Context-dependent learning and causal structure

Samuel J. Gershman

Department of Psychology and Center for Brain Science, Harvard University

KEYWORDS: Bayesian modeling; animal and human associative learning

Address for correspondence:

Samuel Gershman

Department of Psychology

Harvard University

52 Oxford St., room 295.05

Cambridge, MA 02138

Phone: 773-607-9817

E-mail: gershman@fas.harvard.edu

Abstract

The notion of "context" has played an important but complicated role in animal learning theory. Some studies have found that contextual stimuli (e.g., conditioning chamber) act much like punctate stimuli, entering into competition with other cues as would be predicted by standard associative learning theories. Other studies have found that contextual stimuli act more like "occasion setters," modulating the associative strength of punctate stimuli without themselves acquiring associative strength. Yet other studies have found that context is often largely ignored, resulting in transfer of performance across context changes. This article argues that these diverse functions of context arise in part from different causal interpretations of the environment. A Bayesian theory is presented that infers which causal interpretation best explains an animal's training history, and hence which function of context is appropriate. The theory coherently accounts for a number of disparate experimental results, and quantitatively predicts the results of a new experiment designed to directly test the theory.

Introduction

All learning about events in the world occurs in the presence of background stimuli, typically diffuse in space and time, that constitute the "context" of learning. The nature of contextual influences has been studied in animal learning for decades (Balsam & Tomie, 1985; Rosas, Todd, & Bouton, 2013; Urcelay & Miller, 2014), and has played an important role in the neuroscience of learning (Maren, Phan, & Liberzon, 2013). Yet little theoretical consensus has emerged. Indeed, the theoretical literature on context-dependent learning is sparse and incomplete, in part because the empirical data present a formidable challenge. In particular, three functions of context have been proposed, and there is extensive evidence for all three.

First, context sometimes seems to be irrelevant, such that behavior is invariant to changes of context (Bouton & King, 1983; Bouton & Peck, 1989; Kaye, Preston, Szabo, Druiff, & Mackintosh, 1987; Lovibond, Preston, & Mackintosh, 1984). Second, context sometimes acts in a modulatory manner, selecting particular associations between cues and outcomes without itself acquiring associative strength (Bouton & Swartzentruber, 1986; Grahame, Hallam, Geier, & Miller, 1990). This modulatory function of context resembles the action of "occasion setters"—punctate cues that control the associative properties of other cues (Bouton, 1993; Swartzentruber, 1995). Third, context sometimes acts like a punctate cue, leading to competition between cues and contexts (Balaz, Capra, Hartl, & Miller, 1981; Grau & Rescorla, 1984).

The key theoretical question is whether (and how) these diverse functions can be described within a single framework. This paper advances a positive answer: different functions of context correspond to different latent causal structures (cf. Gershman, Norman, & Niv, 2015), and Bayesian inference over these structures determines which function is appropriate given an animal's training history. This theoretical framework takes a step towards resolving the

discrepant experimental data, by providing insight into the factors that govern how context influences learning.

The rest of the paper is organized as follows. First, we present an informal description of the computational model (a technical description is contained in the appendices). We then apply the model to select phenomena from the Pavlovian conditioning literature that illustrate the diversity of context functions. Finally, we present new experimental data from a predictive learning task with humans designed to directly test the model's predictions. In the Discussion, we consider the strengths and limitations of the model, and connect it to a broader family of structure learning models.

Causal structure learning

An animal in a Pavlovian conditioning experiment is exposed to a training history consisting punctate stimuli (cues), contextual stimuli, and outcomes (rewards or punishments). We consider a hypothesis space of causal structures relating these variables, as shown in Figure 1 (technical details can be found in Appendix A). We do not consider this hypothesis space to be exhaustive, but for simplicity we only consider 3 causal structures that capture the main ideas about context as summarized in the Introduction. Each structure is a variation of the linear-Gaussian parametrization that has previously been applied to a number of animal learning phenomena (Dayan & Kakade, 2001; Gershman, 2015; Kruschke, 2008). This parametrization is essentially a probabilistic extension of the Rescorla-Wagner model; it assumes that outcomes are noisy linear functions of the stimuli.

Structure M_1 (irrelevant context) assumes that cues cause outcomes, while context plays no causal role. Structure M_2 (modulatory context) assumes that there is a separate set of cue-outcome contingencies for each context. Structure M_3 (additive context) assumes

that context acts like another cue, combining additively to determine the expected outcome. Thus, M_3 posits that standard cue competition phenomena (e.g., blocking, overshadowing, overexpectation) should apply to context.

Given this hypothesis space, we model the animal as computing the posterior distribution over latent structures given the training history. The posterior distribution encodes the animal's beliefs about what structure generated its training history. The posterior can be updated iteratively when new data are observed, as stipulated by Bayes' rule (detailed in Appendix B). The animal is modeled as emitting a conditioned response (CR) reflecting its outcome expectation. Intuitively, the outcome expectation is computed by taking an average of the expectations for each structure, weighted by the posterior probability of that structure. The details of computing the outcome expectation for each structure are described in Appendix B.

The structure learning model can be viewed as a formalization of the idea that ambiguity is the crucial variable determining the function of context (Bouton, 1993; Rosas et al., 2013). When cues have an unambiguous relationship with the outcome, context exerts a weak influence over learning (i.e., the irrelevant context structure is relatively favored). Context only begins to exert a stronger influence over learning when the relationship is ambiguous but can be disambiguated by taking context into account, such that the modulatory or additive context structures become relatively favored. We explore this idea further in the next section.

Accounting for the diverse functions of context

In this section, we describe several simulations of the model illustrating how it can account for the diverse functions of context observed experimentally.
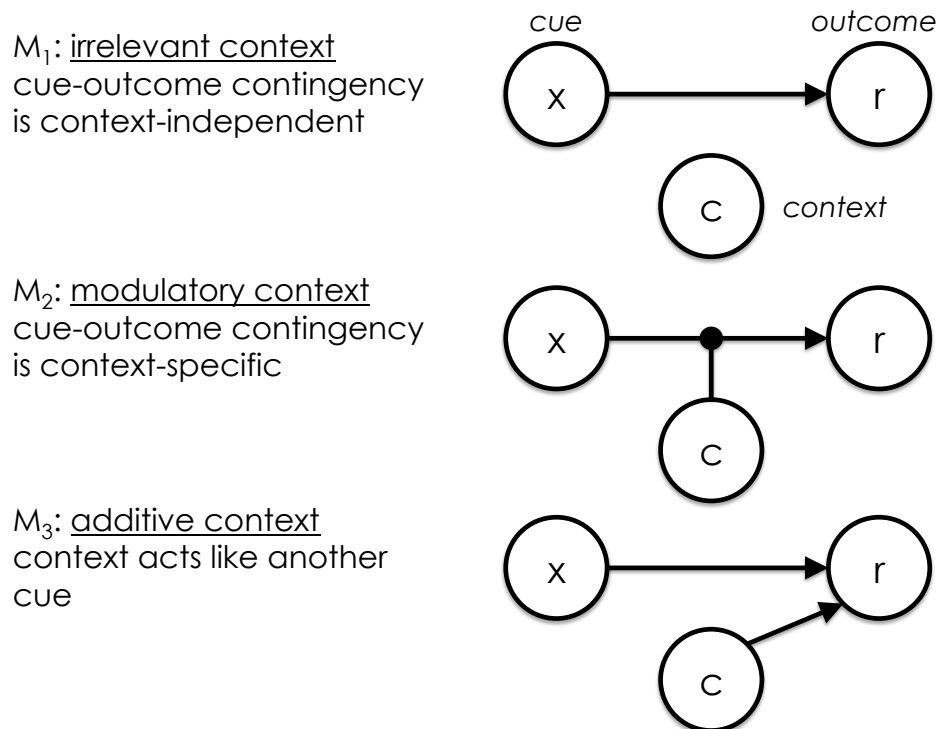
M₁: <u>irrelevant context</u>
cue-outcome contingency
is context-independent

M₂: <u>modulatory context</u>
cue-outcome contingency
is context-specific

M₃: <u>additive context</u>
context acts like another
cue

Figure 1: **The hypothesis space of causal structures**. Each node corresponds to a variable, and arrows denote causal relations. In the case of M_2, the context does not exert a direct causal influence on the outcome, but rather "gates" the causal influence of the cues.

Renewal

We first consider the phenomenon of context-dependent renewal (Bouton & Bolles, 1979; Bouton & King, 1983; Bouton & Swartzentruber, 1986). When an animal is conditioned to expect an outcome following a cue, extinguishing the cue by presenting it in the absence of the outcome results in cessation of conditioned responding. If the training and extinction phases are carried out in two different contexts, then conditioned responding can be renewed by presenting the cue in the training context. A simple explanation of this phenomenon in terms of classical associative learning mechanisms is that the extinction context acquired

inhibitory associative strength—i.e., it became a conditioned inhibitor. This suggests that learning a new cue-outcome association in the extinction context should be retarded relative to a novel context. By the same token, the context should summate with a separately trained cue, thereby depressing the conditioned response. In fact, both of these predictions are false (Bouton & Swartzentruber, 1986): the extinction context does not appear to acquire inhibitory associative strength (but see Polack, Laborda, & Miller, 2012).

These observations led to the idea that contexts act like "occasion setters" (Bouton, 1993; Swartzentruber, 1995), capable of modulating the associative properties of cues without themselves acquiring associative strength. However, as Bouton and others have noted, this occasion-setting function only seems to emerge when cue-outcome relationships are ambiguous (e.g., after extinction); in the absence of ambiguity, cue-specific associations appear to transfer across contexts. For example, conditioned responding is relatively insensitive to context change after initial conditioning (Bouton & King, 1983; Bouton & Peck, 1989; Kaye et al., 1987; Lovibond et al., 1984).

The structure learning model described in the previous section can account for these findings (Figure 2). After conditioning, the posterior probability over causal structures is split primarily between M_1 (irrelevant context) and M_2 (modulatory context), whereas M_3 (additive context) is relatively disfavored. Having only observed a single context, the model does not know yet whether to posit one set of associations for all contexts (M_1) or a different set of associations for each context (M_2). The additive context structure M_3 is disfavored because of a "Bayesian Occam's Razor" effect (MacKay, 2003): it posits 3 weight parameters (one cue-outcome association and two context-outcome associations), whereas the other structures can explain the data equally well with a single parameter (M_1) or two parameters (M_2).[1] The posterior support for M_1 means that conditioned responding will partially

---

[1]There is no Bayesian Occam's Razor penalizing M_2 relative to M_1 because only a single context has been observed by the end of conditioning, and M_2 only explicitly represents the weight parameter for the
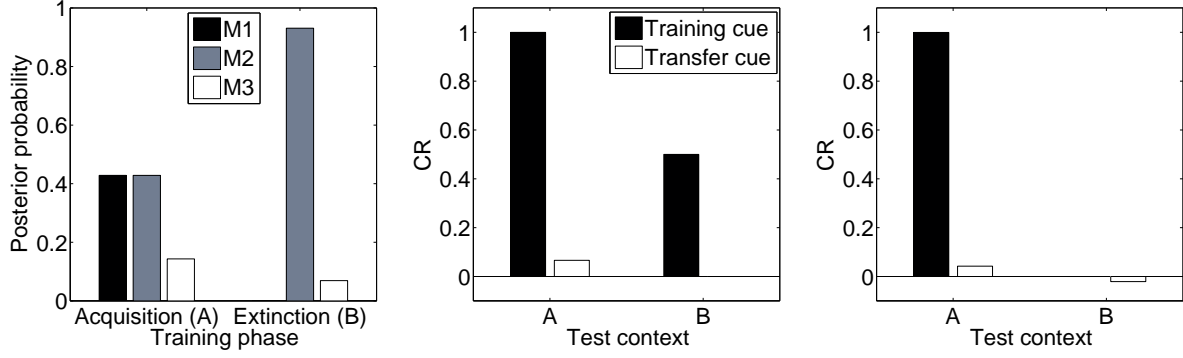
Figure 2: **Renewal after extinction**. (*Left*) Posterior probability over causal structures at the end of the acquisition and extinction phases. By the end of extinction, the modulatory context structure is strongly favored. (*Middle*) Conditioned response (CR) to the training cue or a novel transfer cue when tested after the acquisition phase, which occurs either in the acquisition context (A) or in a novel context (B). The response to the training cue partly generalizes across contexts, but no appreciable response is elicited by the transfer cue, indicating the absence of contextual associations. (*Right*) Conditioned response to the training cue or a novel transfer cue when tested after the extinction phase in context B. Responding to the training cue is renewed upon return to A.

transfer across contexts, as shown in the middle panel of Figure 2.

The story is different after extinction: the modulatory context structure M_2 now has strong support by offering the most parsimonious explanation of the data. Consequently, the cue-outcome association is rendered sensitive to context change (Figure 2, right panel). Notice also that in these simulations a novel cue is never appreciably potentiated or inhibited by context, indicating the lack of support for M_3 (which predicts that context by itself should be capable of supporting conditioned responding).
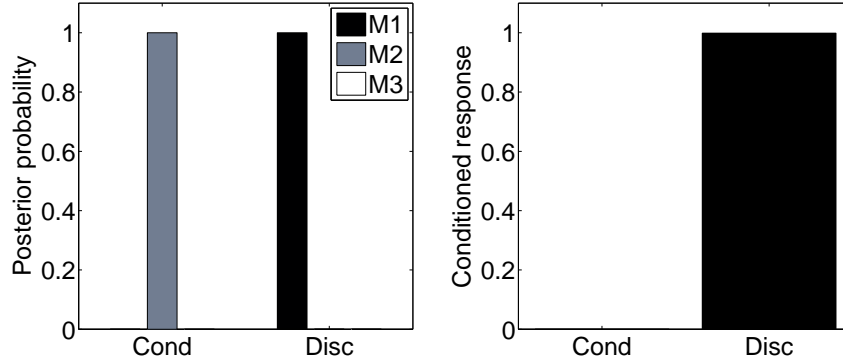
Figure 3: **Making contexts informative**. Following the experimental protocol of Preston et al. (1986), two conditions were simulated. In the discrimination (Disc) condition, one cue was always reinforced while another cue was never reinforced, and these contingencies applied across two different contexts. In the conditional discrimination (Cond) condition, the contingencies reversed across contexts. Animals were then given discrimination training with two novel CSs in a single context, and finally tested with the same CSs in a novel context. (*Left*) At the end of the first phase, the irrelevant context structure (M_1) is favored in the discrimination condition, whereas the modulatory context structure (M_2) is favored in the conditional discrimination condition. (*Right*). Conditional discrimination training leads to context specificity in the final test phase; discrimination training leads to generalization across contexts.

The information value of contexts

An experiment reported by Preston et al. (1986) provides a clear demonstration of how the information value of context controls the context-dependence of learning. In this experiment, animals were initially trained to discriminate between two cues with different outcome probabilities. In the discrimination condition, the outcomes depended only the cues and not on the context. In the conditional discrimination condition, the outcome contingencies reversed across contexts, thus making the contexts informative. As shown in Figure 3 (left panel), the discrimination training strongly favors the irrelevant context structure M_1, whereas the conditional discrimination training strongly favors the modulatory context structure M_2.

The key question is what happens when the animal is subsequently given discrimination training with two novel CSs in a single context and then presented with the same CSs in a novel context. Preston et al. (1986) found that only conditional discrimination led to context-specificity, such that animals in that condition did not generalize responding to the novel CSs from the training context to the novel test context. This finding is reproduced by the structure learning model (Figure 3, right panel) due to the strong support for M_2 following conditional discrimination training.

Effects of outcome intensity and amount of training

Odling-Smee (1978) documented two other factors that govern context-dependence: outcome intensity and amount of training. In particular, Odling-Smee (1978) examined the latency to enter a conditioning chamber following fear conditioning, reasoning that animals would show a longer latency (i.e., stronger CR) if the context had acquired associative strength during training. The CR was found to be stronger for high compared to low intensity shock,

---

second context once it's been observed. By contrast, M_3 explicitly represents the weights for each possible context even before observing those contexts.
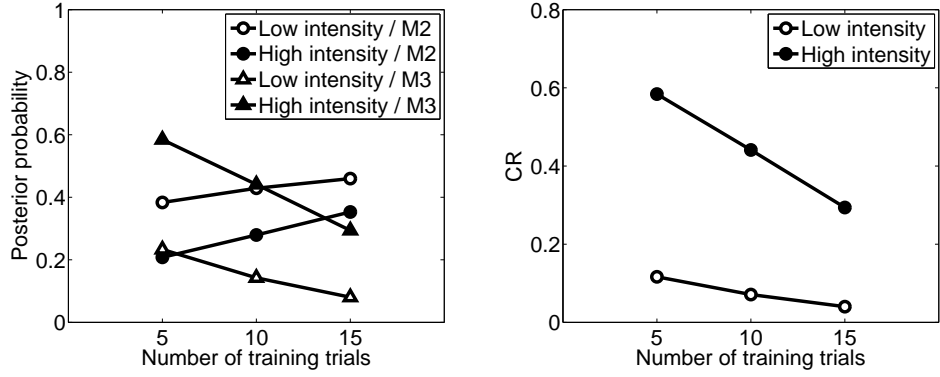
Figure 4: **Manipulations of outcome intensity and number of trials**. Simulations of the experiments reported by Odling-Smee (1978), in which responding to the training context alone was measured following cue-outcome pairings. (*Left*) Posterior probability of M_3 increases with the outcome intensity and decreases with the number of training trials. The probability of M_1 is not shown here because in this case M_1 and M_2 have identical probabilities. (*Right*) Responding to the context increases with the outcome intensity and decreases with the number of training trials.

and lower when the number of training trials increased. Figure 4 shows that simulations of this experiment capture the two basic patterns.

Because there is a prior on the associative weights favoring weights close to 0, the modulatory context structure M_2 cannot easily account for high intensity outcomes, which require a large associative weight for the cue. In contrast, the additive context structure M_3 can explain the same data using a combination of cue and context associations, thereby allowing each to have a weight closer to 0. This leads to a preference for M_3. As the number of trials increases, the preference for M_3 diminishes because sufficient evidence accumulates to support a large weight for the cue in M_2.

An experimental test of the model

The simulations presented in the previous section illustrate how different factors influence the posterior distribution over causal structures. It is however rare to see the operation of all three structures in a single experiment. We now present new data from a human predictive learning task that allows us to uncover all three structures, providing a rigorous experimental test of the structure learning model.

Following other studies of predictive learning in humans (e.g., Abad, Ramos-Álvarez, & Rosas, 2009; Rosas & Callejas-Aguilera, 2006), participants in the experiment were asked to predict whether a particular food (the cue) eaten in a particular restaurant (the context) would produce sickness (the outcome). Each cue was shown in each context an equal number of times; the only difference across conditions was the reinforcement contingencies (Table 1), which were designed to promote particular structural interpretations (Figure 5, top panel). In the test phase, participants were presented with four trials that covered the 2×2 space of old/new cue in an old/new context.

The model makes distinct predictions for test phase behavior across the conditions (Figure 5, bottom panel). When the irrelevant context structure is favored, the old cue should predict the outcome will occur in both old and new contexts, but the new cue should not predict the outcome in either context. When the modulatory context structure is favored, the old cue should predict the outcome will occur only in the old context, while the new cue again should not predict the outcome in either context. Finally, when the additive context structure is favored, both the old and new cues should predict the outcome most strongly in the old context.
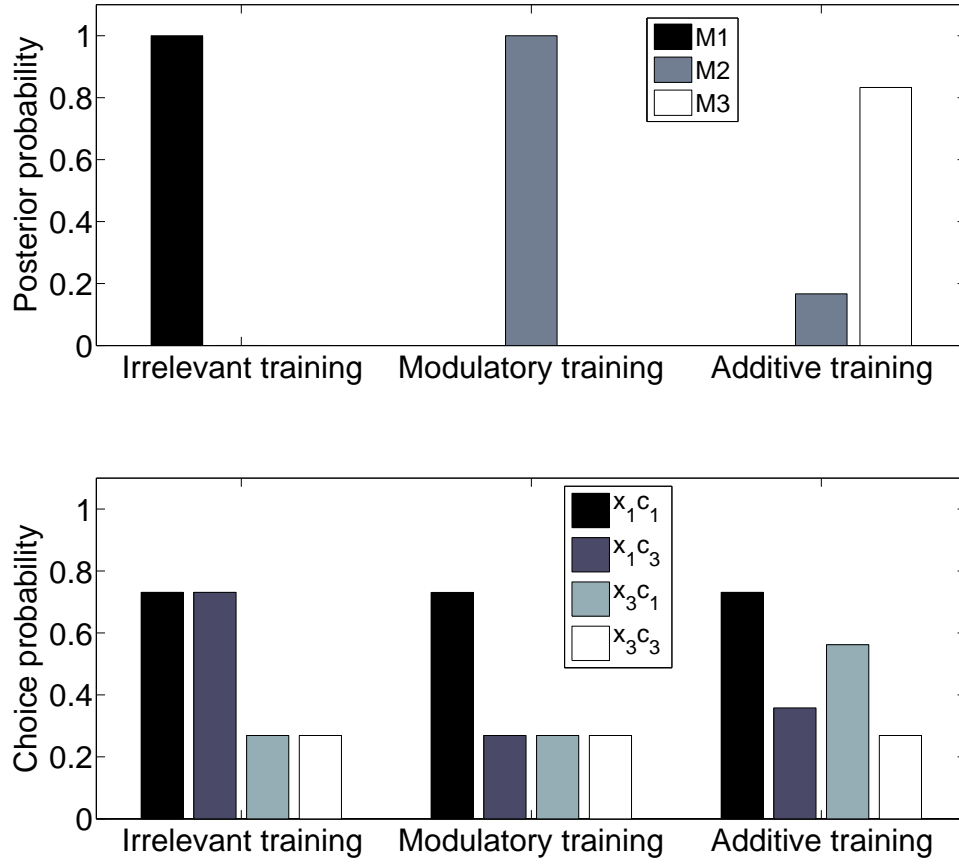
Figure 5: **Experimental predictions**. (*Top*) Posterior probability over causal structures at the end of training for each condition. (*Bottom*) Predicted choice probabilities in the test phase.

	Training phase	Test phase
Irrelevant context group	x_1c_1+ x_2c_1- x_1c_2+ x_2c_2-	x_1c_1 x_1c_3 x_3c_1 x_3c_3
Modulatory context group	x_1c_1+ x_2c_1- x_1c_2- x_2c_2+	x_1c_1 x_1c_3 x_3c_1 x_3c_3
Additive context group	x_1c_1+ x_2c_1+ x_1c_2- x_2c_2-	x_1c_1 x_1c_3 x_3c_1 x_3c_3

Table 1: **Experimental design**. Cues are denoted by (x_1, x_2, x_3) and contexts are denoted by c_1, c_2, c_3). Outcome presentation is denoted by "+" and no outcome is denoted by "−".

Participants

Ninety participants were recruited via the Amazon Mechanical Turk web service ($N = 31$ in the irrelevant training condition, $N = 29$ in the modulatory training condition, and $N = 30$ in the additive training condition). All participants received informed consent and were paid for their participation. The experiment was approved by the Harvard University Internal Review Board.

Materials and procedure

We used a set of 3 Mexican foods (chili, burrito, nachos) and 3 restaurant names (Molina's Cantina, Restaurante Arroyo, El Coyote Cafe). The assignments of foods to cues and restaurant names to contexts were randomized. We label the foods presented to participants as (x_1, x_2, x_3) and the contexts as (c_1, c_2, c_3). On each trial in the training phase, participants were presented with a picture of one food along with the name of a restaurant, and were asked to make a binary prediction about whether an individual would get sick after con-
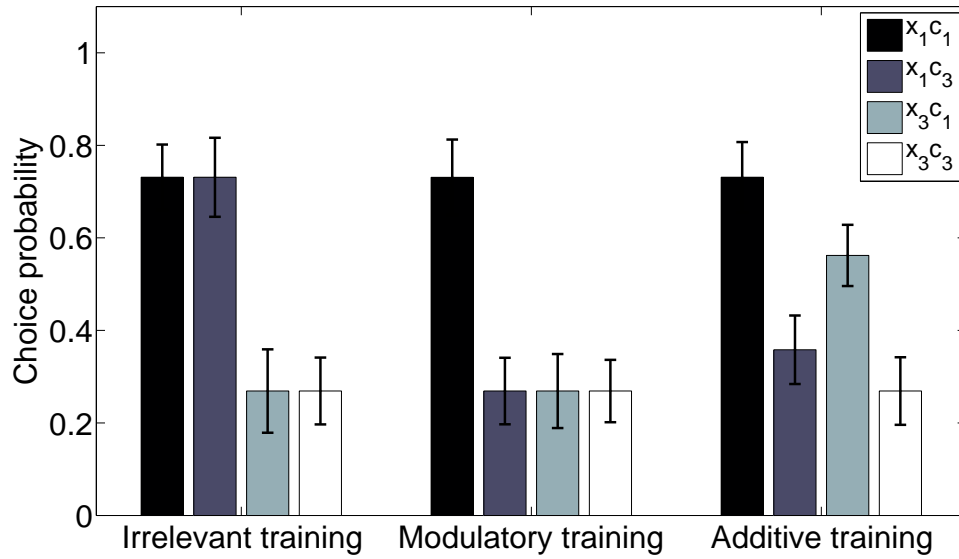
Figure 6: **Experimental results**. Empirical choice probabilities in the test phase. Error bars represent standard error of the mean.

suming the food in that restaurant. Once they entered their response, participants received feedback. The training phase was followed by a test phase, consisting of 4 trials that were procedurally identical to the training phase trials but presented without feedback.

Participants were divided into 3 groups that differed only in terms of the reinforcement contingencies during the training phase, as summarized in Table 1. All participants were presented with the same 4 test trials: an old cue in an old context (x_1c_1), an old cue in a new context (x_1c_3), a new cue in an old context (x_3c_1), or a novel cue in a new context (x_3c_3).

Results

The experimental results are shown in Figure 6. Overall, the model is in excellent quantitative agreement with the choice data $(r = 0.92, p < 0.0001)$, using the same parameters

as in the simulations (i.e., the parameters were not fit to the data). For comparison, the correlation coefficient is smaller if we constrain the model to use a single structure for all conditions: $r = 0.5$ for M_1, $r = 0.7$ for M_2, and $r = 0.88$ for M_3.

We confirmed the qualitative predictions of the model using planned comparisons. In the irrelevant context condition, the old cue was judged to be more predictive of the outcome in both the old and new contexts compared to the new cue $[t(29) = 2.57, p < 0.05]$. In the modulatory context condition, the old cue was judged to be more predictive of the outcome in the old context compared to the other three test trials $[t(27) = 2.73, p < 0.05]$. Finally, in the additive context condition, both the old and new cues were judged as more predictive of the outcome in the old context compared to the new context $[t(29) = 4.04, p < 0.001]$. Importantly, no-single structure model captures this complete pattern of results.

Discussion

This paper has shown how a new structure learning model can account for the diverse functions of context in Pavlovian conditioning and human predictive learning. First, the model explains why different training protocols lead to different forms of context-dependence. Second, the model makes quantitatively accurate predictions about human predictive learning in a task designed to directly test the theory's predictions. No single-structure model could explain the complete pattern of results.

Bayesian structure learning has been proposed as an explanation for many behavioral and neural phenomena that are puzzling from the perspective of classical associative learning theory (Collins & Frank, 2013; Courville, Daw, & Touretzky, 2006; Gershman, Blei, & Niv, 2010; Gershman & Niv, 2012; Gershman et al., 2015; Lloyd & Leslie, 2013; Soto, Gershman, & Niv, 2014). Of particular relevance is the model developed by Gershman et al. (2010),

which explained context-dependent renewal as the consequence of inferring a structure in which the acquisition and extinction trials were generated by distinct latent causes. While this model successfully accounted for the renewal effect, it treated context as a cue; as we have seen, this assumption is invalidated by the numerous observations that context does not appear to acquire associative strength in the same way that cues do. The explanation offered in this paper allows the context to play a modulatory role and hence better explain renewal.

There is, however, a way in which the current model fails at explaining the complete set of renewal phenomena. When an animal is trained in context A, extinguished in context B, and then tested in context C, a renewal effect is also observed (known as "ABC renewal"), though it is typically somewhat weaker than ABA renewal. This is problematic for the model, because the modulatory structure is strongly favored by the end of extinction, and hence the cue should have no associative strength in context C. There are at least two possibilities for dealing with this lacuna. First, we could hypothesize that the context does not have a direct causal link with the outcome, but instead provides information about a latent cause. This would allow some probability that contexts A and C were generated by the same latent cause, and is essentially the explanation proposed by Gershman et al. (2010). Second, we could hypothesize a hierarchical Bayesian model in which the associative strength of the cue is initialized to its "prior" value in context C, and this prior is itself learned from the training history (see Gershman & Niv, 2015). Intuitively, the prior reflects the average expected outcome across all prior phases. Further modeling will be required to decide between these different possibilities.

Conclusion

Context is an important but puzzling concept in animal learning theory. This paper demonstrates that it is possible to provide a systematic analysis of context's multiple functions within a single computational framework. The key idea is that Bayesian inference over causal structures enables an animal to select the appropriate context given its training history. While this idea is unlikely to explain all the existing data on context-dependent learning, it is sufficiently powerful to make new quantitative predictions that we have confirmed experimentally.

Appendix A: causal structures

The animal's training history is represented as $\mathbf{h}_{1:n} = (\mathbf{x}_{1:n}, \mathbf{r}_{1:n}, \mathbf{c}_{1:n})$ for trials 1 to n, consisting of the following variables:

- $\mathbf{x}_n \in \mathbb{R}^D$: the set of D "punctate" stimuli (cues) observed at time n. Typically we will use $x_{nd} = 1$ to indicate that cue d is present and $x_{nd} = 0$ to indicate that it is absent.

- $c_n \in \{1, \ldots, K\}$: the context, which can take on one of K discrete values.[2]

- $r_n \in \mathbb{R}$: the outcome (e.g., reward or punishment).

To keep notation simple, we have omitted any dependence on action, but the treatment below is easily generalized to the instrumental setting where the outcome depends on both the stimuli and the chosen action.

---

[2]More generally, the context could be multidimensional, but for simplicity we do not consider that possibility here.

We consider three specific structures relating the above variables. All the structures have in common that the outcome is drawn from a Gaussian with variance $\sigma_r^2 = 0.01$:

$$r_n \sim \mathcal{N}\left(\bar{r}_n, \sigma_r^2\right), \tag{1}$$

where we have left the dependence on $\mathbf{c}_n$ and $\mathbf{x}_n$ implicit. The structures differ in how the mean $\bar{r}_n$ is computed.

- **Irrelevant context** (M_1):

$$\bar{r}_n = \sum_{d=1}^{D} w_d x_{nd} = \mathbf{w}^\top \mathbf{x}_n. \tag{2}$$

- **Modulatory context** (M_2):

$$\bar{r}_n = \sum_{d=1}^{D} w_{dk} x_{nd} = \mathbf{w}_k^\top \mathbf{x}_n \tag{3}$$

  when $c_n = k$.

- **Additive context** (M_3):

$$\bar{r}_n = \sum_{d=1}^{D} w_d x_{nd} + w_{D+k} = \mathbf{w}^\top \tilde{\mathbf{x}}_n, \tag{4}$$

  again for $c_n = k$. The augmented stimulus $\tilde{\mathbf{x}}_n$ is defined as: $\tilde{\mathbf{x}}_n = [\mathbf{x}_n, \tilde{\mathbf{c}}_n]$, where $\tilde{c}_{nk} = 1$ if $c_n = k$, and 0 otherwise.

We assume each weight is drawn independently from a zero-mean Gaussian prior with variance $\sigma_w^2 = 1$. Each weight can change slowly over time according to a Gaussian random walk with variance $\tau^2 = 0.001$. Finally, we assume that all structures are equally probable

a priori.

Appendix B: probabilistic inference

We can compute the posterior over the weights for a given model M using Bayes' rule:

$$P(\mathbf{w}|\mathbf{h}_{1:n}, M) = \frac{P(\mathbf{h}_{1:n}|\mathbf{w}, M)P(\mathbf{w}|M)}{P(\mathbf{h}_{1:n}|M)}. \tag{5}$$

For M_1, the posterior at time n is $P(\mathbf{w}|\mathbf{h}_{1:n}, M = M_1) = \mathcal{N}(\mathbf{w}; \hat{\mathbf{w}}_n, \Sigma_n)$ with parameters updated recursively as follows:

$$\hat{\mathbf{w}}_{n+1} = \hat{\mathbf{w}}_n + \mathbf{g}_n(r_n - \hat{\mathbf{w}}_n^\top \mathbf{x}_n) \tag{6}$$

$$\Sigma_{n+1} = \Sigma_n' - \mathbf{g}_n \mathbf{x}^\top \Sigma_n', \tag{7}$$

where $\Sigma_n' = \Sigma_n + \tau^2 \mathbf{I}$. These update equations are known as *Kalman filtering*, an important algorithm in engineering and signal processing that has recently been applied to animal learning (Dayan & Kakade, 2001; Gershman, 2015; Kruschke, 2008). The initial estimates are given by the parameters of the prior: $\hat{\mathbf{w}}_0 = 0, \Sigma_0 = \sigma_w^2 \mathbf{I}$. The Kalman gain $\mathbf{g}_n$ (a vector of learning rates) is given by:

$$\mathbf{g}_n = \frac{\Sigma_n' \mathbf{x}_n}{\mathbf{x}_n^\top \Sigma_n' \mathbf{x}_n + \sigma_r^2}. \tag{8}$$

The same equations apply to M_2, but the mean and covariance are context specific: $\hat{\mathbf{w}}_{n+1}^k$ and Σ_n^k. Accordingly, the Kalman gain is modified as follows:

$$\mathbf{g}_{nk} = \frac{\Sigma_{nk}' \mathbf{x}_n}{\mathbf{x}_n^\top \Sigma_{nk}' \mathbf{x}_n + \sigma_r^2} \tag{9}$$

if $c_n = k$, and a vector of zeros otherwise. For M_3, the same equations as M_1 apply, but to the augmented stimulus $\tilde{\mathbf{x}}_n$.

To make predictions about future outcomes, we need to compute the posterior predictive expectation, which is also available in closed form:

$$V_n = \mathbb{E}[r_n|\mathbf{x}_n, c_n, \mathbf{h}_{1:n-1}] = \sum_M \mathbb{E}[r_n|\mathbf{x}_n, c_n, \mathbf{h}_{1:n-1}, M]P(M|\mathbf{h}_{1:n-1}). \tag{10}$$

The first term in Eq. 10 is the posterior predictive expectation conditional on model M:

$$\mathbb{E}[r_n|\mathbf{x}_n, c_n, \mathbf{h}_{1:n-1}, M] = \mathbf{x}_n^\top \hat{\mathbf{w}}_n, \tag{11}$$

where again the variables are modified depending on what model is being considered. The second term in Eq. 10 is the posterior probability of model M, which can be updated according to Bayes' rule:

$$P(M|\mathbf{h}_{1:n}) \propto P(r_n|\mathbf{x}_n, c_n, \mathbf{h}_{1:n-1}, M)P(M|\mathbf{h}_{1:n-1}), \tag{12}$$

where the likelihood is given by:

$$P(r_n|\mathbf{x}_n, \mathbf{c}_n, \mathbf{h}_{1:n-1}, M) = \mathcal{N}(r_n; \mathbf{x}_n^\top \hat{\mathbf{w}}_n, \mathbf{x}_n^\top \Sigma_n' \mathbf{x}_n + \sigma_r^2). \tag{13}$$

To make predictions for the predictive learning experiment, we mapped the posterior predictive expectation onto choice probability (outcome vs. no outcome) by a logistic sigmoid transformation:

$$P(c_n = 1) = \frac{1}{1 + \exp(-2V_n + 1)}, \tag{14}$$

where $c_n = 1$ indicates a prediction that the outcome will occur, and $c_n = 0$ indicates a prediction that the outcome will not occur.

Acknowledgments

References

Abad, M. J., Ramos-Álvarez, M. M., & Rosas, J. M. (2009). Partial reinforcement and context switch effects in human predictive learning. *The Quarterly Journal of Experimental Psychology*, *62*, 174–188.

Balaz, M. A., Capra, S., Hartl, P., & Miller, R. R. (1981). Contextual potentiation of acquired behavior after devaluing direct context-US associations. *Learning and Motivation*, *12*, 383–397.

Balsam, P., & Tomie, A. (1985). *Context and Learning*. Psychology Press.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, *114*, 80–99.

Bouton, M. E., & Bolles, R. (1979). Contextual control of the extinction of conditioned fear. *Learning and Motivation*, *10*, 445–466.

Bouton, M. E., & King, D. A. (1983). Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*, 248–265.

Bouton, M. E., & Peck, C. A. (1989). Context effects on conditioning, extinction, and reinstatement in an appetitive conditioning preparation. *Animal Learning & Behavior*, *17*, 188–198.

Bouton, M. E., & Swartzentruber, D. (1986). Analysis of the associative and occasion-setting properties of contexts participating in a Pavlovian discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, *12*, 333–350.

Collins, A. G., & Frank, M. J. (2013). Cognitive control over learning: creating, clustering, and generalizing task-set structure. *Psychological Review*, *120*, 190–229.

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, *10*, 294–300.

Dayan, P., & Kakade, S. (2001). Explaining away in weight space. In T. Leen, T. Dietterich, & V. Tresp (Eds.), *Advances in neural information processing systems 13* (pp. 451–457). MIT Press.

Gershman, S. J. (2015). A unifying probabilistic view of associative learning. *PLoS Computational Biology*, *11*, e1004567.

Gershman, S. J., Blei, D. M., & Niv, Y. (2010). Context, learning, and extinction. *Psychological Review*, *117*, 197–209.

Gershman, S. J., & Niv, Y. (2012). Exploring a latent cause theory of classical conditioning. *Learning & Behavior*, *40*, 255–268.

Gershman, S. J., & Niv, Y. (2015). Novelty and inductive generalization in human reinforcement learning. *Topics in Cognitive Science*, *7*, 391–415.

Gershman, S. J., Norman, K. A., & Niv, Y. (2015). Discovering latent causes in reinforcement learning. *Current Opinion in Behavioral Sciences*, *5*, 43–50.

Grahame, N. J., Hallam, S. C., Geier, L., & Miller, R. R. (1990). Context as an occasion setter following either CS acquisition and extinction or CS acquisition alone. *Learning and Motivation*, *21*, 237–265.

Grau, J. W., & Rescorla, R. A. (1984). Role of context in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 324–332.

Kaye, H., Preston, G., Szabo, L., Druiff, H., & Mackintosh, N. (1987). Context specificity of conditioning and latent inhibition: Evidence for a dissociation of latent inhibition and associative interference. *The Quarterly Journal of Experimental Psychology*, *39*, 127–145.

Kruschke, J. K. (2008). Bayesian approaches to associative learning: From passive to active learning. *Learning & Behavior*, *36*, 210–226.

Lloyd, K., & Leslie, D. S. (2013). Context-dependent decision-making: a simple Bayesian model. *Journal of The Royal Society Interface*, *10*, 20130069.

Lovibond, P. F., Preston, G., & Mackintosh, N. (1984). Context specificity of conditioning, extinction, and latent inhibition. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 360–375.

MacKay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge university press.

Maren, S., Phan, K. L., & Liberzon, I. (2013). The contextual brain: implications for fear conditioning, extinction and psychopathology. *Nature Reviews Neuroscience*, *14*, 417–428.

Odling-Smee, F. (1978). The overshadowing of background stimuli: some effects of varying amounts of training and UCS intensity. *The Quarterly Journal of Experimental Psychology*, *30*, 737–746.

Polack, C. W., Laborda, M. A., & Miller, R. R. (2012). Extinction context as a conditioned inhibitor. *Learning & Behavior*, *40*, 24–33.

Preston, G., Dickinson, A., & Mackintosh, N. (1986). Contextual conditional discriminations. *The Quarterly Journal of Experimental Psychology*, *38*, 217–237.

Rosas, J. M., & Callejas-Aguilera, J. E. (2006). Context switch effects on acquisition and

extinction in human predictive learning. *Journal of Experimental Psychology: learning, Memory, and cognition*, *32*, 461–474.

Rosas, J. M., Todd, T. P., & Bouton, M. E. (2013). Context change and associative learning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*, 237–244.

Soto, F. A., Gershman, S. J., & Niv, Y. (2014). Explaining compound generalization in associative and causal learning through rational principles of dimensional generalization. *Psychological Review*, *121*, 526–558.

Swartzentruber, D. (1995). Modulatory mechanisms in pavlovian conditioning. *Animal Learning & Behavior*, *23*, 123–143.

Urcelay, G. P., & Miller, R. R. (2014). The functions of contexts in associative learning. *Behavioural Processes*, *104*, 2–12.