

Conditioning and Associative Learning

Alice Mason, Elliot A. Ludvig, and Christopher R. Madan

Abstract

Associative learning is the process whereby humans and other animals learn the predictive relationship between cues in their environment. This process underlies simple forms of learning from rewards, such as classical and operant conditioning. This chapter introduces the basics of associative learning and discusses the role that memory processes play in the establishment and maintenance of this learning. It then discusses the role that associative learning plays in human memory, including through paired associate learning, the enhancement of memory by reward, and the formation of episodic memories. Next, the chapter illustrates how the memory process influences choice in decision-making, where associative learning allows people to learn the values of different options. The chapter concludes with some suggestions about how models of associative learning, memory, and choice can be integrated into a single theoretical framework.

Key Words: associative learning, episodic memory, reinforcement learning, conditioning, reward, decision-making

Introduction

Associative learning is one of the simplest and most ubiquitous forms of learning. In associative learning, humans and other animals learn that one cue or behavior predicts the presence of another stimulus. For example, in the most famous example of associative learning, Pavlov trained dogs that a sound cue predicted the presence of a rewarding stimulus (food), thereby training them to salivate to the cue (Pavlov, 1927). Associative learning, however, is everywhere in our lives, from the marketing campaigns that associate celebrities with famous brands to the gamified reward structures in many video games and apps (Bower & Landreth, 2001; Dean & Biswas, 2001; Deterding et al., 2011; Janiszewski & Warlop, 1993; Peter & Nord, 1982; Rothschild & Gaidis, 1981; Stuart et al., 1987; Walvis, 2008).

Memory plays a crucial role in associative learning. The learning that animals exhibit is a function of what they remember about the learning situation. First, even for the simplest associations, animals form a memory for the relationship between a stimulus and an outcome. In addition, in more complex situations with multiple cues and rewards, memory determines which past associations animals remember for any given cue. These memories are often context-dependent, with specific associations more likely to be recalled in the context in which they originally occurred (Bouton, 1993; Bouton & Moody, 2004). The converse relationship also holds: Memory is influenced by associative learning, perhaps most clearly in cases in which memory for paired items is being tested or when the influence of rewards on

memory is assessed. In the first part of this chapter, we provide an overview of the basics of associative learning. We then examine how associative learning influences memory, leading to a discussion of more recent work on memory-based decision-making.

Basics of Associative Learning

Associative learning occurs when humans and other animals learn the connection between two elements in the world. There are two primary types of associative learning, which can be distinguished based on what is being associated. In classical or Pavlovian conditioning, animals learn an association between two cues. One of the cues is typically a neutral cue, such as a light or tone, and is known as the conditioned stimulus (CS). The other cue is typically a biologically relevant rewarding or aversive cue, such as food or water or an electric shock, and is known as the unconditioned stimulus (US). After repeated pairings of the CS and the US, animals gradually learn to emit an appropriate response (the conditioned response [CR]) in advance of the upcoming stimulus, such as salivating to impending food or blinking before an impending air puff to the eye.

In operant or instrumental conditioning, animals instead learn an association between actions or behaviors and a reward, thereby changing the frequency of those behaviors. These behaviors are not necessarily linked directly to the reward (as with a classical CR), but they can be more arbitrary than a CR. In lab studies of operant conditioning, the responses most often studied are lever presses in rats and mice or key pecks in pigeons. More generally, however, pretty much any response or behavior can be increased or decreased by providing appropriate reinforcement in close association with the action (Ferster & Skinner, 1957), although there are some limitations based on the natural behavioral repertoire of the creature (Breland & Breland, 1961; Staddon & Simmelhag, 1971).

These two types of associative learning share some key properties in terms of how they are learned and any corresponding response acquired. First, the learning is typically incremental—that is, the process is slow, and multiple pairings between the cue/behavior and the reward are usually needed before the association is learned (Figure 12.1). Second, in both cases, the key feature that determines whether learning occurs is the contingency between the stimuli—the degree to which one cue/behavior is predictive of the rewarding outcome. A key difference is

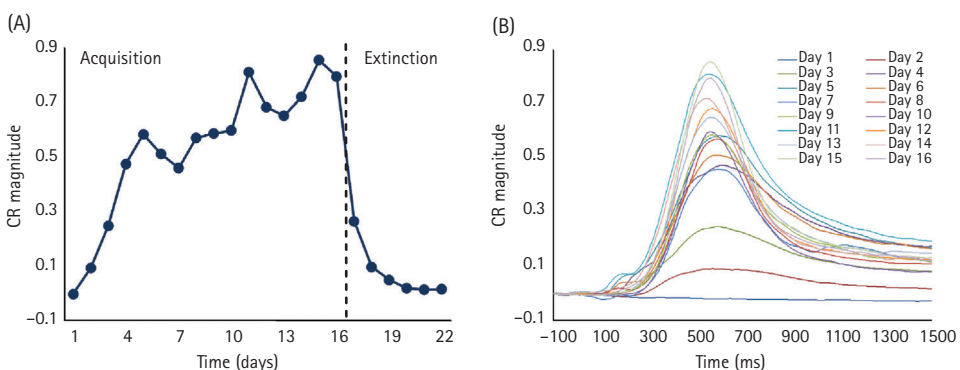


Figure 12.1 Acquisition and extinction in a classical eyeblink conditioning in the rabbit. (A) The time course of learning. During acquisition, as the rabbit learns the association between the predictive cue and outcome, the eyeblink conditioned response (CR) gets stronger. This CR quickly disappears during extinction, when the cue is repeatedly presented without the outcome. (B) The CR is extremely adaptive and well timed: The rabbit blinks at almost exactly the right time from very early in learning. Data adapted from Kehoe et al. (2014).

in the associations that are formed. In classical conditioning, a stimulus–outcome association is formed between the predictive cue and the reward. In operant conditioning, two types of associations may form: a stimulus–response association, whereby the relationship between the available cues and the response are strengthened by the reinforcer, or a response–outcome association directly between the cue and the eventual reinforcer. These three simple types of associations are the fundamental elements of what is learned and remembered in associative learning.

Theories of Associative Learning

Associative learning often appears as a slow, gradual learning process whereby the CR grows with repeated pairings between the cue and reward. Figure 12.1 shows a classic learning curve from classical conditioning, where rabbits learn to blink in response to a tone that predicts upcoming stimulation to their eye region. With each successive pairing, the blink CR becomes more likely, stronger, and better timed (Kehoe et al., 2008, 2014). The initial learning is more rapid, but the growth in the CR slows as the behavior approaches its maximal level. Early work assumed that the association between a CS and the US developed simply due to temporal contiguity, when the two stimuli co-occur (Pavlov, 1927). A similar principle of contiguity also features in many models of episodic memory whereby recall of an item is facilitated by the recall or presentation of another item that occurred close by in time (e.g., Kahana, 1996).

Subsequent work, however, has shown that temporal contiguity is not sufficient for the development of an association in conditioning. This insufficiency is perhaps best illustrated through the phenomenon of blocking (Kamin, 1968). In a blocking experiment, animals are first extensively trained that a given cue (CSA) reliably predicts the US, as in standard conditioning. After that training, animals are then presented with two cues together (i.e., CSA and CSB) paired with the same US. Here, the second cue, CSB, also has a perfect predictive relationship with the US, but animals (and humans) do not show much conditioned responding to this “blocked” CSB cue. When each CS is presented alone, animals still respond to the CSA, but they show greatly attenuated responding to the CSB, despite the perfect temporal contiguity between the CSB and the US.

Instead, what seems to be the key feature for the development of an association in conditioning is the presence of a contingency or predictive relationship between the CS and the US (Rescorla, 1967; Shanks, 2007). In a classic study, Rescorla (1967) showed that presenting the US in the absence of the CS degraded the strength of the association, even while maintaining perfect temporal contiguity between the CS and the US. What mattered most for learning was the relative probability of the US occurring with the CS versus the relative probability of the US occurring without the CS—that is, the contingency. More recent work has shown that relative timing also plays a crucial role, whereby the rate at which the US occurs during the CS versus the background rate determines the speed and strength of conditioning (Gallistel & Gibbon, 2000). Learning can even occur with no contingency or even a negative one and a well-timed stimulus (e.g., Williams et al., 2008, 2017).

Most theories of conditioning assume that animals incrementally learn and remember a stimulus–outcome association or connection between the CS and the US (e.g., Pearce & Hall, 1980; Rescorla & Wagner, 1972; Sutton & Barto, 1990; but see Gallistel et al. [2019] for an opposing information-based viewpoint). These associations effectively act as predictions of the upcoming US. The key shared idea behind these associative theories is that learning proceeds incrementally through an error-correction process. For example, according to the Rescorla–Wagner model, on each trial, animals observe all the available CSs and then generate a prediction of the upcoming US, by summing the associative strengths for the individual CSs.

When the US occurs, the strength of the US is compared to this prediction, thereby generating a (positive or negative) prediction error. Learning then proceeds by adjusting the associative strengths for CSs that were recently present based on the prediction error—making them larger (or smaller) depending on the size (and direction) of the prediction error.

This error-correction process very naturally accounts for the blocking experiment described above: When both CSA and CSB are presented, the US is fully predicted (because of the existing association between CSA and the US); thus, there is no prediction error and no further learning to CSB. Associative models have been incredibly successful in explaining a wide range of conditioning phenomena (for evaluations, see Ludvig et al., 2012; R. Miller et al., 1995), and even form the inspiration for the temporal-difference methods that drive much of modern reinforcement learning in artificial intelligence (Sutton & Barto, 2018). These associative learning models are further bolstered by the observation that dopamine neurons seem to encode exactly the type of reward prediction errors that are core to learning in these models (e.g., Glimcher, 2011; Ludvig et al., 2008; Schultz et al., 1997).

Limitations of Simple Associative Theories

Despite the wide range of successful applications, these simple incremental learning models, however, start to unravel even with very simple extensions of basic conditioning (R. Miller et al., 1995). For example, consider the phenomenon of *spontaneous recovery* (e.g., Rescorla, 2004). Here, animals are first trained that a single CS reliably predicts a US (as in Figure 12.1). Then, in a second, *extinction* phase, animals are again presented the CS, but this time without the US. With this training, animals gradually learn to stop responding, exactly as would be expected from an error-correcting, incremental learning process (see Figure 12.1). According to these models, the stimulus–outcome association that is learned during the initial acquisition phase is gradually eroded as the animal learns that the CS no longer reliably predicts the US. The trouble for the models arises if the animal is left alone for some time after that extinction phase and then presented again with the CS. In this case, the animal will once again exhibit a CR to the CS, even though it has not had any additional training in the interim. A similar recovery of responding occurs if the animal is placed in a new context or primed with the US after the extinction phases (e.g., Bouton, 1993).

Extinction clearly does not simply erase the existing association—animals have memories for the earlier CS–US association that they have acquired, which can be revealed with the right post-extinction manipulations. In basic associative learning models, there is no mechanism for this earlier memory to persist through the extinction phase because they only include a single memory per cue, which is overwritten during the extinction phase. Further mechanisms involving multiple memories, shifting temporal contexts, latent cause inference, forgetting, or relearning are necessary to explain this resurgence of earlier learned associations (e.g., Bouton & Moody, 2004; Estes, 1955; Gershman et al., 2017; Ludvig et al., 2017; but see Ghirlanda, 2005).

Other extensions of classical conditioning protocols highlight similar limitations to the simple associative models, such as the Rescorla–Wagner model (R. Miller et al., 1995). For example, in backwards blocking, animals are trained with the reverse of the blocking protocol described above (e.g., R. Miller & Matute, 1996; Urushihara & Miller, 2010). Animals are first exposed to the combined stimulus (CSA + CSB) paired with the US, until they reliably emit a response to the compound. After this initial training, they are then trained with one of the cues (e.g., CSA) paired alone with the US. As would be expected, animals continue to respond to the CSA, but notably the response to CSB degrades, even though animals get no further exposure to the second stimulus. The animals behave as though they are retrospectively changing

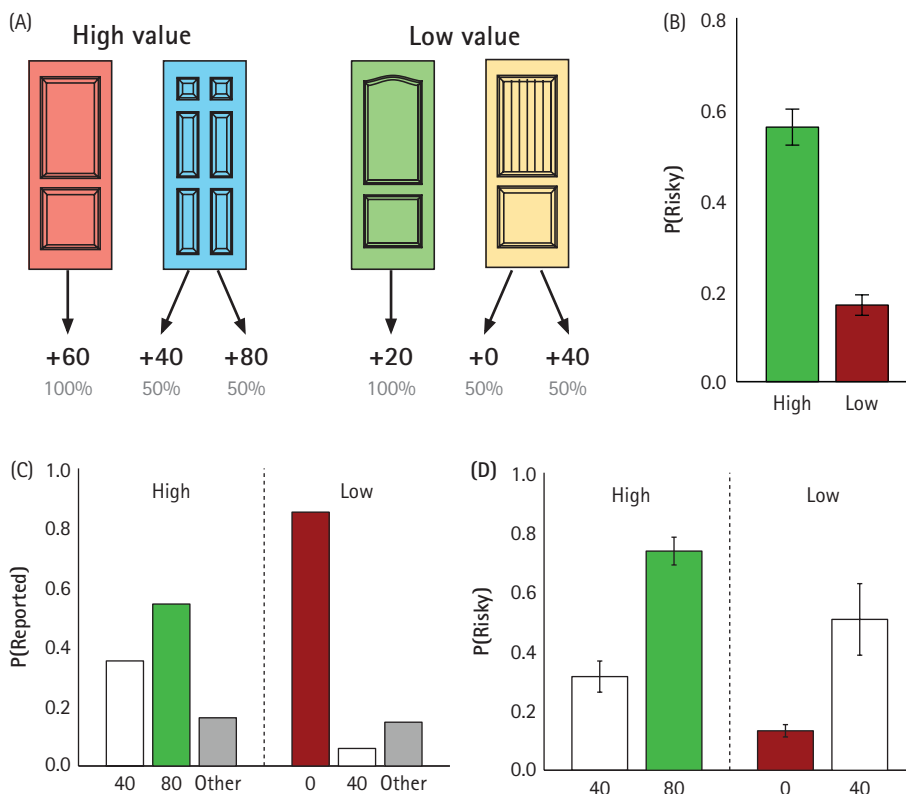


Figure 12.2 Example procedure and results from a doors study. (A) Choices and outcomes for the four possible doors. (B) Proportion of risky choices for high- and low-value outcomes. (C) Proportion of people who reported a specific outcome when asked the first outcome to come to mind. (D) Risky choice conditional on the response to the first-outcome-reported question. Across panels, people overweight the more extreme outcomes (0, +80) in choice by being more risk-seeking for high-value options (which can lead to +80) and in memory by preferentially recalling them. Based on Madan et al. (2014, Experiment 2).

the associative strength for CSB that they had learned in the first, compound-training phase. Humans also exhibit a similar form of backwards blocking in contingency judgments (Shanks, 1985). Again, the simple associative models do not have a mechanism for this change without additional exposure, which would seem to require a memory for earlier experiences (e.g., Le Pelley & McLaren, 2001).

In operant conditioning, animals also exhibit learning and behavior that clearly goes beyond remembering simple stimulus–response or response–outcome associations. In a classic experiment, different groups of rats were trained to navigate a maze over a period of 10 days (Tolman, 1948; Tolman & Honzik, 1930). For one group, the rats were rewarded every time that they completed the maze; in a second group, the rats were allowed to explore the maze but were not rewarded for completing it. During those 10 days, the rewarded rats learned to complete the maze increasingly faster, whereas the unrewarded rats (unsurprisingly) showed no evidence of speeded completion. On the 11th day, a reward was introduced for the previously unrewarded rats for completing the maze—these rats immediately started completing the maze at the same speed as the rats that had been rewarded all along (and perhaps even a touch faster).

This immediate speedup indicates that even in the absence of any reward–outcome associations, the rats had learned the structure of the maze and were able to quickly combine this knowledge/memory with the change in reward contingency. This latent learning points to an important distinction between learning and performance in conditioning (e.g., Durlach, 1989; Lattal, 1999). An animal’s behavior at a given point in time does not necessarily reflect the knowledge/learning that it has accumulated. For example, the apparent slowness in learning during acquisition (see Figure 12.1) may not be due to slow learning but, rather, may be due to the strength of the cue being presented (e.g., Spence, 1953). Similarly, spontaneous recovery is an instance in which time alone is sufficient for a latent association between cue and outcome to be revealed again (see also the discussion of sensory preconditioning, below).

In other tasks, animals also show similar goal-directed sensitivity that goes beyond simple response–outcome associations (Balleine & O’Doherty, 2010; Dolan & Dayan, 2013). Rats, for example, can be readily trained to press a lever for food reward, which they will do quite vigorously. If after this initial training period, animals are given that food in a different context and made ill after receiving it, then, when placed back into the chamber with the lever, they will no longer press the lever even though they have never experienced a negative outcome after pressing the lever (e.g., Adams & Dickinson, 1981). This goal devaluation, however, disappears when animals have been overtrained (for weeks) on the lever press response before the illness-induction manipulation. In this case, the lever press response has become habitual and is no longer sensitive to changes in motivational state or reward value.

This distinction between goal-directed and habitual behavior is fundamental to understanding operant conditioning and seems to reflect two modes of operation of the associative learning system. The distinction has been mapped to a distinction in computational reinforcement learning between model-based and model-free types of reinforcement learning (e.g., Daw et al., 2005; but see K. Miller et al., 2018, 2019). Model-free reinforcement learning corresponds to the simple associative models discussed, whereas model-based reinforcement learning is more cognitively sophisticated and involves making inferences about values/associative strengths based on an explicit model of the world with items stored in memory. A similar distinction has been proposed in human category learning, whereby both implicit, associative and verbal, rule-based processes are involved (e.g., Ashby et al., 1998; Ashby & O’Brien, 2005). These two types of reinforcement learning are also anatomically distinct in the brain, with model-free reinforcement learning dependent on the dorsolateral striatum and model-based reinforcement learning more dependent on the dorsomedial striatum (Daw et al., 2011; Yin et al., 2004, 2005). In this framework, memory acts as a world model, which guides inferences about associative strength or provides samples of remembered experiences for additional learning (Gershman & Daw, 2017; Ludvig et al., 2017; Sutton, 1990).

From Associative Learning to Human Memory

Both model-free and model-based reinforcement learning mirror a distinction in the multiple memory systems approach, which suggests that distinct memory systems exist to support different types of learning. Model-free learning is analogous to the nondeclarative, procedural learning system (Squire & Zola, 1996). This procedural knowledge is implicit and only accessible through behavior; typically, this knowledge is conceptualized as a learned behavioral response to an environmental stimulus (Eichenbaum & Cohen, 2004; see also Chapter 9,). In contrast, model-based reinforcement learning relies on the more flexible declarative memory system, including episodic memories for past events that are formed rapidly, even after a single exposure. A key attribute of episodic memory is that it binds together rich contextual details

and forms associations between features and contexts (see Chapter 6; see also Eichenbaum & Cohen, 2004; Schacter & Tulving, 1994; Wheeler et al., 1997).

As noted above, simple model-free reinforcement learning algorithms, however, are poor predictors of behavior in complex, nonstationary environments. Therefore, the notion that episodic memory may support and complement reinforcement learning in these situations has gained traction (Gershman & Daw, 2017). Episodic memories allow people to predict and plan for future events in situations in which they cannot rely on the extensive experience required by reinforcement learning (Lengyel & Dayan, 2007; Nairne, 2014). Following this line, recent research in reinforcement learning has grown to cover more complex environments and decision-making problems. In doing so, it has identified clear links between simple reinforcement, associative learning, and value-guided memory. In the remainder of this chapter, we focus on the relationship between reward–memory associations in episodic memory and decision-making.

Value-Based Learning in Human Memory

Rewards not only influence what humans and other animals learn about the world around them but also can influence what they remember. Given that memory is a limited-capacity system, the prioritization of motivationally salient events for storage in memory is an important question. As in classical conditioning, to address this question, to-be-remembered items are paired with rewards to examine how people prioritize the learning of specific pieces of information to maximize reward. In one early study, Weiner and Walker (1966) asked participants to learn pairs of letter strings of three consonants (i.e., trigrams) presented on one of four background colors. The background colors conveyed the associated incentive: winning 1¢ or 5¢ for correct recall, avoiding a shock for incorrect recall, or no incentive. The task included 10 repetitions of a list of four pairs of these trigrams, with an interleaved numerical distractor task. Although there were no reliable differences in recall after a short retention interval, after a longer interval, people recalled more trigrams in the 5¢ and shock conditions than the 1¢ and control conditions.

Tarpy and Glucksberg (1966) conducted a follow-up to this study, maintaining the 1¢ and 5¢ incentives but dropping the shock and control conditions. This follow-up study additionally changed the colored-background reward cue from occurring with the to-be-remembered trigram (“pre”) to occurring with the subsequent numerical distractor (“post”). Here, there were effects of incentive and retention interval, with better recall for the 5¢ incentive and the shorter interval, respectively, but no effect of incentive cue position (i.e., pre/post trigram) nor any interaction. Similar procedures involving different incentive levels were used in several other related studies (e.g., Eysenck & Eysenck, 1982; Harley, 1965; Loftus & Wickens, 1970; Tarpy et al., 1969; Weiner, 1966). All of these studies involved explicitly instructed reward cues that corresponded to rewards earned for successful recall or recognition, a commonality shared with many more recent reward–memory studies using either recall (e.g., Castel, 2008; Castel et al., 2002; Elliott et al., 2020) or delayed-recognition tests (e.g., Adcock et al., 2006; Gruber & Otten, 2010; Wolosin et al., 2012).

One potential critique of these studies is that they presented items in mixed lists, which alternate between different reward levels. This alternation could lead to preferential attention or rehearsal for the higher value items. In one experiment, Castel et al. (2002) attempted to address the potential for differences in attention by presenting the reward value after the item rather than concurrently—as did Tarpy and Glucksberg (1966). In both cases, the reward–memory effect was maintained, but this manipulation is not wholly sufficient to rule out the potential for differences in rehearsal. Shigemune et al. (2010) presented high- and low-value items in

different blocks and also observed an effect of reward value on memory. This approach more robustly addresses this potential rehearsal confound. Moreover, Murayama and Kuhbandner (2011) observed a reward–memory enhancement in a between-groups design, but only for trivia answers that were previously rated as uninteresting. Finally, in Shigemune et al. (2017), reward cues were only presented during the recognition test—not during encoding—and a reward-related enhancement of memory was still observed. As such, reward could only influence effort during recognition, but not any aspect of the encoding episode. These findings of reward–memory effects in blocked, between-group manipulations and cues during recognition indicate that at least in some cases, reward effects cannot be explained by preferential attention or rehearsal of items.

Several neuroimaging studies have demonstrated a link between increased activity in reward-related brain areas at encoding and subsequent memory performance. Similar to the reward–memory studies discussed previously, Adcock and colleagues (2006) gave participants small monetary incentives for the correct recall of items. They found an increased correlation in activity between the hippocampus and the ventral tegmental area for rewarded memory items that were subsequently remembered. Items associated with reward anticipation or even curiosity are also better remembered. For example, Wittmann and colleagues (2005) presented participants with images, followed by a probabilistically rewarded number-judgment task. The post-encoding task was designed to manipulate reward anticipation, but it was unrelated to the image apart from temporal proximity. Gruber and colleagues (2014) presented trivia questions and answers with an unrelated image presented in between. The trivia questions had previously been rated by the participant based on how curious the participant would be to know the answer. In subsequent recognition memory tests, rewarded images were better remembered than unrewarded images (Wittmann et al., 2005), and high-curiosity images were better remembered than low-curiosity images (Gruber et al., 2014).

These latter studies were conducted along with functional magnetic resonance imaging (fMRI) data acquisition and found activation of midbrain regions such as the substantia nigra and nucleus accumbens in relation to rewards and that activation in these regions was predictive of subsequent memory success (Gruber et al., 2014; Wittmann et al., 2005). Further studies have used similar methods to elicit reward-value effects on memory (Murayama & Kitagami, 2014; Patil et al., 2017). A particular issue is whether there is a robust one-to-one mapping between reward signaling in the brain and the strength of declarative memory formation. There is some evidence that a prediction error signal, similar to the one that drives associative learning, also guides memory formation (Bunzeck et al., 2010; Diederer et al., 2017; Ergo et al., 2020; Mason et al., 2017; Rouhani et al., 2018; Scimeca et al., 2016; Wimmer et al., 2014).

Most of the reward–memory studies described up to this point have relied on explicitly stated reward instructions. Other studies, however, have used another approach, in which reward values are learned through choices and feedback. In these studies, people need to make a choice between two words and receive a reward associated with the chosen item, with items presented in repeated blocks (e.g., Allen & Estes, 1972; Estes, 1966, 1972; Humphreys et al., 1968; Medin, 1972). Rewards are thus earned in a value-learning task. This approach is more akin to operant conditioning than the previously described procedures, in which there was no choice element. More recently, Madan et al. (2012) adapted this procedure and followed it with unrewarded, incidental lexical-decision and free-recall tasks. People exhibited reward enhancements in both tasks, with faster response times for high-value words, as well as higher free-recall rates. These two effects, however, were negatively correlated across participants and associated with different value-learning behavior; people who took more blocks to reach a

criterion exhibited a stronger free-recall bias toward rewarded items. People thus use different learning strategies in the task, suggesting that multiple memory systems could be involved in successful value learning (see also Chakravarty et al., 2019). A follow-up study by Madan and Spetch (2012) used a near-identical procedure, but with multiple reward values, and found that participants tended to remember the more salient items, both the highest and lowest reward values, following from a U-shaped function.

Although both procedures produce reward–memory effects, there may be differences in their underlying mechanisms. One such difference is the nature of the influence of reward. In the case of instructed rewards (e.g., Adcock et al., 2006; Castel et al., 2002), reward is likely treated as a separate associated property. Such an item–value association could be remembered as a source detail, akin to a location or a font color. In contrast, the more operant-like procedure discussed second (e.g., Estes, 1966; Madan et al., 2012; Medin, 1972) results in incidental encoding and reward values being learned as a motivation-related attribute of the memory itself, more similar to emotional characteristics (for a more comprehensive discussion, see Simonsen & Madan, 2021).

Episodic Memory and Decision-Making

Associative learning often involves more than learning a single stimulus–outcome association. In more complex situations, people can repeatedly encounter many different stimuli and potential actions with long and varying reward histories. To choose what response to emit in these more complex situations, people need to draw on their previously learned experiences and apply them to new contexts, which might require that associations spread across related stimuli, actions, or events. Episodic memory is thought to play a crucial role in this integrative learning and deciding process (Shohamy & Daw, 2015).

One simple example of an environment in which people need to engage in such generalization of association across different stimuli is in sensory preconditioning (Brogden, 1939; Wimmer & Shohamy, 2012). In sensory preconditioning, people are first taught an association between two neutral stimuli (A and B). They then learn that one of the stimuli, B, is associated with a reward. Despite no direct connection between A and reward, people (and other animals) still learn that A is also predictive of reward and will prefer it over another neutral stimulus. This process reflects people’s ability to integrate information across different learning episodes (Shohamy & Daw, 2015), similar to the retrospective revaluation observed in tasks such as backwards blocking (e.g., R. Miller & Matute, 1996) and the ordered associations formed in transitive inference tasks (Howard et al., 2009).

Sensory preconditioning thus influences decision-making by indirectly altering the values of different potential options based on the stimuli with which they are associated. In this vein, Wimmer and Shohamy (2012) designed a memory-based sensory preconditioning task in which items in two sets of images were first paired with each other (A and B), and then some images from one set (B) were paired with small monetary rewards. They then had people choose between the images in Set A that had never been directly paired with reward—but, crucially, had earlier been paired with images in Set B that were later rewarded or not rewarded. Their key prediction was the following: If decision-making is solely controlled by reward processing as coded in the striatum, then people should choose randomly between stimuli (in Set A) that have never been directly paired with rewards. Alternatively, if reward associations can additionally be learned via hippocampal associations, then stimuli that are indirectly linked to rewards should be chosen more often.

The behavioral data indicated much within- and between-participant variability, with participants not consistently preferring the Set A images that were associated with a rewarded

Set B image. Within participants, however, there were some associations in which participants showed a strong level of decision bias, with a preference for the neutral Set A image that was paired with a rewarded Set B image, compared to the Set A images that were associated with unrewarded Set B images. An fMRI analysis indicated that the degree of decision bias was predicted by increased hippocampal activity during reward learning and functional connectivity between the hippocampus and the striatum. This relationship appeared even in the absence of explicit recall of the A → B associations. A similar study, using magnetoencephalography to measure neural activity, also found that the reward-learning stage induced some re-instantiation of the pattern of neural activity evoked by images during the initial learning stage (Kurth-Nelson et al., 2015). These findings support the hypothesis that reward information, used to support decision-making, is integrated into the declarative memory system.

Associative Learning, Memory, and Decision-Making

As noted above, associative learning and conditioning are not restricted to situations in which humans or other animals learn about the value of a single option. In many decision-making situations, people must learn the value of the different options available to them through repeated experience, such as when they try different items on a menu or play slot machines in a casino. In these *decisions from experience* (Herwig & Erev, 2009), people learn to associate different reward values with the different options available to them and select accordingly. If certain associations are stronger and the information is more likely to be prioritized in memory, then it is also more likely to bias choices (Tversky & Kahneman, 1973).

When making these decisions from experience, people tend to exhibit specific biases in their choices. They tend to underweight rare events and overweight recent and extreme events in their choices and affective evaluations (Wulff et al., 2018). For example, when evaluating past experiences, such as a holiday or watching a film, people overweight both the most extreme and the final moments, often termed the peak-end effect (Fredrickson & Kahneman, 1993). This tendency has been linked to the accessibility of the most extreme and most recent items in memory (Aldrovandi et al., 2015; Montgomery & Unnava, 2009), perhaps through enhanced saliency or rehearsal of those items.

Historically, decision-making research has focused on choices where all necessary information is explicitly described (Kahneman & Tversky, 1979), corresponding with, for example, reading the side effects of a medication where the probabilities and outcomes are listed (Hawley et al., 2008). People, however, make different choices when these odds and outcomes are instead learned from experience (Hertwig & Erev, 2009). The way that people remember those past outcomes and sample from memory when making decisions is likely to be a critical factor in understanding this difference. In several models of memory and decision-making, the samples used to predict a choice are drawn from memory according to factors known to affect sampling from memory, such as recency (Plonsky et al., 2015), similarity to current options (Dougherty et al., 1999; Gonzalez et al., 2003), frequency (Stewart et al., 2006), or extremeness (Lieder et al., 2018).

In a series of decision-making experiments, in which people learned the odds and outcome from repeated experience as in associative learning, we found that people exhibit similar biases in both memory and choice (e.g., Ludvig et al., 2018; Madan et al., 2014, 2021). In particular, people choose as if they overweighted the most extreme (highest and lowest) outcomes that they experienced, drawing them toward gambling for higher valued options and toward risk aversion for lower valued ones. These choice biases directly correlated with similar biases in memory recall and frequency judgments after the task (Madan et al., 2014; Mason et al., 2020). Figure 12.2A depicts the basic setup in the experiment, in which people are repeatedly

asked to choose between pairs of doors. People can only learn the reward outcomes and probabilities over time with repeated experience. This method of assessing decision-making adopts methods used in operant conditioning and was inspired by attempts to study risky decision-making in pigeons (Ludvig, Madan, Pisklak, et al., 2014; Madan et al., 2019). When designing a comparative risky decision-making task, inspired by prior work (e.g., Kahneman & Tversky, 1979), we were forced to change from described odds and outcomes to those learned from experience (Ludvig & Spetch, 2011). This innovation led to a series of studies in both pigeons and humans in which we systematically examined the influence of description versus experience, outcome values, a variety of memory biases (e.g., extreme outcomes, recency, and context), time pressure, and outcome salience.

Figure 12.2A illustrates one possible set of values for the four possible choice options in the experiment (Madan et al., 2014): One leads to a fixed low-value win (20), one leads to a fixed high-value win (60), one leads to a risky low-value win (50/50 chance of 0 or 40), and one leads to a risky high-value win (50/50 chance of 40 or 80). The low-value wins can be considered relative losses (i.e., compared to the average), and the high-value wins are relative gains. In this example, the most extreme values (ends of the experienced distribution) are 0 and 80, although the findings generalize to other sets of values, including negative values (i.e., losses) (Ludvig et al., 2018; Ludvig, Madan, & Spetch, 2014), continuous values (Mason et al., 2020), and much larger values (Konstantinidis et al., 2018).

On each trial, participants are shown two doors and asked to choose between them. For example, they might be shown either a high-value or a low-value pair of doors, where the expected value of the choice is controlled. On selecting the door, the outcome is revealed. Over the course of several hundred trials, we can obtain a measure of participants' risk preference for the risky versus the safe (fixed) option. With a standard utility curve, people should be risk averse for relative gains and risk seeking for relative losses, which they typically are with fully described problems (Kahneman & Tversky, 1979). Instead, Figure 12.2B shows how people do the exact opposite: They choose as if they overweight the best and worst options (i.e., the extreme values). As a result, they are more risk seeking for the high-value options where the risky option could lead to the good extreme (+80) than for the low-value options where the risky option could lead to the bad extreme (0; Ludvig et al., 2018; Ludvig, Madan, & Spetch, 2014; Madan et al., 2014).

After completing the choice task, participants' memory for the options and outcomes is then assessed in two ways. They are asked to state the first outcome that comes to mind for each option and to judge how frequently they think each outcome occurred. Figure 12.2C shows how people are more likely to report the extreme as the first to come to mind, and this memory bias correlates with risky choice (Figure 12.2D). Participants also systematically overestimate the frequency of the extreme outcomes, and the degree to which they do this correlates with their risk preferences (Ludvig et al., 2018; Madan et al., 2014, 2017, 2021). Similar correlations emerge even when people are not tasked with choosing between options, but merely to estimate their means (Mason et al., 2020). These results suggest that the accessibility of specific outcomes in memory is crucial when people decide between options where the values were learned through trial-and-error association.

To demonstrate the relationship between memory and choice more directly, Ludvig et al. (2015) introduced priming to the above paradigm (see also Bornstein et al., 2017). By explicitly priming specific outcomes (wins and losses), it was possible to induce more risk seeking. In the choice task, each outcome was initially paired with a neutral image. Then after a period of training, the neutral images were shown prior to participants making a choice (between a risky and a safe option). Images of past wins led participants to select the risky option.

These findings provide further support for the idea that outcomes that are more accessible in memory are more likely to influence choice.

To explain results such as these, reinforcement learning models have begun to accommodate the effects of individual trials on choice. Gershman and Daw (2017) highlight the fact that some models of choice, such as those proposed by Erev and colleagues (e.g., Erev et al., 2008; Plonsky et al., 2015), can be classified as episodic reinforcement learning. In these models, participants evaluate options based on a small set of samples from each option, usually the most recent. Also in these models, individual trials, rather than running averages, are predictive of choice. Relatedly, Bornstein et al. (2017) reanalyzed data from a virtual slot machine task that had previously been examined using a standard incremental model-free reinforcement learning model (Daw et al., 2006). They fitted a sampling model that retrieved individual rewards from past options. The individual rewards were then averaged to compute a final value for each of the options, which matched what people chose. Using the data from the original experiment, however, the authors were only able to estimate and not control the sampling process. Therefore, in an extension of this work, each reward outcome was paired with a “ticket” showing a unique picture. These tickets then served as probes between trials, with the hypothesis that probing a previous trial would alter choice on the next trial. If the ticket had been originally received on a winning option, then following its appearance participants should be more likely to choose the associated option. This prediction is unique to an episodic sampling approach. Incremental reinforcement learning would be unaffected by the probes because the individual trial information is not maintained in such models and therefore the probes cannot provide reward values to update the running averages. In line with their predictions, following a memory probe, participants’ choices were influenced by the outcomes originally linked to the tickets.

Distinguishing Episodic Memory from Associative Learning

In previous sections, we primarily focused on instances in which associative learning provided a mechanism and interpretation that supported our understanding of memory and behavior. There are cases, however, in which associative learning would seem to be a suitable mechanism for a behavior but in which this is *not* supported by behavioral findings. One such case is the learning and recall of verbal paired associates (Calkins, 1986). For example, given the unrelated word pair “RUBBER–ROCKET” within a list of many pairs, associative memory could be tested with cued recall, such that the participant sees or hears the cue “RUBBER _____” and attempts to recall “ROCKET”; this pairing could be considered a stimulus–response association. With minimal additional instruction, the participant could be given the cue “_____ ROCKET” and then would be expected to correctly respond “RUBBER.” Here, the tests differ in their recall direction, with the first consisting of a forward recall and the second consisting of a backward recall. From an associative learning perspective, encoding of this backward-direction association would be separate from the forward-direction association (i.e., backward conditioning; Wagner & Terry, 1975). However, there should be some degree of correlation between the two directions of conditioning, due to common sources of variability (e.g., similarity in temporal context). Taken together, this is a question of associative symmetry.

Generally, forward recall is comparable in difficulty to backward recall (Mandler et al., 1981; Murdock, 1965, 1966), despite initial assumptions to the contrary (see Ekstrand, 1966). These similar mean recall rates for forward and backward recall, however, do not necessitate that they would also be correlated. Within this context, it has been alternatively suggested that forward and backward associations could be learned independently, known as the independent associations hypothesis (Wolford, 1971), or otherwise that associations are learned

as a single, holistic unit, known as the associative symmetry hypothesis (Asch & Ebenholz, 1962). The former hypothesis, but not the latter, would be compatible with associative learning (Kahana, 2002).

To determine whether subjects learn separate unidirectional associations or a single bidirectional association, the same association needs to be tested in both directions. As such, Kahana (2002) implemented a procedure in which association encoding was followed by two blocks of cued-recall tests, each a mixture of forward and backward direction. This “successive testing” approach included having some pairs tested in the same direction each time, either forward or backward, while others were tested in a different direction in each recall test (see also Kahana, 2000; Martin, 1971). These same-direction tests provided an upper bound for comparing the tests where different directions were tested in each block. The results demonstrated that recall was highly correlated when tested in different directions, a finding that has been replicated with object–location associations (Sommer et al., 2007) and instances in which differences in item properties influenced forward versus backward mean recall performance (Madan et al., 2010). Taken together, these studies provide evidence for the associative symmetry hypothesis but are contrary to the stimulus–response view of learning.

In other instances, the learned material exhibits consistency with the independent associations hypothesis and provides support for stimulus–response associations in verbal paired associates. For example, Caplan et al. (2014) used a similar methodology as these prior studies but examined compound words, two-syllable (“semi-compound”) words, and novel compound words. For novel compound words (e.g., BRIEF-TAX), associative symmetry held. For compound words (e.g., ROSE-BUD) and two-syllable words (e.g., SAND-WICH, HUS-BAND, STAP-LER), however, associative symmetry did not hold (see also Kahana, 2002; Yang et al., 2013). As such, recall for these word pairs was more consistent with stimulus–response associative learning. Thus, associative learning does provide a potential underlying mechanism for many, but not all, types of human memory, even in situations in which there is a clear associative structure to the task.

Conclusion and Future Outlook

We have introduced some of the key ideas, findings, and concepts from associative learning and conditioning, relating them to the relevant aspects of human memory, including reward-enhanced memory, episodic memory, and paired associate learning. In doing so, we have drawn together previously disparate areas of research that together begin to map out the relationship between associative learning and memory, in addition to highlighting the relevance of this relationship to decision-making. An emerging theme is how such memory representations guide decision-making. Simple associations between actions and outcomes, often driven by reward predictions errors, are learned over time and can ultimately guide and bias complex decisions.

Many of our everyday choices regarding healthy lifestyles, climate action, and personal finances are made based on remembered information (see, e.g., Chapter 78). To make these decisions, we are required to learn, often from repeated experience, the risk (uncertainty) and reward associated with our actions. How associative learning underpins such decision-making has become a fruitful area of research, and it has become clear that some form of episodic memory is crucial in supplementing the basic associative processes (e.g., Gershman & Daw, 2017). Future work will likely further elucidate how the properties and vagaries of human memory influence what we learn and choose in associative protocols. In addition, future work can help disentangle under what circumstances associative learning depends on episodic memory, which can vary even in seemingly similar tasks such as paired associate learning and value-based decision-making (e.g., Madan, 2013, 2020, 2024; see also Dunsmoor & Kroes, 2019; Palombo et al., 2021).

Acknowledgments

We thank Nick Simonsen for insightful discussions. Preparation of this chapter was supported by a Leverhulme Early Career Fellowship (ECF-2018-408) to Alice Mason.

References

- Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology Section B*, *33*(2b), 109–121.
- Adcock, A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. E. (2006). Reward-motivated learning: Mesolimbic activation precedes memory formation. *Neuron*, *50*, 507–517. doi:10.1016/j.neuron.2006.03.036
- Aldrovandi, S., Poirier, M., Kusev, P., & Ayton, P. (2015). Retrospective evaluations of sequences. *Experimental Psychology*, *62*, 320–334.
- Allen, G. A., & Estes, W. K. (1972). Acquisition of correct choices and value judgements in binary choice learning with differential rewards. *Psychonomic Science*, *27*, 68–72.
- Asch, S. E., & Ebenholtz, S. M. (1962). The process of free recall: Evidence for non-associative factors in acquisition and retention. *Journal of Psychology*, *54*(1), 3–31.
- Ashby, F. G., Alfonso-Reese, L. A., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*(3), 442–481.
- Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Sciences*, *9*(2), 83–89.
- Balleine, B. W., & O'Doherty, J. P. (2010). Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, *35*(1), 48–69.
- Bornstein, A. M., Khaw, M. W., Shohamy, D., & Daw, N. D. (2017). Reminders of past choices bias decisions for reward in humans. *Nature Communications*, *8*(1), 1–9.
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, *114*(1), 80–99.
- Bouton, M. E., & Moody, E. W. (2004). Memory processes in classical conditioning. *Neuroscience & Biobehavioral Reviews*, *28*(7), 663–674.
- Bower, A. B., & Landreth, S. (2001). Is beauty best? Highly versus normally attractive models in advertising. *Journal of Advertising*, *30*, 1–12.
- Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist*, *16*(11), 681–684. <https://doi.org/10.1037/h0040090>
- Brogden, W. J. (1939). Sensory pre-conditioning. *Journal of Experimental Psychology*, *25*(4), 323–332.
- Bunzeck, N., Dayan, P., Dolan, R. J., & Duzel, E. (2010). A common mechanism for adaptive scaling of reward and novelty. *Human Brain Mapping*, *31*, 1380–1394. doi:10.1002/hbm.20939
- Calkins, M. W. (1896). Association. An essay analytic and experimental. *Psychological Review: Monograph Supplements*, *1*(2), i–56. <https://doi.org/10.1037/h0092984>
- Caplan, J. B., Boulton, K. L., & Gagné, C. L. (2014). Associative asymmetry of compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 1163–1171. doi:10.1037/a0036588
- Castel, A. D. (2008). Metacognition and learning about primacy and recency effects in free recall: The utilization of intrinsic and extrinsic cues when making judgments of learning. *Memory and Cognition*, *36*(2), 429–437. doi:10.3758/MC.36.2.429
- Castel, A. D., Benjamin, A. S., Craik, F. I. M., & Watkins, M. J. (2002). The effects of aging on selectivity and control in short-term recall. *Memory & Cognition*, *30*, 1078–1085. doi:10.3758/BF03194325
- Chakravarty, S., Fujiwara, E., Madan, C. R., Tomlinson, S. E., Ober, I., & Caplan, J. B. (2019). Value bias of verbal memory. *Journal of Memory and Language*, *107*, 25–39.
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215.
- Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711.
- Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879.
- Dean, D. H., & Biswas, A. (2001). Third-party organization endorsement of products: An advertising cue affecting consumer prepurchase evaluation of goods and services. *Journal of Advertising*, *30*, 41–57.

- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining “gamification.” In A. Lugmayr, H. Franssila, C. Safran, & I. Hammouda (Eds.), *Proceedings of the 15th MindTrek conference: Envisioning future media environments* (pp. 9–15). Association for Computing Machinery.
- Diederer, K. M. J., Ziauddeen, H., Vestergaard, M. D., Spencer, T., Schultz, W., & Fletcher, P. C. (2017). Dopamine modulates adaptive prediction error coding in the human midbrain and striatum. *Journal of Neuroscience*, *37*(7), 1708–1720. doi:10.1523/JNEUROSCI.1979–16.2016
- Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, *80*(2), 312–325.
- Dougherty, M. R., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review*, *106*(1), 180–209.
- Dunsmoor, J. E., & Kroes, M. C. (2019). Episodic memory and Pavlovian conditioning: Ships passing in the night. *Current Opinion in Behavioral Sciences*, *26*, 32–39. doi:10.1016/j.cobeha.2018.09.019
- Durlach, P. J. (1989). Learning and performance in Pavlovian conditioning: Are failures of contiguity failures of learning or performance? In S. B. Klein & R. B. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of conditioning learning theory* (pp. 19–59). Erlbaum.
- Eichenbaum, H., & Cohen, N. J. (2004). *From conditioning to conscious recollection: Memory systems of the brain* (No. 35). Oxford University Press.
- Ekstrand, B. R. (1966). Backward associations. *Psychological Bulletin*, *65*, 50–64.
- Elliott, B. L., McClure, S. M., & Brewer, G. A. (2020). Individual differences in value-directed remembering. *Cognition*, *201*, 104275.
- Erev, I., Ert, E., & Yechiam, E. (2008). Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions. *Journal of Behavioral Decision Making*, *21*(5), 575–597.
- Ergo, K., De Loof, E., & Verguts, T. (2020). Reward prediction error and declarative memory. *Trends in Cognitive Sciences*, *24*(5), 388–397.
- Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, *62*(3), 145–154.
- Estes, W. K. (1966). Transfer of verbal discriminations based on differential reward magnitudes. *Journal of Experimental Psychology*, *72*, 276–283. doi:10.1037/h0023450
- Estes, W. K. (1972). Reinforcement in human behavior: Reward and punishment influence human actions via informational and cybernetic processes. *American Scientist*, *60*, 723–729.
- Eysenck, M. W., & Eysenck, M. C. (1982). Effects of incentive on cued recall. *Quarterly Journal of Experimental Psychology*, *34*, 489–498. doi:10.1080/14640748208400832
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. Appleton-Century-Crofts. <https://doi.org/10.1037/10627-000>
- Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology*, *65*(1), 45–55. <https://doi.org/10.1037/0022-3514.65.1.45>
- Gallistel, C. R., Craig, A. R., & Shahan, T. A. (2019). Contingency, contiguity, and causality in conditioning: Applying information theory and Weber’s law to the assignment of credit problem. *Psychological Review*, *126*(5), 761–773.
- Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review*, *107*(2), 289–344.
- Gershman, S. J., & Daw, N. D. (2017). Reinforcement learning and episodic memory in humans and animals: An integrative framework. *Annual Review of Psychology*, *68*, 101–128. doi:10.1146/annurev-psych-122414-033625
- Gershman, S. J., Monfils, M. H., Norman, K. A., & Niv, Y. (2017). The computational nature of memory modification. *eLife*, *6*, e23763.
- Ghirlanda, S. (2005). Retrospective revaluation as simple associative learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *31*(1), 107–111.
- Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences of the USA*, *108*, 15647–15654. doi:10.1073/pnas.1014269108
- Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, *27*(4), 591–635.
- Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron*, *84*(2), 486–496. doi:10.1016/j.neuron.2014.08.060
- Gruber, M. J., & Otten, L. J. (2010). Voluntary control over prestimulus activity related to encoding. *Journal of Neuroscience*, *30*(29), 9793–9800.
- Harley, W. (1965). The effect of monetary incentive in paired associate learning using a differential method. *Psychonomic Science*, *2*, 377–378.
- Hawley, S. T., Zikmund-Fisher, B., Ubel, P., Jancovic, A., Lucas, T., & Fagerlin, A. (2008). The impact of the format of graphical presentation on health-related knowledge and treatment choices. *Patient Education and Counseling*, *73*(3), 448–455.

- Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517–523.
- Howard, M. W., Jing, B., Rao, V. A., Probyn, J. P., & Datey, A. V. (2009). Bridging the gap: Transitive associations between items presented in similar temporal contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 391–407. <https://doi.org/10.1037/a0015002>
- Humphreys, M. S., Allen, G. A., & Estes, W. K. (1968). Learning of two-choice, differential reward problems with informational constraints on payoff combinations. *Journal of Mathematical Psychology*, *5*, 260–280. doi:10.1016/0022-2496(68)90075-8
- Janiszewski, C., & Warlop, L. (1993). The influence of classical conditioning procedures on subsequent attention to the conditioned brand. *Journal of Consumer Research*, *20*, 171–189.
- Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*(1), 103–109.
- Kahana, M. J. (2000). Contingency analyses of memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 59–72). Oxford University Press.
- Kahana, M. J. (2002). Associative symmetry and memory theory. *Memory & Cognition*, *30*(6), 823–840.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, *47*(2), 263–291.
- Kamin, L. J. (1968). “Attention-like” processes in classical conditioning. In M. R. Jones (Ed.), *Miami Symposium on the Prediction of Behavior, 1967: Aversive stimulation* (pp. 9–31). University of Miami Press.
- Kehoe, E. J., Ludvig, E. A., Dudeney, J. E., Neufeld, J., & Sutton, R. S. (2008). Magnitude and timing of nictitating membrane movements during classical conditioning of the rabbit (*Oryctolagus cuniculus*). *Behavioral Neuroscience*, *122*(2), 471–476.
- Kehoe, E. J., Ludvig, E. A., & Sutton, R. S. (2014). Time course of the rabbit’s conditioned nictitating membrane movements during acquisition, extinction, and reacquisition. *Learning & Memory*, *21*(11), 585–590.
- Konstantinidis, E., Taylor, R. T., & Newell, B. R. (2018). Magnitude and incentives: Revisiting the overweighting of extreme events in risky decisions from experience. *Psychonomic Bulletin & Review*, *25*(5), 1925–1933.
- Kurth-Nelson, Z., Barnes, G., Sejdinovic, D., Dolan, R., & Dayan, P. (2015). Temporal structure in associative retrieval. *eLife*, *4*, e04919.
- Lattal, K. M. (1999). Trial and intertrial durations in Pavlovian conditioning: Issues of learning and performance. *Journal of Experimental Psychology: Animal Behavior Processes*, *25*(4), 433–450.
- Lengyel, M., & Dayan, P. (2007). Hippocampal contributions to control: The third way. *Advances in Neural Information Processing Systems*, *20*, 889–896.
- Le Pelley, M. E., & McLaren, I. P. L. (2001). Retrospective revaluation in humans: Learning or memory? *Quarterly Journal of Experimental Psychology Section B*, *54*(4b), 311–352.
- Lieder, F., Griffiths, T. L., & Hsu, M. (2018). Overrepresentation of extreme events in decision making reflects rational use of cognitive resources. *Psychological Review*, *125*(1), 1–32.
- Loftus, G. R., & Wickens, T. D. (1970). Effect of incentive on storage and retrieval processes. *Journal of Experimental Psychology: General*, *85*, 141–147. doi:10.1037/h0029537
- Ludvig, E. A., Madan, C. R., McMillan, N., & Spetch, M. L. (2018). Living near the edge: How extreme outcomes and their neighbors drive risky choice. *Journal of Experimental Psychology: General*, *147*(12), 1905–1918.
- Ludvig, E. A., Madan, C. R., Pisklak, J. M., & Spetch, M. L. (2014). Reward context determines risky choice in pigeons and humans. *Biology Letters*, *10*(8), 20140451.
- Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2014). Extreme outcomes sway risky decisions from experience. *Journal of Behavioral Decision Making*, *27*(2), 146–156.
- Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2015). Priming memories of past wins induces risk seeking. *Journal of Experimental Psychology: General*, *144*(1), 24–29.
- Ludvig, E. A., Mirian, M. S., Kehoe, E. J., & Sutton, R. S. (2017). Associative learning from replayed experience. *bioRxiv*, 100800.
- Ludvig, E. A., & Spetch, M. L. (2011). Of black swans and tossed coins: Is the description–experience gap in risky choice limited to rare events? *PLoS One*, *6*(6), e20262.
- Ludvig, E. A., Sutton, R. S., & Kehoe, E. J. (2008). Stimulus representation and the timing of reward-prediction errors in models of the dopamine system. *Neural Computation*, *20*(12), 3034–3054.
- Ludvig, E. A., Sutton, R. S., & Kehoe, E. J. (2012). Evaluating the TD model of classical conditioning. *Learning & Behavior*, *40*(3), 305–319.
- Madan, C. R. (2013). Toward a common theory for learning from reward, affect, and motivation: The SIMON framework. *Frontiers in Systems Neuroscience*, *7*, 59. doi:10.3389/fnsys.2013.00059
- Madan, C. R. (2020). Rethinking the definition of episodic memory. *Canadian Journal of Experimental Psychology*, *74*, 183–192. doi:10.1037/cep0000229

- Madan, C. R. (2024). *Memories that matter: How we remember important things*. Routledge.
- Madan, C. R., Fujiwara, E., Gerson, B. C., & Caplan, J. B. (2012). High reward makes items easier to remember, but harder to bind to a new temporal context. *Frontiers in Integrative Neuroscience*, 6, 61.
- Madan, C. R., Glaholt, M. G., & Caplan, J. B. (2010). The influence of item properties on association-memory. *Journal of Memory and Language*, 63(1), 46–63.
- Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2014). Remembering the best and worst of times: Memories for extreme outcomes bias risky decisions. *Psychonomic Bulletin & Review*, 21, 629–636. doi:10.3758/s13423-013-0542-9
- Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2017). The role of memory in distinguishing risky decisions from experience and description. *Quarterly Journal of Experimental Psychology*, 70, 2048–2059. doi:10.1080/17470218.2016.1220608
- Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2019). Comparative inspiration: From puzzles with pigeons to novel discoveries with humans in risky choice. *Behavioural Processes*, 160, 10–19. doi:10.1016/j.beproc.2018.12.009
- Madan, C. R., & Spetch, M. L. (2012). Is the enhancement of memory due to reward driven by value or salience? *Acta Psychologica*, 139(2), 343–349.
- Madan, C. R., Spetch, M. L., Machado, F. D. S., Mason, A., & Ludvig, E. A. (2021). Encoding context determines risky choice. *Psychological Science*, 32, 743–754. doi:10.1177/0956797620977516
- Mandler, G., Rabinowitz, J. C., & Simon, R. A. (1981). Coordinate organization: The holistic representation of word pairs. *American Journal of Psychology*, 92, 209–222.
- Martin, E. (1971). Verbal learning theory and independent retrieval phenomena. *Psychological Review*, 78, 314–332.
- Mason, A., Farrell, S., Howard-Jones, P., & Ludwig, C. J. (2017). The role of reward and reward uncertainty in episodic memory. *Journal of Memory and Language*, 96, 62–77. doi:10.1016/j.jml.2017.05.003
- Mason, A., Madan, C. R., Simonsen, N., Ludvig, E. A., & Spetch, M. L. (2020, December 13). Biased confabulation in risky choice. *PsyArxiv*. <https://psyarxiv.com/vphgc>
- Medin, D. L. (1972). Partial information and choice behaviour in differential reward magnitude learning. *Psychonomic Science*, 27, 73–76.
- Miller, K. J., Ludvig, E. A., Pezzulo, G., & Shenhav, A. (2018). Re-aligning models of habitual and goal-directed decision-making. In R. Morris, A. Bornstein, & A. Shenhav (Eds.), *Understanding goal-directed decision making: Computations and circuits* (pp. 407–428). Academic Press.
- Miller, K. J., Shenhav, A., & Ludvig, E. A. (2019). Habits without values. *Psychological Review*, 126(2), 292–311.
- Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla–Wagner model. *Psychological Bulletin*, 117(3), 363–386.
- Miller, R. R., & Matute, H. (1996). Biological significance in forward and backward blocking: Resolution of a discrepancy between animal conditioning and human causal judgment. *Journal of Experimental Psychology: General*, 125(4), 370–386.
- Montgomery, N. V., & Unnava, H. R. (2009). Temporal sequence effects: A memory framework. *Journal of Consumer Research*, 36(1), 83–92.
- Murayama, K., & Kitagami, S. (2014). Consolidation power of extrinsic rewards: Reward cues enhance long-term memory for irrelevant past events. *Journal of Experimental Psychology: General*, 143, 15–20. doi:10.1037/a0031992
- Murayama, K., & Kuhbandner, C. (2011). Money enhances memory consolidation—but only for boring material. *Cognition*, 119, 120–124. doi:10.1016/j.cognition.2011.01.001
- Murdock, B. B. (1965). Associative symmetry and dichotic presentation. *Journal of Verbal Learning & Verbal Behavior*, 4, 222–226.
- Murdock, B. B. (1966). Forward and backward associations in paired associates. *Journal of Experimental Psychology*, 71, 732–737.
- Nairne, J. S. (2014). Adaptive memory: Controversies and future directions. In B. L. Schwartz, M. L. Howe, M. P. Toglia, & H. Ortega (Eds.), *What is adaptive about adaptive memory?* (pp. 308–321). Oxford University Press.
- Palombo, D. J., Elizur, E., Tuen, Y. J., Te, A. A., & Madan, C. R. (2021). Transfer of negative valence in an episodic memory task. *Cognition*, 217, 104874. doi:10.1016/j.cognition.2021.104874
- Patil, A., Murty, V. P., Dunsmoor, J. E., Phelps, E. A., & Davachi, L. (2017). Reward retroactively enhances memory consolidation for related items. *Learning and Memory*, 24(1), 65–69. doi:10.1101/lm.042978.116
- Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex* (G. V. Anrep, Trans.). Oxford University Press.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87(6), 532–552.
- Peter, J. P., & Nord, W. R. (1982). A clarification and extension of operant conditioning principles in marketing. *Journal of Marketing*, 48, 102–107.

- Plonsky, O., Teodorescu, K., & Erev, I. (2015). Reliance on small samples, the wavy recency effect, and similarity-based learning. *Psychological Review*, *122*(4), 621–647.
- Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review*, *74*(1), 71–80.
- Rescorla, R. A. (2004). Spontaneous recovery. *Learning & Memory*, *11*(5), 501–509.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.
- Rothschild, M. L., & Gaidis, W. C. (1981). Behavioral learning theory: Its relevance to marketing and promotions. *Journal of Marketing*, *45*, 70–78.
- Rouhani, N., Norman, K. A., & Niv, Y. (2018). Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(9), 1430–1443. doi:10.1037/xlm0000518
- Schacter, D., & Tulving, E. (1994). What are the memory systems of 1994? In D. Schacter & E. Tulving (Eds.), *In memory systems* (pp. 1–38). MIT Press.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599.
- Scimeca, J. M., Katzman, P. L., & Badre, D. (2016). Striatal prediction errors support dynamic control of declarative memory decisions. *Nature Communications*, *7*, 13061.
- Shanks, D. R. (1985). Forward and backward blocking in human contingency judgement. *Quarterly Journal of Experimental Psychology Section B*, *37*(1b), 1–21.
- Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology*, *60*, 291–309.
- Shigemune, Y., Abe, N., Suzuki, M., Ueno, A., Mori, E., Tashiro, M., Itoh, M., & Fujii, T. (2010). Effects of emotion and reward motivation on neural correlates of episodic memory encoding: A PET study. *Neuroscience Research*, *67*, 72–79. <https://doi.org/10.1016/j.neures.2010.01.003>
- Shigemune, Y., Tsukiura, T., Nouchi, R., Kambara, T., & Kawashima, R. (2017). Neural mechanisms underlying the reward-related enhancement of motivation when remembering episodic memories with high difficulty. *Human Brain Mapping*, *38*, 3428–3443. <https://doi.org/10.1002/hbm.23599>
- Shohamy, D., & Daw, N. D. (2015). Integrating memories to guide decisions. *Current Opinion in Behavioral Sciences*, *5*, 85–90. doi:10.1016/j.cobeha.2015.08.010
- Simonsen, N., & Madan, C. R. (2021, October 14). Fractionating the reward–memory literature: Instructed, item-related feedback, and item-unrelated feedback. *PsyArXiv*. doi:10.31234/osf.io/bqz5
- Sommer, T., Rose, M., & Büchel, C. (2007). Associative symmetry versus independent associations in the memory for object–location associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(1), 90–106. <https://doi.org/10.1037/0278-7393.33.1.90>
- Spence, K. W. (1953). Learning and performance in eyelid conditioning as a function of intensity of the UCS. *Journal of Experimental Psychology*, *45*(1), 57–63.
- Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences of the USA*, *93*(24), 13515–13522.
- Staddon, J. E., & Simmelhag, V. L. (1971). The “superstition” experiment: A reexamination of its implications for the principles of adaptive behavior. *Psychological Review*, *78*(1), 3–43. <https://doi.org/10.1037/h0030305>
- Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1–26.
- Stuart, E. W., Shimp, T. A., & Engle, R. W. (1987). Classical conditioning of consumer attitudes: Four experiments in an advertising context. *Journal of Consumer Research*, *14*, 334–349.
- Sutton, R. S. (1990). Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. *International Conference on Machine Learning (ICML)*, *7*, 216–224.
- Sutton, R. S., & Barto, A. G. (1990). Time-derivative models of Pavlovian reinforcement. In M. Gabriel & J. W. Moore (Eds.), *Learning and computational neuroscience* (pp. 497–537). MIT Press.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.
- Tarpy, R. M., & Glucksberg, S. (1966). Effects of incentive and incentive-cue position on short-term retention. *Psychonomic Science*, *5*(8), 313–314.
- Tarpy, R. M., Glucksberg, S., & Lytle, L. D. (1969). Motivation and short-term retention: Evidence for covert rehearsal. *American Journal of Psychology*, *82*(1), 111–116.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208. <https://doi.org/10.1037/h0061626>
- Tolman, E. C., & Honzik, C. H. (1930). Introduction and removal of reward, and maze performance in rats. *University of California Publications in Psychology*, *4*, 257–275.

- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232. [http://doi.org/10.1016/0010-0285\(73\)90033-9](http://doi.org/10.1016/0010-0285(73)90033-9)
- Urushihara, K., & Miller, R. R. (2010). Backward blocking in first-order conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*(2), 281–295. <https://doi.org/10.1037/a0016773>
- Wagner, A. R., & Terry, W. S. (1975). Backward conditioning to a CS following an expected vs. a surprising UCS. *Animal Learning & Behavior*, *3*(4), 370–374. <https://doi.org/10.3758/bf03213461>
- Walvis, T. H. (2008). Three laws of branding: Neuroscientific foundations of effective brand building. *Journal of Brand Management*, *16*(3), 176–194.
- Weiner, B. (1966). Motivation and memory. *Psychological Monographs: General and Applied*, *80*(18), 1–22. doi:10.1037/h0093896
- Weiner, B., & Walker, E. (1966). Motivational factors in short-term retention. *Journal of Experimental Psychology: General*, *71*(2), 190–193.
- Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *121*(3), 331–354. doi:10.1037/0033-2909.121.3.331
- Williams, D. A., Lawson, C. L., Cook, R., Mather, A. A., & Johns, K. W. (2008). Timed excitatory conditioning under zero and negative contingencies. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*, 94–105. doi:10.1037/0097-7403.34.1.94
- Williams, D. A., Todd, T. P., Chubala, C. M., & Ludvig, E. A. (2017). Intertrial unconditioned stimuli differentially impact trace conditioning. *Learning & Behavior*, *45*(1), 49–61.
- Wimmer, G. E., Braun, E. K., Daw, N. D., & Shohamy, D. (2014). Episodic memory encoding interferes with reward learning and decreases striatal prediction errors. *Journal of Neuroscience*, *34*(45), 14901–14912.
- Wimmer, G. E., & Shohamy, D. (2012). Preference by association: How memory mechanisms in the hippocampus bias decisions. *Science*, *338*(6104), 270–273.
- Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H. J., & Düzel, E. (2005). Reward-related fMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron*, *45*, 459–467. doi:10.1016/j.neuron.2005.01.010
- Wolford, G. (1971). Function of distinct associations for paired-associate performance. *Psychological Review*, *78*(4), 303–313. <https://doi.org/10.1037/h0031032>
- Wolosin, S. M., Zeithamova, D., & Preston, A. R. (2012). Reward modulation of hippocampal subfield activation during successful associative encoding and retrieval. *Journal of Cognitive Neuroscience*, *24*, 1532–47. doi:10.1162/jocn_a_00237
- Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description–experience gap. *Psychological Bulletin*, *144*(2), 140–176. <https://doi.org/10.1037/bul0000115>
- Yang, J., Zhao, P., Zhu, Z., Mecklinger, A., Fang, Z., & Li, H. (2013). Memory asymmetry of forward and backward associations in recognition tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(1), 253–269. <https://doi.org/10.1037/a0028875>
- Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Journal of Neuroscience*, *19*(1), 181–189.
- Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2005). Blockade of NMDA receptors in the dorsomedial striatum prevents action–outcome learning in instrumental conditioning. *European Journal of Neuroscience*, *22*(2), 505–512.



OXFORD LIBRARY OF PSYCHOLOGY

The Oxford Handbook of Human Memory

FOUNDATIONS AND APPLICATIONS

Volume 1

Edited by

Michael J. Kahana and
Anthony D. Wagner

OXFORD
UNIVERSITY PRESS