Blanco, F., & Moris, J. (2018). Bayesian methods for addressing long-standing problems in associative learning: The case of PREE. *Quarterly Journal of Experimental Psychology*, *71*(9), 1844-1859.

Bayesian methods for addressing long-standing problems in associative learning: The case of PREE

Fernando Blanco[1] and Joaquín Moris[2]

1-Departamento de Fundamentos y Métodos de la Psicología, Universidad de Deusto, Spain

2-Departamento de Psicología, Universidad de Oviedo, Spain

Correspondence concerning this article should be addressed to Fernando Blanco, Departamento de Fundamentos y Métodos de la Psicología. Universidad de Deusto, 48007, Bilbao, Spain. E-mail: fernandoblanco@deusto.es. Phone: +34 944139000 (Ext. 3238).

Acknowledgement

Joaquín Morís was funded with a Juan de la Cierva Postdoctoral grant. This research was supported by grants PSI2013-43516-R and PSI2016-78818-R of the Spanish Ministry of Economy and Competitiveness, and grant IT955-16 of the Basque Government. We thank Dr. Pedro Cobos for his help in collecting data for the experiment.

Abstract

Most associative models typically assume that learning can be understood as a gradual change in associative strength that captures the situation into one single parameter, or representational state. We will call this view single-state learning. However, there is ample evidence showing that under many circumstances different relationships that share features can be learned independently, and animals can quickly switch between expressing one or another. We will call this multiple-state learning. Theoretically, it is understudied because it needs a different data analysis approach from those usually employed. In this paper, we present a Bayesian model of the Partial Reinforcement Extinction Effect (PREE) that can test the predictions of the multiple-state view. This implies estimating the moment of change in the responses (from the acquisition to the extinction performance), both at the individual and at the group levels. We used this model to analyze data from a PREE experiment with three levels of reinforcement during acquisition (100%, 75% and 50%). We found differences in the estimated moment of switch between states during extinction, so that it was delayed after leaner partial reinforcement schedules. The finding is compatible with the multiple-state view. It is the first time, to our knowledge, that the predictions from the multiple-state view are tested directly. The paper also aims to show the benefits that Bayesian methods can bring to the associative learning field.

Keywords: *Bayesian models*, *Associative learning*, *PREE*.

Bayesian models are becoming increasingly popular in many psychological research areas, both to formalize theories (see examples in Courville, Daw, & Touretzky, 2006; Jacobs & Kruschke, 2011; Kruschke, 2008; O'Reilly, Jbabdi, & Behrens, 2012) and to provide approaches to analyzing data (see the recent reading list of introductory Bayes elaborated by Etz, Gronau, Dablander, Edelsbrunner, & Baribault, 2016). They provide great flexibility and freedom to researchers, because they can be adapted to any experimental situation and to extract information about any parameter of interest, be it at the individual subject level or at the group level. Thus, Bayesian models are "limited only by the user's imagination" (Wagenmakers, Lodewyckx, Kuriyal, & Grasman, 2010, p. 162). Because of this, they can be used to test hypotheses that are hard to study with more traditional models. In this paper, we use Bayesian methods to analyze in a novel way a long-established associative learning finding, the Partial Reinforcement Extinction Effect (PREE).

Throughout its history, the associative learning field (as well as related areas such as contingency or causal learning) has greatly benefited from the use of different formal models and their mathematical implementations. Indeed, these models have been an excellent research heuristic, not only producing potential explanations for well-known learning phenomena, but also leading to the discovery of new ones. For instance, perhaps the most celebrated among associative learning models, the Rescorla-Wagner model, RWM (Rescorla & Wagner, 1972), made predictions that were novel and even counterintuitive at its time and were eventually found to be true (Miller, Barnet, & Grahame, 1995; Siegel & Allan, 1996). In these cases, the assumptions embedded in the model (e.g., that the CSs sum their associative strengths linearly, that inhibitory associations possess negative associative strength, etc.) worked well to advance predictions that were then confirmed empirically.

The RWM is the most successful instance of the large family of associative models, which has dominated the literature of associative learning for decades. These models differ in a number of assumptions and goals, but they share several features that make them distinct and, in fact, might explain their success (Pearce & Bouton, 2001). One core assumption underlying almost all associative models is their conceptualization of learning as a process of gradual change in the strength of the association between mental representations. Learning is, thus, encoded by only one parameter, the associative strength, that changes over training. When the situation is stable (i.e., the contingencies do not change), the parameter adapts gradually to capture it. Likewise, when the situation changes (i.e., when there is a change in the contingencies, such as in extinction), the parameter then adapts accordingly, in a similar way. In sum, a single representation of the learning scenario (encoded by the associative strength value) is updated regularly depending on the situation, either stable or variable. Because of this, we will call this type of models single-state models (admittedly, a few associative models have proposed that multiple associations can be updated simultaneously, but still gradually, as we comment in the Discussion section; see Esber & Haselgrove, 2011).

This central and widespread assumption of modern associative learning leads to a clear prediction: the gradual, negatively accelerated learning curve that now appears in most textbooks (Gallistel, Fairhurst, & Balsam, 2004). This curve is most evident when learning curves of multiple subjects are averaged, because such practice compensates the random noise contributed by each individual participant (Glautier, 2013).

However, although this single-state assumption has been dominant for decades, several authors have pointed out its problems. First, certain phenomena, particularly those related to contextual changes like renewal (Bouton, 1994), suggest that animals are able to simultaneously store multiple representations of the learning setup. Once the animals have captured all the states that characterize the learning situation, they can use (or express) either one state or another depending on the task demands. For instance, a context switch would make subjects replace their "extinction state" by their "acquisition state" (Rosas & Callejas-Aguilera, 2006). This state selection is typically sudden, as it implies no new learning, and contrasts with the traditional single-state assumption, according to which the current representation can only be updated on a trial by trial basis. More importantly, these phenomena imply that, when facing a changing contingency (e.g., extinction), animals can store an intact representation of the previous situation (e.g., acquisition phase) while they are learning the new one, so that they can recover these previous states in the future, as it happens in renewal effects. In sum, they are not bounded to use and update a single parameter, that is, to transform a unique representation of the learning situation, but instead they can learn multiple states (e.g., one for the acquisition, one for the extinction...), perhaps independently, and flexibly express one or another.

Second, those few times when researchers moved their focus from the group-level analysis to the individual-level analysis, they found surprisingly little evidence of the gradual learning curve expected under the single-state assumption of the associative theory. The most common pattern of results resembles more a step-like function, with participants changing suddenly their behavior from the baseline to the final performance level, as if the learning process was all-or-nothing in nature (Gallistel, 2012a; Gallistel et al., 2004; Glautier, 2013). That is, individual-level performance suggests that learning might be better represented by a process in which multiple states are involved, with shifts from one to another. Given that individuals differ in the specific moment when the change in behavior takes place, averaging across subjects produces a seemingly gradual transition. This has led researchers to openly claim that the gradual learning curve is "an artifact of group averaging" (Gallistel et al., 2004). Moreover, if one is interested in examining the learning performance of a given individual participant (as is the case of psychologists in the old-school behavioral tradition), then the group-level averaged curves would be useless, since it is impossible to induct any property (shape, functional form) of the individual curve from the group curve (Estes, 1956; Gallistel, 2012b).

Thus, we have two different types of hypothesis about how behavior changes (single-state vs. multiple-state learning) that make different assumptions and predictions about the shape of the learning curve and the underlying processes that produce it. This debate is particularly important in those situations in which a manipulation is expected to affect the learning rate, or the pace at which learning proceeds, because this parameter is interpreted in completely different ways under the two hypotheses. A very clear example of this is the Partial Reinforcement Extinction Effect (PREE hereafter; Jenkins & Rigby, 1950). For simplicity, imagine a human learning experiment in which two stimuli, a CS and a US (conditioned and unconditioned stimulus, respectively), may occur sequentially. On each trial, the participant's goal is to provide a judgment of the likelihood of the US occurrence. Although the CS is presented on every trial, this design includes groups that differ in the probability with which the US follows the CS. In the continuous reinforcement group, the US is presented in all trials (i.e., 100% reinforcement schedule). However, in the partial reinforcement groups, the US appears only on some of the trials (e.g., 50%, or 75%). After an acquisition phase during which CS and US have been repeatedly paired in the way we described, an extinction phase begins in which only the CS is presented, without being followed by the US in any of the groups. This eventually reduces the subjective expectancy of the US occurrence (as reflected in the judgments), effectively extinguishing the conditioned response. The PREE is defined as a greater resistance to extinction observed when a partial reinforcement schedule was delivered during the acquisition phase (Jenkins & Rigby, 1950), compared to a continuous reinforcement schedule in which all acquisition trials were reinforced. This has been acknowledged as a "paradoxical" effect, because it is those conditions with a lower proportion of reinforced trials that become the more resistant to subsequent extinction.

From the viewpoint of the associative models ascribed to the single-state assumption, according to which learning is a gradual process of modification of the strength of associations, it makes sense to interpret this result as a "slowing down" of learning during the extinction phase. In fact, models such as Mackintosh's (Mackintosh, 1975) propose that experience with imperfect predictors of a US reduces these predictors' learning rate parameters (sometimes called salience or associability), leading to actually slower learning curves. However, other accounts of the PREE are based on the greater difficulty to discriminate between the acquisition and extinction phases when the reinforcement schedule is partial (Baum, 2012; Courville et al., 2006; Mowrer & Jones, 1945): the lower the probability with which the CS was reinforced during the acquisition, the more trials it will take to detect that this probability decreased in the extinction phase. That is, that the conditioned stimulus is no longer being followed by the unconditioned stimulus. Instead of assuming a single state or representation that adapts gradually to the new situation, the multiple-state view suggests that multiple states are represented, with the eventual expression of one of them. This also implies quickly switching between the expression of different states. Notably, this interpretation invites us to focus not on the learning rate during extinction, but on the delay before extinction is observed (Gallistel, 2012a). That is, the multiple-state view implies taking into consideration a different parameter: we would be more interested in detecting the "moment of change", or the "inflection point" in which behavior changes, than in assessing the extinction rate per se, as associative models propose.

According to this view, a change in the level of analysis is also required: we focus on individual subjects instead of groups. This is because the change in behavior is sudden, but the moment varies between subjects, and the averaging across participants or sessions would only blur the current parameter of interest (the moment of change between states), creating an illusory gradual learning curve. As many have pointed out, although the main assumptions of the associative theory (including the gradual strengthening of associations) are formulated at the level of individual subjects, almost all published experiments present aggregated-level data (i.e., averaged data for many participants or many sessions). When the PREE is examined at the individual participant level, learning curves do not reflect a gradual change process, but instead seem to resemble sudden changes of representations or states (Gallistel, 2012a; Gallistel et al., 2004; Glautier, 2013), which contrasts with the traditional single-state assumption.

To sum up, a difficulty in the study of PREE lies in the fact that the two diverging theoretical views (single- vs. multiple-state learning) make predictions at different levels (group vs. individual) and focus on different parameters (learning rate vs. moment of change). In particular, traditional (frequentist) data analysis approaches do not offer a straightforward way to test the predictions of the multiple-state learning view. This motivates our adoption of more flexible methods to analyze our data. In this paper, we do not aim to settle the debate or contrast between the two alternative theoretical views on learning. Rather, we want to encourage further research and theoretical advance by providing a method to test the predictions from the multiple-state view. We will do so by using Bayesian modeling techniques to formalize the assumptions of this view.

A flexible modeling tool: Bayesian methods

If we assume that the performance in many learning paradigms results from switching between multiple states or representations that are stored simultaneously, how could we model this process in the data-analysis stage? Ideally, we need to capture the moment of change in the participants' behavior. We expect this parameter (“moment of change”) to vary due to the experimental manipulation but also between individuals. Therefore, this estimation must be done at the level of each individual participant. Luckily, Bayesian modeling tools offer us a great deal of flexibility to incorporate these and other assumptions.

Before we start describing how to apply Bayesian modeling tools to our example phenomenon, PREE, in this section we provide an overview of Bayesian inference for readers who are not familiar with it. To guide us through the section, we will use the classical example of inferring the success rate for a binary process (Wagenmakers et al., 2010), because it represents one of the simplest scenarios and it is featured in most introductory texts on Bayesian inference (Kruschke, 2010b; Lee & Wagenmakers, 2013). A binary process produces any of two possible outcomes. Thus, it can be used to model many situations, such as coin tossing (heads vs. tails), or scoring in an exam (correct vs. incorrect items). In this context, the data we observe is the count of successful trials, such as the number of heads in a coin-tossing series. We will name this value k, so that if we count three heads in the series, then $k = 3$. Binary processes are usually modeled with a binomial distribution with two parameters: the number of trials (i.e., the number of coin tosses), n; and the success rate (i.e., the probability of obtaining heads), θ. If $\theta = 0.5$, then heads and tails are equiprobable, as in a fair coin. Larger values of this parameter would indicate that the coin is biased towards obtaining more heads than tails. Thus, the goal of the inference is to find out the most likely values for the parameter θ, given the data (i.e., given the number of coin tosses, *n*, and the number of heads, *k*). This will serve to eventually decide whether or not the coin is biased. The problem can be solved with Bayesian modeling tools.

We depict the graphical model we can use for this inference in Figure 1 (adapted from Lee & Wagenmakers, 2013). The nodes represent random variables (e.g., data, parameters), and the edges represent the dependencies between them. In the model, the number of heads obtained, *k*, depends on the number of coin tosses, *n*, and on the probability of obtaining heads, θ. In this graphical notation (Lee, 2008; Wagenmakers et al., 2010), continuous variables (such as θ) are depicted with circular nodes, and discrete variables (such as *k* and *n*) are depicted with square nodes. In addition, shaded nodes correspond to observed variables, whose value we already know, whereas empty nodes (such as θ) represent unobserved variables, whose values we do not know, but want to estimate from the data available.

One key aspect of Bayesian inference is the expression of uncertainty in the form of probability distributions over the parameters. Let us assume that we do not know anything about the binary process we are examining (i.e., we have not tossed any coin yet). Then, a priori, we can specify that all possible values for θ are equally likely. That is, we assume that the "prior distribution" of θ, called $p(\theta)$, is uniform, ranging from 0 to 1, as Figure 2 (left-hand panel) indicates. This assumption might be different in case we had reasons to suspect that some values would be more likely than others. For example, if we suspect that the coin is fair, then the prior distribution would assign a greater probability for $\theta = 0.50$, whereas other values will be assigned lower probabilities, as shown in Figure 2 (right-hand panel).

Suppose we toss the coin several times. Once we observe some data, *d* (i.e., the outcomes of the coin-tosses), we would like to update our knowledge, to obtain a "posterior distribution", or $p(\theta|d)$. Imagine that our data consist of the observation of 8 heads out of 10 coin tosses (i.e., $n = 10$, $k = 8$). Given that we modeled the binary process that generates the data with the binomial distribution, we assume that the number of heads obtained, *k*, are data drawn from the binomial distribution with parameters $n = 10$ and θ (whose value is unknown). That is, as Figure 1 indicates, $k \sim \text{Binomial}(\theta, n)$.[1]

Starting from our initial uniform prior distribution $p(\theta)$, which suggests that all values for θ are equally likely (Figure 2, left-hand panel), now our beliefs should change in light of the new data obtained. The rule to combine the prior distribution with the data to obtain a posterior distribution is the Bayes' theorem:

$$p(\theta|d) = \frac{p(d|\theta)p(\theta)}{p(d)}, \tag{1}$$

where *d* are the data observed in the experiment (coin-tosses), and θ is the parameter whose value we want to infer. For the purpose of this paper, we do not need to discuss all the elements involved in Equation 1. Instead, we refer the reader to the many introductory texts that are currently available (Etz et al., 2016; Kruschke, 2010a; Van de Schoot et al., 2014; Wagenmakers et al., 2010). The important point is that the prior beliefs, expressed as a distribution $p(\theta)$, have been updated in light of the data of the experiment, producing a posterior distribution $p(\theta|d)$, which is the goal of the

---

[1] Note that we are not bounded to assume that the data-generating process is binomial. Rather, we chose to model the process as binomial because it usually represents well this experimental situation. If we realized that more than two outcomes are possible in the experiment (e.g., a game with wins, loses, and ties), then we could use another distribution to model the process (e.g., Poisson and its respective parameter λ), without changing much about the model or the inferential process itself.

inference. This posterior distribution represents the current knowledge, once prior information and actual data have been combined. For example, it would represent our current belief about the bias of a coin after we have seen some coin-tossing experiments. As we advanced, Bayesian inference deals with uncertainty by working with distributions instead of relying on point estimations, as is typical in the frequentist (non-Bayesian) tradition in statistics. The process of solving Equation 1 is schematically represented in Figure 3. An equivalent procedure would be used, for example, to estimate the means of two experimental conditions by combining a prior belief about such values and the data obtained in the experiment.

The final piece of the puzzle that we need to mention in this section before we move on to modeling the PREE concerns the way these posterior distributions are obtained from the combination of prior knowledge and observed data (i.e., how Equation 1 is solved). For simple cases such as the binomial, the problem has a closed-form solution. However, this is rarely the case (this is why, for decades, Bayesian inference was relegated to coin-tossing and similar, very simple, problems). The use of more complex models became possible when random sampling methods such as the Markov Chain Monte Carlo (MCMC) were developed and applied to these situations (Andrieu, De Freitas, Doucet, & Jordan, 2003; Gamerman & Lopes, 2006; Hanson, 2000). MCMC refers to a family of algorithms for sampling sequences of values from a probability distribution. The use of this technique allows us to calculate a density estimate of the posterior distribution, which then we can use for our inferential purposes (e.g., comparing posterior distributions of different groups, or deciding whether or not a meaningful value, such as zero, falls within a credible interval). The technical details about how the algorithm works are omitted here, but the reader can resort to the references above to obtain a deeper description of the process.

Estimating the moment of change between states in the PREE

In our previous depiction of the PREE and the study of extinction, we have explained how, according to several authors, examining curves averaged from multiple subjects or trials would be misleading, and that the interesting parameter to examine in the PREE is the moment of change from the acquisition performance to the extinction performance (Gallistel, 2012a). This parameter, the “moment of change”, becomes central for multiple-state learning, as it can be seen as the moment in which one representation or state is selected and expressed, replacing the previous one. In this section, we present a Bayesian model (inspired by one proposed in Lee & Wagenmakers, 2013, p. 47) that could be used to analyze experimental data from this point of view. The three main assumptions we aim to include in the model are the following: (1) extinction produces a single point of change in behavior (in this example, measured as judgments), (2) participants show inter-individual variability in their specific moments of change, which explains why the curve becomes smoother when data from many subjects are aggregated, and (3) there are also systematic differences in the moment of change between groups, which results in the PREE (i.e., extinction is sensitive to the reinforcement schedule during acquisition).

Thus, one goal of the model is to analyze experimental data to detect the moment of change in the behavior during extinction at the individual level. We will name this individual moment of change with the Greek letter tau, τ. The estimation of "group-level moments of change" will be also possible by studying the group-level parameter τG. Like θ in the example of the previous section, these parameters are initially unknown, and must be inferred from data. In this case, we have specific predictions: we predict that participants in the continuous reinforcement group would show a sudden reduction of their responses, or judgments, soon in the extinction phase, whereas it would take more time for participants in the partial reinforcement groups. That is, $\tau G_{(\text{Continuous})} < \tau G_{(\text{Partial } 75\%)} < \tau G_{(\text{Partial } 50\%)}$. Additionally, we predict more variability in the individual moments of change in the partial reinforcement groups than in the continuous reinforcement group. This is because, in the latter group, most participants would extinguish in the first or second extinction trial (hence there will be low variability), whereas it will take generally more trials in the partial reinforcement groups, creating a wider range of potential moments of change (some participants might still extinguish soon, while others would delay it substantially)[2].

Thanks to the flexibility that Bayesian models feature, we can fully represent a detailed map of the process that (we assumed) generated the data in our experimental situation, which can be seen in Figure 4. The model specifies in a comprehensive way the parameters that were outlined in previous section, and also reflects the actual experimental procedure used to obtain the data. In this model, only a few variables possess known values (to indicate this, they are shaded), the rest will be estimated from the data and prior distributions, using MCMC simulations. For the task at hand, we chose a hierarchical structure, so that the distribution of each variable is described by the variables depicted over it. Thus, we start the description of the model in Figure 4 bottom-up.

In first place, we have the data: in this case, they are the judgments given by each participant *j* on each trial *i*, $D_{i,j}$. Judgments take on values from 0 (meaning that the US was not expected at all) to 100 (meaning that the US was expected with total

---

[2] It must be pointed out that this prediction could be wrong if participants in the partial reinforcement groups systematically delay their behavioral change to a very large extent (i.e., if they postpone the extinction until the last trials in the phase or beyond). This would result in extremely distant, non-overlapping posterior distributions of τG, but of similar (small) variability. However, the extinction phase, with 20 non-reinforced trial was expected to provide ample opportunity to observe an almost complete extinction in all the groups.

certainty). As the data are already known and need not being estimated, they are depicted in grey color in Figure 4. In the model, these data depend on several variables. First, we will focus on the node *t*, which acts like a counter, or an index to keep track of which extinction trial the current judgment corresponds to (i.e., it takes on values 1, 2, 3... for each trial, until it reaches the last extinction trial). Although some theories of the PREE propose that the effect is due to the accumulation of time between US presentations (i.e., time-based account, Gallistel & Gibbon, 2000), the alternative view that PREE is caused by the accumulation of unreinforced trials (i.e., trial-based account) has been more successful in explaining the empirical evidence (Bouton, Woods, & Todd, 2014; Haselgrove, Aydin, & Pearce, 2004). Consequently, we chose to model the phenomenon in terms of trials rather than in terms of accumulated time. The plate surrounding the variables *D* and *t* can be read as a "loop" which, in this case, iterates over trials (see Lee & Wagenmakers, 2013, for more examples featuring the "plate notation"). That is, for the participant *j*, $D_{j,1}$ is the judgment given in the first trial (i.e., when $t = 1$), $D_{j,2}$ is the judgment given in the second trial ($t = 2$), and so on. So far, this structure represents succinctly the sequence of judgments for each trial in a given participant.

Then, the hierarchical structure of Figure 4 unfolds upwards within another plate that corresponds to the individual participant level, with indices $j = 1$ (participant 1), $j = 2$ (participant 2), etc., which contains the whole sequence of trials. Within each participant *j*, we propose that the sequence of data *D* is drawn from a normal (Gaussian) distribution described by the nodes that are depicted over it (or, put it another way, we propose that a normal process described by these latter variables generated the data *D*). A normal distribution can be fully specified by two parameters, mean and standard deviation. In our model, the distribution from which the data *D* were drawn is described by μ1, μ2, and σ. The rationale for this choice is as follows: from the moment when the extinction phase starts, and until it reaches a given trial τ (i.e., from $t = 1$ to $t = \tau$), *D* is drawn from a normal distribution with mean μ1 and standard deviation σ. But, at one point τ, the process changes suddenly (i.e., extinction is expressed), thus producing different judgments. Therefore, from the trial τ onwards, the data *D* are distributed differently, also according to a normal distribution, but with mean μ2 (which we expect to be smaller than μ1, given that judgments should be lower in the extinction phase), and keeping a similar standard deviation σ (Equation 2).

$$D_{i,j} \sim \begin{cases} \text{Gaussian}(\mu 1, \sigma) & \text{if } t < \tau \\ \text{Gaussian}(\mu 2, \sigma) & \text{if } t \geq \tau \end{cases} \tag{2}$$

Thus, τ is presented as the moment of the extinction phase that identifies the change in the judgments for a given participant: from the acquisition performance (mean μ1) to the extinction performance (mean μ2), as Figure 5 illustrates. The model will just find the value of τ that breaks the series of data into two parts so that the difference between their means is maximal.

Here, we made assumptions that could have been different, such as that the nature of the process generating *D* is normal, and that the standard deviation is constant during the whole phase. In these two particular cases, these choices were made for the sake of simplicity (e.g., if the standard deviation is constant, there is one less parameter in the model), but what matters is that these modeling decisions are explicitly and transparently acknowledged in the model and, thus, may be discussed or changed if the researcher believes it is necessary. This speaks again to the flexibility of this type of modeling.

We continue our description of the model. In Bayesian methods, as implied by Equation 1, inference about a parameter starts from an initial distribution of beliefs about its possible values (i.e., a prior distribution). For instance, in our coin-tossing example, we might assume, before observing any data, that the coin is fair, meaning that the distribution is centered over $\theta = .50$. Alternatively, we could acknowledge that we know nothing about the process that generates the data, thus assuming that all values of the parameter are equally likely (these two options were depicted in Figure 2). Likewise, we need to specify prior distributions for the parameters of our PREE model. First, as is usual with parameters that are not the focus of the hypothesis testing, we assume rather uninformative prior distributions for $\mu1$, $\mu2$, and σ. Thus, $\mu1$ is assumed to be drawn from a normal distribution with mean equal to the mean of the last five acquisition trials (which would in this case represent the acquisition performance, or state) and large standard deviation, 100 (i.e., a flat distribution). Likewise, $\mu2$ is assumed to be drawn from a normal distribution with mean 0 (because this is the expected value at the end of the extinction phase) and standard deviation 100. Finally, we define the prior distribution for $1/\sigma^2$ as a Gamma distribution with very small shape and mean parameters (both are 0.01)[3]. The Gamma distribution is chosen because it samples only skewed positive values. Uninformative priors like these are usually preferred because they do not normally affect Bayesian estimations to a great extent (Gelman, Jakulin, Pittau, & Su, 2008; see some unusual exceptions in Gelman, 1996), so that the conclusions are not too determined by the assumption of a specific form of prior knowledge, and are more strongly driven by data.

---

[3] The prior is defined for $1/\sigma^2$ instead of for σ directly, because the software WinBUGS defines the normal distribution in terms of mean and precision, instead of in terms of mean and standard deviation, as is usual in most applications. It is easy to convert between the two parameters, because precision is 1/variance.

Finally, the model has one more layer in the hierarchy, corresponding to the group-level. So far, all the variables (τ, μ1, μ2...) were defined at the individual-subject level, that is, they may vary freely from participant to participant. But even acknowledging the individual variability, one still can predict that participants would differ in their "moment of change" between-groups (as we made explicit above). Thus, once the individual "moments of change", τ, have been estimated, we can then obtain an estimation of the underlying group-level distributions from which these individual τ were drawn. To model this, we assumed that the individuals' τ are distributed normally with mean τG and standard deviation σG. These two parameters are unknown and, according to our predictions, they will vary between groups, since they would be affected by the reinforcement schedule presented during the acquisition phase. This is why they depend on the observed (i.e., known by the experimenter) variable *G*, which represents the group to which the participant belongs to. In Figure 4, the group-level parameters τG and σG are surrounded by a plate with index *k* that could take on one of three values (*k* = 1: Continuous reinforcement group; *k* = 2: 75% Partial reinforcement group; *k* = 3: 50% Partial reinforcement group), to indicate that each of the three groups in our experiment will have different values for the group-level parameters. This layer of the structure allows for predictions at the group level that mirror the ones we planned when we conducted the behavioral experiment. Specifically, the predictions are that: (a) Participants in the continuous reinforcement group would extinguish their behavior soon, in the first extinction trials. This means that their τ would be generally low, and thus the mean of their distribution, τG, would be low too. In addition, there would be little variability in their moments of change (most of them would extinguish in the first or second extinction trials), hence σG would be small. (b) Participants in the two partial reinforcement groups would take longer before they change their judgments (as such, their τG would be larger than in the continuous reinforcement group). Also, there will be more variability in the individual τ in these latter groups, because there is a wider range of possible values in which extinction would be expected to appear: some people might still show a change in the first or second extinction trials, but probably others would change their judgments on later trials, or even at the end of the extinction phase. That is, in the partial reinforcement groups we expect more between-individuals variability in τ, and therefore their σG will be larger, as long as there is not a ceiling effect (see Footnote 2). Notably, this is a prediction that concerns the shape of the posterior distribution, which are naturally evaluated by Bayesian estimation methods.

The prior distributions for these two last parameters, τG and σG, need also to be specified before we can estimate their values. In this case, we chose again to assume uninformative priors, although other options are also possible. Thus, in each group, τG was drawn from an uniform distribution, which means that all the possible values are equally probable a priori. The prior for σG was also uninformative: a Gamma distribution with small mean and shape parameters (both were 0.01). As we argued before, uninformative priors are generally a way to prevent the prior knowledge from affecting the estimations too much, and this is the strategy we chose to follow.

All these parameters and the relationships between them, summarized in Figure 4, convey the full experimental situation with its structure and all the elements of interest, thus respecting the complexity of the actual data. Also, because the Bayesian model outputs a posterior distribution instead of a point estimation of the parameters, we can make predictions not only about the values of the parameters, but also about their variability (e.g., the distribution of τG will be narrower in a continuous reinforcement group than in partial reinforcement groups, as we explained above). Finally, which is most important for our current paper, the model captures the assumptions that the researcher wants to incorporate, in particular the notion that the result of learning can be expressed as a quick replacement and switching between multiple states or representations, instead of the gradual modification of a single state. To do this, the data are examined at the individual (in addition to the group) level, and the aim is to detect a moment of change in behavior rather than estimating the extinction rate of a learning curve that only makes sense when pooling many participants together. This illustrates how Bayesian tools are flexible enough to fully incorporate the experimental situation regardless of the parameters one aims to study or the level of analysis intended.

Behavioral Experiment

Before implementing the above outlined model, we first obtained data from a behavioral experiment featuring a PREE design. Our aim was to replicate this effect and then, using these data, to show how Bayesian modeling allows making inferences about the parameters of interest (i.e., the moment of change in extinction) at both individual and group levels, so that the predictions by the multiple-states learning view can be tested.

Method

Participants and apparatus. Fifty-five Psychology students of the University of Malaga took part in the experiment in exchange for course credits. They were randomly distributed as follows: 18 participants in the continuous reinforcement group, 18 in the 75% partial reinforcement group, and 19 in the 50% partial reinforcement group. The experiment was programmed in E-Prime for Windows and conducted in a large computer room with ten individual cubicles.

Procedure. Participants were told that a cartoon character was trying to turn on a light bulb. On a series of 42 trials, the character was shown on the screen, ready to operate a light switch, and participants were asked to rate how likely the light bulb would come on at that time. The ratings were given by clicking with a mouse on a continuous scale from 0 to 100. The instructions indicated that clicking on the rightmost part of the scale meant that it was very likely that the light would come on, whereas clicking on the leftmost part of the scale meant that it was very unlikely that the light would come on. A visual marker appeared on the point of the scale where the participant clicked. Although the program assigned a value (from 0 to 100) to each rating, proportional to its position in the scale, this number was not made visible to the participant. There was no time pressure to respond. Once the judgement was provided, the screen showed the character pressing the switch and the outcome corresponding to that trial: either the light came on or it stayed off. Both outcomes were accompanied by a sound and the facial expression of the character, which had the corresponding positive or negative emotional value of the outcome. Participants could then press the space bar to move to the next trial.

The first 20 trials corresponded to the acquisition phase, in which the light bulb was turned on (i.e., the US was shown) on the 50%, 75% or 100% of the trials, depending on the group. The trial order was random. Two additional reinforced trials were included at the end of the acquisition phase, to make the groups more comparable in their experience right before moving to the extinction phase. The subsequent extinction phase was identical for all participants and consisted of 20 unreinforced trials. There was no indication to the participants of the change of phase through any means.

Results

First, we report the traditional, frequentist, data analysis that is similar to those conducted on most PREE studies to date. The trial-by-trial prediction judgments were our dependent variable. As we expected, the three groups differed in the evolution of these judgments during the acquisition phase (left-hand of Figure 6). A one-way ANOVA conducted on the last block of five acquisition trials revealed significant between-groups differences: $F(2, 52)= 59.93$, $p < .001$, $\eta_p^2 = .697$. The continuous reinforcement group gave higher judgments than the 75% partial reinforcement group, $t(34) = 5.12$, $p < .0001$, and this latter, in turn, gave higher judgments than the 50% partial reinforcement group, $t(35) = 4.397$, $p < .0001$. In associative terms, this is commonly interpreted as the three groups reaching different learning asymptotes that depend on the reinforcement schedule.

The PREE was then studied at the aggregated level by examining the extinction curves (right side of Figure 6). The trial-by-trial data were submitted to a 3 (Group) x 20 (Trial) ANOVA, yielding a significant main effect of Trial, $F(19, 988) = 15.26$, $p < .001$, $\eta_p^2 = .227$, no main effect of Group, $F(2, 52) = 1.89$, $p = .165$, $\eta_p^2 = .07$, and a Trial x Group interaction, $F(38, 988) = 2.32$, $p < .001$, $\eta_p^2 = .082$. This interaction suggests that the extinction curves were different between groups, with the continuous reinforcement group showing a faster decline in the responses than the two partial reinforcement groups. As simple effect contrasts show, the tendency becomes evident in the fourth, fifth and sixth extinction trials: $F(2, 52) = 3.370$, $p = .042$, $\eta_p^2 = .082$; $F(2, 52) = 4.008$, $p = .024$, $\eta_p^2 = .134$; and $F(2, 52) = 3.184$, $p = .049$, $\eta_p^2 = .109$ (Table 1 contains simple-effect between-group contrasts, included here for completeness sake). This result is compatible with the PREE.

Table 1

Simple-effect between-group contrasts for the three extinction trials in which ANOVA was significant (i.e., fourth, fifth, and sixth trials).

Contrast	Fourth extinction	Fifth extinction	Sixth extinction

	trial	trial	trial
Continuous vs. Partial 75%	$p = 0.668$	$p = 0.062$	$p = 0.038$
Continuous vs. Partial 50%	$p = 0.019$	$p = 0.008$	$p = 0.029$
Partial 75% vs. Partial 50%	$p = 0.052$	$p = 0.408$	$p = 0.924$

A Bayesian model for PREE: The multiple-state assumption made explicit

The previous analyses examined the possibility that the extinction curves proceeded at different pace depending on the reinforcement schedule presented during the acquisition phase, consistent with the single-state assumption of the standard associative models. The results, studied by means of an analysis of variance, were in line with the PREE. However, the predictions made from the multiple-state learning view focus on different parameters. Consequently, we present the application of the model described in earlier sections to make inferences about the data collected in the behavioral experiment.

Method

The model in Figure 4 was programmed in WinBUGS, a software for Bayesian analysis of complex statistical models that uses MCMC algorithms to produce an estimation of the posterior distributions (Thomas, 1994). The BRUGS package for R (Thomas, O’Hara, Ligges, & Sturtz, 2006) served to handle the experimental data and the results of the simulations. Also, we used scripts derived from those by Kruschke (2010b) to plot the posterior distributions and compute highest density intervals (see below).

In our model, we were interested in estimating the posterior distributions of τ, τG and the rest of the parameters, so that we can test our predictions. These estimations were achieved by using MCMC sampling methods. Because we wanted to ensure that the samples were taken from the stationary distributions and were not affected by the random starting points of the chains, we sampled five separate Markov chains for each parameter, and took 51,000 samples. The first 1,000 samples of each chain were discarded to get more stability (a technique commonly known as "burn-in"). These are standard practices in Bayesian parameter estimation (Kruschke, 2010b).

Results

Before we report the results of our simulations, we stop briefly to look at the data more closely. In the previous section, we observed a seemingly gradual extinction curve that appeared at the group-level (see Figure 6). As we explained in the Introduction, the data may look very different at the individual level. Our code available at the Open Science Framework produces plots of the learning curves for each individual participant. The examination of these curves reveals great variability: some participants extinguish readily and permanently (e.g., participants 16 and 37), whereas others seem to extinguish progressively (e.g., participant 45), others show erratic behavior (e.g., participants 1 and 14), or even seem to not extinguish at all. This speaks once again to the importance of not assuming that a feature that emerges at the group level (e.g., smooth curve, extinction, etc.) necessarily applies to each and every individual in the sample (Gallistel, 2012b). Acknowledgement of individual variability can improve the modeling of the actual underlying process. For Figure 7, we have selected a few representative participants that show an apparent drop in their judgments during the extinction phase, as the model assumes.

Looking at Figure 7, and focusing on the extinction phase, it is possible to figure out when a certain participant changed their behavior (from the acquisition performance to the extinction performance). The Bayesian model is designed to capture the moment when an individual participant made this change in the judgments. As mentioned above, we called this parameter τ, and its value is the one that splits the extinction phase into two parts so that the difference between their mean judgments is maximal (Equation 2). Figure 8 represents the output of the model, in the form of the posterior distributions for the individual τ of the same participants represented in Figure 7. The horizontal axis of Figure 8 represents the extinction phase, and the posterior distribution informs us about the degree of credibility of each possible value for the individual τ of a participant. For instance, consider participants 16, 28 and 37 (all from the Continuous reinforcement group): they seem to extinguish abruptly and quickly in Figure 7 (top row), as soon as the extinction phase starts. This impression largely converges with the estimations depicted in Figure 8 (top row), because their posterior distributions are narrow (indicating little uncertainty about the moment of change τ) and they are centered over the initial extinction trials (which means that they changed their behavior soon in the extinction phase). The participants in the partial reinforcement groups (Figure 8, second and third rows) show similarly narrow distributions (i.e., little uncertainty in the estimations), but they are shifted towards the right part of the plot, indicating a quick, but delayed extinction (i.e., τ values are larger in these groups). This aligns with the predictions we made before building our model.

An important fact is that the model produced posterior distributions that inform us about the most likely values for the individual τ of each participant, and also convey the uncertainty associated to such estimations (e.g., narrow distributions indicate little uncertainty). To help with the interpretation of the posterior distributions, Figure 8 includes a thick black line superimposed onto the horizontal axes, corresponding to the 95% highest density interval (HDI). The HDI is the interval that contains the 95% of the mass of the posterior distribution (Kruschke & Vanpaemel, 2015). This is the Bayesian equivalent of a confidence interval, although its interpretation is more intuitive (i.e., we can assume with 95% confidence that the true value of the parameter lies between the limits of a 95% HDI). The varying widths of the intervals displayed by certain participants (e.g., participants 32 versus 47) are indicative of different degrees of uncertainty: the estimation indicates that it is not always as clear when the judgments of these participants actually changed during the extinction phase, and this is represented accordingly in the posterior distribution. Wide HDIs suggest that the estimation was uncertain, and the shape of the distribution can help us identify the most likely values (e.g., as when it is multimodal, see participant 32). Additionally, values that fall out of the HDI are highly unlikely (e.g., none of these participants seemed to delay the extinction beyond the 15th trial, which is never included in the intervals). The use of posterior distributions and HDIs instead of point estimations is, then, very helpful for the researcher, to better understand the data observed (see Kruschke & Liddell, in press, for a more detailed explanation about how to use HDIs for Bayesian hypothesis-testing).

So far, we have discussed the estimations of individual participants. However, our main predictions concerned the differences between groups, in particular those represented by the parameter τG (i.e., the moment of extinction for groups). Thus, Figure 9 depicts the posterior distributions of the τG for each group. In line with the predictions, the distributions of τG are centered in larger values and have larger variances as a function of the amount of unreinforced trials received in the acquisition phase. That is, in the continuous reinforcement group, the mean of τG is small (3.46, *SD* = 0.29, 95% HDI = 3 to 3.98) and the distribution is narrow. The 95% HDI does not overlap with those of the 75% partial reinforcement group (*M* = 5.75, *SD* = 0.81, 95% HDI = 4.27 to 7.45) and the 50% partial reinforcement group (*M* = 8.03, *SD* = 0.92, 95% HDI = 6.3 to 9.95), suggesting that the group-level moment of change, τG, was indeed smaller in the Continuous reinforcement group than in the other two groups, as expected. The difference between the two partial reinforcement groups is less evident, as there is some overlap in their 95% HDIs, although the results were expected: (1) The values of τG were larger (i.e., extinction was delayed) as a function of the amount of unreinforced trials, and (2) the distributions of τG become wider with lower rates of reinforcement, as a consequence of individuals showing more variability in their individual τ.

Discussion

An important part of the associative tradition, exemplified by theories such as the Rescorla-Wagner model (Rescorla & Wagner, 1972), rests on the assumption that learning can be captured as a gradual change in one parameter that typically represents the associative strength between two stimuli in a given moment. If contingencies change, then the parameter must be updated just as it happens when contingencies are stable. That is, animals form a single representation, or state, of the learning situation, and transform it when necessary to adapt to changes. Most associative models share this single-state assumption, according to which only one representation or state can be learned at a time, and consequently predict the gradual, negatively accelerated, learning curves to which most of us have got used. However, some evidence suggests that the single-state assumption is wrong (Gallistel, 2012a; Gallistel et al., 2004). Consequently, a few associative models included the notion that reinforced and nonreinforced trials could be learnt as two different associations (one with the US representation, and another one with the no-US representation) that co-exist and compete. This idea is based on uncertainty-guided learning models (Pearce & Hall, 1980), and was adopted to create hybrid associative models that, unlike the original RWM, take into account both prediction error and uncertainty (Esber & Haselgrove, 2011). Nonetheless, these hybrid models still use gradual acquisition functions, in line with the associative tradition. One the other hand, according to the multiple-state learning assumption that we describe in this paper, animals could be able to acquire and store multiple representations of the learning situation (e.g., the one before and the one after a contingency change), and eventually use the one that better suits the current task goals and demands. These diverging views on learning (traditional single-state vs. multiple-state) produce different predictions for certain well-established phenomena, such as the PREE, which is the focus of this paper. This should naturally stimulate researchers to conduct experiments aimed at discriminating between the theories. However, the multiple-state learning view has not been extensively studied in the PREE. One of the reasons is that testing this view requires a data-analysis approach that allows for simultaneously studying both the individual- and the group-levels, and that examines a different parameter of interest (the moment of change in behavior, instead of the learning rate). In this paper, we try to fill this methodological gap by developing a Bayesian model that allows a more direct test of the predictions of the multiple-state learning view.

Our Bayesian model is designed to analyze series of data (judgments) in a PREE design. It incorporates a switch between two states (or performance levels), and the parameter of interest is the point in the series of data when individuals' behavior switches from the acquisition performance level to the extinction performance level (i.e., the "moment of change"). The model allows for both group and individual variability in this parameter, and it provides estimations at the individual, as well as the group level.

We applied the model to actual data obtained in a behavioral experiment with a PREE design. The results indicated that, indeed, the model can provide a satisfactory fit at the individual and group levels. In line with the predictions derived from a multiple-state view of PREE (Gallistel, 2012a), the moment of change varied with the probability of reinforcement. Participants who were trained under a continuous reinforcement schedule during the acquisition phase showed an abrupt reduction in the judgments early during the extinction phase. By contrast, participants in the two partial reinforcement groups delayed this reduction in their judgments for at least a few more trials.

Moreover, since our Bayesian model estimates full posterior distributions for the parameters, we were able to address another (novel) prediction that concerned individual and between-group variability in the moment of change: participants in the continuous reinforcement group showed little variability in this parameter, while the distributions were much wider in the other two groups. This is again in line with a two-state view of PREE. If the two phases were less discriminable in the partial reinforcement groups, we would expect the moment when the judgments changed to be also more variable. This prediction could be tested because of one feature that seems unique of Bayesian models, the use of probability distributions to represent uncertainty during the whole inference process.

Thanks to our Bayesian model, we have been able to document a recurrent pattern that many individuals show in learning tasks: they tend to exhibit large, sudden changes in their behavior after a contingency chance (Gallistel, 2012b). Although the experiment cannot rule out the possibility of the single-state approach, our findings are compatible with the multiple-state view. Importantly, the predictions from the multiple-state approach were rarely tested in the past. This is due to the inherent difficulty in analyzing a parameter such as “the moment of change”, especially when we focus on the individual level. Whereas traditional data-analysis approaches are too rigid to allow testing these predictions, our Bayesian model succeeded in finding evidence in line with the multiple-state assumption. From this point, researchers should make an effort to formalize and develop an actual theory built on the assumption of the multiple-state approach, which is currently missing, although it is implicit in some current theories of occasion-setting and renewal (Rosas & Callejas-Aguilera, 2006). These future theories should find ways to combine the predictions at the group and at the individual levels.

Notably, some of the researchers who previously pointed the problems of analyzing group-level learning curves also proposed their own alternative solutions. One of the obvious remedies to the perils of aggregating data is to resort to cumulative records and individual curve-fitting (Gallistel, 2012a; Gallistel et al., 2004). We believe that our Bayesian approach possesses important advantages, such as a more direct representation of the process that generated the data and the estimation of the parameter that is theoretically relevant, and the novel prediction concerning the variability in the moment of change. On the other hand, the study of cumulative records seems more neutral (data-driven, not theory-driven). Because of this, they can be complementary approaches.

It is not the goal of our paper to provide evidence to discriminate between the two perspectives of learning in the PREE (single-vs. multiple-state). Rather, we noticed that, while the predictions of the traditional single-state view have been often studied, the predictions of the multiple-state view were tested only indirectly (e.g., both the effect of context-switch manipulations and the existence of abrupt changes in individual learning curves are evidence supporting this view). Our contribution fills this gap by providing a method that allows a more direct test of predictions that were largely unstudied in the past. But, how could Bayesian methods help researchers in the associative learning field discriminate between the predictions of different theoretical views (e.g., single- vs. multiple-state)? Flexible model comparison is another one of the advantages of Bayesian inference. Thus, one could develop two models, each one representing the process under a different theoretical approach. Then, their ability to model and predict the data could be quantified and compared. This goal is far beyond the scope of our paper, but these two hypothetical models could be designed to differ to a great extent if needed. Fortunately, unlike other techniques, Bayesian hypothesis testing can take into account the complexity of each model, and penalize it accordingly, which is sometimes known as "the Bayesian Ockham's Razor" (Jefferys & Berger, 1992; see also Wagenmakers et al., 2010, for more details about Bayesian model comparison). In sum, comparing between Bayesian models could be a feasible way to advance in our knowledge on associative learning, something already quite common in other areas of experimental psychology.

The finding that learning procedures sometimes produce rapid switches between levels of responding is not completely incompatible with traditional associative models of learning. However, to account for these patterns of data, that seem better characterized by the moment of change than they are by other parameters such as learning rate, these associative models should incorporate additional mechanisms that allow for this behavior. Just to mention one possibility, learning could still proceed according to an associative strength update rule such as that of RWM, but then a nonlinear response rule could translate this strength into behavior so that it produces step-like changes in the data (e.g., it could resemble a logistic function), without the need to postulate the tracking of multiple states, or the replacement of one state by another, as the alternative view assumed. However, this type of response rules is rarely mentioned in associative theories. When researchers aim to test the prediction of an associative model, they typically use the original model, without incorporating any extra response rule. A clear example is the analysis of learning curves in contingency learning (López, Almaraz, Fernández, & Shanks, 1999), which is often presented as evidence in favour of associative processes, although, as argued before, this could be misleading because gradual learning curves only appear at the aggregated level. Admittedly, we have some learning theories that do incorporate response rules that could lead to extreme, rapid changes in behavior, specially for phenomena such as occasion setting and renewal. The most notable exponent of this is the comparator hypothesis (Arcediano, Escobar, & Miller, 2005). Then again, this theory is at least partially silent about the learning mechanism itself, as it is principally the formalization of a response rule (but see the more recent SOCR, or Sometimes Competing Retrieval model, which formalizes in detail the comparator hypothesis; Stout & Miller, 2007). Using a very different perspective, Frey and Sears (1978) extended the Rescorla and Wagner model (1972) by incorporating an attentional mechanism as well as a complex, non-monotonic, response rule based on Zeeman's catastrophe theory (1973). With these two additions, the model can explain sigmoidal learning curves, as well as non-gradual effects like reacquisition. In sum, associative theories could be endowed with mechanisms to produce step-like patterns of responding like the ones seen in the individual learning curves of our experiment, but normally, either such mechanisms are not explicitly formalized into the associative theory (as is the case of RWM), or these response mechanisms are modeled independent of the association formation mechanism (e.g., comparator hypothesis), or even lack the level of formalization that associative models typically achieve (Bouton, 2004).

One important point is that our Bayesian model aims at making sensible inferences from experimental data, rather than proposing a new theoretical approach to challenge previous ones. As long as the three assumptions embedded in the Bayesian model remain plausible (i.e., there is one change point in the data, there is individual variability in the moment of change, and there are between-group differences in the moment of change) the model can be used regardless of the theoretical interpretation. For example, if associative models, by including an extra response rule or by other means, are capable of producing step-like changes in behavior at the individual level, then the data could still be analyzed under the assumption of a change-point in the series of data, and not under the assumption of a gradual change. This possibility shows that it is a versatile analysis tool.

When examining the Bayesian model that we have presented in this paper, the reader might discover features that could have been implemented differently. In fact, many of the decisions made when modeling any situation are, in part, arbitrary. After all, modeling implies making assumptions, and Bayesian models are not different in this concern. Just to mention one example, in our model τ and τG (representing the moments of change at the individual and group levels, respectively) are continuous variables. This conveys the idea that they represent a moment in the subjective flow of time that participants perceive, the moment where they switch states[4]. However, in our experimental procedure, time was discretized into trials. Thus, a different but also sensible modeling decision would have been to assume τ and τG were drawn from a discrete distribution, such as a Poisson distribution. This can be readily modified in the model: instead of mean and standard deviation, the parameter that defines the Poisson distribution is λ (i.e., the expected number of occurrences of an event per unit of time). Furthermore, we could even compare the performance of the two models (the one with continuous τ and τG, and the one that uses a Poisson process to model them), to see

---

[4] Associative models have made similar assumptions in the past, proposing that time is divided into micro-trials (e.g., Mclaren & Mackintosh, 2000).

which of them better fits the data. Interestingly, some experiments on the PREE found results favoring a trial-based account over a time-based account of the phenomenon (Bouton et al., 2014; Haselgrove et al., 2004). This empirical evidence can be used to guide modeling decisions such as whether time should be discretized into trials or considered in a more continuous fashion. Moreover, the model can be adapted to different experimental situations. For example, the PREE has been found also in within-subject designs (Pearce, Redhead, & Aydin, 1997). By changing the boundaries of the "Participants" plate in our Figure 4, we could examine PREE in these conditions too without substantially modifying the model's assumptions.

Finally, one additional goal of this paper was to constitute an invitation to consider Bayesian methods in the field of associative learning, just as it is becoming usual in other research areas. Their flexibility and the informational richness of the output they produce (e.g., posterior distributions instead of point estimations, which accounts for uncertainty) make them ideal tools for associative learning theorists to create better implementations of their hypothesis-testing, and a great chance to expand the field to new horizons.

References

Andrieu, C., De Freitas, N., Doucet, A., & Jordan, M. I. (2003). An introduction to MCMC for machine learning. *Machine Learning, 50*(1–2), 5–43. doi: 10.1023/A:1020281327116

Arcediano, F., Escobar, M., & Miller, R. R. (2005). Bidirectional associations in humans and rats. *Journal of Experimental Psychology: Animal Behavior Processes, 31*(3), 301–18. doi: 10.1037/0097-7403.31.3.301

Baum, W. M. (2012). Extinction as discrimination: the molar view. *Behavioural Processes, 90*(1), 101–10. doi: 10.1016/j.beproc.2012.02.011

Bouton, M. E. (1994). Context, ambiguity, and classical conditioning. *Current Directions in Psychological Science, 3,* 49–53. doi: 10.1111/1467-8721.ep10769943

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory, 11*(5), 485–94. doi: 10.1101/lm.78804

Bouton, M. E., Woods, A. M., & Todd, T. P. (2014). Separation of time-based and trial-based accounts of the partial reinforcement extinction effect. *Behavioural Processes, 101,* 23–31. doi: 10.1016/j.beproc.2013.08.006

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences, 10*(7), 294–300. doi: 10.1016/j.tics.2006.05.004

Esber, G. R., & Haselgrove, M. (2011). Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning.

Proceedings of the Royal Society B: Biological Sciences, *278*, 2553–2561. doi: 10.1098/rspb.2011.0836

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, *53*(2), 134–140. doi: 10.1037/h0045156

Etz, A., Gronau, Q. F., Dablander, F., Edelsbrunner, P. A., & Baribault, B. (2016). *How to become a Bayesian in eight easy steps: An annotated reading list*. Retrieved from osf.io/8wkpd

Frey, P. W., & Sears, R. J. (1978). Model of conditioning incorporating the Rescorla-Wagner associative axiom, a dynamic attention process, and a catastrophe rule. *Psychological Review*, *85*(4), 321–340. doi: 10.1037//0033-295X.85.4.321

Gallistel, C. R. (2012a). Extinction from a rationalist perspective. *Behavioural Processes, 90*(1), 66–80. doi: 10.1016/j.beproc.2012.02.008

Gallistel, C. R. (2012b). On the evils of group averaging: Commentary on Nevin's "Resistance to extinction and behavioral momentum." *Behavioural Processes, 90*(1), 98–99. doi: 10.1016/j.beproc.2012.02.013

Gallistel, C. R., Fairhurst, S., & Balsam, P. (2004). The learning curve: implications of a quantitative analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(36), 13124–31. doi: 10.1073/pnas.0404965101

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review*, *107*(2), 289–344. doi: 10.1037//0033-295X.107.2.289

Gamerman, D., & Lopes, H. F. (2006). *Markov chain Monte Carlo: Stochastic simulation for Bayesian inference*. Boca Raton, FL: Chapman & Hall.

Gelman, A. (1996). Bayesian model-building by pure thought: some principles and examples. *Statistica Sinica*. Retrieved from http://www3.stat.sinica.edu.tw/statistica/oldpdf/A6n113.pdf%5Cnpapers3://publication/uuid/A364DE2F-F7D7-4021-BE04-5FDA29AF6E65

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics, 2*(4), 1360–1383. doi: 10.1214/08-AOAS191

Glautier, S. (2013). Revisiting the learning curve (once again). *Frontiers in Psychology*, *4*, 982. doi: 10.3389/fpsyg.2013.00982

Hanson, K. M. (2000). Tutorial on Markov Chain Monte Carlo. In *Proceedings of the 29th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Technology*. Gif-sur-Yvette, France.

Haselgrove, M., Aydin, A., & Pearce, J. M. (2004). A Partial Reinforcement Extinction Effect Despite Equal Rates of Reinforcement During Pavlovian Conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*(3), 240–250. doi: 10.1037/0097-7403.30.3.240

Jacobs, R. a., & Kruschke, J. K. (2011). Bayesian learning theory applied to human cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*, 8–21. doi: 10.1002/wcs.80

Jefferys, W. H., & Berger, J. O. (1992). Ockham’s Razor and Bayesian Analysis. *American Scientist*, *80*(1), 64–72.

Jenkins, W. O., & Rigby, M. K. (1950). Partial (periodic) versus continuous reinforcement in resistance to extinction. *Journal of Comparative and Physiological Psychology*, *43*, 30–40.

Kruschke, J. K. (2008). Bayesian approaches to associative learning: From passive to active learning. *Learning & Behavior*, *36*(3), 210–226. doi: 10.3758/LB.36.3.210

Kruschke, J. K. (2010a). Bayesian data analysis. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(5), 658–676. doi: 10.1002/wcs.72

Kruschke, J. K. (2010b). *Doing Bayesian Data Analysis: A Tutorial with R and BUGS*. London, UK: Elsevier.

Kruschke, J. K., & Liddell, T. M. (n.d.). The Bayesian New Statistics : Hypothesis testing , estimation , meta-analysis , and planning from a Bayesian perspective. *Psychonomic Bulletin & Review*. doi: 10.3758/s13423-016-1221-4

Kruschke, J. K., & Vanpaemel, W. (2015). Bayesian estimation in hierarchical models. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *The Oxford Handbook of Computational and Mathematical Psychology*. Oxford, UK: Oxford University Press.

Lee, M. D. (2008). Three case studies in the Bayesian analysis of cognitive models. *Psychonomic Bulletin & Review*, *15*(1), 1–15. doi: 10.3758/PBR.15.1.1

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge, UK: Cambridge University Press.

López, F. J., Almaraz, J., Fernández, P., & Shanks, D. R. (1999). Adquisición progresiva del conocimiento sobre relaciones de predictivas: Curvas de aprendizaje en juicios de contingencia. *Psicothema*, *11*(2), 337–349.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276–298. doi: 10.1037/h0076778

Mclaren, I. P. L., & Mackintosh, N. J. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. *Animal Learning & Behavior*, *26*(3), 211–246. doi: 10.3758/BF03200258

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological Bulletin*, *117*(3), 363–86. doi: 10.1037/0033-2909.117.3.363

Mowrer, O. H., & Jones, H. M. (1945). Habit strength as a function of the pattern of rein- forcement. *Journal of Experimental Psychology*, *35*, 293–311.

O'Reilly, J. X., Jbabdi, S., & Behrens, T. E. J. (2012). How can a Bayesian approach inform neuroscience? *The European Journal of Neuroscience*, *35*(7), 1169–79. doi: 10.1111/j.1460-9568.2012.08010.x

Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology*, *52*, 111–139. doi: 10.1146/annurev.psych.52.1.111

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.

Pearce, J. M., Redhead, E. S., & Aydin, A. (1997). Partial Reinforcement in Appetitive Pavlovian Conditioning with Rats. *The Quarterly Journal Of Experimental Psychology*, *50*(4), 273–294.

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). New York: Appleton-Century-Crofts.

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*(3), 461–74. doi: 10.1037/0278-7393.32.3.461

Siegel, S., & Allan, L. G. (1996). The widespread influence of the Rescorla-Wagner model. *Psychonomic Bulletin & Review*, *3*(3), 314–321. doi: 10.3758/BF03210755

Stout, S. C., & Miller, R. R. (2007). Sometimes-Competing Retrieval (SOCR): A Formalization of the Comparator Hypothesis. *Psychological Review*, *114*(3), 759–783. doi: 10.1037/0033-295X.114.3.759

Thomas, A. (1994). BUGS: a statistical modelling package. *RTA/BCS Modular Languages Newsletter*, *2*, 36–38.

Thomas, A., O'Hara, B., Ligges, U., & Sturtz, S. (2006). Making BUGS Open. *R News 6*, *6*(1), 12–17.

Van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & van Aken, M. A. G. (2014). A Gentle Introduction to Bayesian Analysis: Applications to Developmental Research. *Child Development*, *85*(3), 842–860. doi: 10.1111/cdev.12169

Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage-Dickey method. *Cognitive Psychology*, *60*(3), 158–189. doi: 10.1016/j.cogpsych.2009.12.001

Zeeman, E.C. (1976) Catastrophe theory. *Scientific American*, *234(4)*, 65–83.

Figure captions

Figure 1. Graphical model for inferring the success rate of a binary process.

Figure 2. Two possible distributions for representing our knowledge about the success rate of a binomial process. On the left panel, we choose a uniform prior distribution, according to which all values for the parameter are equiprobable, to acknowledge our complete lack of information about the process. On the right panel, we represent another possible assumption: that, a priori, the two outcomes (e.g., heads and tails) appear with the same probability, and thus the more likely value for θ is 0.5 (this is implemented as a Beta prior distribution). These a priori beliefs are chosen by the researcher, and reflect her knowledge on the matter. Note that the figure represents probability density functions, which means that the area under the function integrates to 1, hence densities higher than 1 are possible.

Figure 3. Bayesian inference implies a transition from the prior distribution of the parameter θ (left panel) to a new posterior distribution (right panel) that incorporates the newly acquired data. In this example, we start from a uniform prior distribution, which is updated after observing eight heads in a series of ten coin-tossing experiments.

Figure 4. Our Bayesian model for inferring the moment of change in the series of judgments.

Figure 5. Schematic representation of the rationale behind the model. The plot represents fictitious data (judgments) of an individual participant during the extinction phase (trial $t = 1, 2...$). The participants' judgments change from the acquisition performance to the extinction performance. In the acquisition performance, the judgments are distributed with mean μ1, but at some point they change suddenly to the extinction performance, in which they are distributed with mean μ2. This change takes place in the trial τ, which is initially unknown and must be inferred.

Figure 6. Group-level learning curves in our experiment. The vertical line separates the two training phases, acquisition and extinction.

Figure 7. Individual-level learning curves for some selected participants in our experiment. The first, second, and third row of panels contain participants from the Continuous, partial 75%, and Partial 50% reinforcement groups, respectively. The vertical line separates the two training phases, acquisition and extinction.

Figure 8. Density of the posterior distributions of the individual τ estimated for the participants that were presented in Figure 7. Horizontal axes represent the extinction trials (from 1 to 20). The thick horizontal lines below the distributions represent the 95% HDI (see main text) for the posterior distribution of the individual τ.

Figure 9. Density of the posterior distributions estimated for the group-level parameter τG, which represents the moment of change in the judgments for each group. The thick lines close to the axis represent the 95% HDIs.

Figure 1.

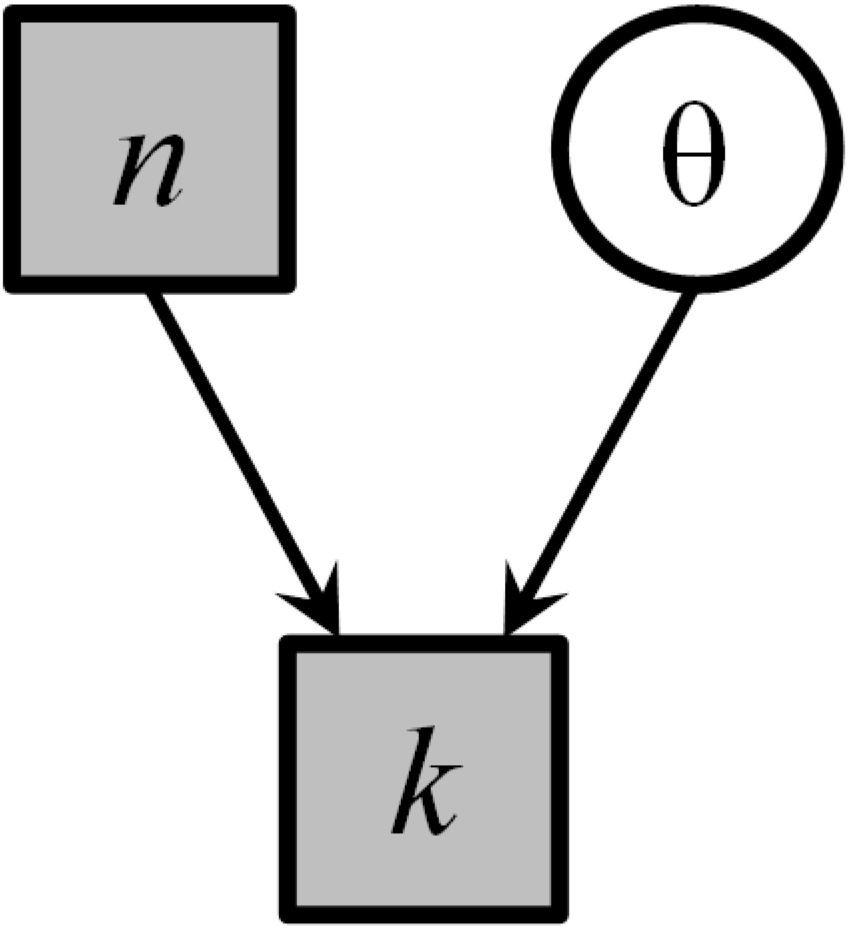

Figure 2.

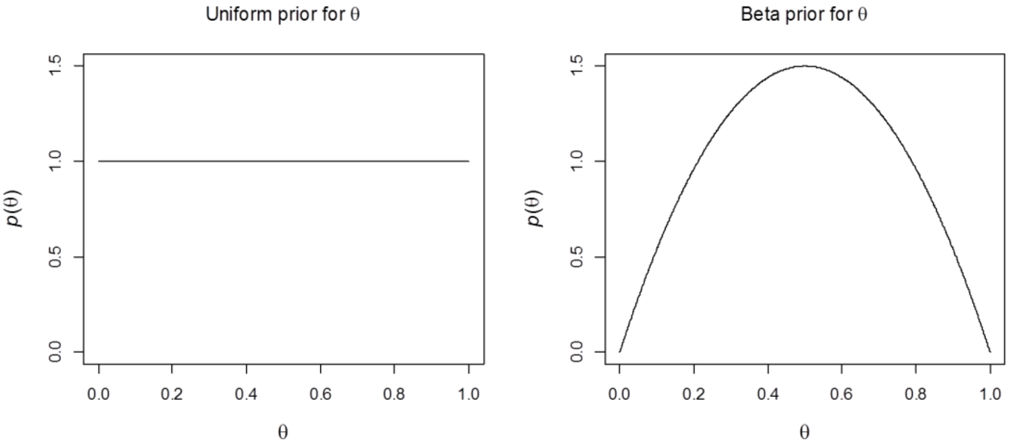
Uniform prior for θ
Beta prior for θ
p(θ)
θ
0.0
0.5
1.0
1.5
0.0
0.2
0.4
0.6
0.8
1.0

Figure 3.

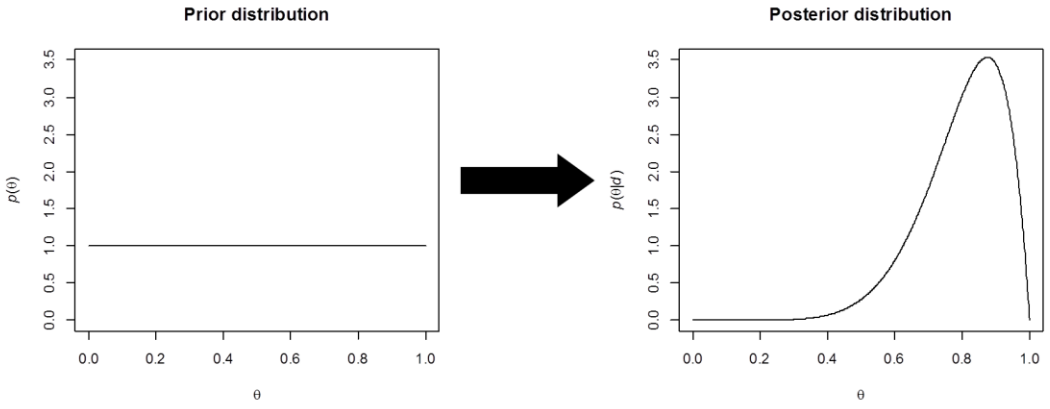
Prior distribution
p(θ)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.0
0.2
0.4
0.6
0.8
1.0
θ
Posterior distribution
p(θ|d)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.0
0.2
0.4
0.6
0.8
1.0
θ

Figure 4.

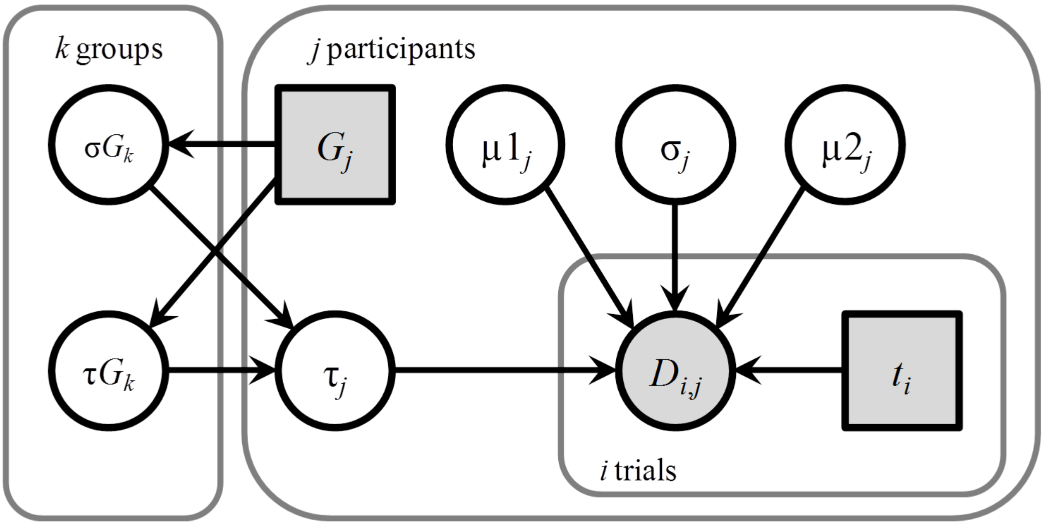
k groups
j participants
σG_k
G_j
$\mu 1_j$
σ_j
$\mu 2_j$
τG_k
τ_j
$D_{i,j}$
t_i
i trials

Figure 5.

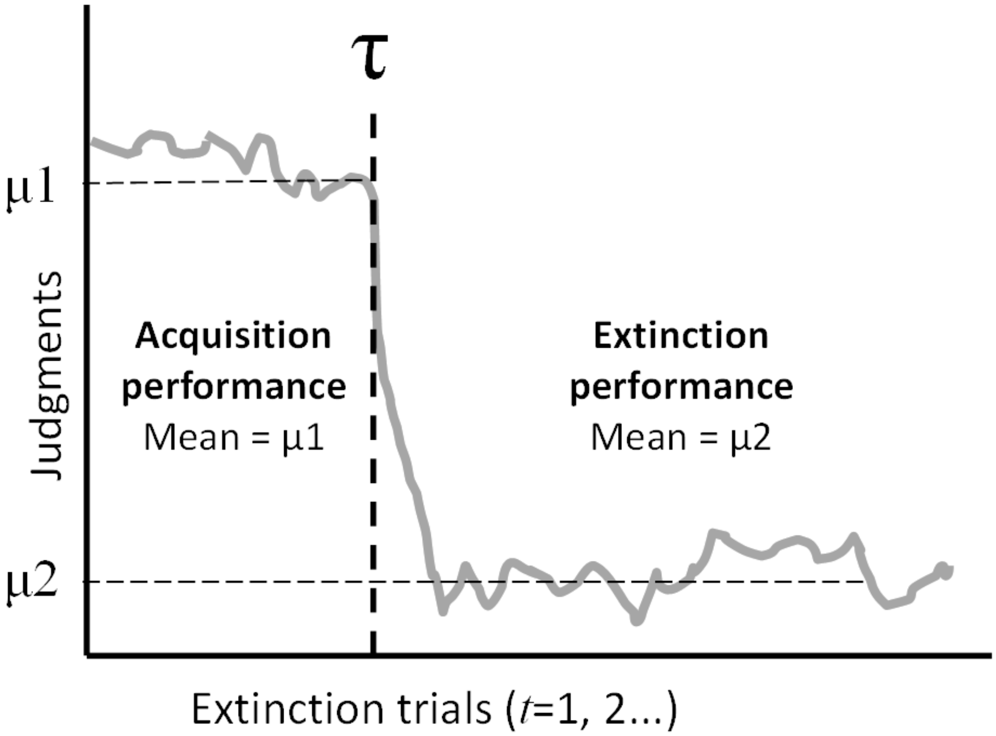
τ
μ1
μ2
Judgments
Acquisition performance
Mean = μ1
Extinction performance
Mean = μ2
Extinction trials (t=1, 2...)

Figure 6.

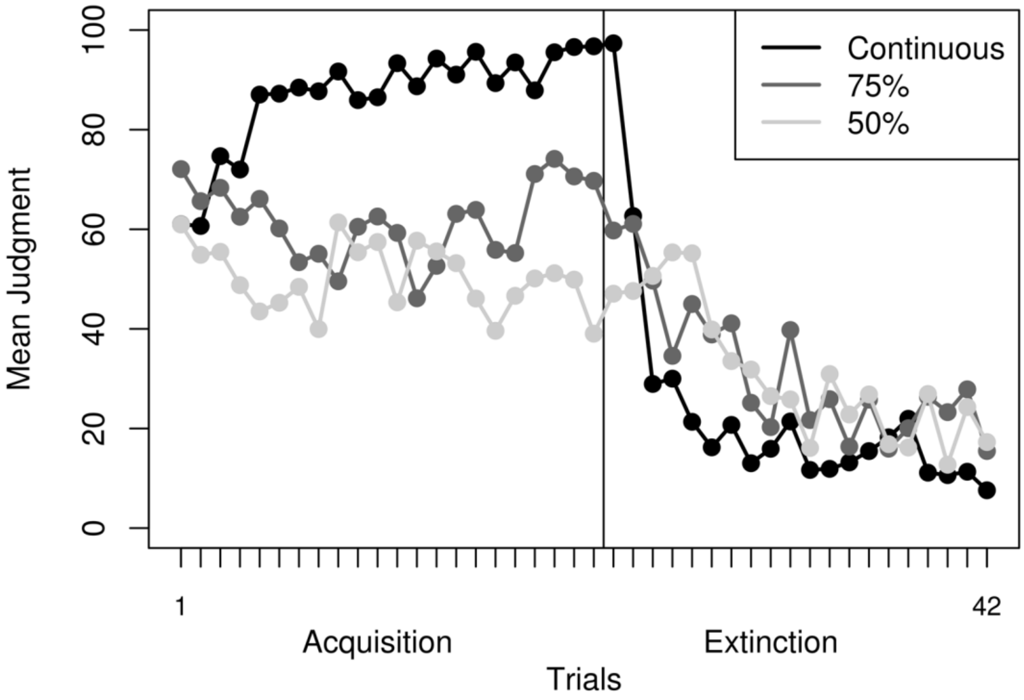
Mean Judgment
0
20
40
60
80
100
Continuous
75%
50%
1
42
Acquisition
Extinction
Trials

Figure 7.

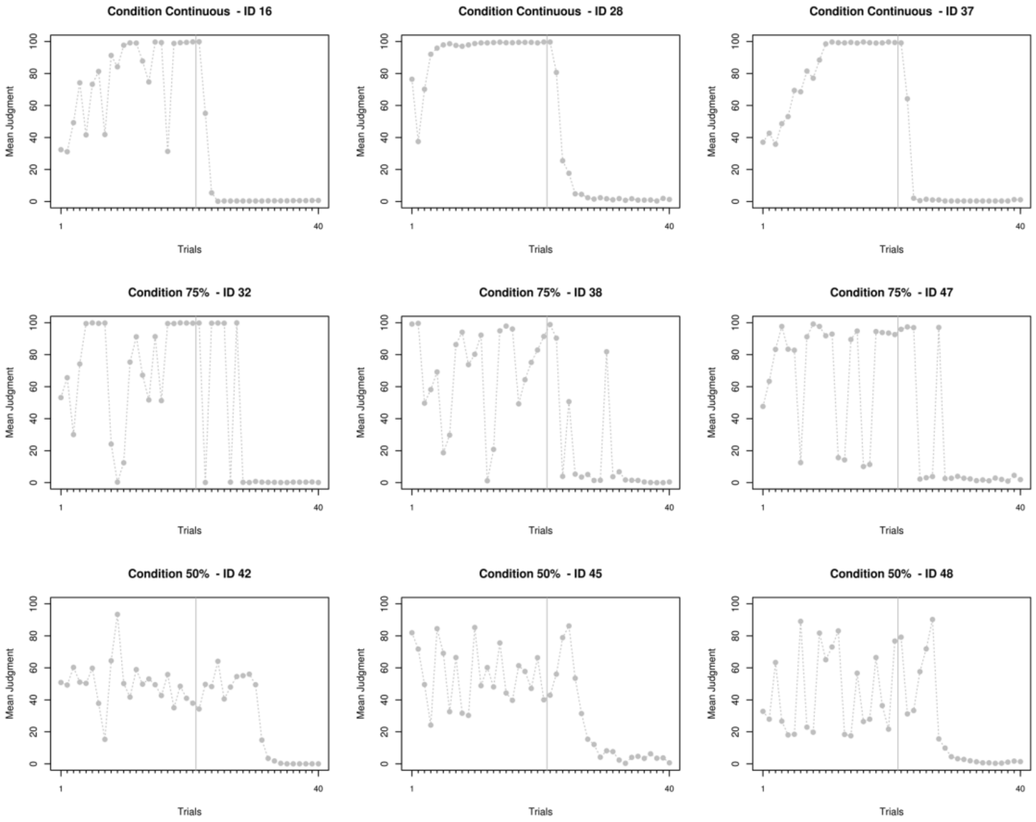
Condition Continuous - ID 16
Condition Continuous - ID 28
Condition Continuous - ID 37
Condition 75% - ID 32
Condition 75% - ID 38
Condition 75% - ID 47
Condition 50% - ID 42
Condition 50% - ID 45
Condition 50% - ID 48
Mean Judgment
0
20
40
60
80
100
Trials
1
40

Figure 8.

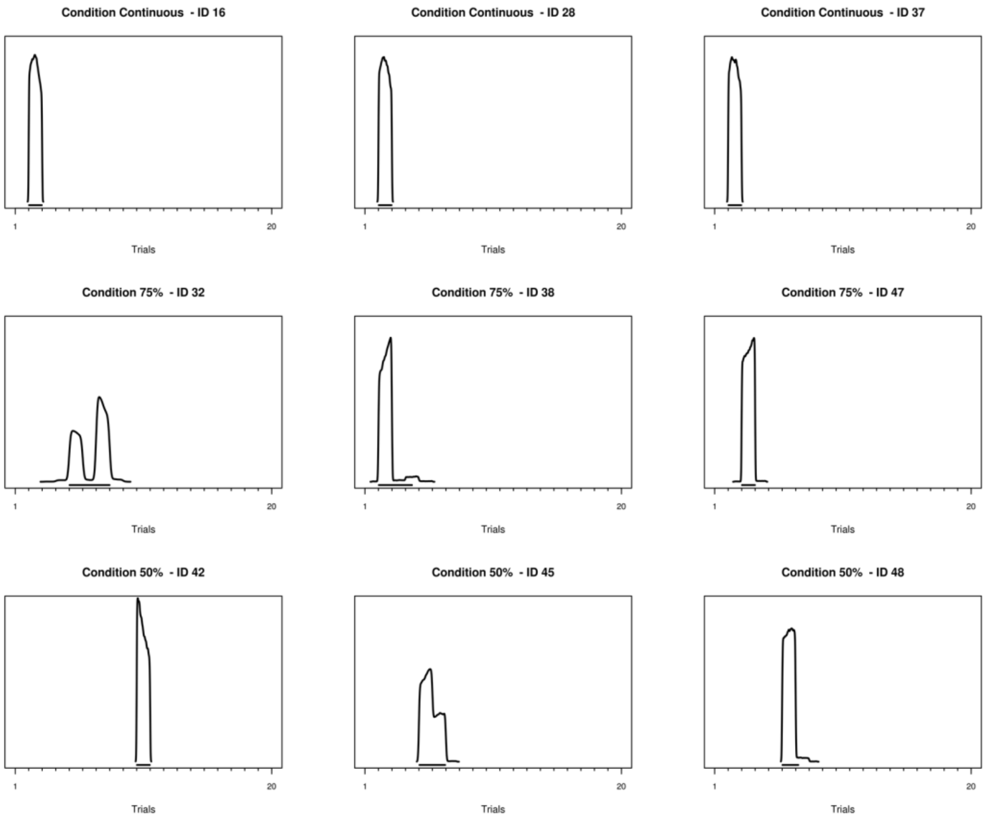
Condition Continuous - ID 16
Condition Continuous - ID 28
Condition Continuous - ID 37
Condition 75% - ID 32
Condition 75% - ID 38
Condition 75% - ID 47
Condition 50% - ID 42
Condition 50% - ID 45
Condition 50% - ID 48
1
20
Trials

Figure 9.