

Bayesian Analysis of Interleaved Learning and Response Bias in Behavioral Experiments

Anne C. Smith, Sylvia Wirth, Wendy A. Suzuki and Emery N. Brown

J Neurophysiol 97:2516-2524, 2007. First published Dec 20, 2006; doi:10.1152/jn.00946.2006

You might find this additional information useful...

This article cites 28 articles, 8 of which you can access free at:

<http://jn.physiology.org/cgi/content/full/97/3/2516#BIBL>

Updated information and services including high-resolution figures, can be found at:

<http://jn.physiology.org/cgi/content/full/97/3/2516>

Additional material and information about *Journal of Neurophysiology* can be found at:

<http://www.the-aps.org/publications/jn>

This information is current as of April 30, 2007 .

Bayesian Analysis of Interleaved Learning and Response Bias in Behavioral Experiments

Anne C. Smith,¹ Sylvia Wirth,² Wendy A. Suzuki,³ and Emery N. Brown^{4,5}

¹Department of Anesthesiology and Pain Medicine, University of California, Davis, California; ²Institut des Sciences Cognitives, Bron, France; ³Center for Neural Science, New York University, New York, New York; ⁴Neuroscience Statistics Research Laboratory, Department of Anesthesia and Critical Care, Massachusetts General Hospital, Boston; and ⁵Department of Brain and Cognitive Sciences, Harvard–MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts

Submitted 6 September 2006; accepted in final form 16 December 2006

Smith AC, Wirth S, Suzuki WA, Brown EN. Bayesian analysis of interleaved learning and response bias in behavioral experiments. *J Neurophysiol* 97: 2516–2524, 2007. First published December 20, 2006; doi:10.1152/jn.00946.2006. Accurate characterizations of behavior during learning experiments are essential for understanding the neural bases of learning. Whereas learning experiments often give subjects multiple tasks to learn simultaneously, most analyze subject performance separately on each individual task. This analysis strategy ignores the true interleaved presentation order of the tasks and cannot distinguish learning behavior from response preferences that may represent a subject's biases or strategies. We present a Bayesian analysis of a state-space model for characterizing simultaneous learning of multiple tasks and for assessing behavioral biases in learning experiments with interleaved task presentations. Under the Bayesian analysis the posterior probability densities of the model parameters and the learning state are computed using Monte Carlo Markov Chain methods. Measures of learning, including the learning curve, the ideal observer curve, and the learning trial translate directly from our previous likelihood-based state-space model analyses. We compare the Bayesian and current likelihood-based approaches in the analysis of a simulated conditioned T-maze task and of an actual object–place association task. Modeling the interleaved learning feature of the experiments along with the animal's response sequences allows us to disambiguate actual learning from response biases. The implementation of the Bayesian analysis using the WinBUGS software provides an efficient way to test different models without developing a new algorithm for each model. The new state-space model and the Bayesian estimation procedure suggest an improved, computationally efficient approach for accurately characterizing learning in behavioral experiments.

INTRODUCTION

Accurate characterizations of behavior in learning experiments are essential for understanding how we acquire and retain new information. In typical behavioral learning experiments subjects are presented with two or more tasks to solve simultaneously. The level of difficulty of the experiment can be controlled by the number of task presented. A common paradigm is to present the tasks to the subject by interleaving them in random order (Jog et al. 1999; Law et al. 2005; Paton et al. 2006; Williams and Eskandar 2006; Wirth et al. 2003). The most frequently recorded behavioral data are the trial-by-trial sequences of correct and incorrect responses. Whereas learning experiments often give a sub-

ject multiple tasks to learn simultaneously, analyses of learning behavior often characterize subject performance on each individual task separately. This analysis strategy ignores the interleaved presentation order of the tasks and makes it difficult to distinguish performance changes ascribed to learning from performance changes that may be associated with a bias or a strategy the subject has adopted.

A wide range of data analysis methods have been applied to determine when learning occurs for a single task. Such methods include the consecutive correct response criterion (Stefani et al. 2006), the change-point test (Gallistel et al. 2004; Paton et al. 2006), and stochastic models applied to both binary data (Smith et al. 2004, 2005; Wirth et al. 2003) and to reaction time data (Dayan et al. 2000; Smith 1995; Yu and Dayan 2003). Although complex stochastic models of learning multiple tasks have been proposed (Bussemeyer and Townsend 1993; Ditterich 2006; Estes 1978; Luce et al. 1965; Ratcliff and Rounder 2000; Suppes 1959, 1990; Usher and McClelland 2001; Verguts et al. 2002; Verhelst and Glas 1995), these models are not used routinely by experimentalists in the analysis of binary response data and are not capable of handling specific response biases. There is new interest in stochastic models for data analysis because of a need to relate behavioral measures of learning to changes in neural activity (Gallistel et al. 2004; Paton et al. 2006; Suzuki and Brown 2005; Wirth et al. 2003; Wolbers and Büchel 2005; Yoshida and Ishii 2006). Of the stochastic models being considered in current behavioral analyses, the flexibility of state-space models makes them well suited for characterizing interleaved learning experiments and correcting for response biases.

By extending in two ways current likelihood-based state-space models of learning (Smith et al. 2004, 2005), we present an approach to analyzing a learning experiment in which the tasks presented are interleaved and the subject may have a behavioral bias. First, we augment the univariate state-space model for learning a single task to a multivariate state-space model that represents the cognitive states of the multiple tasks and the cognitive state of the subject's bias. Second, we introduce a Bayesian approach using Monte Carlo Markov Chain methods for estimating the model parameters and the unobserved cognitive states. We illustrate our method in the analysis of a simulated experiment of a rat executing an alternating T-maze task with an initial left-turn bias and in the

Address for reprint requests and other correspondence: A. C. Smith, Department of Anesthesiology and Pain Medicine, TB-170, One Shields Ave., University of California, Davis, CA 95616 (E-mail: annesmith@ucdavis.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

analysis of an actual learning experiment in which a monkey executes an object–place association task (Wirth et al. 2005).

METHODS

A state-space model for interleaved learning and bias

We assume that the learning experiment can be modeled using a state-space framework (Durbin and Koopman 2001; Kitagawa and Gersh 1996; Smith and Brown 2003; Smith et al. 2004, 2005). The state-space model consists of two equations: a state equation and an observation equation. We define a state equation that allows us to disambiguate the subject’s cognitive state regarding each task being learned from his/her possible response bias. Therefore in this analysis, the state equation will define the temporal evolution of the cognitive state of each task the subject is learning and the temporal evolution of the subject’s response bias.

The observation equation defines how the observed data relate to the unobservable cognitive state process for each task and the cognitive state process for the subject’s response bias. The data we observe in the interleaved learning experiment are the series of correct and incorrect responses as a function of trial number for each of the tasks the subject is learning. In addition, we observe the sequence of specific responses on each trial. Used together in the state-space analysis, the series of correct and incorrect responses and the series responses can be used to distinguish learning of each task from a response bias.

In this analysis, the learning state for each task will be defined as the cognitive state corrected for the subject’s bias. As in our previous learning analyses (Smith et al. 2004, 2005; Wirth et al. 2003), we compute from the learning state process the learning curve that defines the probability of a correct response as a function of trial number. We define the *learning curve* as a function of the learning state process so that an increase in the learning state process increases the probability of a correct response and a decrease in the learning process decreases the probability of a correct response.

For clarity, we present the state-space model in the context of a simple conditioned T-maze experiment (Barnes et al. 2006; Jog et al. 1999). In this experiment, a rat is placed on the longest or start arm of a T-shaped maze apparatus and is trained to associate an auditory cue (i.e., either a high or low tone) with entering the left or right arm for a food reward. The response data constitute whether the animal makes a correct turn at a given trial. In this experiment the number of possible tasks (associations) to be learned is two—that is, high tone associated with a left turn and low tone associated with a right turn. In a noninterleaved analysis of this experiment the responses would be divided into two separate binary series corresponding to the initial tone presentation and each series would be analyzed separately. For our interleaved analysis, we make use of the additional information of which direction the animal actually turned on a given trial. In this example, we assume these are also binary data such that a one indicates the animal turned left and a zero indicates the animal turned right. If the presentation order of the two tasks is pseudorandom, the cognitive state relating to bias will be near zero both when the animal responds correctly and when the animal responds randomly. When the animal exhibits a left (right) response bias, this state will be above (below) zero and can be used to modify the assessment of learning estimated from the binary incorrect/correct responses alone.

To define the observation model for an interleaved learning experiment, we assume that J tasks (associations) are presented over K trials. Let $n_{k,j}$ be 1 if the response on trial k is correct for task j and 0 otherwise, where $j = 1, \dots, J$ and $k = 1, \dots, K$. Let $n_{k,J+1}$ be a 1 if the animal turns left on trial k and 0 if it turns right. Let $n_k = \{I_{k,1}n_{k,1}, \dots, I_{k,J}n_{k,J}, n_{k,J+1}\}$ be the responses observed on trial k , where $I_{k,j}$ is the indicator function that is 1 if task j is presented at trial k and 0 otherwise. We let $N = \{n_1, \dots, n_K\}$ be the observed responses from all K trials. We define $p_{k,j}$ as the probability of a

correct response on trial k to task j , $p_{k,J+1}$ as the probability that the animal chooses to turn left on trial k and we define $p_k = (p_{k,1}, \dots, p_{k,J+1})$. It follows that the observation model for trial k is

$$\Pr(n_k|p_k) = (p_{k,J+1})^{n_{k,J+1}}(1 - p_{k,J+1})^{1-n_{k,J+1}} \times \prod_{j=1}^J I_{k,j} [(p_{k,j})^{n_{k,j}}(1 - p_{k,j})^{1-n_{k,j}}] \quad (1)$$

To relate performance on trial k to performance on prior and subsequent trials, we define a two-component state-space model—one component describes the propensity of the animal to give a correct response and the second component describes the propensity of the animal to make a left turn as its response. Let $x_{k,j}$ be the subject’s cognitive state about task j on trial k . We assume that the cognitive state on trial k for task j is related to the cognitive state at trial k by the Gaussian random-walk state-space model

$$x_{k,j} = x_{k-1,j} + \varepsilon_{k,j} \quad (2)$$

where $\varepsilon_{k,j}$ is Gaussian error with zero mean and variance σ_j^2 for $j = 1, \dots, J$. Let $x_{k,J+1}$ be the subject’s cognitive state about choosing left on trial k , which is related to the subject’s cognitive state about choosing left on trial $k - 1$ by the Gaussian random-walk state-space model

$$x_{k,J+1} = x_{k-1,J+1} + \varepsilon_{k,J+1} \quad (3)$$

where $\varepsilon_{k,J+1}$ is Gaussian error with zero mean and variance σ_{J+1}^2 . If we let $x_k = (x_{k,1}, x_{k,2}, \dots, x_{k,J+1})$ and $\varepsilon_k = (\varepsilon_{k,1}, \varepsilon_{k,2}, \dots, \varepsilon_{k,J+1})$ then we can express the two components of the state-space model given in Eqs. 2 and 3 as the vector equation

$$x_k = x_{k-1} + \varepsilon_k \quad (4)$$

We take $x = (x_1, \dots, x_K)$ to be the vector of cognitive states across the entire experiment.

To relate the cognitive state model in Eq. 4 to the observation model in Eq. 1, we define $p_{k,j}$ in terms of values of $x_{k,j}$ as

$$p_{k,j} = [1 + \exp(x_{k,j})]^{-1} \exp(x_{k,j}) \quad (5)$$

for $j = 1, \dots, J + 1$. Expressing $p_{k,j}$ as a logistic function of $x_{k,j}$ ensures that these probabilities are constrained to lie between zero and one. As $x_{k,j}$ increases (decreases) to positive (negative) infinity $p_{k,j}$ increases (decreases) to 1 (0). We note that if $x_{k,J+1} = 0$ then $p_{k,J+1} = 0.5$ and the animal is equally likely to choose left or right. In this case, there is no bias.

To determine the subject’s cognitive state regarding learning, we must disambiguate the propensity to respond correct from the propensity to respond in a biased manner. We accomplish this separation by using the state-space model components and assuming that directional bias has an additive effect on the cognitive state. Thus we define the learning state as

$$z_{k,j} = x_{k,j} \pm x_{k,J+1} \quad (6)$$

where the sign in front of $x_{k,J+1}$ is positive (negative) for the low (high) tone–right (left) turn reward trial.

A Bayesian analysis of the learning state-space model

We can express the unknown parameters in this model as $\theta = (x_0, \sigma_1^2, \dots, \sigma_{J+1}^2)$, where $x_0 = (x_{0,1}, \dots, x_{0,J+1})$ is the cognitive state of the animal about the J tasks and turn propensity at the outset of the task. In our previous state-space models of learning we used the Expectation–Maximization algorithm to compute maximum-likelihood estimates of θ and the unobserved cognitive or learning state process x (Smith et al. 2004, 2005; Wirth et al. 2003). Although a similar approach would be possible here, we introduce instead a Bayesian approach to computing θ and x . The goal of the Bayesian

analysis is to compute the posterior probability density of θ and x , defined from Bayes' rule as

$$p(\theta, x|N) = \frac{p(\theta)p(x|\theta)p(N|x, \theta)}{p(N)} \quad (7)$$

where $p(\theta)$ is a prior probability density for θ and $p(x|\theta)$ is the joint probability density of the cognitive state process defined by Eq. 4 as follows

$$p(x|\theta) = \prod_{k=1}^K p(x_k|x_{k-1}, \theta) = (2\pi|\Sigma|)^{-K/2} \times \exp\left\{-\frac{1}{2}\sum_{k=1}^K (x_k - x_{k-1})\Sigma^{-1}(x_k - x_{k-1})\right\} \quad (8)$$

where Σ is a $(J + 1) \times (J + 1)$ diagonal matrix with the j th diagonal element σ_j^2 for $j = 1, \dots, J + 1$ and $p(N|x, \theta)$ is the joint probability density or likelihood of the data defined from Eq. 1 as

$$p(N|x, \theta) = \prod_{k=1}^K \Pr(n_k|p_k) = \prod_{k=1}^K (p_{k,J+1})^{n_{k,J+1}}(1 - p_{k,J+1})^{1-n_{k,J+1}} \times \sum_{j=1}^J I_{k,j}[(p_{k,j})^{n_{k,j}}(1 - p_{k,j})^{1-n_{k,j}}] \quad (9)$$

The prior probability density $p(\theta)$ is defined as

$$p(\theta) = p(x_0) \prod_{j=1}^{J+1} p(\theta_j) \quad (10)$$

where $p(x_0)$ is a uniform probability density on the interval $[-a, a]$ and $p(\theta_j) = p(\sigma_j^{-2})$ is a gamma probability density with parameters α and β for $j = 1, \dots, J + 1$.

For inference purposes, we compute the marginal posterior probability density of each component of θ , defined from

$$p(\theta_j|N) = \int p(\theta_{[j]}, x|N) d\theta_{[j]} dx \quad (11)$$

where $\theta_{[j]}$ denotes the elements of θ excluding θ_j . We compute Eqs. 7 and 11 using Monte Carlo Markov Chain (MCMC) methods (Congdon 2003; Gilks et al. 1996). In Bayesian analyses, MCMC methods are widely used Monte Carlo techniques for evaluating joint and marginal posterior probability densities by simulating stationary Markov chains. Because the Bayesian analysis provides an approximate posterior probability density for each parameter θ_j in the form of a set of Monte Carlo samples, we can use any summary statistic of the set of Monte Carlo samples, such as the mean or median, as the Bayes' estimate of the parameter. Similarly, 100%(1 - α) confidence (Bayesian credibility) intervals can be computed directly by taking the $\alpha/2$ and the $1 - (\alpha/2)$ quantiles of the Monte Carlo sample probability density.

We conduct the MCMC computations using the software WinBUGS (Lunn et al. 2000; Spiegelhalter et al. 2004). Given specifications of the prior and joint probability density of the data or likelihood models, WinBUGS chooses a Monte Carlo scheme to simulate the desired posterior probability densities. It is possible for the user to select the Monte Carlo scheme. In our simulations we use the default schemes chosen by WinBUGS. For the analyses we present here, we provide the WinBUGS code and interface to run it using Matbugs (Murphy and Mahdavian 2005) from Matlab (The MathWorks, Natick, MA) at our website <http://www.ucdmc.ucdavis.edu/anesthesiology/research/asmith.html>.

We assessed convergence of our MCMC simulation by first analyzing graphically the stationarity and mixing of three Monte Carlo

chains. Second, we tracked the Brooks–Gelman–Rubin statistic, which compares between- and within-chain variance (Brooks and Gelman 1998; Gelman and Rubin 1992), and required that it be <1.2 for all parameters (Kass et al. 1998). For the tasks we consider in RESULTS, $<30,000$ Monte Carlo iterations per chain (including 1,000 burn-in iterations) were needed to achieve convergence in <5 min of CPU time on a Pentium IV desktop computer.

Specification of initial conditions in interleaved learning experiments

In experiments in which the subject is believed to start with an initial response bias, we estimate the initial probability of a correct response under the Bayesian formulation by assigning an uninformative prior to the mean of each initial state $x_{0,j}$ for all tasks, $j = 1, \dots, J$. A second approach, which we use in our full Bayesian-interleaved analyses, is to use knowledge of the structure of the experiment. This is particularly useful in binary response experiments in which a correct response for one task corresponds to an incorrect response for a second task. For example, in the T-maze task, if the animal has an initial left-turn tendency, then high-tone associations will appear all correct and low-tone associations will appear all incorrect. In this case, we assume at trial zero that the probability of a correct response to the high tone and the probability of a correct response to a low tone sum to one. In the state-space domain on $[-\infty, \infty]$, this means that the sign of the cognitive state for the high tone is opposite in sign to the sign of the cognitive state for the low-tone association at trial zero.

Analysis of learning

The learning curve is the estimate of the probability of a correct response as a function of trial number. We report three estimates of the learning curve. For each task (association) j the first learning curve is computed without bias correction from the Bayesian analysis using Eq. 5, defined as

$$\hat{p}_{k,j}^B = [1 + \exp(\hat{x}_{k,j})]^{-1} \exp(\hat{x}_{k,j}) \quad (12)$$

for tasks $j = 1, \dots, J$ where the circumflex accent (hat) denotes the estimate. The second learning curve estimate is computed with the bias correction from the Bayesian analysis by evaluating the estimates computed in the Bayesian analysis in Eq. 6, defined as

$$\hat{p}_{k,j}^{BI} = [1 + \exp(\hat{z}_{k,j})]^{-1} \exp(\hat{z}_{k,j}) \quad (13)$$

The third learning curve estimate is the maximum-likelihood estimate described previously in Smith et al. (2004), which does not account for either the interleaved nature of the learning or the response bias, and is defined as

$$\hat{p}_{k,j}^{EB} = [1 + \exp(\hat{x}_{k,j}^{EB})]^{-1} \exp(\hat{x}_{k,j}^{EB}) \quad (14)$$

As in our previous analyses (Smith et al. 2004, 2005), we define the learning trial for each estimation procedure in terms of the ideal observer (IO). We chose a level of certainty of 0.95 and defined the ideal observer learning trial with a level of certainty 0.95 [IO(0.95)] as the earliest trial r , such that the probability of a correct response is >0.95 for the all trials $k \geq r$.

Experimental protocol: object–place association task

As a second more complex example, we also consider data from an actual experiment in which a monkey was trained to associate four different object–place combinations viewed on a computer screen with either a late or early bar release response (Fig. 1; object–place associative learning task; Wirth et al. 2005). In this task, the animal initiated each trial by fixating on a central plus shape on a computer monitor. One of two possible visual objects was then shown in one of two possible places on the monitor for 500 ms. Each day, two novel

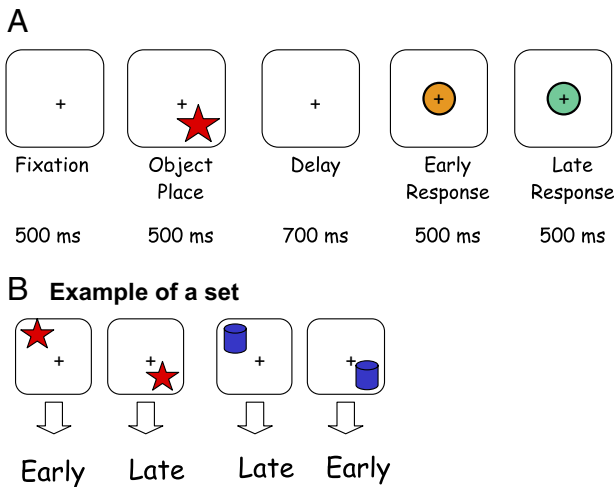


FIG. 1. A: schematic of the object–place association task described in the text. B: example set of 4 possible associations learned in one experiment.

objects and two distinct spatial locations on the computer monitor were used. After a delay interval of 700 ms, an orange circle was shown for 500 ms followed immediately by a green circle for another 500 ms. Each object–place combination was associated with either an early bar release during the orange circle (early release) or a late bar release during the green circle (late release). An example learning set is shown in Fig. 1B. A correct early or late bar release response resulted in a liquid reward. Previous analysis showed that monkeys commonly exhibit early/late response biases on this task (Wirth et al. 2005).

RESULTS

Analysis of a single learning task using empirical Bayes and full Bayesian approaches

We first compared the learning curves estimated by the full Bayesian (FB) MCMC implementation for a single learning task with our previously described likelihood-based, empirical Bayes (EB) approach (Smith et al. 2004). As an example sequence, we simulated a 30-trial sequence of correct and incorrect responses that represent, say, the responses to a low-tone–right-turn association in the T-maze task described in METHODS. The correct/incorrect responses are shown as black/gray squares above Fig. 2, A and B. The data suggest that the animal may have a bias at the start of the experiment because there are initially 10 consecutive incorrect responses. After trial 20, the task appears to be learned because there are 10 consecutive correct responses.

For both the EB and FB approaches we assume the unobserved cognitive state process follows the random walk given by $x_{k,1} = x_{k-1,1} + \epsilon_{k,1}$ for $k = 1, \dots, K$, where $\epsilon_{k,1} \sim N(0, \sigma_1^{-2})$ with $x_{0,1} = 0$ (EB approach) and $x_{0,1} \sim N(0, \sigma_1^{-2})$ (FB approach). Fixing the initial mean of $x_{0,1}$ at zero, we implicitly assume the probability of a correct response at the time step before the first observation is chance at 0.5. For the EB approach, we use the EM algorithm to estimate unknown variance parameter (σ_1^{-2}) and the cognitive state process. For the FB approach, we use MCMC with gamma priors for σ_1^{-2} to ensure that the variance values are always positive. The learning curve is computed from the state estimates using Eq. 12.

The EB approach learning curve (Fig. 2A, median and 90% confidence bounds) starts with a probability close to 0.2 at trial

1, declines, shows a slight increase from trials 9 to 11, and then monotonically increases from trial 14 onward. The IO(0.95) learning trial from this analysis is trial 22. The FB learning curve shows a similar structure (Fig. 2B, green dotted and red solid curves with corresponding 90% confidence bounds). We show FB learning curves estimated with two different choices of a gamma prior, with parameters (5, 5) and (10, 10). Both of these priors have a mean of 1 with respective variances of 0.2 and 0.1. In this analysis, the confidence bounds are slightly narrower, resulting in IO(0.95) learning trial estimate of 21, one trial earlier than the EB learning trial estimate.

This analysis shows that for learning curves estimated for a single task, the EB and FB approaches give similar solutions. The discrepancy between estimates of the confidence bounds results from slight differences in model specification and estimation.

Analysis of simulated interleaved learning: a conditioned T-maze task

As our first illustration of the FB analysis applied to an experiment in which tasks are presented in an interleaved manner, we simulated binary data of a rat performing the conditioned T-maze task described in METHODS. We assume the animal starts the 60-trial experiment with a left-turn bias (Fig. 3, A and B, top blue/red arrowheads indicate left/right turns, lower black/gray squares indicate correct/incorrect responses, respectively). We constructed the data such that the animal initially followed the strategy of turning left for the first 20 trials, chose randomly for the next 20 trials, and then performed correctly for both associations for the remaining 20 trials. For simplicity in simulating these data, we assumed the high-tone–left-turn and low-tone–right-turn associations were

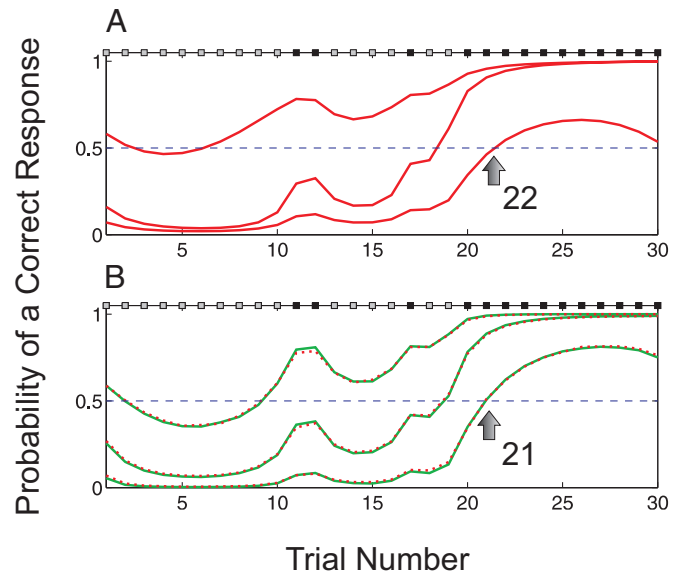


FIG. 2. A: empirical Bayes (EB) estimate of the learning curve (median and 90% confidence bounds) for a sequence of binary data corresponding to the responses to the low-tone–right-turn association in a simulated T-maze task. Correct/incorrect responses are shown above each panel as black/gray squares, respectively. B: full Bayesian (FB) estimates of learning curves and 90% confidence bounds for the same data set as in A. Choice of prior on the precision of the random walk can result in slightly different learning curves. Priors used, shown in B, result in similar learning curves and are gamma with parameters (5, 5) (green dashed curve) and (10, 10) (red solid curve).

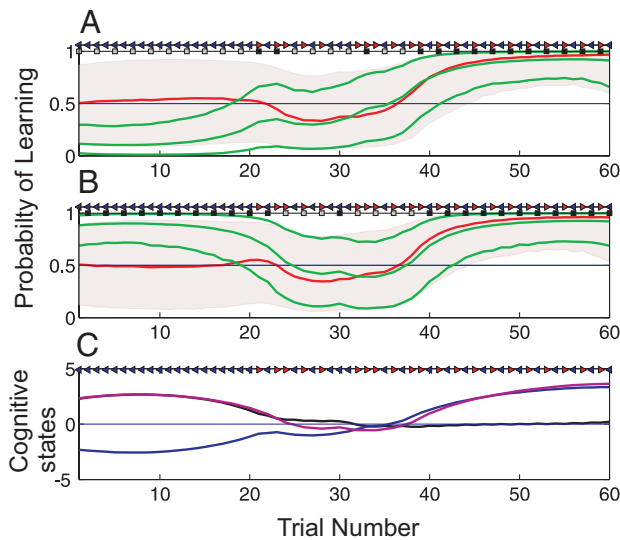


FIG. 3. Learning curves computed for simulated T-maze experiment for the low-tone-right-turn trials (A) and high-tone-left-turn trials (B). Correct/incorrect responses are shown as black/gray squares above each panel. Blue/red arrowheads above each panel indicate the left/right turns, respectively. Animal has a significant left-turn bias at the start as indicated by the run of 20 blue arrowheads at the start of the experiment. Median learning curves computed without taking into account bias (FB approach) are shown in green. Median learning curves and 90% confidence limits from the FB-interleaved approach are shown in red. Estimated cognitive processes ($x_{k,1}$, $x_{k,2}$, and $x_{k,3}$ for $k = 1, \dots, K$) are shown in C (blue, purple, and black curves, respectively).

tested on alternating trials. This is not necessary as long as the presentation order is pseudorandom with equal probability for both auditory cues. Therefore our data consisted of 60 responses for the bias estimation and 30 responses for each high- and low-tone association.

A FB learning curve analysis computed for the low-tone-right-turn portion of the task without taking into account behavioral bias (Fig. 3A, green curves) indicates that performance is below chance (a probability of 0.5 for this task) at the start and rises above chance in the second half of the experiment. The learning curve for the high-tone-left-turn association (Fig. 3B, green curve) starts close to 1, drops below chance, and then rises back up to 1 by the end of the experiment. The time course of the cognitive process corresponding to each of these learning curves closely mirrors the time course of its corresponding learning curve (Fig. 3C, blue and purple curves for low- and high-tone responses, respectively).

We now consider the cognitive state for the response bias (Fig. 3C, black curve). Because these data contain 20 consecutive ones at the start of the experiment, the cognitive process for the response bias is initially positive and does not decline to zero until the response behavior becomes more variable after trial 20. To correctly identify the learning behavior, we follow Eq. 6 and add the cognitive state process for the bias to the cognitive state process for low-tone responses and subtract it from the cognitive state process for the high-tone responses. After correcting for the response bias, the estimates of learning curves for low- and high-tone trials (Fig. 3, A and B, red curves and red-shaded 90% confidence bounds) are similar. With the bias correction, both learning curves are close to chance for the first 20 trials, fall below chance for trials 22–35, and increase almost monotonically from trial 36 to the end of the experiment.

For this particular example by including the cognitive state related to response bias, the position of the IO(0.95) learning trial changed by only one trial for each association. However, the shape and width of the learning curves distributions did change. This new analysis alters our interpretation of the state of learning over the initial third of the experiment. First, if we consider the low-tone-right-turn trials (green curves, Fig. 2A), our initial analysis would have indicated a run of nine trials at the start of the task where the animal was performing significantly below chance, possibly leading to the conclusion that the animal knew the association but was deliberately avoiding a reward. The addition of a term representing the cognitive bias state critically increased the width of the learning curve confidence bounds at the start, making this conclusion less credible. Second, for the high-tone-left-turn trials if we ignore turn bias (Fig. 2B, green curves), the learning curve is U-shaped and the animal appears to have learned, then forgotten, and then learned again. The addition of a bias correction lowers the learning curve and widens the confidence bounds in the initial 20 trials. Although it is impossible to be certain that the animal did not learn the high-tone-left-turn association at the start and then forget it, the lack of variability in its responses suggests that it is highly plausible to subtract out the “perseverative” behavior in the first 20 trials.

Analysis of actual interleaved task learning: object-place association task

To illustrate the FB analysis applied to an actual interleaved learning experiment, we consider data from the object-place association task described in METHODS (Fig. 1; Wirth et al. 2005). In this experiment, the animal was presented with four object-place associations over 157 total trials. Associations 1 through 4, also known as *conditions*, were presented for 41, 41, 35, and 40 trials, respectively, and their correct/incorrect responses are shown as black/gray squares above the panels in Fig. 4. The correct response for conditions 1 and 3 (Fig. 4, A

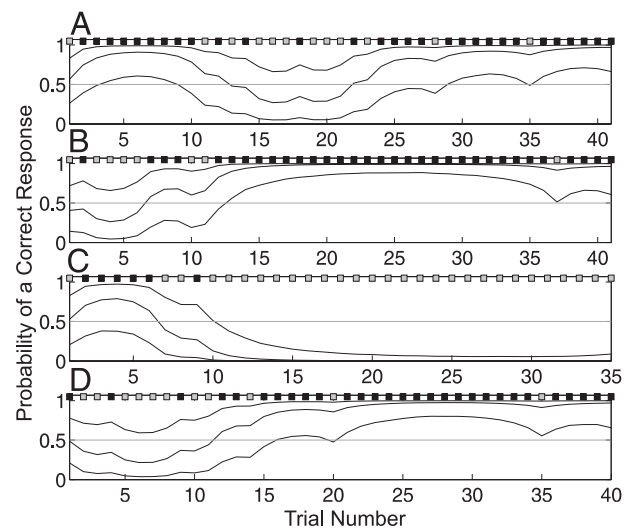


FIG. 4. Learning curves and 90% confidence bounds computed using the FB approach for real data from 4 interleaved associations in an object-place association task. Correct/incorrect responses for conditions 1–4 are indicated as black/gray squares above A–D, respectively. In this example, the data are fitted assuming each association is learned in isolation and assuming that the probability of a correct response at the start of the experiment is chance (0.5).

and C) was an early bar release, whereas the correct response for conditions 2 and 4 (Fig. 4, B and D) was a late bar release. Figure 4, A–D (black curves) illustrates the FB learning curves and the 90% confidence bounds for the set of four object–place associations analyzed as if each task were learned separately. We conclude that conditions 1, 2, and 4 are all learned during the experiment with IO(0.95) learning trials of 36, 13, and 21, respectively.

Figure 5, A–D shows the FB learning curves for the four object–place associations in their true presentation order. The top row of colored squares shown in Fig. 5 displays the early (blue squares) and late (red squares) releases. The second row of colored squares are the correct (black squares) or incorrect (gray squares) responses for these conditions. These response data are the same response data as shown in Fig. 4. The release data suggest that the animal may have an early release bias at the outset of the experiment and a late release bias at the end of the experiment.

If the animal has an early release bias at the start of the experiment, the FB model should lower the learning curve for associations with early release reward until it is clear that the animal's responses to all four associations vary from trial to

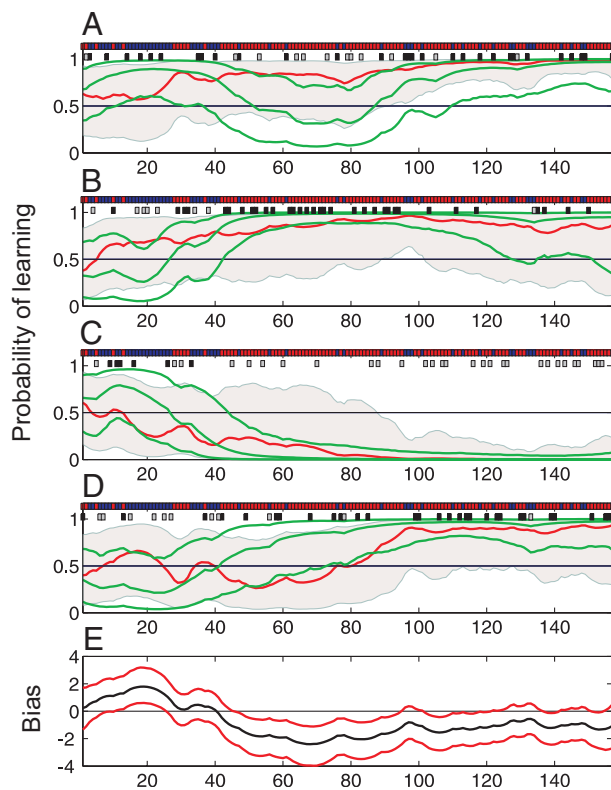


FIG. 5. Learning curves computed for real data from 4 conditions in an object–place association task (A–D, respectively). Correct/incorrect response data are shown as black/gray squares above A–D, respectively. Late/early responses at each trial are also shown as red/blue squares, respectively, above each panel. Conditions 1 and 3 (A and C) were rewarded for an early bar release. Conditions 2 and 4 (B and D) were rewarded for a late bar release. Median learning curves and 90% confidence bounds computed using the FB approach are shown in green. Median learning curves with 90% confidence bounds computed using the FB-interleaved approach are shown in red. E: estimated cognitive process for behavioral bias with 90% confidence bounds. Large number of early responses in the data at the start causes this curve to be significantly above zero early in the experiment. From roughly trial 45 onward the responses are biased toward late (negative values).

trial. Once the release responses show variability we can be more certain that, either the animal has no bias or, assuming the presentation is pseudorandom, it is responding correctly to the presented associations. We require the magnitude of the bias term to be high when a larger number of similar release responses are made and low when responses are switching between late and early in an interleaved manner.

To apply the interleaved state-space model with bias to this task we take $J + 1 = 5$ and assume that there are four cognitive states and a fifth cognitive state representing the bias. Each of the four cognitive processes for one of the four association tasks is only partially observed because a different task is given at each trial. As in the simulated example, we used Eq. 6 to compute the bias-corrected learning curves where the sign in front of $x_{k,5}$ is negative for early-release associations (conditions 1 and 3, Fig. 5, A and C) and positive for late-release associations (conditions 2 and 4, Fig. 5, B and D).

As in the previous example, we first plot the learning curves computed without explicitly taking into account possible response bias (FB approach, Fig. 5, A–D, green curves are median and 90% confidence limits). These are the learning curves computed solely from the cognitive state processes $x_{k,j}$, without considering either the interleaved structure in the experiment or the possible response bias. The performance for conditions 1 and 4 is above chance, i.e., the lower 90% confidence bound is >0.5 , and remains >0.5 , respectively, from trials 112 and 69 until the end of the experiment. The performance on condition 2 (Fig. 5B) surpasses chance at trial 42, but falls below chance at the end of the experiment and would therefore be designated not learned. Performance on condition 3 shows little to no indication of learning because performance is below chance from trial 44 onward.

Figure 5E (black curve) shows the estimated cognitive state for the response bias. The binary response data (Fig. 5, A–D above) suggest a tendency for early release up to nearly trial 40 (multiple blue squares in the top row of Fig. 5A) and a tendency for late response bias after that (multiple red squares in the top row of Fig. 5A). This same pattern is reflected quantitatively in the estimated cognitive state for the response bias (Fig. 5E, black curve with red 90% confidence bounds). There is a clear early-response bias for the first part of the experiment and an overall late-response bias for the balance. Applying the FB-interleaved method with the estimated bias correction (Fig. 5A, red curve and shaded 90% confidence bounds) moves the learning trial for condition 1 (early-reward) from 112 to 93. It has the effect of lowering the learning curve at the start and raising the learning curve at the end of the experiment. For the other early-reward condition (3), the point at which the learning curve is below chance moves from trial 44 to trial 88 (Fig. 5C) because of the additional uncertainty introduced by including the bias correction. For the late-release conditions (2 and 4, Fig. 5, B and D), the late-release bias at the end of the experiment has the effect of lowering the learning curves. This effect is particularly noticeable for condition 4, which is not learned according to the FB-interleaved method, but which is learned at trial 69 with the FB approach.

Consideration of the true presentation order and the possible response bias in our model has reduced the number of associations estimated from three (FB approach applied to binary series from each task separately) to two (FB approach) to 1 (FB-interleaved). The difference between the isolated FB anal-

yses (Fig. 4) and the FB approach (Fig. 5, green curves) is first in the inclusion of true presentation order resulting in gaps between observations and second in the specification of the initial conditions. For the FB approach applied separately (Fig. 4), we assumed the starting probability was at chance and equal to 0.5. For the FB approach in Fig. 5 we estimated the initial conditions from the data assuming the initial distributions of the probability of a correct response for late-release conditions and the probability of a correct response for early-release associations summed to one. Finally, inclusion of the tendency to keep making late releases in the FB-interleaved approach had the effect of lowering the late-release association learning curves (associations 2 and 4). That is, because the subject tended to make late releases across all tasks more often than chance, the model indicated that the experimenter should be less certain that the association was truly learned.

DISCUSSION

We have presented a state-space model for analyzing learning experiments consisting of binary time series in which two or more tasks are presented in an interleaved manner and the subject may have a response bias. This research builds on our previous state-space framework for modeling learning from binary measurements in behavioral experiments. In simulated and actual data analyses we demonstrated the ability of our methods to disambiguate bias from actual learning. We introduced a Bayesian approach for model estimation and showed that all our previous definitions of learning criteria translate directly into the Bayesian framework. For the interleaved association task, we demonstrated that the monkey had an early-release bias at the start and a late-release bias at the end of the experiment. This finding altered our interpretation of this experiment in that when the analyses of the individual response time series were analyzed separately we concluded that the animal learned three of the four conditions. However, by considering all of the tasks simultaneously and considering the animal's response bias we can be certain only that the animal learned one of the conditions.

State-space modeling of interleaved learning and bias

To construct a state-space model that allowed us to represent the cognitive state of each task the subject was learning along with the state of its response bias we augmented the state equation for the learning process to include a component for each cognitive state and a component for the response bias. This differed from our previous work in which each interleaved task was treated as if it was being learned in isolation and the model analyses were conducted separately. We used the augmented state-space model previously to compute simultaneously individual and population learning estimates (Smith et al. 2005). In this case, the learning curve for a given task depends only on the cognitive state variable for that process. In our new model, the learning state for a given task is defined as the difference or sum between the learning state for that task and the state of the subject's response bias (Eq. 6). The cognitive state of the subject's bias tracks whether the response behavior favors a particular response or occurs at random. To accurately characterize the subject's learning state we have to consider four cases. If the response behavior is random, then

the cognitive state process for the bias should be close to zero and have little effect on the learning state and thus on the estimate of the learning curve. If the response behavior were not random and biased toward a particular response then subtracting the bias state from the cognitive state of the particular task provides a more accurate characterization of the subject's learning state for that task. On the other hand, if the response behavior were not random and biased away from the reward or response the bias-corrected estimate of the learning state is in this case the cognitive state for the task plus the cognitive state for the bias. In the final case, the response behavior is all correct in which case, assuming the presentation order of the tasks is pseudorandom, the cognitive state process for the bias should again be close to zero and have little effect on the learning state. Taking these four possibilities into account, the bias-corrected learning curve for each task is defined as a function of the learning state from Eq. 13.

The observation component of our new state-space model places the response data in the proper temporal sequence in which they are observed and uses as a second observation process the subject's sequence of actual responses on each trial. This is different from previous state-space models of learning in which the response data for each task are analyzed separately and the response behavior of the subject is not considered.

Bayesian model fitting

In addition to introducing a more detailed model for learning, we have also introduced use of a Bayesian approach to model parameter estimation. The parameters in the new state-space model could have been estimated as in our previous work by maximum likelihood using the EM algorithm (Smith et al. 2004, 2005). Despite the similar structure between our previous and current state-space models of learning, an important drawback to this approach is that it requires the design of a new EM algorithm for each new model formulation. This makes it more challenging to provide broadly useful software that neuroscientists may use to analyze their behavioral data. In contrast, the Bayesian formulation of the task allows us to conduct the model fitting using Monte Carlo Markov Chain methods implemented in the WinBUGS (Lunn et al. 2000; Spiegelhalter et al. 2004) software package. An important advantage of WinBUGS is that it suffices to specify the state-space model and appropriate prior distributions for the parameters and WinBUGS will implement an efficient Monte Carlo procedure to simulate the exact posterior densities of the parameters. We found that for the analyses presented here simply using the default settings in WinBUGS and specifying prior distributions for the parameters as described in RESULTS yielded a robust approach to model parameter estimation. We found that currently accepted criteria for evaluating convergence of the Markov Chain worked well for deciding when the Monte Carlo procedures had accurately computed the posterior densities.

An important improvement of the Bayesian approach is that it provides estimates of the exact posterior densities for the state processes, whereas the EM algorithms we previously implemented provided Gaussian approximations to the state processes. As is standard, the trade-off between use of the likelihood-based approach and the Bayesian approach to estimate model parameters is the trade-off between specifying in

the Bayesian case a prior distribution for the model parameters and in the likelihood case specifying plausible starting values for the EM algorithm. We found that the insights we had gained in specifying starting values for the EM algorithm could be easily translated into plausible prior distributions for the MCMC algorithms.

Future directions

Several extensions of the state-space model analysis paradigm are possible. First, we can include nonbinary response data such as reaction and response times to provide a more refined analysis of a subject's performance. Second, we can include more complex behavioral response biases in behavioral experiments. For example, in the object-place task, the animal might have shown an object bias, responding only on trials in which one of the objects was presented, but not the other. Once identified, this kind of bias can be easily modeled using our state-space framework. Third, in the current state-space model we have assumed that the experiment is designed such that all tasks are presented pseudorandomly and with equal probability. The state-space model can be adjusted when the tasks are presented with unequal probabilities by including additional terms in the state and observation models.

Finally, this state-space model can also be extended to allow for other types of interaction among learning of tasks. Following Usher and McClelland (2001), we can rewrite Eq. 4 as follows

$$x_k = Ax_{k-1} + \varepsilon_k \quad (15)$$

where in the current analysis we have $A = I$. The off-diagonal elements of matrix A can then be used to assess the level of competition or enhancement of learning among the interleaved tasks.

For the applications we consider in which the data are relatively short sequences of binary responses (<100 trials per task) the large number of parameters in A (Eq. 15) makes simultaneous estimation of the model parameters and the cognitive state more challenging. This is a problem we are currently studying.

Our results suggest that modeling the interleaved structure in the learning experiment and making use of data on the subject's response behavior through a new state-space model coupled with an efficient MCMC procedure for model parameter estimation using WinBUGS provides both an accurate and practical approach to characterize learning in complex behavioral experiments.

GRANTS

This work was supported by National Institutes of Health Grants MH-071847 to E. N. Brown and A. C. Smith, DA-015644 to E. N. Brown and W. A. Suzuki, and MH-58847 to W. A. Suzuki; McKnight Foundation grant to W. A. Suzuki; and Fondation pour la Recherche Médicale, France grant to S. Wirth.

REFERENCES

- Barnes TD, Kubota Y, Hu D, Jin DZZ, Graybiel AM. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* 437: 1158–1161, 2005.
- Busemeyer JR, Townsend JT. Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychol Rev* 101: 446–469, 1993.
- Congdon P. *Applied Bayesian Modeling*. Chichester, UK: Wiley, 2003.
- Dayan P, Kakade S, Montague PR. Learning and selective attention. *Nat Neurosci* 3: 1218–1223, 2000.
- Ditterich J. Stochastic models of decisions about motion direction: behavior and physiology. *Neural Netw* 19: 981–1012, 2006.
- Durbin J, Koopman SJ. *Time Series Analysis by State-Space Methods*. Oxford, UK: Oxford Univ. Press, 2001.
- Estes WK. (Editor). *Handbook of Learning and Cognitive Processes*. New York: Wiley/Halsted Press, 1978.
- Gallistel CR, Fairhurst S, Balsam P. The learning curve: implications of a quantitative analysis. *Proc Natl Acad Sci USA* 101: 12124–13131, 2004.
- Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences [with discussion]. *Stat Sci* 7: 457–511, 1992.
- Gilks WR, Richardson S, Spiegelhalter DJ. *Monte Carlo Markov Chain in Practice*. New York: Chapman & Hall/CRC, 1996.
- Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM. Building neural representations of habits. *Science* 286: 1745–1749, 1999.
- Kakade S, Dayan P. Acquisition and extinction in auto-shaping. *Psychol Rev* 109: 533–544, 2002.
- Kass RE, Carlin BP, Gelman A, Neal RM. Markov Chain Monte Carlo in practice: a roundtable discussion. *Am Stat* 52: 93–100, 1998.
- Kitagawa G, Gersh W. *Smoothness Priors Analysis of Time Series*. New York: Springer-Verlag, 1996.
- Law JR, Flanery MA, Wirth S, Yanike M, Suzuki WA, Smith AC, Frank LM, Brown EN, Stark CEL. fMRI activity during the gradual acquisition and expression of paired-associate memory. *J Neurosci* 25: 5720–5729, 2005.
- Luce RD, Bush RR, Galanter E. (Editors). *Handbook of Mathematical Psychology*. New York: Wiley, 1965.
- Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS—a Bayesian modeling framework: concepts, structure, and extensibility. *Stat Comput* 10: 325–337, 2000.
- Murphy K, Mahdaviani M. MATBUGS software. <http://www.cs.ubc.ca/~murphyk/Software/MATBUGS/matbugs.html>, 2005.
- Paton JJ, Belova MA, Morrison SE, Salzman CD. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439: 865–870, 2006.
- Ratcliff R, Rouder JN. A diffusion model account of masking in two-choice letter identification. *J Exp Psychol Hum Percept Behav* 26: 127–140, 2000.
- Smith AC, Brown EN. Estimating a state-space model from point process observations. *Neural Comput* 15: 965–991, 2003.
- Smith AC, Frank LM, Wirth S, Yanike M, Hu D, Kubota Y, Graybiel AM, Suzuki WE, Brown EN. Dynamic analysis of learning in behavioral experiments. *J Neurosci* 24: 447–461, 2004.
- Smith AC, Stefani MR, Moghaddam B, Brown EN. Analysis and design of behavioral experiments to characterize population learning. *J Neurophysiol* 93: 1776–1792, 2005.
- Smith PL. Psychophysically principled models of visual simple reaction time. *Psychol Rev* 102: 567–593, 1995.
- Spiegelhalter DJ, Thomas A, Best N, Lunn D. WinBUGS v. 1.4.1. Imperial College and Medical Research Council (MRC), United Kingdom. <http://www.mrc-bsu.cam.ac.uk/bugs/> [latest update], 2004.
- Stefani MR, Moghaddam B. Rule learning and reward contingency are associated with dissociable patterns of dopamine activation in the rat prefrontal cortex, nucleus accumbens, and dorsal striatum. *J Neurosci* 23: 1–9, 2006.
- Suppes P. A linear model for a continuum of responses. In: *Studies in Mathematical Learning Theory*, edited by Bush RR, Estes WK. Stanford, CA: Stanford Univ. Press, 1959, p. 400–414.
- Suppes P. On deriving models in the social sciences. *Math Comput Model* 14: 21–28, 1990.
- Suzuki WA, Brown EN. Behavioral and neurophysiological analysis of dynamic learning processes. *Behav Cogn Neurosci Rev* 4: 67–95, 2005.
- Usher M, McClelland JL. The time course of perceptual choice: the leaky, competing accumulator model. *Psychol Rev* 108: 550–592, 2001.
- Usher M, McClelland JL. Loss aversion and inhibition in dynamical models of multialternative choice. *Psychol Rev* 111: 757–769, 2004.
- Verguts T, De Boeck P. A Rasch model for detecting learning while solving an intelligence test. *Appl Psychol Measure* 24: 151–162, 2000.
- Verhelst ND, Glas CAW. A dynamic generalization of the Rasch model. *Psychometrika* 58: 395–415, 1995.
- Williams ZM, Eskandar EN. Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat Neurosci* 9: 562–568, 2006.

- Wirth S, Chiu C, Sharma VS, Avsar E, Smith AC, Scalon J, Brown EN, Suzuki WA.** Analysis of hippocampal signals during learning of selective object–place associations. Program No. 776.1. *Abstract Viewer/Itinerary Planner*. Washington, DC: Society for Neuroscience, 2005, Online.
- Wirth S, Yanike M, Frank LM, Smith AC, Brown EN, Suzuki WA.** Single neurons in the monkey hippocampus and learning of new associations. *Science* 300: 1578–1584, 2003.
- Wolbers T, Büchel C.** Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *J Neurosci* 25: 3333–3340, 2005.
- Yoshida W, Ishii S.** Resolution of uncertainty in prefrontal cortex. *Neuron* 50: 781–789, 2006.
- Yu A, Dayan P.** Expected and unexpected uncertainty: ACH and NE in the neocortex. In: *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, vol. 15, 2003.