



A social insect perspective on the evolution of social learning mechanisms

Ellouise Leadbeater^{a,1} and Erika H. Dawson^{b,2}

^aSchool of Biological Sciences, Royal Holloway University of London, Egham TW20 0EX, United Kingdom; and ^bLaboratoire Evolution, Génomes, Comportement et Ecologie, CNRS, 91198 Gif-sur-Yvette, France

Edited by Kevin N. Laland, University of St. Andrews, St. Andrews, United Kingdom, and accepted by Editorial Board Member Andrew G. Clark May 29, 2017 (received for review January 14, 2017)

The social world offers a wealth of opportunities to learn from others, and across the animal kingdom individuals capitalize on those opportunities. Here, we explore the role of natural selection in shaping the processes that underlie social information use, using a suite of experiments on social insects as case studies. We illustrate how an associative framework can encompass complex, context-specific social learning in the insect world and beyond, and based on the hypothesis that evolution acts to modify the associative process, suggest potential pathways by which social information use could evolve to become more efficient and effective. Social insects are distant relatives of vertebrate social learners, but the research we describe highlights routes by which natural selection could coopt similar cognitive raw material across the animal kingdom.

social learning | associative learning | observational conditioning | social insects | *Bombus*

An expanding body of work now shows that social learning (1), once considered the preserve of vertebrate species, is a feature of insect behavioral repertoires (2, 3). Insects not only learn about foraging skills, food preferences, brood hosts, and potential mates by responding to information provided inadvertently by others (4–11), but also transmit these behaviors further, such that they propagate through groups (8, 12) and possibly even through wild populations (13). Some of these phenomena appear similar to socially learned behavior patterns that have been described in vertebrates (11, 14) at least outside the context of imitation, and as such they are interesting extensions to the taxonomic distribution of social learning. However, there is more to be gained from these findings than the claim that “insects are smarter than we thought.” An insect perspective encourages questions about the basic cognitive raw material that underlies social learning.

Associative learning theory describes the rules that govern the strength of connections between neural representations of stimuli, and there is little doubt that these rules are taxonomically widespread (15–18). We now know that associative learning is a means to build a complex, representation-based picture of the world that is far from the simple stimulus–response/reinforcement paradigms through which it was originally characterized (19–22). Correspondingly, it has been repeatedly argued that associative explanations are powerful enough to encompass many cognitive phenomena (23–29), and social learning is no exception. Heyes in particular has highlighted that social learning mechanisms may be fundamentally associative (28–30; see also ref. 31). But the contention that no major qualitative leaps are required to distinguish at least some social learning processes from asocial learning is not an argument against evolutionary change. It has been argued that a likely pathway for cognitive evolution may be the accumulation of small quantitative changes that render domain-general processes a closer fit to the specific cognitive demands of a particular niche (32, 33). Because almost all animals interact with others at some point during their lifetime, these demands most likely include effective processing of social information. In other words, whereas there may be little evidence to support a polarized view of asocial- and social learning processes, it does not follow that

the ability to learn socially is simply a useful byproduct of the fundamental ability to learn asocially that has remained untouched by natural selection. There is clear empirical evidence that natural selection can fine-tune associative processes to particular tasks (34–36).

For certain social learning processes that typify human behavior, such as imitation or learning through language or instruction, this mechanistic middle ground between adaptation- and preadaptation-based explanations for social learning is the subject of substantial empirical attention (37–40). However, this is not the case for those social learning mechanisms that underlie the majority of social information use outside the primates, such as local or stimulus enhancement, social facilitation, or goal emulation (41). The effects produced by these “simpler” processes are well-labeled (1), but these definitions often tell us little about how the social stimuli upon which they are based come to control behavior. For example, imagine that an individual observes a demonstrator using a tool to get food, and then uses the same tool to extract food for him or herself, without using the same exact actions. This fits a classic definition of “emulation” (1), because observation of a demonstrator interacting with objects in its environment has rendered the observer more likely to perform any actions that have a similar effect on those objects. An associative hypothesis might put forward that perhaps, through the demonstrator’s actions, the observer came to associate the tool with food, and thus manipulated the tool in its own manner (1). This is not an alternative option, but rather a different level of explanation: the label describes the effect, and associative learning theory puts forward a testable hypothesis as to why that effect occurs, on a mechanistic level (39). Understanding those mechanisms is critical to forming a picture of how natural selection has acted upon cognitive raw material, but as yet, such an understanding is lacking. As Galef (41) puts it, “an entire field of local enhancement awaits exploration.”

Addressing this issue first requires exploration of the basic mechanisms through which associative learning could produce adaptive social information use, and here we review a series of case studies carried out by ourselves and by other researchers that specifically illustrate how social learning effects can sit within an associative framework. In the light of claims that students of animal cognition might sometimes be “association-blind” (24), we hope that this approach proves useful not only in highlighting how associative learning theory can be applied to

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, “The Extension of Biology Through Culture,” held November 16–17, 2016, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and video recordings of most presentations are available on the NAS website at www.nasonline.org/Extension_of_Biology_Through_Culture.

Author contributions: E.L. and E.H.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. K.N.L. is a guest editor invited by the Editorial Board.

¹To whom correspondence should be addressed. Email: Elli.Leadbeater@rhu.ac.uk.

²Present address: Unité de Modélisation Mathématique et Informatique des Systèmes Complexes, Unité Mixtes Internationales, Institut de Recherche pour le Développement/Université Pierre et Marie Curie 209, 93143 Bondy, France.

empirical work on animal social learning, but also in stimulating consideration of how natural selection might modify associative mechanisms. Social learning involves upstream input mechanisms that define how social information is received, and downstream parameters that define how associations are weighted, processed, stored, retrieved, and applied, and all may be potential targets for selection (29, 32, 33). Given the phylogenetic distance between insects and vertebrates (and even between the diverse vertebrate groups within which social learning abilities are manifest), we do not infer that any evolutionary modifications would necessarily be conserved across lineages through common descent; such inference would have little foundation. Rather, our aim is to illustrate the cognitive toolbox from which social learning can be built. We hope to provide some food for thought for those researchers from both the invertebrate and the vertebrate world that are interested in the question of how associative learning could be shaped by natural selection to create something specifically, if not exclusively, social.

An Associative Framework

Bumblebees (*Bombus* spp.) and honey bees (*Apis* spp.) forage for nectar and pollen in rapidly changing floral landscapes, where floral reward levels vary not only between flower species but also with season, time of day, local pollinator abundance, and even shade patterns (42). Foragers visit thousands of flowers per day, and quickly learn about the stimuli that predict where to find rewards in the particular flower species and patches on which they forage. Information provided inadvertently by other foraging bees influences this learning process, and a simple example whereby social bees learn from their conspecifics about rewarding flower types provides a good introduction to an associative framework for studying social learning.

Worden and Papaj (5) allowed bumblebees (*Bombus impatiens*) to observe their foraging conspecifics through a screen. Those conspecifics foraged on only one of two available flower colors, and when the observers were later permitted to forage alone, they “copied” the color preferences of the demonstrators (5, 43, 44). This type of learning initially appears conceptually opaque, because learning theory is based upon the fundamental concept of prediction error, which requires a difference between predicted and experienced outcomes (45). In this observational paradigm, observers do not directly experience any rewarding outcome, nor had these laboratory bees ever had the chance to learn that matching the color choices of conspecifics was rewarding. Thus, the bees in this situation seem to have learned about flower color simply by observing the behavior of their conspecifics.

Second-order conditioning is an associative phenomenon that can potentially explain why animals respond to certain stimuli as though they have been directly associated with food rewards, when they have not. This process is best known for its use in psychological research to study the contents of learning (46), but perhaps the most illustrative example derives from the work of Pavlov (47), who famously trained dogs that the tick of a metronome (a conditioned stimulus, CS) predicted the arrival of food (an appetitive unconditioned stimulus, US+). When Pavlov later trained the same dogs that presentation of a black square (a second conditioned stimulus, CS2) predicted the sound of the metronome in the absence of food, he found that subsequent presentation of the square alone evoked salivation. In other words, the black square (CS2) had come to elicit the same response as the food (US+), despite the two never having been experienced together. An association had formed between the black square and either the food itself or the appetitive state induced by the conditioned state of the bell (it remains unclear which) (20, 48), and this association constitutes a second-order conditioned relationship. Whereas in classic conditioning paradigms, it is an unconditioned response to a stimulus that comes to be elicited by a new stimulus, in second-order conditioning paradigms, it is a conditioned response that undergoes what is functionally the same effect.

Second-order conditioning is relevant to social learning because it provides a mechanism by which a learned response to a social stimulus, such as conspecific behavior or its products, could come to be elicited by a new asocial stimulus (49). Hence, any biologically important stimulus (US) that an animal has learned to associate with conspecific presence, or with a particular conspecific behavior pattern, or a vocalization—indeed, with any social stimulus at all—could become secondarily conditioned to a new asocial stimulus (a CS2). In Worden and Papaj’s bumblebee example (5), observers might “copy” the color preferences of demonstrators because they have previously learned to associate conspecifics (a CS) with food rewards (US+), perhaps because other bees are found at rewarding feeders or flowers. On subsequent observation of conspecifics on a particular flower color (a CS2), that flower color should be rendered attractive through second-order conditioning, even though it has not been directly paired with food.

We recently carried out an experiment to test this hypothesis (Fig. 1) (43), in which we controlled the previous social foraging experience of observer bees (*Bombus terrestris*). Individuals that had learned to associate conspecifics (CS) with sucrose (US+) behaved similarly to bees in previous experiments (5), “copying” the color preferences of the demonstrators that they had observed through a screen. However, bees that had learned to associate conspecifics (CS) with an aversive substance (US−) actively avoided those same colors, and bees that had never foraged with conspecifics were not influenced by conspecific choices. In other words, the results support that observing conspecifics through a screen influenced forage preferences through second-order conditioning.

Second-order conditioning is an associative mechanism, and the use of an associative framework to explain empirical results in social learning research is not new. Perhaps the best-known example comes from the work of Cook and Mineka (50–52), who demonstrated almost 30 years ago that juvenile rhesus monkeys (*Macaca mulatta*) can acquire a fear of snakes through observation of a frightened conspecific interacting with a snake, a phenomenon that they termed “observational conditioning.” As the name implies, these results are considered to reflect a classic conditioning process, whereby the sight of a frightened conspecific (US) elicits an unconditioned fear response in observers. This affective state becomes conditioned to stimuli that are experienced at the same time, in this case a snake model (CS), which thus also acquires the ability to elicit fear (28, 52). In other words, the snake is simply a conditioned stimulus that comes to elicit the fear response. Cook and Mineka (50) found that fear could be conditioned to snakes in this way, but not to flowers, implying that fear cannot be socially conditioned to any randomly chosen stimulus. However, this is not an argument against an associative explanation, because in primates snake stimuli are generally particularly easily conditioned to fear responses, whereas flower stimuli are not (53–56). For example, in humans, pairing of snake pictures with electric shock leads those pictures to later elicit heart-rate acceleration (indicating fear), whereas the same protocol involving flower pictures leads to heart-rate deceleration (indicating arousal of attention) (54, 56). It is thus not surprising that a fear response invoked by a social stimulus can be easily conditioned to snakes but was not detectable for flowers.

What is the link between the second-order conditioning and the observational conditioning process that we describe above? In both cases, a response to a social stimulus becomes conditioned to a new, asocial stimulus. In observational conditioning, an unconditioned response (here fear, elicited by seeing a frightened conspecific) becomes conditioned to a new stimulus. In second-order conditioning, a conditioned appetitive response becomes conditioned to a new stimulus. The two mechanisms are functionally analogous, and the difference lies in whether the initial response to the social stimulus is acquired through learning (a conditioned response) or unlearned

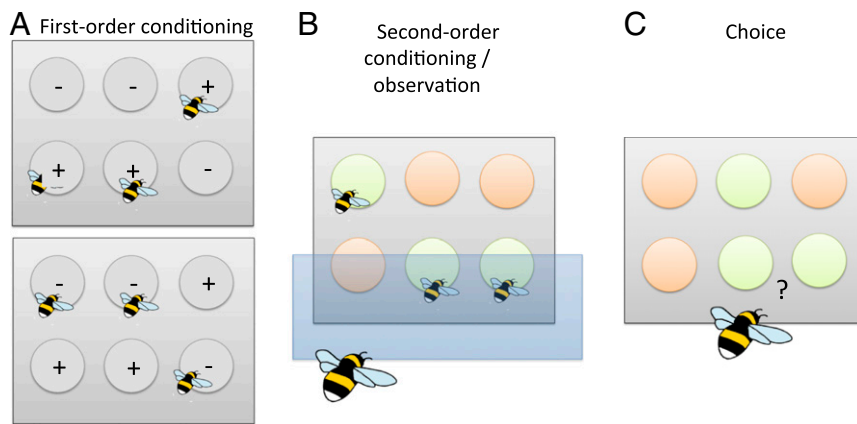


Fig. 1. Second-order conditioning of flower color preferences in bumblebees (43). (A) Observer bees were initially allowed to forage on a floral array in which the presence of conspecifics (CS1) predicted either sucrose (US+; Upper) or quinine (US-; Lower). A third group completed this training in the absence of any conspecifics (not pictured). (B) All bees then observed conspecifics (CS1) foraging on one flower color (CS2) and ignoring an alternative, through a glass screen. (C) Finally, each observer was permitted to forage alone on the colored array.

predisposition [an unconditioned response; although Heyes (29) highlights that presumed predispositions may in fact reflect learning].

Note that although we have framed this argument in terms of classic conditioning, the same case can be made for operant conditioning (57). Matched-dependent behavior (58) describes instances whereby matching of another animal's behavior (i.e., a response to a social stimulus) is rewarded (e.g., by finding food; a reinforcer) and thus increases in frequency, and Church (59) has shown that this response can become conditioned to the asocial stimulus that initially elicited the demonstrator's behavior. The key point is that responses to social stimuli—be they acquired through classic conditioning, operant conditioning, or a history of natural selection—are conditioned to new asocial stimuli. As an illustration, consider the following example (60). Social information from injured conspecifics, in the form of volatiles from stressed conspecifics, typically elicits aversion responses in

foraging honey bees (*Apis mellifera*) (61). We recently found that bees not only avoid areas where they currently detect such volatiles, but also later avoid colored lights that were experienced at the same time (60) (Fig. 2). Thus, the avoidance response, which is initially elicited by the social stimulus, becomes conditioned to a new asocial stimulus. In the laboratory, we used colored lights as asocial stimuli; in the wild, floral features that predict the presence of sit-and-wait predators, such as crab spiders (*Thomisidae* spp.) could fulfill the same role.

It is important to reiterate that this associative framework complements, rather than adds to, the collection of processes that are labeled as social learning mechanisms, such as local or stimulus enhancement, or social facilitation (1). These labels describe effects, and learning theory offers an explanation for why these effects occur, rather than an additional alternative effect (39). In some cases, an associative framework is already

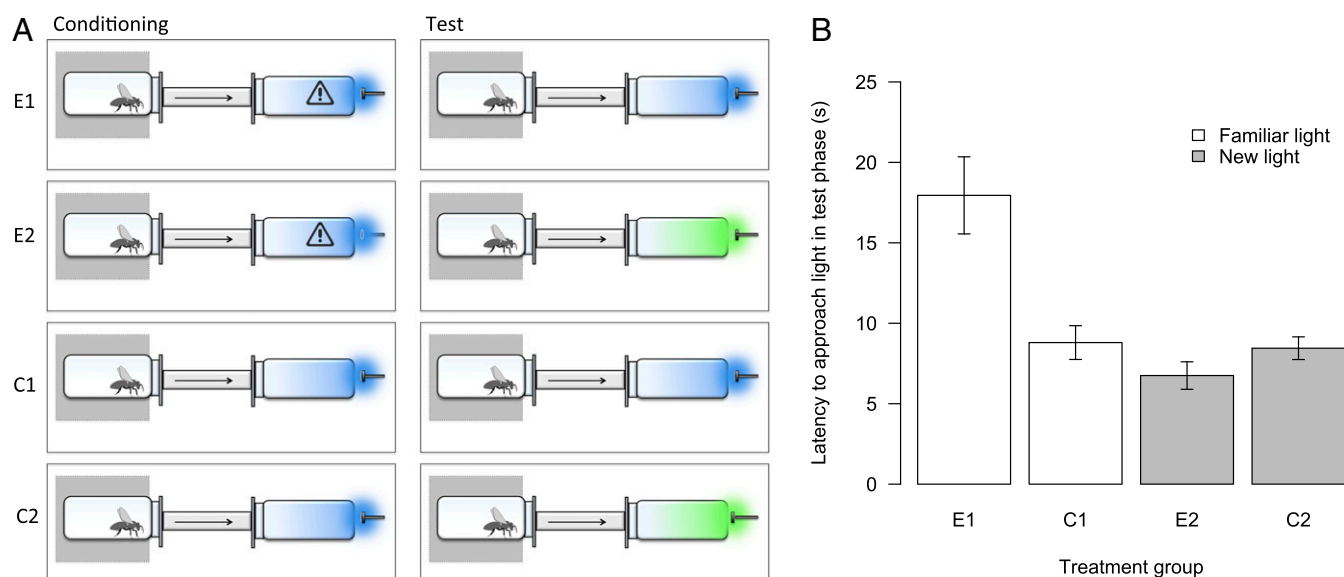


Fig. 2. Honey bees learn to respond to colored lights through exposure to alarm volatiles (60). (A) We used an assay whereby highly phototactic subjects walked up a dark tube toward a colored light (balanced blue/green design; only blue is pictured). Warning triangle indicates the presence of volatiles from a stressed conspecific. (B) In the training phase, bees in groups E1 (experimental) and E2 (control for sensitization/habituation effects) were slower to approach the light, but only group E1 were slower in the testing phase. Thus, responses were conditioned to the specific stimuli that had been contiguous with stress volatiles in the training phase.

part of the definition of a particular process (e.g., observational conditioning or matched-dependent behavior) and for others (e.g., local enhancement, stimulus enhancement, emulation, or imitation) it is not. Furthermore, our focus here is on mechanisms that are taxonomically widespread within the animal kingdom, but similar arguments have been made for processes that are typically associated with the primate lineage [32, 39, 40; see also Lotem et al. (33), who make such an argument for imitation within this issue]. Moreover, this associative framework is relevant to any form of social stimulus, irrespective of whether it is visual, olfactory, or auditory, or whether it is a specific behavior pattern, bias, or expression, or simply physical presence. The key concept is that social stimuli elicit affective states, neural representations, or motor patterns, which can then be associated with contiguous asocial stimuli.

Saliency of Social Stimuli

Above, we argued that responses to social stimuli can produce a functionally analogous outcome irrespective of whether they arise through unlearned predisposition or through learning, because they can be conditioned to new, asocial stimuli. However, the distinction between such unconditioned and conditioned responses to social stimuli is important for questions about evolved traits. A clear pathway for natural selection might be to implement small quantitative changes that render social associations more likely to be learned, such that information deriving from informative cues is acquired particularly rapidly. One means (but not the only means, as we discuss later) to do so could involve changes to upstream mechanisms that determine whether animals notice or pay attention to social stimuli. We begin with a focus on the question of whether social stimuli are more salient than less-ecologically informative alternatives.

The saliency of a stimulus—the property that renders it conspicuous or noticeable, and thus likely to be attended to—is a key determinant of the speed with which it can be associated with other stimuli (62). Saliency depends on the species-specific characteristics of the receiver's sensory system, such that a stimulus that is salient for one taxonomic group may be less so for another (63). For example, carbon disulphide (a component of rat breath) is a particularly salient stimulus for young rats, and is quickly associated with any food flavors that are encountered at the same time (64). Experimental evolution paradigms provide evidence that environments where particular cue types are reliably useful for decision-making select for changes to the attentional or perceptual mechanisms that determine stimulus saliency. For example, *Drosophila* lines for which learning about olfactory stimuli is the best method of identifying a good brood host develop enhanced sensitivity to such cues and ignore visual alternatives, and vice versa (36). In a social context, the question of whether social cues are especially salient has rarely been directly addressed [with the exception of Galef et al. (64)], but a number of bee studies touch upon the topic, with mixed results.

When two conditioned stimuli (e.g., a social and an asocial stimulus) are simultaneously paired with an unconditioned stimulus, the more salient stimulus overshadows learning about the less salient stimulus (47). Accordingly, an experiment by Dunlap and colleagues initially appears to suggest that conspecifics present might be a particularly salient stimulus for bumblebees (65). In this study, subjects (*B. impatiens*) were trained to find sucrose rewards in floral arrays where both asocial and social cues (floral color and pinned conspecifics, respectively) provided some indication of which flowers were rewarding, before assessing which of the two cue types the bees preferentially used on a test array. When both cue types had been equally reliable at predicting sucrose, the bees disproportionately favored social cues in the tests, and most surprisingly, they also used social cues even when floral cues had been more reliable. Bees only resorted to using floral cues if social cues had been entirely useless predictors of reward. These results would seem to indicate that the social stimulus is the more salient alternative, overshadowing any

learning about the floral cue, and correspondingly, the authors discuss why the ecological niche of pollinators might favor reliance on social cues. However, given that floral color is also likely a very useful cue for bees, it seems surprising that social cues should be more salient, and consideration of these findings in the context of blocking (66)—another mainstay of associative learning theory—raises a testable alternative explanation.

Association between a CS (e.g., flower color) and a US (e.g., sucrose) is typically impaired if that CS is presented together with another CS that has previously been associated with the US. This phenomenon could potentially explain Dunlap et al.'s (65) results, because the bees have prior experience of social foraging but not of the flower colors in question. As members of laboratory colonies that forage ad libitum from feeders placed in flight arenas when not participating in experiments, individuals would have had an opportunity to associate conspecific presence with reward, but no similar experience of flower colors. During the training, bees should thus learn little about the reliability of the floral CS, because the simultaneous presence of the social CS blocks learning about the asocial cue. The only situation where bees should learn about the floral CS would occur if the social CS became unreliable. In this situation, the association between conspecific presence and sucrose rewards should move toward extinction, allowing for new learning of the association between floral color and sucrose, exactly as Dunlap et al.'s results illustrate.

It may well be the case that bees treat asocial and social cues differently, and post hoc associative explanations should generate testable predictions to be useful. Here, one simple means to rule out a blocking explanation would be to closely control the previous foraging experience of the bees. In this study system, where individuals never leave the flight arena and the colony's foraging needs can be met by providing sucrose and pollen directly into the nest, it is entirely possible to create individual foragers with no experience of social foraging (43, 67, 68). Finding the same effect under these circumstances would render the case for enhanced saliency of social stimuli compelling, and a further experiment by Dawson and Chittka (68) employs such an approach. These authors compared the speed at which bumblebees (*B. terrestris*) learned to associate either a social CS (the presence of dead, pinned conspecifics) or an asocial CS (a coin/plastic disk/black wooden cuboid) with sucrose, in a free-flying choice paradigm. Bees were more likely to associate the social than asocial CS with reward, and were also more likely to use the social CS to identify rewarding flowers in a transfer test involving a novel target flower color. On realizing that their results could reflect the fact that subjects had previous experience of social foraging from laboratory feeders, Dawson and Chittka repeated their experiment using bees that had never foraged with others, and found that the effect was maintained. Interestingly, pinned honey bee (*A. mellifera*) demonstrators, which visit comparable but not identical floral resources, elicited similar results, suggesting that natural selection might lead to increased saliency of cues that derive from useful heterospecifics, as well as conspecifics.

Social Associations

Saliency is an important determinant of the ease with which an association is acquired, but it is by no means the only contributing factor. Whereas saliency affects the extent to which a stimulus is made available for learning, learning itself requires that associations between neural representations of stimuli, affective states, or motor patterns, are formed around that stimulus. Natural selection might act upon the parameters that determine the number and timing of exposures that are required before a particular association is committed to memory (69). The key feature of such prepared learning is that certain combinations of stimuli, rather than any particular stimulus alone, elicit rapid learning. For example, consider the oft-cited "Garcia effect," whereby rats rapidly learn to associate tastes but not

audible tones with gastrointestinal illness. The combination of taste and illness is critical here, and the same effect is not observed when tastes are associated with shock (35, 70), ruling out the suggestion that tastes are simply more salient stimuli than tones. In fact, Garcia and Koelling (35) found that tones were more easily associated with shock than taste was, potentially because loud noises are likely to be a relevant cue to imminent pain, whereas taste is not.

In a social learning context, there are clear candidate hypotheses concerning stimulus combinations that await investigation. For example, above we discussed evidence that bees might acquire associations between nectar rewards and social stimuli more rapidly than associations between nectar rewards and asocial stimuli (65, 68). A prepared learning hypothesis would predict that social stimuli might be particularly easy to associate with sucrose reward levels, but not with an aversive stimulus. Or alternatively, perhaps the sign of the CS-US relationship is also important, such that bees learn positive sucrose/social cues relationships easily but not negative ones. These latter two alternatives—that bees are particularly sensitive to social CSs when learning about where to find food, or that they are particularly likely to learn positive social CS-sucrose combinations—are qualitatively different traits. Theory correspondingly predicts that they should evolve under different circumstances (36, 71). To visualize the difference, consider again the Garcia effect (35). Although many stimuli might precede a feeling of illness, the true cause will often be related to recently eaten food, so an a priori prioritization of associations between taste and illness, rather than sound and illness, makes evolutionary sense. Now consider the situation where a particular flavor always predicts illness. Here, theory would predict the fixation of an aversion to that flavor, rather than prepared learning (71). Thus, a priori expectations and prepared learning about social stimuli are both means by which natural selection could facilitate social learning processes, but the ecological conditions that favor their evolution are distinct. Teasing apart the roles of stimulus salience, prepared learning, and a priori expectation is a task that invites empirical exploration in our social insect system and in others.

Retrieval and Implementation of Learned Associations

We have discussed potential pathways by which natural selection could modify stimulus salience, or the downstream learning parameters that influence memory formation, and suggested means by which such hypotheses could be explored. However, we have not yet touched upon the possibility that selection could produce modifications to the final retrieval or implementation of learned information. This is particularly important in light of the large volume of literature on “social learning strategies” that describes how animals use social information most often in those situations in which it is most beneficial (72, 73). We begin with an example that illustrates how associative processes could bring about such context-specificity.

As we have alluded to in an earlier example in this paper, foraging bees suffer predation by camouflaged sit-and-wait predators that ambush individuals as they land on flowers to feed. Dawson and Chittka (74) allowed bees to forage in environments that mimicked high or low risk of such predation, and found that bumblebees (*B. terrestris*) appear to use social information to identify safe flowers specifically in dangerous environments. Their subjects were initially trained on an array where landing on flowers of one color morph led to brief capture in a pressure trap (simulating spider attack), whereas an alternative color morph was safe. Note that the color morphs simulate flower species with different levels of spider occupancy, rather than spiders themselves, which are typically cryptic. When subsequently tested on an array containing only flowers of the dangerous morph, bees strongly preferred to land on the single flower where a live demonstrator could be seen feeding, an effect that was entirely absent when tested on flowers of the safe morph (Fig. 3). In other words, bees seemed to use social information adaptively when foraging

on flower types where predation was a real threat. Bees could not have simply learned that the social cue was useful on dangerous flowers but irrelevant on safe ones because individuals were trained on the mixed array entirely alone.

How might associations between stimuli produce such context-specific and potentially adaptive behavior? Conditioned suppression is an associative phenomenon widely used to study the acquisition and extinction of fear, and it predicts exactly the effect that Dawson and Chittka (74) demonstrated in bumblebees (30). In the presence of a cue predicting an aversive stimulus that an animal would usually avoid, learned food-seeking behaviors are typically repressed, although they do not disappear altogether (75). Thus, a rat that has learned to press a lever for food typically reduces the frequency of pressing when a light that predicts foot shock is turned on (76). Similarly, when under predation pressure from a threat that is hard to directly detect, bumblebees increase both their latency to probe and rejection rate of all flowers of the color morph associated with danger (77, 78). With this in mind, picture a bee that enters a flight arena filled with flowers of the safe color morph, and happens to first come across an unoccupied flower. Because the bee has learned that the holes in the flowers reliably contain sucrose rewards, it is very likely to land and feed. If it instead first comes across an occupied flower where a demonstrator is foraging, it is also very likely to land and feed. Even if the presence of the demonstrator renders flowers much more attractive, the difference in acceptance rates between unoccupied (very attractive) and occupied (extremely attractive) flowers will be hard to detect experimentally, because all flowers are very likely to be accepted. Now consider an environment where the floral color morph predicts danger. All conditioned responses will be suppressed, such that flowers in general are less likely to be accepted. In this situation, if the presence of a demonstrator bee is attractive, the effect is much more likely to be detected experimentally (Fig. 4), because it is no longer the case that a bee almost invariably accepts the first flower that it encounters.

This associative hypothesis does not require that selection deriving from risky contexts has influenced the weight that bees ascribe to social information. It is simply an alternative to the suggestion that individuals strategically respond to the circumstances in which they find themselves by computing the likely pay-offs of to social information use. However, our hypothesis again generates a testable prediction. If dangerous environments simply render the effect of an attractive social stimulus more detectable, the same should be true for an attractive asocial stimulus. Thus, replacing conspecific demonstrators with asocial stimuli that have previously been conditioned to sucrose should produce analogous results.

A study by Smolla et al. (79) employs exactly this approach in a different social learning context. Based on the premise that bees should use a “copy-when-uncertain” strategy, which follows from an agent-based model that they develop, these authors pretrained bees (*B. terrestris*) that either a social (model bee) or an asocial (green rectangle) cue predicted reward in a floral array. Half of the bees in each group were then trained that the floral array contained highly variable rewards and half learned that rewards were constant, in the absence of both cue types. When subsequently presented with a nonrewarding test array, those bees that had learned that rewards were variable used the social cue to find food, but those that had experienced the constant array did not. Their results thus support a “copy-when-uncertain” interpretation, but as pointed out above, the fact that more variable rewards render flowers less attractive (80, 81) means that the use of any cue, not just social ones, should be rendered more detectable. However, crucially, the difference between contexts was much less evident for the asocial cue and did not reach statistical significance. This difference seems unlikely to be attributed to greater salience or associability of the social cue, because Smolla et al. (79) state that pretraining was similarly successful for both cue types. Smolla et al.’s results invite further exploration that has not yet

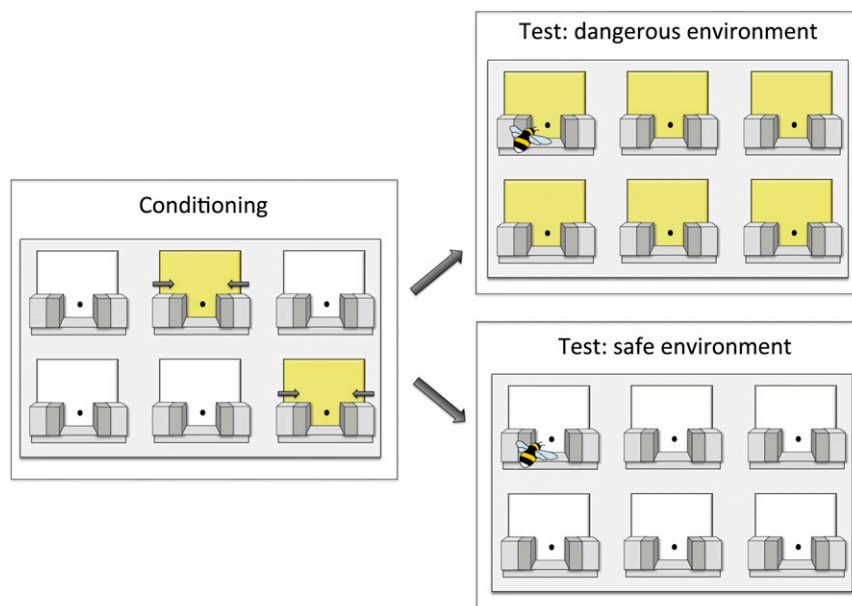


Fig. 3. Social information use varied with context (74). Bees were trained that sucrose could be obtained by extending the proboscis through holes demarcated by colored squares. During training, one color (here yellow) predicted brief capture in a pressure trap. Bees were then tested in either a “dangerous” or a “safe” environment, where one live demonstrator was foraging at a single location.

been carried out, but their approach of comparing context-specific responses to social and asocial stimuli is a promising one that seems a useful way to evaluate the specific characteristics that govern learning about social stimuli.

Competing Hypotheses

As a whole, the insect-based studies that we have discussed illustrate the power of associative learning to generate adaptive behavior. There is nothing new about associative explanations for social learning phenomena; Heyes in particular has long championed this approach (24, 28–30, 39, 40). However, the fact that a hypothesis has explanatory power is not evidence of its truth, and consequently, debate arises over whether associative

learning should be accredited with the status of a “default” explanation for social learning phenomena, to be assumed true unless proven false (24). Arguments that favor this approach in animal cognition are typically based on Morgan’s canon, which states that “no animal activity should be interpreted in terms of higher psychological processes if it can be fairly interpreted in terms of processes that stand lower in the scale of psychological evolution and development” (82), but this has led to debate over whether Morgan’s canon itself fit up to the task (24).

Perhaps such a polarized perspective is less productive than it could be. If there were clear evidence that social learning processes were more efficient in species where social information repeatedly presents itself than in more solitary species (for

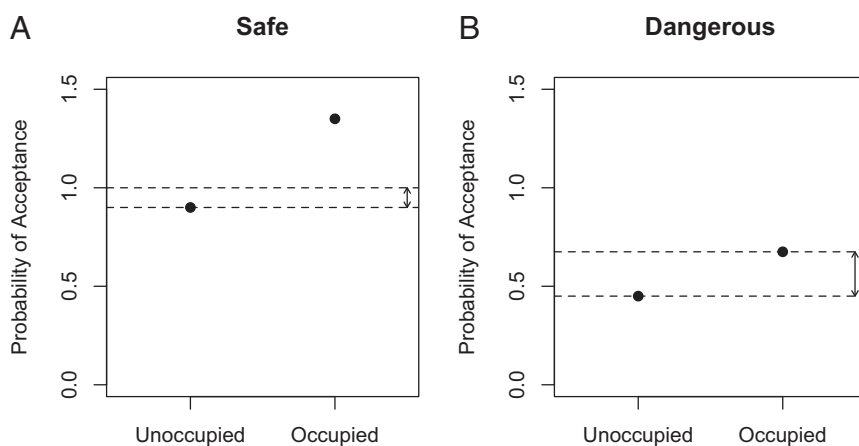


Fig. 4. Conditioned suppression should render social effects more detectable when aversive stimuli are present. (A) Safe environment. We assume that bees have learned to associate the feeding holes in artificial flowers with sucrose rewards. Therefore, in the absence of aversive stimuli, the chances of acceptance on encountering a flower are high (here set at 0.9, for illustration purposes). The presence of a demonstrator is also attractive (here set at 0.45), perhaps because bees have previously learned to associate conspecifics with sucrose. If two appetitive conditioned stimuli are presented together, the subject’s expectation is equal to their combined strength (62), so occupied flowers are very attractive, but a probability of acceptance cannot exceed 1. Thus, the detectable effect of demonstrator presence is small (arrow). (B) Dangerous environment. The presence of an aversive stimulus (the dangerous flower color) reduces all responses for food (here, suppression ratio has been set at 0.5). Thus, the difference in the probability of acceptance between the unoccupied and occupied flowers is now relatively more detectable (arrow).

example, in the social vs. solitary insects, which span multiple origins of sociality), it might be argued that the two are sufficiently different traits to justify discussion of evolutionary parsimony. However, such evidence is sparse. Most studies that compare learning efficiency across social and asocial contexts conclude that the two are highly correlated (83–85; but see also ref. 86). Here, we have suggested that a productive way forward might be to search for small quantitative differences between associative processes in asocial and social contexts, rather than qualitative leaps. Heyes (29) has highlighted that the pathway of least resistance for natural selection might be to modify the input mechanisms that determine which aspects of the world are made available for learning, contrasting such processes with learning mechanisms that determine how stimuli are linked and committed to memory. Increased salience of social stimuli is one such input mechanism, but increased salience of social stimuli could be evolutionarily advantageous for many reasons other than acquiring information, such as recognizing kin, selecting mates, or defending a territory (87). Perhaps more convincing evidence that natural selection has shaped responses to social information would be evidence of prepared learning about social stimuli, of unconditioned responses to social stimuli that were specific to a social learning context, or of social learning strategies that cannot be accounted for by associative learning theory.

Our choice of study taxon—the social insects—has meant that we have made little mention of processes that characterize mainly primate behavior, such as imitation and (to a lesser extent) emulation. Nonetheless, very recent work has begun to focus on potentially emulative behavior even in bees (8, 9), and it is not our intention to imply that such processes follow a different evolutionary pathway to other forms of social learning. In fact, associative explanations for imitation are prominent in the psychology literature [40, 88; see also Lotem et al. (33)]. Explaining how individuals copy a novel sequence of actions through imitation invokes a “correspondence problem” (89, 90) because the seen movements of others must somehow be matched to motor representations of self-movements. Associative models of imitation propose that such links could arise through previous experience of contingency between performing an action and seeing it performed (91). For example, an infant might often observe others smiling when she or he smiles, leading to association between the visual and motor representations of smiling. She or he will also typically observe a raised arm each time that she raises her own arm, or react to an unexpected surprise with the same expressions as others nearby. This elegant idea generates both theoretical debate and extensive empirical exploration. However, imitation is but one social learning mechanism, and this intensive exploration

of mechanism is not the norm. It is surprising that for the vast majority of cases in which animals respond to social information, we understand very little about underlying mechanisms at all (41). To return to Galef’s statement that “an entire field of local enhancement awaits exploration” (41), we suggest that this exploration begins with a full understanding of how reported “simpler” social learning phenomena fit into an associative framework.

Conclusion

We became interested in insect social learning based on a thought-provoking anecdotal observation that the acceptance of a laboratory feeder by a single bumblebee seemed to render that feeder popular, rather than through careful choice of a study system. It was surprising to subsequently find that insects seemed capable of what we considered a relatively sophisticated cognitive process. However, questions about cognition in invertebrates inevitably involve detailed dissection of mechanisms, and a critical feature of this system has been the ability to closely control the previous social foraging experience of experimental subjects in a way that would not be possible for many species. We hope that the body of research presented above shows that there was value in the novel perspective offered by a rather unlikely model. The studies that we have described are based on a small-brained study organism, but one that excels at accomplishing the cognitive tasks that are relevant to its own ecological niche (92), that is highly social, and that offers exceptional experimental tractability. Other insect systems offer similar advantages for investigating social learning contexts that we have not mentioned in detail here; for example, a large literature now documents the existence of remarkably rapid and generalizable mate-choice copying in *Drosophila* fruit flies (3, 11, 12). It may well be the case that social learning abilities have traveled further along some evolutionary routes in other lineages, and less far in others, but the rules that govern associative learning are taxonomically widespread within the animal kingdom, and natural selection may well coopt the same cognitive raw material across multiple evolutionary lineages.

ACKNOWLEDGMENTS. We thank the organizers and funders of the Arthur M. Sackler Colloquium on “The Extension of Biology Through Culture,” from which this paper derives, the many participants at the colloquium who provided informative feedback and discussion, and Simon Reader for comments on an earlier draft of the manuscript. Several of the empirical studies presented here were coauthored by Lars Chittka, and the ideas that we have discussed owe much to his direction and guidance. E.L. is funded by European Research Council Starting Grant BeeDanceGap, and the empirical work described here also derives from an Early Career Fellowship from The Leverhulme Trust. E.H.D.’s current position is funded by a Fyssen foundation postdoctoral fellowship.

- Hoppitt W, Laland K (2013) *Social Learning: An Introduction to Mechanisms, Methods and Models* (Princeton Univ Press, Princeton, NJ).
- Grüter C, Leadbeater E (2014) Insights from insects about adaptive social information use. *Trends Ecol Evol* 29:177–184.
- Leadbeater E, Chittka L (2007) Social learning in insects—From miniature brains to consensus building. *Curr Biol* 17:R703–R713.
- Mery F, et al. (2009) Public versus personal information for mate copying in an invertebrate. *Curr Biol* 19:730–734.
- Worden BD, Papaj DR (2005) Flower choice copying in bumblebees. *Biol Lett* 1: 504–507.
- Leadbeater E, Chittka L (2007) The dynamics of social learning in an insect model, the bumblebee (*Bombus terrestris*). *Behav Ecol Sociobiol* 61:1789–1796.
- Leadbeater E, Chittka L (2008) Social transmission of nectar-robbing behaviour in bumble-bees. *Proc Biol Sci* 275:1669–1674.
- Alem S, et al. (2016) Associative mechanisms allow for social learning and cultural transmission of string pulling in an insect. *PLoS Biol* 14:e1008529; erratum in 14(12): e1008529.
- Loukola OJ, Perry CJ, Coscos L, Chittka L (2017) Bumblebees show cognitive flexibility by improving on an observed complex behavior. *Science* 355:833–836.
- Battesti M, Moreno C, Joly D, Mery F (2015) Biased social transmission in *Drosophila* oviposition choice. *Behav Ecol Sociobiol* 69:83–87.
- Dagaëff A-C, et al. (2016) *Drosophila* mate copying correlates with atmospheric pressure in a speed-learning situation. *Anim Behav* 121:163–173.
- Battesti M, Moreno C, Joly D, Mery F (2012) Spread of social information and dynamics of social transmission within *Drosophila* groups. *Curr Biol* 22:309–313.
- Goulson D, Park KJ, Tinsley MC, Bussiere LF, Vallejo-Marin M (2013) Social learning drives handedness in nectar-robbing bumblebees. *Behav Ecol Sociobiol* 67:1141–1150.
- Danchin E, Giraldeau LA, Valone TJ, Wagner RH (2004) Public information: From noisy neighbors to cultural evolution. *Science* 305:487–491.
- Thompson R, McCONNELL J (1955) Classical conditioning in the planarian, *Dugesia dorotocephala*. *J Comp Physiol Psychol* 48:65–68.
- Kemenes G, Benjamin PR (1989) Appetitive learning in snails shows characteristics of conditioning in vertebrates. *Brain Res* 489:163–166.
- Walters ET, Carew TJ, Kandel ER (1981) Associative learning in *Aplysia*: Evidence for conditioned fear in an invertebrate. *Science* 211:504–506.
- Spatz HC, Emanns A, Reichert H (1974) Associative learning of *Drosophila melanogaster*. *Nature* 248:359–361.
- Rescorla RA (1988) Pavlovian conditioning. It’s not what you think it is. *Am Psychol* 43: 151–160.
- Holland PC, Sherwood A (2008) Formation of excitatory and inhibitory associations between absent events. *J Exp Psychol Anim Behav Process* 34:324–335.
- Timberlake W (1994) Behavior systems, associationism, and Pavlovian conditioning. *Psychon Bull Rev* 1:405–420.
- Pearce JM, Bouton ME (2001) Theories of associative learning in animals. *Annu Rev Psychol* 52:111–139.
- Dickinson A (2012) Associative learning and animal cognition. *Philos Trans R Soc Lond B Biol Sci* 367:2733–2742.

24. Heyes C (2012) Simple minds: A qualified defence of associative learning. *Philos Trans R Soc Lond B Biol Sci* 367:2695–2703.
25. Dwyer D, Burgess KV (2011) Rational accounts of animal behaviour? Lessons from C. Lloyd Morgan's canon. *Int J Comp Psychol* 24:349–364.
26. Guilford T, Burt De Perera T (2016) An associative account of avian navigation. *J Avian Biol* 48:191–195.
27. Shettleworth SJ (2010) Clever animals and killjoy explanations in comparative psychology. *Trends Cogn Sci* 14:477–481.
28. Heyes CM (1994) Social learning in animals: Categories and mechanisms. *Biol Rev Camb Philos Soc* 69:207–231.
29. Heyes C (2012) What's social about social learning? *J Comp Psychol* 126:193–202.
30. Heyes C, Pearce JM (2015) Not-so-social learning strategies. *Proc Biol Sci* 282:7.
31. Galef BG (1992) The question of animal culture. *Hum Nat* 3:157–178.
32. Lotem A, Halpern JY (2012) Coevolution of learning and data-acquisition mechanisms: A model for cognitive evolution. *Philos Trans R Soc Lond B Biol Sci* 367:2686–2694.
33. Lotem A, Halpern JY, Edelman S, Kolodny O (2017) The evolution of cognitive mechanisms in response to cultural innovations. *Proc Natl Acad Sci USA* 114:7915–7922.
34. Mery F, et al. (2007) Natural polymorphism affecting learning and memory in *Drosophila*. *Proc Natl Acad Sci USA* 104:13051–13055.
35. Garcia JK, Koelling RA (1966) Relation of cue to consequence in avoidance learning. *Psychon Sci* 4:123–124.
36. Dunlap AS, Stephens DW (2014) Experimental evolution of prepared learning. *Proc Natl Acad Sci USA* 111:11750–11755.
37. Truskanov N, Lotem A (2017) Trial-and-error copying of demonstrated actions reveals how fledglings learn to 'imitate' their mothers. *Proc Biol Sci* 284:20162744.
38. Ho MK, MacGlashan J, Littman ML, Cushman F (March 21, 2017) Social is special: A normative framework for teaching with and learning from evaluative feedback. *Cognition*, 10.1016/j.cognition.2017.03.006.
39. Heyes C (2017) When does social learning become cultural learning? *Dev Sci* 20:e12350.
40. Cook R, Bird G, Catmur C, Press C, Heyes C (2014) Mirror neurons: From origin to function. *Behav Brain Sci* 37:177–192.
41. Galef BG (2013) Imitation and local enhancement: Detrimental effects of consensus definitions on analyses of social learning in animals. *Behav Processes* 100:123–130.
42. Heinrich B (1979) *Bumblebee Economics*. (Harvard Univ Press, Cambridge, MA).
43. Dawson EH, Avargués-Weber A, Chittka L, Leadbeater E (2013) Learning by observation emerges from simple associations in an insect model. *Curr Biol* 23:727–730.
44. Avargués-Weber A, Chittka L (2014) Observational conditioning in flower choice copying by bumblebees (*Bombus terrestris*): Influence of observer distance and demonstrator movement. *PLoS One* 9:e88415.
45. Schultz W, Dickinson A (2000) Neuronal coding of prediction errors. *Annu Rev Neurosci* 23:473–500.
46. Rescorla R (1980) *Pavlovian Second-Order Conditioning: Studies in Associative Learning* (Laurence Erlbaum Associates, Hillsdale, NJ).
47. Pavlov IP (1927) *Conditioned Reflexes* (Oxford Univ Press, Oxford).
48. Winterbauer NE, Balleine BW (2005) Motivational control of second-order conditioning. *J Exp Psychol Anim Behav Process* 31:334–340.
49. Chittka L, Leadbeater E (2005) Social learning: Public information in insects. *Curr Biol* 15:R869–R871.
50. Cook M, Mineka S (1989) Observational conditioning of fear to fear-relevant versus fear-irrelevant stimuli in rhesus monkeys. *J Abnorm Psychol* 98:448–459.
51. Mineka S, Cook M (1988) Social learning and the acquisition of snake fear in monkeys. *Social Learning: Psychological and Biological Perspectives*, eds Zentall TR, Galef BG (Laurence Erlbaum Associates, Hillsdale, NJ), pp 51–75.
52. Mineka S, Cook M (1993) Mechanisms involved in the observational conditioning of fear. *J Exp Psychol Gen* 122:23–38.
53. Ohman A (2009) Of snakes and faces: An evolutionary perspective on the psychology of fear. *Scand J Psychol* 50:543–552.
54. Cook EW, 3rd, Hodes RL, Lang PJ (1986) Preparedness and phobia: Effects of stimulus content on human visceral conditioning. *J Abnorm Psychol* 95:195–207.
55. Tomarken AJ, Sutton SK, Mineka S (1995) Fear-relevant illusory correlations: What types of associations promote judgmental bias? *J Abnorm Psychol* 104:312–326.
56. Ohman A, Mineka S (2003) The malicious serpent: Snakes as a prototypical stimulus for an evolved module of fear. *Curr Dir Psychol Sci* 12:5–9.
57. Zentall TR, Galef BG, eds (1988) *Social Learning: Psychological and Biological Perspectives* (Laurence Erlbaum Associates, Hillsdale, NJ).
58. Miller NE, Dollard J (1941) *Social Learning and Imitation* (Yale Univ Press, New Haven, CT).
59. Church RM (1968) Applications of behaviour theory to social psychology. *Social Facilitation and Imitative Behavior*, eds Simmel EC, Hoppe RA, Milton GD (Allyn & Bacon, Boston).
60. Dawson EH, Chittka L, Leadbeater E (2016) Alarm substances induce associative social learning in honeybees, *Apis mellifera*. *Anim Behav* 122:17–22.
61. Balderrama N, et al. (1996) A deterrent response in honeybee (*Apis mellifera*) foragers: Dependence on disturbance and season. *J Insect Physiol* 42:463–470.
62. Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, eds Black AH, Prokasy WF (Appleton-Century-Crofts, NY), pp 64–99.
63. Shettleworth S (2010) *Cognition, Evolution and Behavior* (Oxford Univ Press, Oxford).
64. Galef BG, Jr, Mason JR, Preti G, Bean NJ (1988) Carbon disulfide: A semiochemical mediating socially-induced diet choice in rats. *Physiol Behav* 42:119–124.
65. Dunlap AS, Nielsen ME, Dornhaus A, Papaj DR (2016) Foraging bumble bees weigh the reliability of personal and social information. *Curr Biol* 26:1195–1199.
66. Kamin LJ (1969) Predictability, surprise, attention, and conditioning. *Punishment and Aversive Behavior*, eds Campbell BA, Church RM (Appleton-Century-Crofts, New York), pp 279–296.
67. Leadbeater E, Chittka L (2011) Do inexperienced bumblebee foragers use scent marks as social information? *Anim Cogn* 14:915–919.
68. Dawson EH, Chittka L (2012) Conspecific and heterospecific information use in bumblebees. *PLoS One* 7:e31444.
69. Domjan M, Galef BG (1983) Biological constraints on instrumental and classical-conditioning—Retrospect and prospect. *Anim Learn Behav* 11:151–161.
70. Dwyer DM (2015) Experimental evolution of sensitivity to a stimulus domain alone is not an example of prepared learning. *Proc Natl Acad Sci USA* 112:E385.
71. Dukas R (1998) Evolutionary ecology of learning. *Cognitive Ecology*, ed Dukas R (Univ of Chicago Press, Chicago), pp 129–174.
72. Laland KN (2004) Social learning strategies. *Learn Behav* 32:4–14.
73. Kendal RL, Coolen I, Laland KN (2009) Adaptive trade-offs in the use of social and personal information. *Cognitive Ecology II*, eds Dukas R, Ratcliffe JM, pp 249–271.
74. Dawson EH, Chittka L (2014) Bumblebees (*Bombus terrestris*) use social information as an indicator of safety in dangerous environments. *Proc Biol Sci* 281:20133174.
75. Annau Z, Kamin LJ (1961) Conditioned emotional response as a function of intensity of US. *J Comp Physiol Psych* 54:428–432.
76. Estes WK, Skinner BF (1941) Some quantitative properties of anxiety. *J Exp Psychol* 29:390–400.
77. Ings TC, Chittka L (2008) Speed-accuracy tradeoffs and false alarms in bee responses to cryptic predators. *Curr Biol* 18:1520–1524.
78. Lenz F, Ings TC, Chittka L, Chechkin AV, Klages R (2012) Spatio-temporal dynamics of bumblebees foraging under predation risk. *Phys Rev Lett* 108:098103.
79. Smolla M, Alem S, Chittka L, Shultz S (2016) Copy-when-uncertain: Bumblebees rely on social information when rewards are highly variable. *Biol Lett* 12:20160188.
80. Shafir S, Wiegmann DD, Smith BH, Real LA (1999) Risk-sensitive foraging: Choice behaviour of honeybees in response to variability in volume of reward. *Anim Behav* 57:1055–1061.
81. Seefeldt S, De Marco RJ (2008) The response of the honeybee dance to uncertain rewards. *J Exp Biol* 211:3392–3400.
82. Lloyd Morgan C (1909) *An Introduction to Comparative Psychology* (Walter Scott Publishing, London).
83. Lefebvre L, Giraldeau L-A (1996) Is social learning an adaptive specialization? *Social Learning in Animals: The Roots of Culture*, eds Heyes CM, Galef BG, Jr (Academic, London), pp 107–128.
84. Reader SM, Hager Y, Laland KN (2011) The evolution of primate general and cultural intelligence. *Philos Trans R Soc Lond B Biol Sci* 366:1017–1027.
85. Reader SM (2003) Innovation and social learning: Individual variation and brain evolution. *Animal Biology* 53:147–158.
86. Templeton JJ, Kamil AC, Balda RP (1999) Sociality and social learning in two species of corvids: the pinyon jay (*Gymnorhinus cyanocephalus*) and the Clark's nutcracker (*Nucifraga columbiana*). *J Comp Psychol* 113:450–455.
87. Leadbeater E (2015) What evolves in the evolution of social learning? *J Zool (Lond)* 295:4–11.
88. Heyes C (2016) *Homo imitans?* Seven reasons why imitation couldn't possibly be associative. *Philos Trans R Soc Lond B Biol Sci* 371:20150069.
89. Brass M, Heyes C (2005) Imitation: Is cognitive neuroscience solving the correspondence problem? *Trends Cogn Sci* 9:489–495.
90. Nehaniv CL, Dautenhahn K (2002) The correspondence problem. *Imitation in Animals and Artifacts*, eds Dautenhahn K, Nehaniv CL (MIT Press, Cambridge, MA).
91. Catmur C, Walsh V, Heyes C (2009) Associative sequence learning: The role of experience in the development of imitation and the mirror system. *Philos Trans R Soc Lond B Biol Sci* 364:2369–2380.
92. Chittka L, Thomson JD (2001) *Cognitive Ecology of Pollination* (Cambridge Univ Press, Cambridge, UK).