian creatures might not coincide, and one might go far down the Carnapian road without becoming a Pearlian creature, or vice versa.

Here is an example used by Taylor to make a point like this (drawing on Tomasello and Call, 1997).[26] Suppose you see the wind move a tree branch, and some fruit falls. Humean creatures might learn an association between wind and fallen fruit, which could be useful. The next time they feel wind, they expect to find fruit. The extraction of patterns from what is seen might be very sophisticated (they might track how much wind tends to predict which particular kinds of fruit will fall, and so on). But a different sort of creature could see the fruit fall and realize, Aha! I can just move the tree branch myself! No need to wait for the wind.

If you just happened to *do* a tree shake and fruit resulted, you could succeed in this context as a Skinnerian creature. You could learn to go around shaking trees. That requires that you have some reason to perform the action the first time. As a Skinnerian, you might do things of this kind at random. But that is not very efficient, and a Pearlian creature does not need to do this. From the experience of seeing the wind shaking the tree, the animal can work out that the same effect can be achieved by an intervention. The creature who does this might not be thinking very abstractly—might not be far down the Carnapian road—but is thinking causally, and that's a powerful thing.

Which animals are Pearlian creatures? Taylor, working with a team of bird people and causal reasoning people (2014), did an experiment with New Caledonian crows that was designed to probe this question, and he did not find Pearlian behaviors. In contrast, two-year-old children passed the test. Their experiment set up a situation analogous to the wind–fruit scenario. The crows were trained on a puzzle box in a way that enabled them to sometimes get food but in a way that

25. See Pearl (2000); Spirtes et al. (2000); Gopnik and Schulz (2007); Woodward (2005),

26. See Tomasello and Call (1997).

should have made it clear to a Pearlian creature that some novel manipulations (akin to shaking the branch) were available. Children picked up on the shortcuts and crows did not. The crows, in contrast, seemed able to approach the problem with instrumental conditioning—rewarded behaviors were repeated—but that's all they could manage.[27]

I've distinguished two ways of becoming more sophisticated than an ordinary Skinnerian creature, the Carnapian and the Pearlian roads. Perhaps some facility with abstraction is needed to become a Pearlian, perhaps not. Either way, I suggest that once an animal has the Pearlian capacity, there's a natural path to becoming a *Popperian creature* in Dennett's sense. The crucial step to this sort of intelligence is the ability to run an experiment involving action offline, to work out what would happen if I were to do *X*. If an animal can choose the most efficacious action in this way without performing it first, then the trial-and-error mechanism of Darwin and Skinner has been entirely internalized. As Popper said, this capacity enables our theories to "die in our stead."[28] Not only need the Popperian organism not die for its experiments, but it need not encounter aversive experiences, either. The aversive experiences need only be envisaged.

Being a Popperian creature in this sense is quite demanding. People are such creatures. I take it that it's not completely clear that any other animals are, though some mammals and birds are possibilities. Dennett, in some of his discussions, sets a much lower bar for being a Popperian creature. He says that you're a Popperian as long as you don't emit new behaviors completely blindly, but use information from the environment to filter which actions are chosen and exposed to the perils of reinforcement. If this is the bar, then as Dennett says, fish, primates, and many other animals are Popperian. But here I think Dennett mishandles his own hierarchy. If the bar to clear in order to become Popperian is just that there be some experiential filtering of which behaviors are initially performed, then even classical conditioning suffices. In these discussions, Dennett's "filter" on behaviors apparently need not be a consequence-assessor. But there *does* have to be an internal consequence-assessor in order for the mechanism to retain a variation-and-selection character. It's probably quite difficult for an animal to fully internalize the "test" part of the generate-and-test mechanism, as opposed to just having *some* way of shaping its other actions other than reinforcement of actual performance. And having a separation between a generate phase and a test

27. Taylor et al. (2014). "Aesop's fable" cases are also possibilities (Bird & Emery, 2009).

28. Popper said this in his Darwin lecture, "Natural Selection and the Emergence of the Mind," November 8, 1977, Cambridge University, England.

phase, where the test phase is truly internalized, may require a capacity to run internal experiments using causal reasoning.[29]

An objection that might be raised at this point is that a Pearlian creature, in the sense I have described, *is* a Popperian one. There is no way to achieve a "grasp" of causal relations and put this knowledge to use in an intervention if not through a Popperian generate-and-test mechanism. Perhaps that's true, but I leave open the possibility that there can be causal cognition of a substantial and definite kind without the offline experimentation characteristic of the Popperian. Another objection that might be made is that one can get to the Popperian mechanism without using a strong, intervention-based notion of causation, via a purely Carnapian road. Again, perhaps that's true, but it might not be. I suspect that the ability to try out behavioral options offline has no obvious relation to abstraction—it might be done while only noting simple relations between objects. We see that in Taylor's experiment with babies and crows.

I'll make one more point about this trio of categories: Carnap, Pearl, Popper. Some readers might have wondered before now about Wolfgang Köhler's (1924) notion of *insight*. Köhler, working in the early twentieth century, thought that some of the behaviors he saw in chimps showed a grasp of the causal properties of objects, and an ability to use those objects to achieve novel solutions to problems (such as stacking boxes to climb on and reach a high banana). The idea of insight has often been seen as suggestive but problematically vague. I think insight might best be viewed as some sort of combination of the things I've discussed so far in this section—causal cognition, abstraction, offline experimentation. It's not a definite psychological kind—there are no "Köhlerian creatures"—but various Köhlerian behaviors arise in different ways from this family of mechanisms.[30]

I'll now introduce another path to the roads beyond Skinner. All the mechanisms discussed in this section are individualistic. Another piece of the picture is social learning. In Dennett's tower, *Gregorian creatures* are found above Popperians. They improve their generate-and-test abilities with socially based tools for thinking, such as language. But once social learning is put on the table, we need to consider it in its own right. Social learning is not something that arises "after" Popperian mechanisms, the next level of a tower. It's more widespread, and reasonably seen as a distinct step away from purely Skinnerian creatures, with its

---

29. A literature that must bear on this—but whose exact bearing I've not yet been able to work out—is the literature on "counterfactual" thinking in nonhumans, including rats, and also a literature on the role of "fictive," as opposed to actual, rewards gained by an animal. See, respectively, Laurent and Balleine (2015); and Kim et al. (2015). This is a topic for discussion on another occasion.

30. See Shettleworth (2012).

own role in behavioral adaptation. One thing an animal can do is make use of another's history of trial and error, taking on that animal's behavioral adaptations as its own.

The role of simple forms of social learning was illuminated by an elegant model due to Alan Rogers (1989). Assume a population in which everyone is learning a behavior by trial and error of a Skinnerian kind. You can economize on the process, and behave just as adaptively as the others, by copying the behavior that is widespread around you. A population of Skinnerian learners can be invaded by an imitating mutant. I'll call this a *mimetic creature*. These creatures, who blindly copy, can invade Skinnerians, who pay the cost of trial and error. But when the mimetic creatures become very common, their copying becomes less effective, as they are no longer reliably copying someone who really knows what to do—the model assumes that the environment changes from time to time, and with it the appropriate behavior changes, too. Eventually, Skinnerians will bounce back. They pay the cost of learning, but when they are rare, they are the only ones likely to be performing the right behaviors. An equilibrium will be reached. Rogers also showed that (given some other reasonable assumptions) the average fitness in the population when the Skinnerians and mimetics are at equilibrium is the same as it is when the population consists only of Skinnerians.

Recent work has shown the presence of imitation in some surprising cases—animals without notable social lives. Examples include the red-footed tortoise (Wilkinson, Kuenstner, Mueller, & Huber, 2010), a stingray (Thonhauser, Gutnick, Byrne, Kral, Burghardt, & Kuba, 2013), and more controversially, an octopus (Fiorito & Scotto, 1992). These findings are used, with other arguments, by Cecilia Heyes (2011) to argue that "social" learning is not as distinctive in its mechanisms as people often suppose; she thinks that in at least many cases, it arises from the use of ordinary associative processes in a social context. If that is true, setting up a distinct category of "mimetic" creatures would be a bit misleading.[31]

Distinctions can be made between simpler and more sophisticated forms of imitation—you can merely copy what is common or visible around you, or instead copy only what seems to be successfully employed by others. The evolutionary consequences of sophisticated forms of imitation can be enormous, as Michael Tomasello has argued in his account of human cognitive change (1999). The distribution of simple forms of imitation above shows that this is another trait that has almost certainly evolved independently several times. That's undoubtedly true if the octopus results stand up but also likely given the presence

31. She does think that social learning may often involve specializations on the input side—some animals pay special attention to the behavior of others.

of this behavior in stingrays, whose common ancestor with mammals dates from 450 million years ago or earlier.[32]

I'll add one more broad category. In 1948, E. C. Tolman introduced the idea of "cognitive maps," internal representations used by animals in tasks like navigation. Evidence for them was found in behaviors that do not conform to standard behaviorist models of conditioning (especially Skinnerian processes). Examples include shortcut behaviors (which can be seen as spatial versions of "insight"). Map-making in this sense is not restricted to dealing with physical space, and though this is controversial, the construction of some map-like structures is probably not explainable with instrumental learning of the kind discussed here, or other standard associationist algorithms (Gallistel & King, 2010). The discovery of "place cells" and their role in the rat hippocampus has vindicated Tolman's original insights (O'Keefe & Nadel, 1978; Ólafsdóttir, Barry, Saleem, Hassabis, & Spiers, 2015). There's also now evidence for capacities of this kind in some invertebrate groups. The experimental challenge is to rule out explanations of navigation in terms of the use of chemical cues, using a single landmark near the destination, or "path integration." Honeybees, however, can find their way home after several disruptions that would affect these simpler methods (Cheeseman et al., 2014). Some suggestive, though less systematic, work on octopus foraging has been done by Jennifer Mather (1991). Octopuses follow long, looping paths when they forage, and they reliably arrive home even in turbulent waters, often approaching their den from a different direction from the one they went out on. It's not known how this is done, but it might well indicate a capacity for mental mapping of a Tolmanian kind.

When the idea of "mapping" is understood very abstractly, the distinction between the Tolmanian and Skinnerian categories may become blurred in another way. Much work on both classical and instrumental conditioning now looks at "model-based" strategies, in which experience is used not only to associate one event with another but to build a minimal inner model of the dependency relations in the environment that give rise to experienced events.[33] Inner mapping or modeling might thus underly some behaviors usually associated with associationism.

4. Two Trees

More categories might be added, especially in the area of social cognition. A longer discussion might next consider the role of language and its internalization as

32. See also Moore (2004). He counts many independent origins.

33. Doll, Simon, and Daw (2012); Dayan and Berridge (2014).

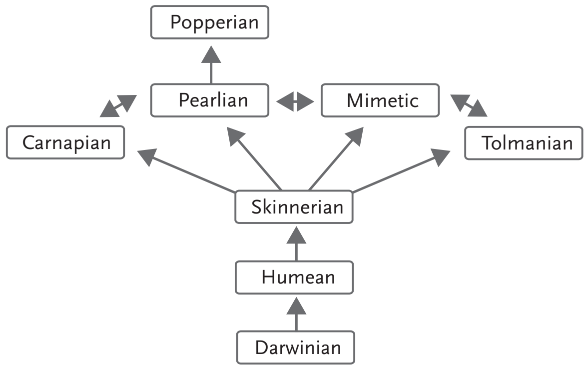

FIGURE 2. A "tower," dissolving into a network structure with both branchings and rejoinings, making use of the categories discussed in the text. The paths leading out of the "Skinnerian" node are not exclusive; an animal might treat two or more of them at once, and I assume that "lateral" movement between the post-Skinnerian categories is also possible (including movement that bypasses nodes that happen to be adjacent in the figure—those adjacency relationships are not supposed to be significant).

a cognitive tool (*Vygotskian creatures*). But I'll stop the story here, sum up, and consider some morals.

Figure 2 charts the relationships discussed in the last few sections. The main role of the figure is to distinguish between a tower-like structure at the early stages and a branching structure later on. I say "branching," but the shape is not strictly a tree, because some paths can rejoin after diverging.

There is empirical support for some of the relations represented, though it's not clear which are merely de facto and which are a matter of principle. It's likely that there are no Skinnerian creatures who are not also Humean and Darwinian. So initially, there is a nesting of categories. At present, though, it appears likely that the categories beyond the Skinnerian level are related in a less cumulative way. It would be hard to show this decisively, because it depends on the absence of capacities in crucial cases. But a possible picture may include relationships like these: bees can map the world, abstract, and show a form of imitation, but they don't use a rich notion of cause.[34] Octopuses can imitate and map but may have little use for abstraction. Stingrays and tortoises can imitate. Perhaps they can map and abstract, but it would not be surprising if they can't. I accept, on the other hand, that further work might uncover a single crucial innovation that underlies

34. For bee social learning, see Avarguès-Weber and Chittka (2014).

several of the branches I picture emerging from the Skinnerian category.[35] It's also true that my "Carnapian" category is especially vague; I introduced it partly to make vivid a contrast between mere pattern-recognition and cognition that makes use of a rich concept of cause. The use of this concept of cause is both powerful and uncertain in its relationships to precursor forms.

A further set of questions concern how the structure in Figure 2 relates to the tree of animal evolution represented in Figure 1. Classical conditioning is so widespread that it might have a single origin, or perhaps a couple, in early neural animals. It is also simple in its demands in a way that makes it look relatively easy to invent, however—at least in *some* form—and that weakens the case for a single origin. Instrumental conditioning, in contrast, appears to be more scattered on the tree, though absences of evidence should once again be handled with caution. I've not made much use of molecular evidence in this paper, and this can also bear on questions about single versus multiple origins; Barron et al. (2010) note, for example, that dopamine plays a role in reward-seeking behavior in a great many animal groups. The traits in Figure 2 that go beyond instrumental learning, such as mapping and imitation, also appear to be scattered with respect to the tree. Imitation probably arose independently, for example, in stingrays, tortoises, and perhaps octopuses. Spatial mapping is seen in bees and rats, and has an obvious rationale there, but it may well be absent in many animals who don't need it.

It seems likely that Darwinian evolution built Humean creatures before it built Skinnerian ones, and it built those on a Humean base. Things *might* not be this way—it's possible that they evolved in the other order, or together, and then instrumental learning was lost, diluted, or hidden in a range of animals. But that looks a bit unlikely. Instead, it appears that evolution first built a form of within-generation adaptation that is not a generate-and-test mechanism, but one that works by other means—by "instructive" rather than selective processes, as people used to say.[36] Selection does have a kind of overall primacy in the story—I agree with Dennett about that. But when selection built the first of this sequence of psychological adaptations, it built not more selection but something else. After that, evolution did build another selection process—instrumental learning. Then it built several other things, eventually including the third, internalized, Popperian form of selection. In filling out the story, reward seeking that involves the moment-to-moment shaping of behavior instead of than learning also has

35. Heyes, again, argues that merely fine-tuning the input mechanisms can make a Skinnerian creature look quite different (2011).

36. Godfrey-Smith (1996).

to come in somewhere. That is probably older, part of the basis for instrumental learning, and something with its own glimmer of generate-and-test.

Continuing work may also bring more dramatic changes to the picture. I'll mention one example to finish. Brian Dias and Kerry Ressler reported in 2014 that a classically conditioned fear response to an odor in mice was transmitted, through sperm, to children and grandchildren of the conditioned animals. The mechanism appears to involve DNA methylation. Assuming the finding holds up, this is pretty remarkable, and appealingly disruptive with respect to the picture sketched in this paper. Classical conditioning, again, is an "instructive" mechanism, in the terms of the old selective/instructive distinction, and here classical conditioning is part of a means by which a behavior is shaped transgenerationally by a Lamarckian process. (I think it's common to use the term "Lamarckian" too broadly and freely, but this case surely counts.) Two "instructive" mechanisms for change then combine, and a Humean capacity, one step up on the tower, reaches down and affects intergenerational adaptation, otherwise the province of Darwin. Perhaps we are also Escherian creatures, as Dennett's collaborator Douglas Hofstadter has, in a different sense, long argued.[37]

Acknowledgments

Thanks to Alex Taylor, Bryce Huebner, Tony Dickinson, John Greenwood, and Celia Heyes for help with this paper.

Works Cited

Ardiel, E. L., & Rankin, C. H. (2010). An elegant mind: Learning and memory in *Caenorhabditis elegans*. *Learning and Memory*, *17*, 191–201.

Avarguès-Weber, A., Dyer, A. G., Combe, M., & Giurfa, M. (2012). Simultaneous mastering of two abstract concepts by the miniature brain of bees. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 7481–7486.

Avarguès-Weber, A., d'Amaro, D., Metzler, M., & Dyer, A. (2014). Conceptualization of relative size by honeybees. *Frontiers in Behavioral Neuroscience*, *8*. doi:10.3389/fnbeh.2014.00080

Avarguès-Weber, A., & Chittka, L. (2014). Local enhancement or stimulus enhancement? Bumblebee social learning results in a specific pattern of flower preference. *Animal Behaviour*, *97*, 185–191.

Bain, A. (1859). *Emotions and the will.* London: J. W. Parker and Son.

37. Hofstadter (1979).

Balleine, B. W., Espinet, A., & Gonzalez, F. (2005). Perceptual learning enhances retrospective revaluation of conditioned flavor preferences in rats. *Journal of Experimental Psychology: Animal Behavior Process*, *31*, 341–350.

Barron, A. B., Søvik, E., & Cornish, J. L. (2010). The roles of dopamine and related compounds in reward-seeking behavior across animal phyla. *Frontiers in Behavioral Neuroscience*, *4*, 163. doi:10.3389/fnbeh.2010.00163

Bird, C. D., & Emery, N. J. (2009). Rooks use stones to raise the water level to reach a floating worm. *Current Biology*, *19*, 1410–1414. doi:10.1016/j.cub.2009.07.033

Budd, G. E., & Jensen, S. (2015). The origin of the animals and a "savannah" hypothesis for early bilateral evolution. *Biological Reviews.* Online publication. Retrieved from doi:10.1111/brv.12239

Campbell, D. T. (1974). Evolutionary epistemology. In P. A. Schlipp (Ed.), *The philosophy of Karl R. Popper* (pp. 412–463). La Salle, IL: Open Court.

Cheeseman, J. F., Millar, C. D., Greggers, U., Lehmann, K., Pawley, M. D., Gallistel, C., . . . & Menzel, R. (2014). Way-finding in displaced clock-shifted bees proves bees use a cognitive map. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 8949–8954. doi:10.1073/pnas.1408039111

Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cognitive, Affective, and Behavioral Neuroscience*, *14*, 473–492. doi:10.3758/s13415-014-0277-8

Dennett, D. C. (1969). *Content and consciousness*. New York, NY: Routledge.

Dennett, D. C. (1975). Why the law of effect will not go away. *Journal for the Theory of Social Behaviour*, *5*, 169–188.

Dennett, D. C. (1978). *Brainstorms: Philosophical essays on mind and psychology*. Montgomery, VT: Bradford Books.

Dennett, D. C. (1995). *Darwin's dangerous idea: Evolution and the meanings of life*. New York, NY: Simon and Schuster.

Dennett, D. C. (1996). *Kinds of minds: Toward an understanding of consciousness*. New York, NY: Basic Books.

Dias, B. G., & Ressler, K. J. (2014). Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nature Neuroscience*, *17*, 89–96.

Dickinson, A. (2012). Associative learning and animal cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 2733–2742. doi:10.1098/rstb.2012.0220

Dickinson, A., & Balleine, B. (1994). Motivational control of goal-directed action. *Learning and Behavior*, *22*, 1–18. doi:10.3758/BF03199951

Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology*, *22*, 1075–1081. doi:10.1016/j.conb.2012.08.003

Fiorito, G., & Scotto, P. (1992). Observational learning in *Octopus vulgaris*. *Science*, *256*, 545–547.

Gagliano, M., Renton, M., Depczynski, M., & Mancuso, S. (2014). Experience teaches plants to learn faster and forget slower in environments where it matters. *Oecologia*, *175*, 63–72. doi:10.1007/s00442-013-2873-7

Gallistel, C. R., & King, A. P. (2010). *Memory and the computational brain: Why cognitive science will transform neuroscience.* Chichester, England: Wiley-Blackwell.

Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence in avoiding learning. *Psychonomic Science*, *4*, 123–124.

Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge, UK: Cambridge University Press.

Gopnik, A., & Schultz, L. (Eds.). (2007). *Causal learning: Psychology, philosophy, and computation*. Oxford, UK: Oxford University Press.

Greenwood, J. D. (2009). *A conceptual history of psychology*. Boston, MA: McGraw-Hill.

Haralson, J. V., Groff, C. I., & Haralson, S. J. (1975). Classical conditioning in the sea anemone, *Cribrina xanthogrammica*. *Physiology and Behavior*, *15*, 455–460.

Heyes, C. (2011). What's social about social learning? *Journal of Comparative Psychology*, *126*, 193–202.

Hofstadter, D. R. (1979). *Gödel, Escher, Bach: An eternal golden braid.* New York, NY: Basic Books.

Hume, D. (1739/2000). *A treatise of human nature.* Oxford, UK: Oxford University Press.

Kasumovic, M. M., Elias, D. O., Sivalinghem, S., Mason, A. C., & Andrade, M. C. B. (2010). Examination of prior contest experience and the retention of winner and loser effects. *Behavioral Ecology*, *21*, 404–409. doi:10.1093/beheco/arp204

Kim, K. U., Huh, N., Jang, Y., Lee, D., & Jung, M. W. (2015). Effects of fictive reward on rat's choice behavior. *Scientific Reports*, *5* [Article No. 8040]. Retrieved from doi:10.1038/srep08040

Köhler, W. (1924). *The mentality of apes.* New York, NY: Harcourt Brace.

Lamarck, J. B. (1809/1914). *Zoological philosophy* (Hugh Elliot, Trans.). London: Macmillan.

Laurent, V., & Balleine, B. W. (2015). Factual and counterfactual action-outcome mappings control choice between goal-directed actions in rats. *Current Biology*, *25*, 1074–1079. doi:10.1016/j.cub.2015.02.044

Lyon, P. (2015). The cognitive cell: Bacterial behavior reconsidered. *Frontiers in Microbiology*, *6*, 264. doi:10.3389/fmicb.2015.00264

Mather, J. (1991). Navigation by spatial memory and use of visual landmarks in octopuses. *Journal of Comparative Physiology A*, *168*, 491–497.

Moore, B. R. (2004). The evolution of learning. *Biological Reviews*, *79*, 301–335.

Nuttley, W. M., Atkinson-Leadbeater, K. P., & van der Kooy, D. (2002). Serotonin mediates food-odor associative learning in the nematode *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 12449–12454.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford, UK: Oxford University Press.

Ólafsdóttir, H. F., Barry, C., Saleem, A. B., Hassabis, D., & Spiers, H. J. (2015, June). Hippocampal place cells construct reward-related sequences through unexplored space. *ELife*, *4*, e06063. Retrieved from http://dx.doi.org/10.7554/eLife.06063

Pearl, J. (2000). *Causality: Models, reasoning, and inference.* Cambridge, England: Cambridge University Press.

Perry C., Barron A., & Cheng, K. (2013). Invertebrate learning and cognition: Relating phenomena to neural substrate. *WIREs Cognitive Science*. doi:10.1002/wcs.1248

Peterson, K., Cotton, J., Gehling, J. G., & Pisani, D. (2008). The Edicaran emergence of bilaterians: Congruence between the genetic and the geological fossil records. *Philosophical Transactions of the Royal Society B*, *363*, 1435–1443.

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review*, *74*, 71–80.

Rogers, A. R. (1989). Does biology constrain culture? *American Anthropologist*, *90*, 819–831.

Rudin F. S., & Briffa, M. (2012). Is boldness a resource-holding potential trait? Fighting prowess and changes in startle response in the sea anemone, *Actinia equine*. *Proceedings of the Royal Society B*, *279*, 1904–1910.

Ryan, J. F., Pang, K., Schnitzler, C E., Nguyen, A., Moreland, R T., Simmons, D K., . . . & Baxevanis, A D. (2013). The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution. *Science*, *342*, 124259. doi:10.1126/science.1242592

Shanks, A. (2002). Previous agonistic experience determines both foraging behavior and territoriality in the limpet *Lottia gigantea* (Sowerby). *Behavioral Ecology*, *13*, 467–471. doi:10.1093/beheco/13.4.467

Shettleworth, S. J. (2012). Do animals have insight, and what is insight anyway? *Canadian Journal of Experimental Psychology*, *66*, 217–226. doi:10.1037/a0030674

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search.* Cambridge, MA: MIT Press.

Taylor, A. H., Cheke, L. G., Waismeyer, A., Meltzoff, A. N., Miller, R., Gopnik, A., . . . & Gray, R. D. (2014). Of babies and birds: Complex tool behaviours are not sufficient for the evolution of the ability to create a novel causal intervention. *Proceedings of the Royal Society B*, *281*, 20140837. doi:10.1098/rspb.2014.0837

Terao, K., Matsumoto, Y., & Mizunami, M. (2015). Critical evidence for the prediction error theory in associative learning. *Scientific Reports*, *5* [Article No. 8929]. Retrieved from doi:10.1038/srep08929

Thonhauser, K. E., Gutnick, T., Byrne, R. A., Kral, K., Burghardt, G. M., & Kuba, M. J. (2013). Social learning in cartilaginous fish (stingrays *Potamotrygon falkneri*). *Animal Cognition*, *16*, 927–932.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M, & Call, J. (1997). *Primate Cognition*. Oxford: Oxford University Press.

Wells, M. J. (1968). Sensitization and the evolution of associative learning. *Symposium on the Neurobiology of Invertebrates 1967*, 391–411.

Wilkinson, A., Kuenstner, K., Mueller, J., & Huber, L. (2010). Social learning in a non-social reptile (*Geochelone carbonaria*). *Biology Letters*, *6*(5): 614–616. doi:10.1098/rsbl.2010.0092

Woodward, J (2005). *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.