E-Article

The Evolutionary Origin of Associative Learning

Anselmo C. Pontes,[1,2,*] Robert B. Mobley,[1,3] Charles Ofria,[1,2] Christoph Adami,[1,4] and Fred C. Dyer[1,3]

1. BEACON Center, Michigan State University, East Lansing, Michigan 48824; 2. Department of Computer Science and Engineering, Michigan State University, East Lansing, Michigan 48824; 3. Department of Integrative Biology, Michigan State University, East Lansing, Michigan 48824; 4. Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan 48824

Submitted October 23, 2018; Accepted July 15, 2019; Electronically published November 26, 2019

Online enhancements: supplemental PDF. *Dryad data:* https://doi.org/10.5061/dryad.f45gh6s.

ABSTRACT: Learning is a widespread ability among animals and, like physical traits, is subject to evolution. But how did learning first arise? What selection pressures and phenotypic preconditions fostered its evolution? Neither the fossil record nor phylogenetic comparative studies provide answers to these questions. Here, we take a novel approach by studying digital organisms in environments that promote the evolution of navigation and associative learning. Starting with a nonlearning sessile ancestor, we evolve multiple populations in four different environments, each consisting of nutrient trails with various layouts. Trail nutrients cue organisms on which direction to follow, provided they evolve to acquire and use those cues. Thus, each organism is tested on how well it navigates a randomly selected trail before reproducing. We find that behavior evolves modularly and in a predictable sequence, where simpler behaviors are necessary precursors for more complex ones. Associative learning is only one of many successful behaviors to evolve, and its origin depends on the environment possessing certain information patterns that organisms can exploit. Environmental patterns that are stable across generations foster the evolution of reflexive behavior, while environmental patterns that vary across generations but remain consistent for periods within an organism's lifetime foster the evolution of learning behavior. Both types of environmental patterns are necessary, since the prior evolution of simple reflexive behaviors provides the building blocks for learning to arise. Finally, we observe that an intrinsic value system evolves alongside behavior and supports associative learning by providing reinforcement for behavior conditioning.

Keywords: associative learning, origin of learning, evolution of behavior, digital evolution, evolutionary transitions, artificial intelligence.

* Corresponding author; email: pontes@msu.edu.

ORCIDs: Pontes, https://orcid.org/0000-0003-0671-4437; Mobley, https://orcid.org/0000-0001-9114-1076; Ofria, https://orcid.org/0000-0003-2924-1732; Adami, https://orcid.org/0000-0002-2915-9504.

Am. Nat. 2020. Vol. 195, pp. E1–E19. DOI: 10.1086/706252

Introduction

Associative learning has long been considered fundamental to the adaptability of behavior and development of knowledge about the world (Hume 1738). It is also widely assumed that associative learning emerged as animal behavior evolved greater complexity and may have provided new avenues for this complexity to increase (Godfrey-Smith 1996; Weber and Depew 2003; Duckworth 2009; Ginsburg and Jablonka 2010; Brown 2013). The general fitness advantage of learning in living organisms seems clear: learning enables an organism to adapt its behavior during its lifetime without requiring genetic changes across generations (as with evolution), and, unlike other forms of behavioral plasticity that occur during development, learning can result in very rapid rather than gradual behavioral modifications (Dennett 1996; Dukas 2013). Most research on the evolution of learning has focused on the adaptive specialization of learning—how the speed of learning, biases to learn certain things better than others, and capacity to store learned information correlate with the reliance on learning in an organism's natural environment (Seligman 1970; Stephens 1991; Mery and Kawecki 2002; Dukas and Ratcliffe 2009; Shettleworth 2010; Domjan 2012). Little is known, however, about the historical question of what selection pressures and evolutionary precursors facilitated the emergence of learning from ancestors incapable of doing so or about the processes that allowed more complex forms of learning to evolve from simpler ones (Moore 2004; Dunlap et al. 2019).

Most people assume that complex behavior evolves in response to complex challenges; however, the evolution of behavioral complexity need not entail the emergence of learning (Godfrey-Smith 2002). Rather, learning evolves under specific environmental dynamics: where conditions that are relevant to the organism's fitness change on the timescale of generations but remain relatively stable within an individual's lifetime (Stephens 1991). Furthermore, in the particular case of the evolution of associative learning, there must also be learnable cues that reliably correlate with the state of the

environment (Dunlap and Stephens 2009). In this situation, organisms may benefit if they use those cues to track current conditions and map them to appropriate responses. Since the environment and cues may change between generations, the mapping cannot be encoded genetically and must be learned during the organism's lifetime.

Researchers have explored the factors of environmental dynamics and cue availability that are necessary for the evolution of associative learning using both mathematical and empirical approaches (Stephens 1991; Dunlap and Stephens 2009). However, it is still an open question whether these factors are sufficient for associative learning to emerge during the evolution of an organism's behavioral repertoire. Here, we propose:

> Hypothesis 1. The initial evolution of associative learning depends on the scaffolding provided by the prior evolution of a repertoire of instinctual behaviors that exploit stable environmental patterns.

Skinner and others speculated that complex behavioral traits do not evolve independently of each other but build on preexisting ones according to a characteristic evolutionary sequence that starts with simple movement, then sensing, followed by tropisms and reflexes, and finally learning (Skinner 1984; Miller and Todd 1991). Similarly, it has been suggested that different forms of learning are not independent but evolve from one another in a specific sequence, where more complex forms build on the mechanisms of simpler ones and subsume them (Wells 1968; Razran 1971; Hawkins and Kandel 1984*a*, 1984*b*). For example, associative learning would have evolved from sensitization (Wells 1968; Razran 1971; Hawkins and Kandel 1984*a*, 1984*b*), a simpler, nonassociative form of learning where an organism increases its response to a repeated stimulus (van Duijn 2017). Therefore, we propose:

> Hypothesis 2. Complex behaviors, including learning, do not arise and function independently from one another. Instead, as more complex cognitive processes arise, they do so in a modular and stepwise manner, where early instinctual behaviors (such as moving and sensing) are co-opted and integrated into increasingly more complex ones (such as error recovery or path prediction) before eventually reaching associative learning.

It has also been speculated that the emergence of associative learning required only minor modifications in preexisting memory mechanisms (Hawkins and Kandel 1984*a*, 1984*b*; Ginsburg and Jablonka 2010), enabling it to evolve in parallel in different species (Ginsburg and Jablonka 2010). Thus, we propose:

> Hypothesis 3. Associative learning can arise suddenly, as a result of small modifications in preexisting cognitive mechanisms, as opposed to arising gradually and independently by accumulating incremental changes under selection.

Finally, given the expectation that environmental characteristics, such as stability and cue availability, shape the type of learning that evolves (Stephens 1991; Dunlap and Stephens 2009; Domjan 2012), we investigate an additional hypothesis on the flexibility of the associative learning mechanism that evolves in a particular environment. We propose:

> Hypothesis 4. Organisms that evolve associative learning will not be able to change established associations (e.g., reversal learning) unless such changes were necessary for success during evolution.

Our research focuses on a definition of associative learning that emphasizes its consequences for behavior rather than the mechanisms by which it works. We think this approach is justified because associative learning is traditionally defined in operational rather than mechanistic terms—for example, as "a behavioral modification, dependent on reinforcement, involving new associations between different sensory stimuli, or between sensory stimuli and responses" (Ginsburg and Jablonka 2010, p. 13)—and may not even be a unitary behavioral trait with consistent properties across species. For example, it is by no means clear that associative learning involves distinct mechanisms from those underlying simpler, nonassociative forms of learning, such as habituation and sensitization. In *Drosophila*, mutants incapable of associative learning also show reduced habituation and sensitization (Duerr and Quinn 1982), and in *Aplysia*, sensitization and associative learning share many of the same molecular elements (Roberts and Glanzman 2003). It is also not clear whether there is only one way of implementing associative learning mechanistically. All animals in which associative learning has been well established have a central nervous system (i.e., brains)—although many animal groups have not yet been tested (Ginsburg and Jablonka 2010)—but having a brain is not necessary for associative learning: plants are capable of it (Gagliano et al. 2016), and single-cell organisms may be as well (Armus et al. 2006; Fernando et al. 2009). These observations suggest that associative learning has evolved independently, acquiring different properties in different lineages (Moore 2004; Ginsburg and Jablonka 2010; Shettleworth 2010). Hence, they also justify the assumption that we can study the evolution of associative learning as a phenotypic attribute of behavior that is independent of a particular mechanistic implementation.

Major challenges arise in studying the evolutionary origin of learning. One challenge is the utter lack of fossil

evidence, especially from periods as remote as the Precambrian, when associative learning behavior is believed to have first evolved (Ginsburg and Jablonka 2010). Another is the difficulty of performing phylogenetic comparisons to study the origin, as opposed to the adaptive function, of behavioral traits. Although phylogenies are valuable to infer ancestral character states, sequences, and timing of evolution of traits, this approach is virtually silent on the selective forces and mechanisms involved (Ord and Martins 2010) and may suggest patterns of evolution that could result from multiple different processes (Losos 2011). In addition, associative learning is such a widespread and likely ancient behavior that it is particularly challenging to reconstruct an accurate phylogeny because of the lack of out-groups and because its origin presumably predates the rapid adaptive radiation of the Cambrian explosion (van Duijn 2017). The ubiquity of associative learning behavior among extant species is also a challenge for experimental evolution, which has been very successful in studying the adaptive modification of existing learning mechanisms in animals but can reveal little about the origins and early evolution of learning (Dunlap et al. 2019).

To overcome these limitations, here we study the origins of learning behavior in populations of self-replicating computer programs that undergo open-ended evolution in a virtual environment (Grabowski et al. 2010). These digital organisms are selected for their ability to cope with behavioral challenges in which associative learning may confer a fitness advantage; specifically, the environment provides alternative courses of action and cues that reliably correlate with the correct action, although these cues vary across generations (Dunlap and Stephens 2009).

This approach allows ample opportunities for a wide range of behaviors to evolve and enables the discovery of evolutionary principles that are potentially independent of the cognitive machinery that is undergoing evolution. We emphasize that digital evolution is not a simulation of evolution but rather an instantiation of it (Pennock 2007): although digital organisms are evaluated in simulated environments, their behavioral control algorithm undergoes actual Darwinian evolution. Specifically, (i) organisms reproduce and pass on their evolved traits, including their behavioral algorithm, to their offspring; (ii) inheritance is subject to mutations, producing variation; and (iii) individual fitness depends on an organism's performance at specific behavioral tasks and determines the outcome of the competition for space in a size-limited population. This approach enables true experimental study of evolutionary history across multiple replicate lineages evolving under different conditions, providing insights not only on the outcomes of evolution but also on the transitions that occur in different lines of descent. Digital evolution has a proven track record of expanding evolutionary theory (Wilke et al. 2001; Lenski et al. 2003; Chow et al. 2004) with supporting evidence often collected later in biological systems (Codoñer et al. 2006). Previous studies in Avida have also demonstrated the evolution of instinctive navigation, such as gradient ascent and trail-following behavior (Grabowski et al. 2008), including the genetically encoded use of memory to dictate subsequent behavior (Grabowski et al. 2010). Here, we extend this work beyond reflexive behaviors to study the evolution of associative learning where each individual organism must discover a mapping between environmental cues and the optimal response.

Our results support the aforementioned hypotheses and, moreover, provide a rich picture of the circumstances that favor—or disfavor—the evolution of learning, including the critical role played by historical contingency. Learning is a rare outcome of evolution in our system, not because of any intrinsic difficulty in the underlying computation but rather because oftentimes lineages evolve highly flexible behavioral strategies over which learning does not provide a strong selective advantage. When learning does evolve, it emerges via an almost stereotypical sequence, as proposed by Skinner and others (Skinner 1984; Miller and Todd 1991). Finally, we find that the evolution of behavior is inseparable from the evolution of an intrinsic value system, the innate gauge of an organism's experiences that provides positive or negative feedback on its actions.

Experimental System

We used the Avida digital evolution platform for all of our experiments (Ofria et al. 2009, 2015). Avida is a linear genetic programming platform, meaning that each organism's genome consists of an ordered sequence of computer instructions in a machine-like language. Instructions are simple, self-contained operations, such as adding two values, storing a value in memory, or skipping to another instruction if one value is greater than another. During evolution, random mutations occur that can insert, remove, or replace instructions in offspring. Note that any sequence of Avida instructions can be executed; as such, mutations will always produce valid programs even if their functionality may be meaningless.

In addition to the instructions for arithmetical and logical operations described above, we used a single instruction that caused the organism to reproduce as well as a set of instructions that acted as simple sensors and effectors to interact with the environment (described in the next section). Using Avida provided key benefits for the experimental study of evolution. For example, the set of instructions we used formed a Turing-complete programming language that, in theory, can represent any algorithm—including any behavioral control algorithm—given the necessary sensors and effectors. In addition, it is easy to analyze an Avida organism to dissect and study the behavioral control algorithms

that evolve. Furthermore, we can archive all ancestors and their evolutionary lines of descent to examine the evolutionary transitions that occurred along any lineage, allowing us to study patterns that reveal how one set of behaviors might potentiate another.

An Avida organism is defined by a sequence of instructions (its genome), and each particular sequence defines its genotype. In our experiments, each population was seeded with a "naive" organism that lacked any instruction for behavioral control other than the one necessary to reproduce. Such an organism's genome consisted of a sequence of null instructions that acted as placeholders for future behavioral "genes" and a single "reproduce" instruction. To reproduce, an organism had to execute a minimum number of instructions, that is, spend a minimum amount of time in the environment in order to mature. At the same time, an organism also had an upper limit in the number of instructions it could execute before it tried to reproduce, essentially creating a maximum age. If an organism failed to reproduce by the time this limit was reached, it was eliminated from the population. Reproduction was asexual and resulted in the production of two offspring, both inheriting a copy of the parent's genome. However, only one of the offspring was subject to mutation, while the other remained identical to the parent and essentially replaced it.

Populations were capped at 3,600 organisms. Once that limit was reached, every organism that was born resulted in an existing one being randomly removed. Organisms did not interact with each other in the environment; however, the age limit and the competition for space in the size-limited population created a strong selection pressure for fast reproduction.

How well an organism performed the behavioral task determined the rate at which its offspring's instructions were executed and consequently how quickly they could reproduce. Therefore, the better an organism performed on the behavioral task, the faster its offspring executed their behavioral algorithm and reproduced. Thus, behaviors evolved in this digital system in a purely Darwinian fashion.

The Behavioral Task

Bees, ants, and other insects are known to use local and distant landmarks for navigation (Dyer 1991, 1998; Collett and Collett 2002; Grabowski 2009; Grabowski et al. 2010). For example, experiments have shown that bees can learn visual cue associations to successfully navigate complex mazes (Zhang et al. 1996, 1999). Inspired by these experiments, the behavioral task that we presented to evolving Avida organisms consisted of navigating a trail of nutrients in a virtual arena (fig. 1), where nutrients provided cues that indicated the direction to follow—if organisms evolved the ability to sense and use them. An organism's task was to complete as much of the trail as possible and then reproduce before the end of its life. The system kept track of the organism's cumulative performance by counting the number of new nutrient locations it visited and subtracting the number of empty (i.e., off-trail) locations encountered. This count was then divided by the total number of nutrients in the trail to compute the organism's "task quality," which ranged from 0 to 1 (negative values were set to zero). Nutrient locations were counted only on the first visit; subsequent visits to the same nutrient location would not affect an organisms'

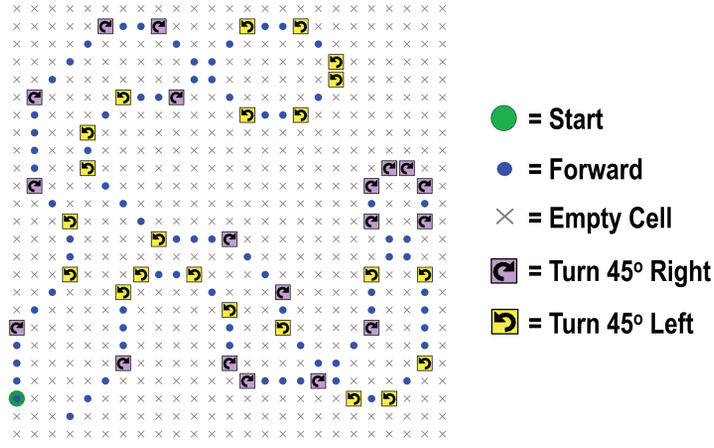


Figure 1: Sample arena and nutrient trail. Shown is one of four virtual arenas from an environment. Each virtual arena contained a single trail of nutrients laid out in a unique configuration. At the beginning of its life, each organism was placed alone at the start of the trail (green circle) in a randomly selected arena and oriented in the direction of the next nutrient.

task quality. However, visits to empty locations were always deducted. The organism had no sensory feedback about its task quality (i.e., cumulative performance), similar to the way a natural organism cannot sense its own fitness. Nevertheless, our organisms are limited relative to natural ones that may be able to measure payoffs of their foraging decisions by the rate of some physiological condition, such as gut fullness (Charnov 1976).

Each environment in our evolutionary experiments consisted of four virtual arenas, each with a different trail configuration (fig. 1). Every time an organism was born, it was randomly assigned to one of the four arenas, placed at the beginning of the trail on a nutrient location, and oriented in the direction of the next nutrient. The use of four trail configurations reduced the likelihood of an organism evolving a rigid control algorithm tailored to a single nutrient trail (genetically hardwiring a sequence of actions) instead of a flexible control algorithm that captures the principles of trail navigation.

Each of our experiments consisted of between 50 and 900 replicates. At the end of an experiment, we selected the predominant (most abundant) genotype from each replicate's final population for behavioral analysis. Given the large population size, the predominant genotype typically represented dozens of organisms, implying that they, on average, outperformed the rest of the population on all four trail configurations. Indeed, in these experiments we found that the predominant genotype typically had the highest task quality scores on each of the four trails; thus, we measured its performance by computing the population's average maximum task quality (AMTQ) scores across all trail configurations.

An organism's interaction with the environment depended on sensor and effector instructions acquired through mutation and maintained during evolution. These instructions conferred the abilities to sense the nutrient content of the current location ("sense current"), rotate right by 45 degrees ("rotate right"), rotate left by 45 degrees ("rotate left"), take one step ahead ("move ahead"), and take one step back while facing forward ("move back").

The execution of a sense current instruction provided feedback, in the form of an integer, about the nutrient content of the location the organism occupied. Empty locations and nutrients, in both straight portions of the trail and at turn points, were each sensed as different values. Therefore, the numeric value of the nutrients could cue the organism to the direction of the trail once they evolved the ability to interpret the sensed values correctly. There were four types of cues: right turn (45 degrees), left turn (45 degrees), forward, and empty location (fig. 1). Nutrients that indicated forward (forward cue) and empty location were always represented by the integers 0 and -1, respectively. Meanwhile, nutrients that indicated turns (turn cues) were each assigned a distinct random number between 1 and 100 every time an organism started on a trail, and this assignment persisted only during the organism's lifetime. Since forward cues and empty locations had persistent values between generations, organisms could evolve to use them to predict optimal future moves. However, the environmental uncertainty represented by randomized turn cues created an additional challenge for the organism: it not only had to move and follow the trail, but it also had to identify the direction represented by the turn cues. The optimal way to overcome this challenge was for an organism, within its lifetime, to associate either the right turn cue or the left turn cue with the correct action, thus identifying the opposite turn cue by exclusion.

Although one cannot predict the course of evolution, if associative learning evolved in our experiments, we expected to recognize it by observing the path of the organism along the trail. When placed on a new trail and allowed a period of exposure to the different turn cues, an organism capable of associative learning should be able to consistently turn to the correct direction every time it encounters a turn cue, something that would not be possible if the organism were using heuristics or choosing randomly.

Experimental Conditions

Experiment 1

In experiment 1, we tested four different environments, each with four possible trail configurations (table 1; figs. S1–S4; figs. S1–S19 are available online). In three of the environments, the trails of nutrients started with a simple (and presumably predictable) pattern (table 1). In the fourth environment, which served as a control, the trails provided nutrients in an unpredictable pattern—that is, each of the first two turns had equal probability of being to the right or to the left. This setup allowed us to test our first three hypotheses (as presented above). We performed 50 evolutionary replicates for each of the four conditions listed in table 1. See section S.1 of the supplemental PDF (available online) for additional details on methods.

Experiment 2

In experiment 2, we applied an additional selection pressure aimed at the evolution of reversal learning. We used only the nutrient cued environment but reversed the turn cues at approximately the 85% mark of each trail (fig. S5). In a complementary experiment, we tested different cue reversal positions ranging between 10% to 90% in 2.5% increments and found that it did not affect the results significantly (supplemental PDF, sec. S.4; fig. S8). Therefore, we report only the results for the 85% mark.

Table 1: Environments for experiment 1

	Predictable-start environments			Control environment
	One fixed turn	Two fixed turns	Nutrient cued	Random start
Trail start pattern	The first turn was always to the right in all four trails (fig. S1)	The first turn was always to the left and the second was always to the right in all four trails (fig. S2)	The direction of the first and second turns was random in all four trails but could be predicted by counting the number of forward cues preceding the first turn (an odd number meant left, while even meant right; fig. S3)	The direction of the first and second turns was random, and the number of forward cues preceding the first turn was the same in all four trails (fig. S4)

Note: Each environment contained four different trail configurations. An organism experienced only one trail configuration in its lifetime. See section S.1 of the supplemental PDF for images of each environment.

We performed 900 evolutionary replicates in experiment 2. The reason for the larger number of replicates than in experiment 1 was to generate a sufficient number of phenotypes for lineage studies, especially to explore the ancestry of the rare organisms that evolved reversal learning. A lineage study consists of singling out the final predominant organism of a population and reconstructing its line of descent, testing every ancestral genotype on the behavioral task to uncover how the behavior evolved over time. Although this experiment was designed to test hypothesis 4, it also enabled us to obtain additional evidence relevant to hypothesis 3. See section S.1 of the supplemental PDF for additional methodological details.

Raw data, code, and a video associated with this research are available in the Dryad Digital Repository (https://doi.org/10.5061/dryad.f45gh6s; Pontes et al. 2020). The custom version of Avida used in this study is available at https://github.com/mercere99/Avida-AssociativeMemory (Ofria et al. 2015).

Results

Repeated Evolution of Adaptive Behaviors: Error Recovery, Imprinting, and Reversal Learning

Our experiments resulted in the evolution of organisms capable of adapting to unpredictable environments by using a variety of strategies, including associative learning. We also observed the evolution of flexible strategies that did not rely on learning (table 2). We called the most successful nonlearning strategy "error recovery," in which an organism attempted to follow the nutrient trail and, on stepping off the trail, performed the necessary actions to return to it but did not modify its future behavior based on the error. A particularly notable result was the repeated evolution of associative learning, including both a rigid form that we called "imprinting" and a more flexible form that we called "relearning" (described in table 2).

We also found recurrent patterns in the behavioral strategies that evolved. Organisms from different evolutionary replicates, which inevitably had genotypes producing distinct behavioral control algorithms, generated a consistent set of behavioral phenotypes. We analyzed more than 300 out of 1,100 replicates across all experimental conditions and found, notably, that they all fell into five easily recognizable categories, including relearning, imprinting, and error recovery (previously mentioned), plus "searching" and "path predicting" (see table 2). We found some hybrids of these strategies as well.

The specific type of associative learning that evolved in our experiments was "instrumental conditioning," in which an organism forms an association between a stimulus and a behavior from its repertoire (Ginsburg and Jablonka 2010). Organisms that performed imprinting formed an association early in their lives that was used for future decisions but could never be modified. Organisms that performed relearning also formed associations between cues and actions early in their lives but were able to form new associations if the cues changed, regardless of whether they were swapped or replaced with novel ones. Additionally, we identified environmental factors and historical constraints that strongly influence whether associative learning evolves.

The ability to relearn when cues are swapped is called "reversal learning," a learning ability that is sometimes regarded as cognitively complex (Hadar and Menzel 2010; Bissonette and Powell 2012; Xue et al. 2013). A typical organism capable of reversal learning followed the trail of nutrients until it encountered a turn cue. Since the integers representing turn cues were randomly assigned for each generation, the organism then attempted to turn 45 degrees in a default direction and move forward one step. If this step led to a nutrient-containing location, the organism continued to follow the trail. However, if the organism turned in the "wrong" direction and found itself on an empty location, it engaged in a corrective reaction by taking one step back

Table 2: Behavioral strategies found in all experiments

Behavioral strategy	Description	Typical AMTQ
Learning behaviors:		
Relearning	A flexible and generalizable strategy, based on instrumental conditioning, that allowed organisms to navigate any trail configuration regardless of the starting pattern. Organisms using this strategy were able to re-form the cue-response association multiple times, even when the cues were reversed (fig. S15).	95%–97% (experiment 2; 97%–99% in non-cue-reversal environment)
Imprinting	A somewhat rigid strategy, based on instrumental conditioning, where organisms made the cue-response association early, and only once, in their lifetimes (fig. 2). These organisms were not able to relearn cues that were changed or reversed. This strategy was further classified into two subtypes: "generalizable" and "nongeneralizable." The generalizable subtype enabled organisms to imprint and navigate on any trail configuration regardless of the starting pattern, while the nongeneralizable subtype limited organisms to imprint on trails with the same starting pattern in which the organisms evolved.	97%–99% (experiment 1; 87%–92% in cue reversal environment)
Reflexive behaviors:		
Error recovery	A flexible and generalizable strategy that allowed organisms to navigate any trail configuration regardless of the starting pattern. These organisms did not discriminate between right and left turn cues. Instead, they reacted to turn cues by turning in one direction (each organism had its own default direction) and, if this direction led an organism off the trail, returning and trying the other direction (fig. 2).	78%–84% (experiments 1 and 2)
Searching	A strategy that appeared only in hybrid combinations with others, especially error recovery and path predicting. It was triggered if the organism stepped off the trail and typically involved performing several moving and turning steps to try to find another segment of the trail and rejoin it (fig. S7).	48%–60% (experiments 1 and 2)
Path predicting	A strategy where organisms encoded behavioral sequences in their genomes that matched the initial portion of different trails (fig. S12). This strategy enabled them to successfully navigate the first few segments of any trail but not the entire trail.	7%–14% (experiments 1 and 2)

Note: We analyzed more than 300 organisms across all experimental conditions. Although the details of their behavior differed among experiments, all behavioral phenotypes could be classified into five strategies or hybrids of two or more. The performance results for each behavioral strategy fell into a typical performance range, measured by average maximum task quality (AMTQ), and ranked from highest to lowest.

and turning 45 degrees twice (90 degrees) in the opposite direction (as if recoiling and turning away). The organism then made the association between the turn cue and correct action, such that the turn cue alone was sufficient to trigger the correct action in subsequent encounters. If the turn cues remained consistent, the organism navigated the remainder of the trail without error. Alternatively, if the turn cues changed further along the trail (including cue reversals) the organism again exhibited the corrective reaction and updated its association to the new cue, resuming the navigation without further error. Cues could be changed any number of times with the organism always relearning the turn cue and navigating without error until the cue changed again or it reached the end of the trail (fig. S15).

This serial reversal learning behavior evolved repeatedly, although it was a rare outcome, evolving in only 10 out of 900 replicates in experiment 2, where we specifically selected for reversal learning (and not at all in the 200 replicates in experiment 1, where reversal learning was not actively selected). Nevertheless, many replicates that did not result in the evolution of reversal learning still produced organisms that were able to efficiently navigate the entire trail using either imprinting or error recovery.

Early Trail Predictability Produces Behavioral Building Blocks for Learning

Although all environments could promote the evolution of simple controlled movement, not all of them could lead to the evolution of learning. All environments were constructed in a way that could potentially select for behavioral biases, such as moving along the trail of nutrients and avoiding

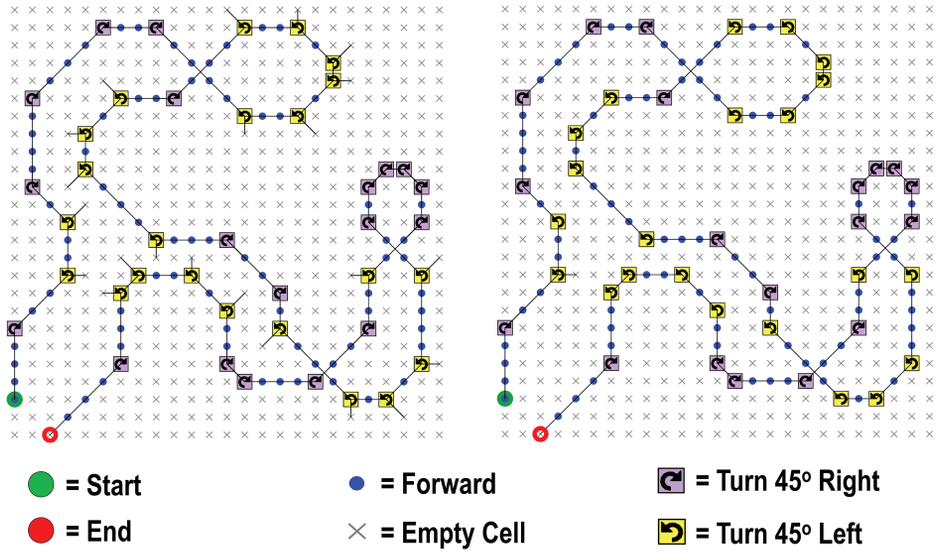


Figure 2: Two top-performing strategies in experiment 1. Shown are the paths of the final predominant organisms from two different replicates that evolved in the nutrient cued environment in experiment 1. Both were tested in the same trail configuration to facilitate comparison. In the left panel, an organism using an error recovery strategy achieved a task quality score of 81% of the maximum. Starting from the green circle, it moved straight while sensing forward cues but always tried to turn right (45 degrees) when sensing a turn cue. If turning right led the organism into an empty cell, it would retreat to the previous position and turn toward the left (90 degrees). It continued to repeat this behavior at every turn cue without ever learning from its error. In the right panel, an organism from a separate replicate using a generalizable imprinting strategy achieved a task quality score of 98% of the maximum. It also tried to turn right when sensing a turn cue. However, it stepped off the path only once at the first left turn. It learned the correct cue-response association and navigated the remainder of the trail without error.

empty locations, that contributed to the overall task performance. Indeed, the very first behaviors to evolve were simple forms of controlled movement, such as oscillatory behavior (moving back and forth) and moving to an edge of the path segment and stopping (see supplemental PDF, sec. S.4.1 for an example).

In addition, all environments provided organisms with the features thought necessary for learning to evolve: frequent choices of actions (move straight, turn right, or turn left) and cues that change each generation but reliably indicate the best choice within a generation (Dunlap and Stephens 2009, 2016; Dunlap et al. 2019). However, while these features were present in all environments, they proved insufficient to evolve learning. Specifically, no replicates in the random start environment produced organisms capable of learning (or even error recovery). In fact, none of the organisms from this environment were able to navigate past the first turn, and their task quality remained at or below 4% of the maximum across all 50 replicates. The environments in which learning did evolve (i.e., one fixed turn, two fixed turns, and nutrient cued) all had a property that the random start environment lacked: trails providing a high initial degree of predictability across generations that enable organisms to evolve behavioral building blocks and navigate the trail reflexively before evolving learning (Skinner 1984). These building blocks include moving repeatedly, sensing the current cue, distinguishing the different cues and reacting to them, turning to either side, retreating to the trail when an empty location is sensed, storing a cue in memory, and comparing the current cue with the one in memory. This result supports hypothesis 1.

The predictable-start environments (one fixed turn, two fixed turns, and nutrient cued; table 1; figs. S1–S3) were the only ones to evolve complex behaviors, including learning. These environments also produced a wider range of navigational strategies and organisms that reached substantially higher task quality than any organism in the random start environment (fig. 3; table 3). The nutrient cued environment produced the largest proportion of organisms that could navigate the entire trail, followed by the one fixed turn and the two fixed turns environments. These organisms used imprinting, error recovery, or a hybrid strategy (table 3). The organisms that achieved at least 25% AMTQ but did not complete the trail used the same strategies but performed more slowly, or simply reproduced before reaching the end of the trail (table 3).

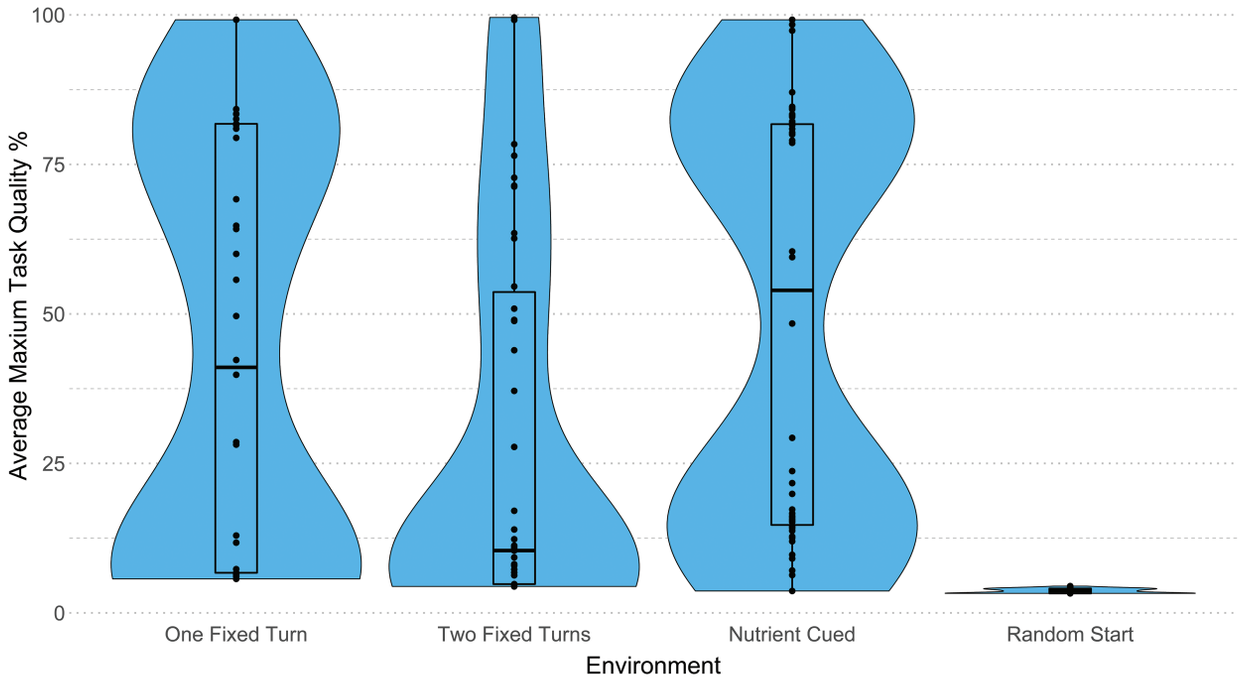


Figure 3: Distribution of average maximum task quality (AMTQ) per environment for experiment 1. Each violin plot represents the distribution of AMTQ across replicates for a given environment. Only the environments that started with a predictable pattern (one fixed turn, two fixed turns, and nutrient cued) evolved organisms that could finish the trail. They also produced a wider range of navigational strategies and organisms that reached much higher task quality than the control environment (random start).

Learning May Not Generalize to Novel Environments

Both imprinting and error recovery were successful strategies in experiment 1, but they differed in how well organisms could generalize to novel trails. Organisms that used error recovery did not depend on the pattern at the start of the trail for their navigation and could finish any trail configuration that we tested (fig. 2). In contrast, most of the organisms that used imprinting depended on the specific start pattern from the environment in which they had evolved to form the cue association. When tested in trails with a different start pattern, these organisms were not able to navigate far and scored poorly in task quality. However, two replicates in the nutrient cued environment evolved a generalizable version of imprinting that allowed the organisms to navigate any trail configuration independently of the starting pattern. These organisms began navigating the trail and, when sensing a turn cue, turned to a default direction. On making their first wrong turn and stepping off the trail, these organisms used error recovery to step back onto the trail and turn to the other direction. At this point, they imprinted the turn cue that led them astray and used the learned association to navigate the remainder of the trail (fig. 2). However, when tested in trails containing cue reversals or replacements, these organisms were not capable of coping with such changes and made wrong turns and stepped off the trail. They then resorted to using error recovery to get back on the trail and continue navigating until the end. This result led us to propose hypothesis 4, namely, that the environment has to present cue reversals along the trail to foster the evolution of more "complex" learning abilities, such as relearning and reversal learning (Hadar and Menzel 2010; Bissonette and Powell 2012; Xue et al. 2013), a hypothesis that we tested as part of experiment 2.

Cue Reversals during Evolution Foster the Ability to Relearn during a Lifetime

In experiment 2, we used only the nutrient cued environment because it was the only one where generalizable imprinting evolved in experiment 1. At approximately the 85% mark of each trail we swapped (reversed) the values associated with the turning cues, requiring the organism to learn to turn in the opposite direction of the one it learned at the beginning of the trail. We named this condition the "cue reversal" environment (fig. S5). In a complementary experiment, we tested varying the cue reversal position between

the 10% and the 90% mark, without any significant effect on the results (supplemental PDF, sec. S.4; fig. S8).

The results support hypothesis 4, although, as in previous experiments, the evolution of a complex learning ability proved to be a rare occurrence. Of 900 replicates, only 18 evolved the capacity for any form of reversal learning (fig. 4). In 10 of these 18 replicates, organisms also evolved the capacity for serial reversal learning, even though their ancestors had only experienced a single cue reversal in their lifetimes. In a serial reversal learning trial, the agent is confronted by a repeated reversal of a two-symbol combination. Organisms from the 10 replicates that could perform this task exhibited behavior that generalized to any trail configuration. In the other eight replicates, organisms evolved at least some capacity for reversal learning and relearning. However, their behavior had limitations, such as (i) being able to learn certain pairs of cues and not others, (ii) generalizing their behavior to some novel trail configurations and not others, or (iii) having a "short memory," which is "forgetting" the association after a while and needing to learn it anew on making a wrong turn. These limitations led to failures in staying on the trail, and in these cases organisms got lost or stuck outside the trail or resorted to navigating by error recovery or searching.

As in experiment 1, the fittest organisms (based on task quality) that evolved in experiment 2 were those that used learning strategies. The organisms capable of relearning scored as high as 97% of the maximum and were the fittest overall. Their behavior was similar to the generalizable imprinting that evolved in experiment 1 in that they made the association between the cue and the correct action on stepping off the trail, but they were also capable of relearning if a cue reversal led them off the trail. Intriguingly, these organisms could also relearn when tested in environments where an initial pair of turn cues was replaced by a completely different pair as well as when they were reversed or changed multiple times along the trail, even though we did not specifically select for this form of flexibility.

The next fittest organisms that were capable of learning employed various hybrid strategies involving imprinting, error recovery, and path predicting to reach task quality scores as high as 93% of the maximum. Although incapable of relearning per se (i.e., replacing a cue association with another), they were able to form temporary associations

Table 3: Experiment 1: summary of results

	Predictable-start environments			Control environment
	One fixed turn	Two fixed turns	Nutrient cued	Random start
	Replicates in which organisms finished the trail			
Proportion of replicates	18/50	13/50	23/50	0/50
Strategies evolved (no. replicates)	Imprinting (3) Error recovery (15)	Imprinting (5) Error recovery (7) Hybrid of path predicting and error recovery (1)	Imprinting (3[a]) Error recovery (20)	NA
Highest AMTQ observed (strategy)	99% (imprinting)	99.7% (imprinting)	99% (imprinting)	NA
	Replicates in which organisms did not finish the trail (AMTQ ≥ 25%)			
Proportion of replicates	9/50	4/50	4/50	0/50
Strategies evolved (no. replicates)	Imprinting (1) Error recovery (8)	Imprinting (2) Error recovery (1) Hybrid of error recovery and searching (1)	Error recovery (1) Path predicting (1) Hybrid of error recovery and searching (1) Hybrid of path predicting, searching, and imprinting (1)	NA

Note: Shown are the performance and strategies of the organisms with average maximum task quality (AMTQ) equal to or higher than 25%, organized by environment. We examined only a sample of organisms that had less than 25% AMTQ. Those that were examined displayed previously described strategies and did not travel far on the trail. NA = not applicable.

[a] Two of these organisms performed a generalizable version of the imprinting strategy that allowed them to navigate any trail configuration independently of the starting pattern.

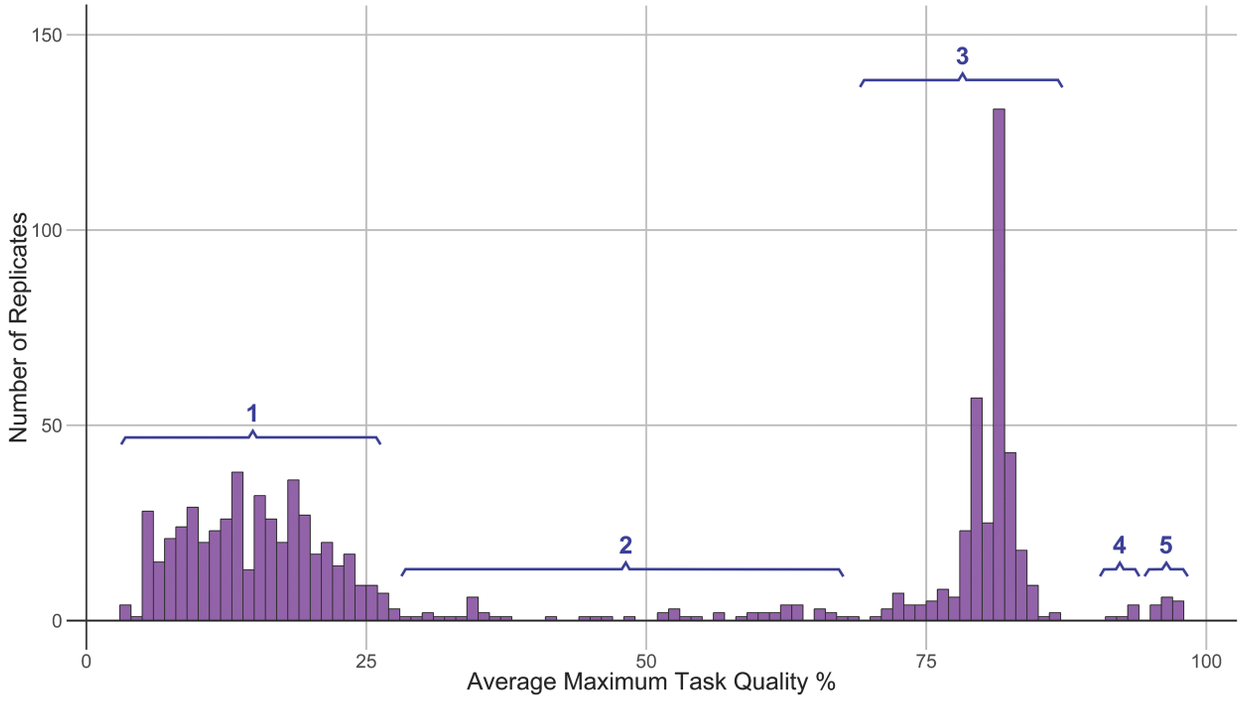


Figure 4: Distribution of average maximum task quality across 900 replicates. The performance histogram of all final predominant organism in experiment 2 reveals a marked grouping by behavioral strategy. Organisms in groups 1 and 2 did not finish the trails, while those in groups 3–5 did. Group 1 consisted mainly of organisms that navigated by path predicting and its hybrids. Group 2 consisted mainly of organisms that navigated by error recovery, imprinting, and their hybrids. Group 3 consisted mainly of organisms that navigated by more effective forms of error recovery. Group 4 consisted mainly of organisms that employed imprinting hybrids. Group 5 consisted mainly of organisms capable of relearning. The behaviors from groups 1–3 were assessed from a sample of organisms. Those of groups 4 and 5 were assessed from all organisms.

(short-term imprinting). This "short memory" gave the organisms the opportunity to form a new association after the previous one had extinguished. This hybrid strategy turned out to score higher in task quality than imprinting or error recovery alone. For additional results and a "bestiary" of evolved behaviors, see sections S.3 and S.4 of the supplemental PDF.

The Stepwise Evolution of Learning

We found a discernible pattern in the evolutionary trajectories of the organisms that evolved learning strategies (relearning and imprinting). Despite the organisms having evolved completely independently, these lineages passed through a characteristic sequence of phenotypic stages corresponding to two or more of the categories we described in table 2.

We analyzed the ancestral lineages of all of the final predominant organisms that evolved imprinting in experiment 1 and ten of the final predominant organisms capable of relearning in experiment 2. Starting from a sessile common ancestor, all lineages first evolved the capacity for moving, then sensing, followed by reflexive navigation and then learning, a result that supports hypothesis 2. In addition, error recovery preceded the evolution of associative learning in all of the lineages where the final predominant organism made the cue-response association by stepping off the trail (generalizable imprinting and relearning). In lineages where imprinting evolved directly from path predicting, the final predominant organisms were not capable of error recovery, and their behavior did not generalize to other trail configurations (nongeneralizable imprinting; figs. 5, 6).

Learning Can Evolve Suddenly

Finally, we found that during evolution, the transitions from one strategy to another could occur abruptly, often as a result of a single mutation. This is not to say that a single mutation was sufficient to produce a new strategy but rather that the new strategy often evolved silently over a great number of generations until one or a few mutations triggered the transition in behavior, a result that supports hypothesis 3. Sometimes this evolutionary transition would give the organism a large fitness advantage, and its descendants would sweep

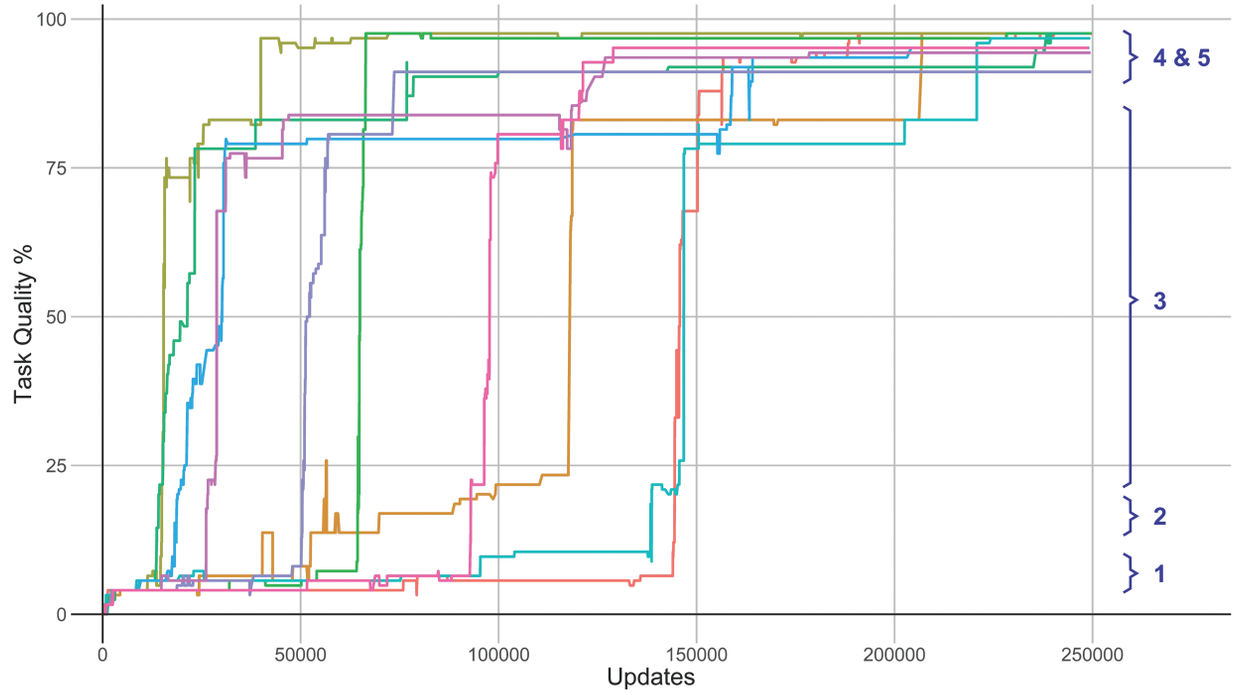


Figure 5: Evolutionary history: 10 lineages. Shown is the evolution of task quality over time in each of the 10 lineages that were ultimately capable of serial relearning from experiment 2. As they transitioned to a new strategy, some lineages had great gains in task quality, while others had more gradual ones. All lineages, however, went through occasional periods of fitness loss. Different task quality ranges often corresponded to specific behavioral strategies. Range 1 corresponded to path predicting, range 2 corresponded to hybrid strategies that included searching, range 3 corresponded to error recovery, and ranges 4 and 5 corresponded to imprinting and relearning.

the population. For example, the transition between error recovery and associative learning (imprinting or relearning) always occurred in one generation (we never observed any instance of a behavior that would be intermediary, such as a simpler form of learning). In one of the lineages we analyzed, the transition from error recovery to relearning raised the AMTQ from 81% to 98% in a single generation. This strategy transition was triggered by a single mutation that changed the flow of the algorithm so that after an error recovery event, the value of the currently sensed cue would be stored in memory (figs. S16, S17). The remainder of the error recovery process stayed intact and was subsumed by the newly acquired relearning capacity. Other components of the relearning algorithm, such as the module for storing the cue in memory, had already been part of the ancestor for many generations but were not used or did not affect the organism's task quality. This result represents a clear case of historical contingency, where one or more modules had to be in place before new mutations could lead to a fitness gain (Blount et al. 2008; Lenski 2017). See section S.4.1 of the supplemental PDF for figures and phenotypic descriptions of the major evolutionary transitions in this lineage.

Discussion and Conclusions

Emergence of Learning Depends on the Prior Evolution of Reflexive Behaviors

Most studies of the evolution of learning have focused on the selection pressures that may act to increase or decrease an organism's reliance on learning (Mery and Kawecki 2002; Dunlap and Stephens 2009; Dunlap et al. 2019). Our study complements and extends this work by examining how learning may have first arisen. As Dunlap demonstrated (Dunlap and Stephens 2009, 2016; Dunlap et al. 2019), learning is favored in environments that present alternative courses of action, where the best action cannot be predicted at the beginning of an organism's life but environmental cues exist that reliably correlate with the best action. However, we found that although all of the environments possessed those presumably necessary qualities, they were not sufficient for learning to arise, as evidenced by results from the random start environment. Instead, as hypothesis 1 predicts, for organisms to initially evolve the capacity for learning, they must first accumulate simple behavioral building blocks to cope with the environment reflexively. In the cases presented

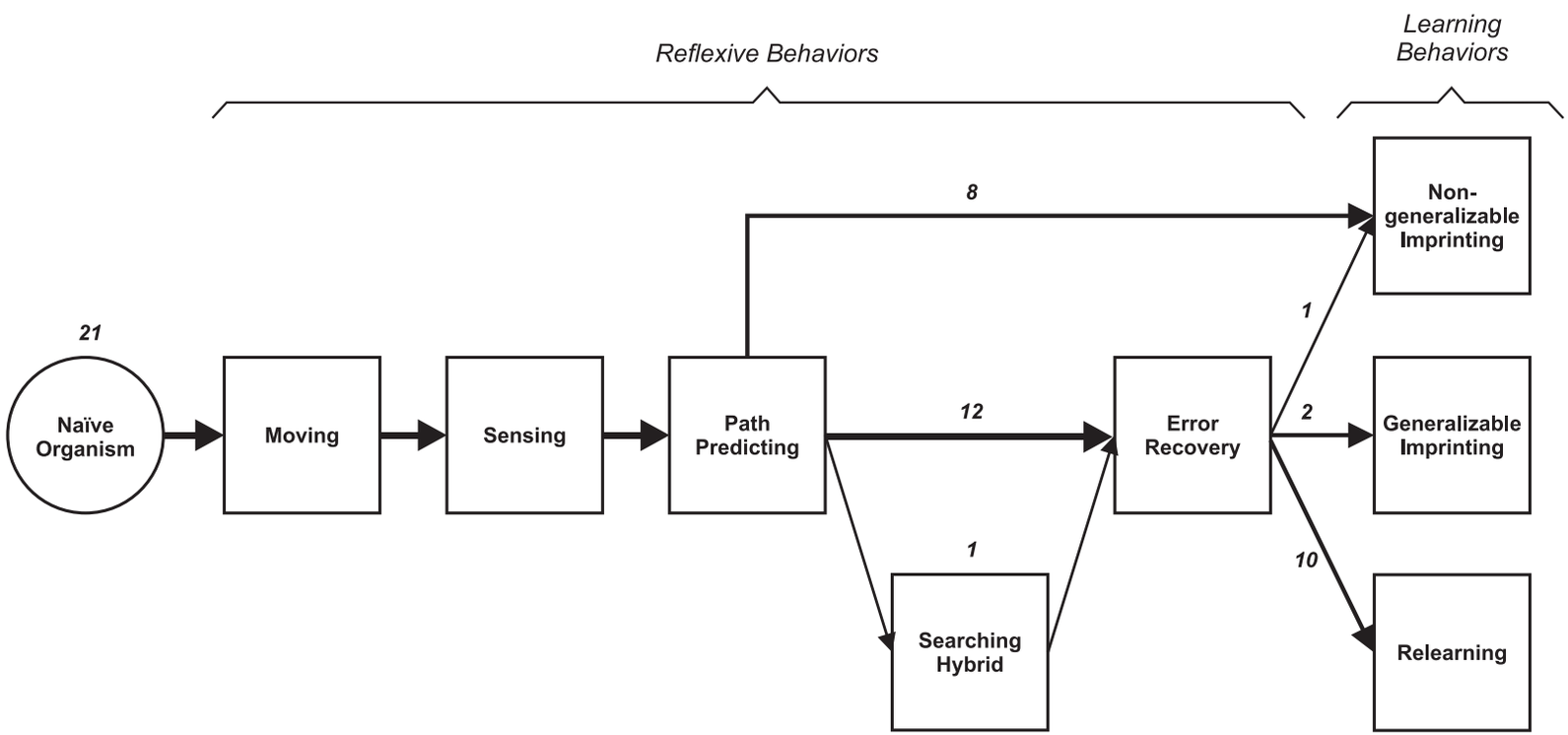


Figure 6: Commonly observed evolutionary sequences. Shown are the evolutionary trajectories of the 11 lineages that evolved associative learning in experiment 1 and the 10 lineages that evolved serial relearning in experiment 2. Behaviors evolved in a characteristic sequence of phenotypic stages. Starting from a naive and sessile common ancestor, all 21 lineages evolved the capacity for moving, then sensing, followed by reflexive navigation and then learning. The numbers next to the arrows indicate how many lineages followed a particular pathway, with thicker lines indicating more common evolutionary pathways in relation to alternatives.

here, these building blocks include an ability to move, to sense different cues, and to perform a range of actions (move forward, turn left or right, step back) in response to different cues. Crucially, generalizable learning (generalizable imprinting and relearning; table 2) arose only in lineages that first evolved a reflexive ability to correct for missteps and return to a trail of resources (error recovery; table 2). With these reflexive behaviors in place, associative learning can then evolve because it confers an advantage by enabling an organism to modulate its behavior based on experience. Moreover, we find that reflexive and learning behaviors are shaped by different characteristics of the environment—the former by regularities that are stable across generations, and the latter by patterns that vary across generations but persist for periods within an organism's lifetime. The most flexible learning ability—relearning new cue associations multiple times during an organism's life—depends on specific selection for it (i.e., swapping cues within the individual's lifetime as in experiment 2), as we proposed in hypothesis 4.

Stepwise and Modular Evolution of Complex Behaviors

Across many replicate evolutionary runs in several experimental conditions, we found an almost stereotypical historical sequence leading to the ability to learn. Furthermore, our results are consistent with the idea that behavioral control algorithms evolve modularly (Soyer and Goldstein 2011), where more complex behaviors evolve by building on simpler ones and sharing their mechanisms. For example, learning mechanisms incorporated previously evolved error recovery behavior (figs. S16, S17). This result supports our hypothesis 2, originally proposed by Skinner and others for natural organisms (Skinner 1984; Miller and Todd 1991; Moore 2004), that learning abilities evolve by building on previously evolved reflexive behaviors.

However, in contrast to Skinner's model, we found that not all intermediate modules have an immediate survival value. Such is the case with the previously mentioned organism that evolved relearning from error recovery in a sudden transition triggered by a single mutation (figs. S16, S17) and whose error recovery ancestors had already acquired the capacity to store the cue in memory but never used this ability and, therefore, gained no fitness benefit. Only when a mutation connected the memory-storing with the error recovery module did the organism acquire the capacity to learn, thus gaining fitness.

It is important to clarify that no single Avida instruction or even specific set of instructions could bestow learning on an arbitrary nonlearning organism. All associative learning algorithms we observed were assemblies of many instructions that had to be executed in the proper order for learning behavior to manifest (see sample learning organisms; fig. S18). That a single mutation could activate this behavior in an offspring only demonstrates that the remainder of the mechanism was already in place either as part of the existing behaviors or as neutral instructions (Lenski et al. 2003; Ofria et al. 2008).

In the eleven lineages leading to imprinting in experiment 1 and in the ten lineages leading to serial relearning in experiment 2 (fig. 5), we routinely found that complex abilities evolved from simpler ones in sudden transitions triggered by just a few mutations. This finding supports our hypothesis 3, that learning may arise through minor modification of existing mechanisms, and also lends credence to the proposition that something similar could have happened among natural organisms leading up to the Cambrian explosion (Ginsburg and Jablonka 2010).

More generally, these sharp transitions in phenotype are a consequence of the modular evolution of behavior. Modularity inherently reduces the requirements for evolving a new trait if it can build on existing ones, increasing phenotypic complexity with relatively modest genetic modifications (McAdams et al. 2004; Kashtan and Alon 2005; Wagner et al. 2007), which can build up silently and, once completed, cause a sudden shift in phenotype.

Why Learning Was Rare

Despite striking regularities during the course of evolution, associative learning was actually a rare outcome even in environments that fostered it (7% of lineages in experiment 1, 2% of lineages in experiment 2). Our results suggest some possible explanations. First, as mentioned above, complex behaviors can be hard to evolve, in part because they may depend on the preexistence of reusable intermediary modules—including features without survival value—and are therefore subject to the stochasticity of historical contingencies in general. Another possible explanation is that a reflexive strategy involving error recovery may already confer high fitness, such that the fitness gain associated with a learning strategy may not be enough for learning to arise and spread in the population. Across evolutionary replicates, we found organisms that could solve the problem in surprisingly different ways and obtain high levels of fitness, even in these simple environments. Furthermore, there can be implicit costs on more complex algorithms, including greater mutational fragility and longer processing times. Even making more mistakes on the trail, a shorter, sufficiently faster algorithm could reproduce more quickly and thus outcompete more complex algorithms that made fewer mistakes but executed too slowly. Surprisingly, we found in a follow-up experiment (supplemental PDF, sec. S.5) that the amount of computational memory available to an organism is not a constraint on the evolution of learning in our system as long as the minimum amount necessary to solve the task

is provided. We performed a version of experiment 2, where we reduced the amount of memory available in the organism's CPU from 26 integers to 2, which is the minimum necessary to solve the learning task, but did not see a significant difference in the frequency of evolution of the relearning strategy or in the average task quality and distribution of task quality in the final population compared with experiment 2 (fig. S19; table S3). Overall, the same conditions that explain the rarity of solutions involving learning were also responsible for the variety of solutions and evolutionary paths we observed, better resembling natural evolution, where learning typically entails some kind of cost and is not always adaptive (Johnston 1982; Miller and Todd 1991; Godfrey-Smith 2002; Dunlap and Stephens 2016; Dunlap et al. 2019).

Interestingly, the stepwise succession of behaviors observed in our lineage studies, in conjunction with the diversity of final strategies from different replicates, are reminiscent of how behaviors appeared on trace fossils from the Ediacaran and early Cambrian, becoming more complex and diverse over time (Carbone and Narbonne 2014).

The Scientific Value of an Open-Ended Evolutionary Model

In comparison with prior studies of the evolution of learning using computational methods (Todd and Miller 1991; Izquierdo and Harvey 2007; Izquierdo et al. 2008), ours is striking in the open-endedness of the evolutionary process, which parallels that of biological evolution. Avida (Ofria et al. 2009) employs relatively neutral genetic building blocks consisting primarily of algebraic and logic instructions, which do not constrain or favor the evolution of any particular behavioral algorithm. Thus, we were able to explore a large solution space and gain insights into the evolutionary dynamics that are also likely to occur in natural open-ended systems, even though nature uses very different building blocks. In Avida, the sheer number of potential solutions creates evolutionary dynamics and patterns that are not possible to observe using simpler digital evolution systems. For example, Izquierdo's groundbreaking work on the evolution of associative learning using neural networks consisted of evolving only the connections between preexisting neurons (Izquierdo and Harvey 2007; Izquierdo et al. 2008). Although many insights were gained from that experiment, the limited number of potential solutions also led to a smaller diversity of outcomes. In fact, simply using neural networks, which are intrinsically designed to form associations, means that fewer mutational steps are needed from a starting point to evolve appropriate connections compared with the enormous search space in Avida. In our experiments, the behavioral algorithms evolved from scratch, using the most basic computer programming language elements, from ancestors incapable of sensation, movement, or navigation of any kind.

Furthermore, in our system even a basic behavioral building block, such as the move back instruction, could evolve from an assembly of simpler actions, as we demonstrate in a preliminary experiment (supplemental PDF, sec. S.2). We performed a version of experiment 1 without the move back instruction, and even without it many organisms evolved the capacity to navigate the entire trail using either imprinting or error recovery. These organisms evolved behavior functionally equivalent to the move back instruction by assembling other instructions from the basic instruction set.

Early Evolution of an Intrinsic Value System

An unexpected outcome of this study was that it provided insights into the evolution of motivational mechanisms, which are thought to be integral to adaptive decision-making (Miller and Todd 1991; Breazeal 2004; Panksepp 2004; Damasio 2005; Singh et al. 2009, 2010). Some of the earliest building blocks to evolve across all of our experiments were those responsible for evaluating experiences. In our system, evaluations were implicit features of the evolved controller and not distinct modules for deciding "good" or "bad." They were also essential to behavior control, since organisms could not sense their own task quality scores to determine whether an action was beneficial or harmful. Early in evolution, values started as arbitrary biases, such as moving constantly or favoring turning one way or another, but biases that proved adaptive (e.g., preferring continuous movement while avoiding empty or previously visited locations) would fix, excluding less fit alternative biases. Over time, an intrinsic value system evolved that ensured appropriate behavior in response to specific inputs, and when associative learning arose, this value system provided reinforcement for behavior conditioning.

We can thus reinterpret the associative learning mechanism we have observed in the light of a value system: when an organism capable of learning senses an empty location, it displays the avoidance behavior because, in effect, it negatively values the experience. It associates this negative experience with the cue that led it to the empty location, and from then on, experiencing the cue alone is sufficient to activate the avoidance behavior.

Reversal Learning Seems No More Complex than Initial Learning

Reversal learning is often deemed more challenging cognitively than initial learning (Hadar and Menzel 2010; Bissonette and Powell 2012; Xue et al. 2013). However, in our experiments organisms that evolved the ability for reversal

learning showed no difference in the capacity or speed of learning between the initial and subsequent learning events. Thus, once reversal learning evolves, it does not seem cognitively more complex than associative learning itself, at least in this system. Our point is not to undercut the study of how serial reversal learning in animals becomes faster with experience and correlates with cognitive flexibility (Bond et al. 2007; Cauchoix et al. 2017; Buechel et al. 2018) but to call for a refinement of the ideas around what is required for reversal learning to occur.

How Evolution Continues to Shape Associative Learning

In a follow-up analysis (supplemental PDF, sec. S.6), we looked into how evolution continued to shape learning after it appeared in a lineage. In the lineages that produced associative learning in experiments 1 and 2, we found that learning ability would become attuned to the environment of evolution in a variety of ways. For example, (i) ancestral organisms that could learn some cue combinations but not others would eventually give rise to descendants that could learn any cue combination; (ii) ancestral organisms that required multiple exposures to learn the cue-response association gave rise to descendants that required fewer exposures and, ultimately, final organisms that required only a single exposure; and (iii) in environments without cue reversals, ancestral organisms that could re-form associations multiple times gave rise to final organisms that could imprint only once.

These adaptations are consistent with the literature on preparedness and other so-called constraints on learning (Seligman 1970; Shettleworth 2010; Domjan 2012; Dunlap 2017) as well as the literature on sensitive periods of plasticity (Bateson 1979; Cashdan 1994; Fawcett and Frankenhuis 2015). The key themes are that evolution produces learning mechanisms that are optimized for the needs of an animal in the environment where it evolved, and since learning is costly, evolution will often restrict the periods of an animal's life when it is most capable of learning (sensitive periods). An example of learning optimization is when an animal that relies on odors for foraging can learn more quickly to associate odors with good or bad foods than visual cues with the same foods (Dunlap and Stephens 2014). The phenomenon of sensitive periods for learning is illustrated by filial imprinting in birds, where a chick learns who its mother is early in life and that association does not change (Cashdan 1994).

Consistent with this literature, the imprinting strategy was adaptive in experiment 1, where there were no cue reversals. In that environment, ancestral organisms that were capable of re-forming the cue association multiple times eventually gave rise to organisms that could form the association only once, presumably becoming more efficient. The sensitive period for learning in those lineages became restricted to the beginning of an organism's life. Meanwhile in experiment 2, where the environment contained cue reversals, the ability to re-form the cue associations (relearning and short-term imprinting) was adaptive, and the sensitive period for learning lasted an organism's entire life.

Although our experiments were not designed to investigate these topics, the patterns we found suggest a future area of study in which Avida is used to systematically explore how the evolutionary environment can constrain and optimize learning abilities.

Implications for Artificial Intelligence

The insights of this study are relevant to the field of artificial intelligence, where lifetime learning has long been a challenge. We demonstrated that adaptive autonomous agents, capable of learning and navigation, can be produced by evolutionary methods, using biologically consistent scenarios where the environment fosters the evolution of learning and decision making, instead of traditional methods based on human design, which are difficult to scale up and to apply to novel tasks. One of our future goals is to extend this study and test whether we can evolve more complex forms of learning, such as contextual learning and rule learning (the learning of rules and concepts), and see whether their evolution follows the same sequence suggested in the literature (Wells 1968; Razran 1971; Hawkins and Kandel 1984*a*, 1984*b*; Skinner 1984; Miller and Todd 1991; Moore 2004). We could test this hypothesis by introducing additional cue types and requiring the organism to perform additional tasks in more intricate trails.

Implications for the Evolution of Behavior

Finally, we believe that the evolution of learning in a digital environment would be useful to investigate the effect that learning behavior has on evolvability and rate of adaptive evolution. Some researchers have proposed that learning increases evolvability, since behavioral flexibility shields organisms from some selective pressures, allowing the population to maintain its diversity to cope with future selective events (Brown 2013; Dukas 2013). Others have proposed that learning could either drive evolution by helping organisms adapt to different niches, where they would experience different selective pressures leading to change, or inhibit it by protecting them from selective pressures, leading to stasis (Duckworth 2009). It has even been suggested that the emergence of learning drove the diversification of complex behavior during the Cambrian explosion (Ginsburg and Jablonka 2010). Overall, we agree with the remarkable

assertion by B. F. Skinner (1984, p. 220) that understanding "the conditions under which [learning] evolved are helpful in understanding its nature."

Acknowledgments

We thank Annat Haber for reviewing the manuscript and making numerous suggestions to improve it, Laura Grabowski for inspiring our study and helping us get started, Ali Tehrani-Saleh for proposing and writing the step back instruction and participating in our initial discussions and exploratory experiments, Andrew Mitchell for writing the software to plot the path of the organisms in the behavioral arena, David M. Bryson for helping us set up Avida and advising on the choice of instruction set and configurations, and the BEACON Center for the Study of Evolution in Action for bringing biologists and computer scientists together to work on experimental evolution on an unprecedented scale while supporting our efforts, financially, logistically, and otherwise. Finally, many thanks to the editors (Daniel Bolnick and Renée Duckworth) and reviewers at *The American Naturalist*, whose thoughtful feedback and constructive comments substantially improved our article.

Literature Cited

Armus, H. L., A. R. Montgomery, and J. L. Jellison. 2006. Discrimination learning in paramecia (*P. caudatum*). Psychological Record 56:489–498.

Bateson, P. 1979. How do sensitive periods arise and what are they for? Animal Behaviour 27:470–486.

Bissonette, G. B., and E. M. Powell. 2012. Reversal learning and attentional set-shifting in mice. Neuropharmacology 62:1168–1174.

Blount, Z. D., C. Z. Borland, and R. E. Lenski. 2008. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. Proceedings of the National Academy of Sciences of the USA 105:7899–7906.

Bond, A. B., A. C. Kamil, and R. P. Balda. 2007. Serial reversal learning and the evolution of behavioral flexibility in three species of North American corvids (*Gymnorhinus cyanocephalus*, *Nucifraga columbiana*, *Aphelocoma californica*). Journal of Comparative Psychology 121:372–379.

Breazeal, C. 2004. Designing sociable robots. MIT Press, Cambridge, MA.

Brown, R. L. 2013. Learning, evolvability and exploratory behaviour: extending the evolutionary reach of learning. Biology and Philosophy 28:933–955.

Buechel, S. D., A. Boussard, A. Kotrschal, W. van der Bijl, and N. Kolm. 2018. Brain size affects performance in a reversal-learning test. Proceedings of the Royal Society B 285:20172031.

Carbone, C., and G. M. Narbonne. 2014. When life got smart: the evolution of behavioral complexity through the Ediacaran and early Cambrian of NW Canada. Journal of Paleontology 88:309–330.

Cashdan, E. 1994. A sensitive period for learning about food. Human Nature 5:279–291.

Cauchoix, M., E. Hermer, A. S. Chaine, and J. Morand-Ferron. 2017. Cognition in the field: comparison of reversal learning performance in captive and wild passerines. Scientific Reports 7:12945.

Charnov, E. L. 1976. Optimal foraging: attack strategy of a mantid. American Naturalist 110:141–151.

Chow, S. S., C. O. Wilke, C. Ofria, R. E. Lenski, and C. Adami. 2004. Adaptive radiation from resource competition in digital organisms. Science 305:84–86.

Codoñer, F. M., J.-A. Darós, R. V. Solé, and S. F. Elena. 2006. The fittest versus the flattest: experimental confirmation of the quasispecies effect with subviral pathogens. PLoS Pathogens 2:e136.

Collett, T. S., and M. Collett. 2002. Memory use in insect visual navigation. Nature Reviews Neuroscience 3:542–552.

Damasio, A. R. 2005. Descartes' error: emotion, reason, and the human brain. Penguin, London.

Dennett, D. C. 1996. Darwin's dangerous idea: evolution and the meanings of life. Touchstone, New York.

Domjan, M. 2012. Biological or evolutionary constraints on learning. Pages 461–463 *in* N. M. Seel, ed. Encyclopedia of the sciences of learning. Springer, Boston.

Duckworth, R. A. 2009. The role of behavior in evolution: a search for mechanism. Evolutionary Ecology 23:513–531.

Duerr, J. S., and W. G. Quinn. 1982. Three *Drosophila* mutations that block associative learning also affect habituation and sensitization. Proceedings of the National Academy of Sciences 79:3646–3650.

Dukas, R. 2013. Effects of learning on evolution: robustness, innovation and speciation. Animal Behaviour 85:1023–1030.

Dukas, R., and J. M. Ratcliffe, eds. 2009. Cognitive ecology II. University of Chicago Press, Chicago.

Dunlap, A. S. 2017. Biological preparedness. Pages 1–7 *in* J. Vonk and T. Shackelford, eds. Encyclopedia of animal cognition and behavior. Springer, Cham.

Dunlap, A. S., M. W. Austin, and A. Figueiredo. 2019. Components of change and the evolution of learning in theory and experiment. Animal Behaviour 147:157–166.

Dunlap, A. S., and D. W. Stephens. 2009. Components of change in the evolution of learning and unlearned preference. Proceedings of the Royal Society B 276:3201–3208.

———. 2014. Experimental evolution of prepared learning. Proceedings of the National Academy of Sciences 111:11750–11755.

———. 2016. Reliability, uncertainty, and costs in the evolution of animal learning. Current Opinion in Behavioral Sciences 12:73–79.

Dyer, F. C. 1991. Bees acquire route-based memories but not cognitive maps in a familiar landscape. Animal Behaviour 41:239–246.

———. 1998. Cognitive ecology of navigation. Pages 201–260 *in* Cognitive ecology: the evolutionary ecology of information processing and decision making. University of Chicago Press, Chicago.

Fawcett, T. W., and W. E. Frankenhuis. 2015. Adaptive explanations for sensitive windows in development. Frontiers in Zoology 12:S3.

Fernando, C. T., A. M. Liekens, L. E. Bingle, C. Beck, T. Lenser, D. J. Stekel, and J. E. Rowe. 2009. Molecular circuits for associative learning in single-celled organisms. Journal of the Royal Society Interface 6:463–469.

Gagliano, M., V. V. Vyazovskiy, A. A. Borbély, M. Grimonprez, and M. Depczynski. 2016. Learning by association in plants. Scientific Reports 6:38427.

Ginsburg, S., and E. Jablonka. 2010. The evolution of associative learning: a factor in the Cambrian explosion. Journal of Theoretical Biology 266:11–20.

Godfrey-Smith, P. 1996. Complexity and the function of mind in nature. Cambridge Studies in Philosophy and Biology. Cambridge University Press, New York.

———. 2002. Environmental complexity and the evolution of cognition. Pages 233–249 *in* R. Sternberg and J. Kaufman, eds. The evolution of intelligence. Lawrence Erlbaum, London.

Grabowski, L. M. 2009. The evolutionary origins of memory use in navigation. PhD diss. Michigan State University, East Lansing.

Grabowski, L. M., D. M. Bryson, F. C. Dyer, C. Ofria, and R. T. Pennock. 2010. Early evolution of memory usage in digital organisms. Pages 224–231 *in* Artificial life XII: Proceedings of the Twelfth International Conference on the Synthesis and Simulation of Living Systems. Presented at the ALIFE. MIT Press, Cambridge.

Grabowski, L. M., W. R. Elsberry, C. Ofria, and R. T. Pennock. 2008. On the evolution of motility and intelligent tactic response. Pages 209–216 *in* Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation. GECCO '08. ACM, New York.

Hadar, R., and R. Menzel. 2010. Memory formation in reversal learning of the honeybee. Frontiers in Behavioral Neuroscience 4:186.

Hawkins, R. D., and E. R. Kandel. 1984*a*. Is there a cell-biological alphabet for simple forms of learning? Psychological Review 91:375–391.

———. 1984*b*. Steps toward a cell-biological alphabet for elementary forms of learning. Pages 385–404 *in* G. Lynch, J. L. McGaugh, and N. M. Weinberger, eds. Neurobiology of learning and memory. Guilford, New York.

Hume, D. 1738. A treatise of human nature: being an attempt to introduce the experimental method of reasoning into moral subjects. Oxford University Press, London.

Izquierdo, E., and I. Harvey. 2007. The dynamics of associative learning in an evolved situated agent. Pages 365–374 *in* F. Almeida e Costa, L. M. Rocha, E. Costa, I. Harvey, and A. Coutinho, eds. Advances in artificial life. Presented at ECAL 2007. Springer, Berlin.

Izquierdo, E., I. Harvey, and R. D. Beer. 2008. Associative learning on a continuum in evolved dynamical neural networks. Adaptive Behavior 16:361–384.

Johnston, T. D. 1982. Selective costs and benefits in the evolution of learning. Pages 65–106 *in* J. S. Rosenblatt, R. A. Hinde, C. Beer, and M.-C. Busnel, eds. Advances in the study of behavior. Vol. 12. Academic Press, New York.

Kashtan, N., and U. Alon. 2005. Spontaneous evolution of modularity and network motifs. Proceedings of the National Academy of Sciences 102:13773–13778.

Lenski, R. E. 2017. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. ISME Journal 11:2181–2194.

Lenski, R. E., C. Ofria, R. T. Pennock, and C. Adami. 2003. The evolutionary origin of complex features. Nature 423:139–144.

Losos, J. B. 2011. Seeing the forest for the trees: the limitations of phylogenies in comparative biology. American Naturalist 177: 709–727.

McAdams, H. H., B. Srinivasan, and A. P. Arkin. 2004. The evolution of genetic regulatory systems in bacteria. Nature Reviews Genetics 5:169–178.

Mery, F., and T. J. Kawecki. 2002. Experimental evolution of learning ability in fruit flies. Proceedings of the National Academy of Sciences 99:14274.

Miller, G. F., and P. M. Todd. 1991. Exploring adaptive agency I: theory and methods for simulating the evolution of learning. Pages 65–80 *in* D. S. Touretzky, J. L. Elman, T. J. Sejnowski, and G. E. Hinton, eds. Connectionist models. Morgan Kaufmann, San Mateo, CA.

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

Ofria, C., C. T. Brown, and C. Adami. 2015. Avida.

Ofria, C., D. M. Bryson, and C. O. Wilke. 2009. Avida. Pages 3–35 *in* M. Komosinski and A. Adamatzky, eds. Artificial life models in software. Springer, London.

Ofria, C., W. Huang, and E. Torng. 2008. On the gradual evolution of complexity and the sudden emergence of complex features. Artificial Life 14:255–263.

Ord, T. J., and E. P. Martins. 2010. Evolution of behaviour: phylogeny and the origin of present day diversity. Pages 108–128 *in* D. Westneat and C. W. Fox, eds. Evolutionary behavioral ecology. Oxford University Press, New York.

Panksepp, J. 2004. Affective neuroscience: the foundations of human and animal emotions. Oxford University Press, Oxford.

Pennock, R. T. 2007. Models, simulations, instantiations, and evidence: the case of digital evolution. Journal of Experimental and Theoretical Artificial Intelligence 19:29–42.

Pontes, A. C., R. B. Mobley, C. Ofria, C. Adami, and F. C. Dyer. 2020. Data from: The evolutionary origin of associative learning. American Naturalist, Dryad Digital Repository, https://doi.org/10.5061/dryad.f45gh6s.

Razran, G. 1971. Mind in evolution: an east-west synthesis of learned behavior and cognition. Houghton Mifflin, Boston.

Roberts, A. C., and D. L. Glanzman. 2003. Learning in *Aplysia*: looking at synaptic plasticity from both sides. Trends in Neurosciences 26:662–670.

Seligman, M. E. 1970. On the generality of the laws of learning. Psychological Review 77:406–418.

Shettleworth, S. J. 2010. Cognition, evolution and behavior. Oxford University Press, Oxford.

Singh, S., R. L. Lewis, and A. Barto. 2009. Where do rewards come from? Pages 2601–2606 *in* N. Taatgen and H. van Rijn, eds. Proceedings of the Annual Meeting of the Cognitive Science Society. Vol. 31. Cognitive Science Society.

Singh, S., R. L. Lewis, A. G. Barto, and J. Sorg. 2010. Intrinsically motivated reinforcement learning: an evolutionary perspective. IEEE Transactions on Autonomous Mental Development 2:70–82.

Skinner, B. F. 1984. The evolution of behavior. Journal of the Experimental Analysis of Behavior 41:217–221.

Soyer, O. S., and R. A. Goldstein. 2011. Evolution of response dynamics underlying bacterial chemotaxis. BMC Evolutionary Biology 11:240.

Stephens, D. W. 1991. Change, regularity, and value in the evolution of animal learning. Behavioral Ecology 2:77–89.

Todd, P. M., and G. F. Miller. 1991. Exploring adaptive agency II: simulating the evolution of associative learning. Pages 306–315 in J. A. Meyer and S. W. Wilson, eds. From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior. MIT Press, Cambridge, MA.

van Duijn, Marc. 2017. Phylogenetic origins of biological cognition: convergent patterns in the early evolution of learning. Interface Focus 7:20160158.

Wagner, G. P., M. Pavlicev, and J. M. Cheverud. 2007. The road to modularity. Nature Reviews Genetics 8:921–931.

Weber, B. H., and D. J. Depew. 2003. Evolution and learning: the Baldwin effect reconsidered. MIT Press, Cambridge, MA.

Wells, M. J. 1968. Sensitization and the evolution of associative learning. Pages 391–411 *in* J. Salánki, ed. Neurobiology of invertebrates: Proceedings of the Symposium Held at the Biological Research Institute of the Hungarian Academy of Sciences (Tihany) September 4–7, 1967. Springer, Boston.

Wilke, C. O., J. L. Wang, C. Ofria, R. E. Lenski, and C. Adami. 2001. Evolution of digital organisms at high mutation rates leads to survival of the flattest. Nature 412:331–333.

Xue, G., F. Xue, V. Droutman, Z.-L. Lu, A. Bechara, and S. Read. 2013. Common neural mechanisms underlying reversal learning by reward and punishment. PLoS ONE 8:e82169.

Zhang, S. W., K. Bartsch, and M. V. Srinivasan. 1996. Maze learning by honeybees. Neurobiology of Learning and Memory 66:267–282.

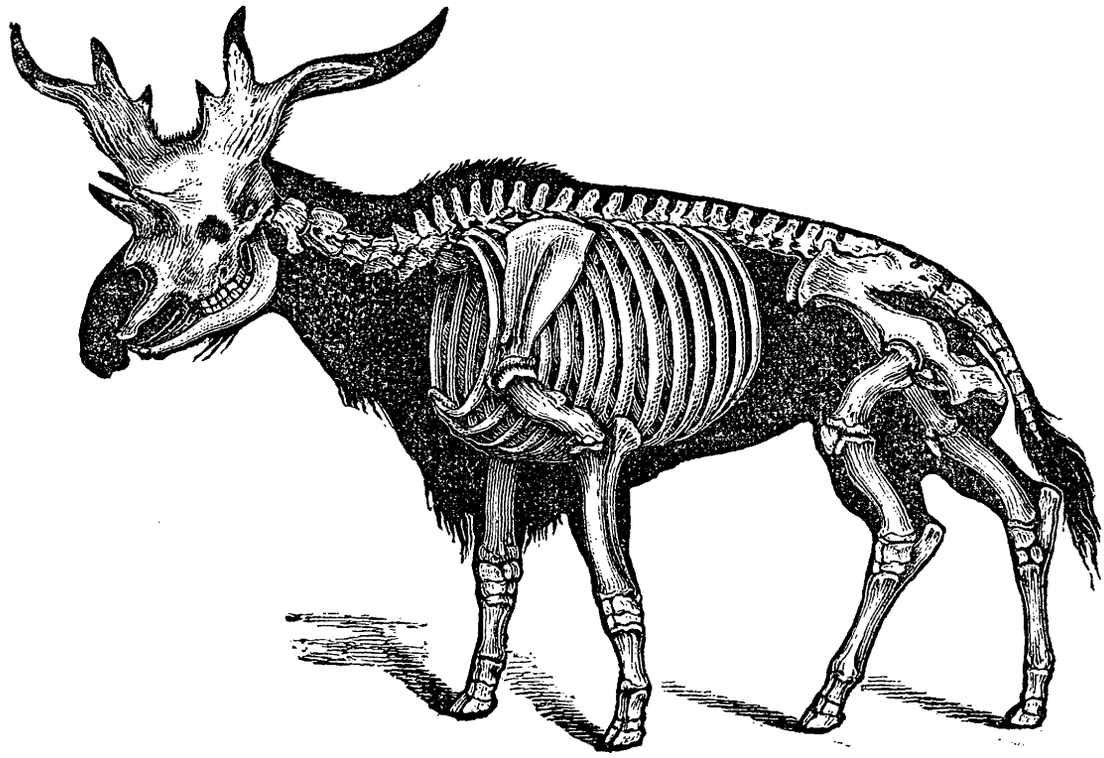
Zhang, S. W., M. Lehrer, and M. V. Srinivasan. 1999. Honeybee memory: navigation by associative grouping and recall of visual stimuli. Neurobiology of Learning and Memory 72:180–201.

References Cited Only in the Online Enhancements

Bryson, D. M., and C. Ofria. 2013. Understanding evolutionary potential in virtual CPU instruction set architectures. PLoS ONE 8: e83242.

Ofria, C., and C. Adami. 2002. Evolution of genetic organization in digital organisms. Pages 296–313 *in* L. F. Landweber and E. Winfree, eds. Evolution as computation. Springer, Berlin.

Associate Editor: Renée A. Duckworth
Editor: Daniel I. Bolnick



"This was a Tertiary, probably Miocene, 'elephantine stag, having four horns and probably a long proboscis, being in some points between the stags and the Pachyderms. It is supposed to have had the bulk of an elephant and greater height.' (Dana's Manual.)" From the review of "Restoration of the *Sivatherium*" (*The American Naturalist*, 1877, 11:434–436).