

# Characterizing Learning by Simultaneous Analysis of Continuous and Binary Measures of Performance

M. J. Prerau,<sup>1,7</sup> A. C. Smith,<sup>1</sup> Uri T. Eden,<sup>6,7</sup> Y. Kubota,<sup>2,3</sup> M. Yanike,<sup>4</sup> W. Suzuki,<sup>4</sup>  
A. M. Graybiel,<sup>2,3</sup> and E. N. Brown<sup>1,2,5</sup>

<sup>1</sup>Department of Anesthesia and Critical Care, Massachusetts General Hospital, Boston, Massachusetts; <sup>2</sup>Department of Brain and Cognitive Sciences and <sup>3</sup>McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, Massachusetts; <sup>4</sup>Center for Neural Science, New York University, New York City, New York; <sup>5</sup>Division of Health Sciences and Technology, Harvard Medical School/Massachusetts Institute of Technology, Cambridge, Massachusetts; <sup>6</sup>Department of Mathematics and Statistics and <sup>7</sup>Program in Neuroscience, Boston University, Boston, Massachusetts

Submitted 24 November 2008; accepted in final form 17 August 2009

**Prerau MJ, Smith AC, Eden UT, Kubota Y, Yanike M, Suzuki W, Graybiel AM, Brown EN.** Characterizing learning by simultaneous analysis of continuous and binary measures of performance. *J Neurophysiol* 102: 3060–3072, 2009. First published August 19, 2009; doi:10.1152/jn.91251.2008. Continuous observations, such as reaction and run times, and binary observations, such as correct/incorrect responses, are recorded routinely in behavioral learning experiments. Although both types of performance measures are often recorded simultaneously, the two have not been used in combination to evaluate learning. We present a state-space model of learning in which the observation process has simultaneously recorded continuous and binary measures of performance. We use these performance measures simultaneously to estimate the model parameters and the unobserved cognitive state process by maximum likelihood using an approximate expectation maximization (EM) algorithm. We introduce the concept of a reaction-time curve and reformulate our previous definitions of the learning curve, the ideal observer curve, the learning trial and between-trial comparisons of performance in terms of the new model. We illustrate the properties of the new model in an analysis of a simulated learning experiment. In the simulated data analysis, simultaneous use of the two measures of performance provided more credible and accurate estimates of the learning than either measure analyzed separately. We also analyze two actual learning experiments in which the performance of rats and of monkeys was tracked across trials by simultaneously recorded reaction and run times and the correct and incorrect responses. In the analysis of the actual experiments, our algorithm gave a straightforward, efficient way to characterize learning by combining continuous and binary measures of performance. This analysis paradigm has implications for characterizing learning and for the more general problem of combining different data types to characterize the properties of a neural system.

## INTRODUCTION

Learning is a dynamic process that can be defined as a change in behavior that occurs as a result of experience. Learning is often studied by demonstrating that after a series of trials, a subject can perform a previously unfamiliar task more reliably than what would be predicted by chance. The ability of a subject to learn a new task is commonly tested to study the effect of learning on genetic manipulations (Rondi-Reig et al. 2001) brain lesions (Dias et al. 1997; Dusek and Eichenbaum

1997; Fox et al. 2003; Roman et al. 1993; Whishaw and Tomie 1991; Wise and Murray 1999), pharmacological interventions (Stefani et al. 2003), and electrical stimulation (Williams and Eskandar 2006). Accurate characterizations of learning are also needed to link behavioral changes to changes in neural activity in target brain regions (Barnes et al. 2005; Jog et al. 1999; Law et al. 2005; Wirth et al. 2003).

Binary measurements, such as the sequence of correct and incorrect responses, are the measurements most typically recorded to track behavioral changes in learning experiments. Performance is characterized by estimating a learning curve that defines the probability of a correct response as a function of trial number and/or by identifying the learning trial, i.e., the trial on which the change in behavior suggestive of learning can be established using a statistical criterion (Gallistel et al. 2004; Jog et al. 1999; Siegel and Castellan 1988; Smith et al. 2004, 2005; Wirth et al. 2003).

Continuous measurements such as the reaction time (the time it takes the subject to react to a stimulus), the response time, and the total amount of time it takes the subject to execute a task or the task velocity, are also frequently recorded in learning experiments. A separate analysis of the continuous measurements is often presented to support the findings of, or apparent behavior identified in, the analysis of the binary observations.

A frequent assumption in learning studies is that with repeated exposure to the task the subject will show an improvement in performance (Usher and McClelland 2001). This phenomenon is frequently represented by assuming a monotonic relation between exposure to the task (trial number) and the measure of performance (Gallistel et al. 2004). As a consequence, deterministic logistic or sigmoid functions are often fit to each type of performance measure separately.

In no case have the continuous and the binary measurements been analyzed together to characterize learning. Combined analysis of continuous and binary data may be a valuable way to study formally whether it is plausible to assume a common mechanism underlying both measures of performance. When plausible, a combined analysis may improve the statistical power of the study. In addition, for cases in which a subject has reached a plateau in his/her binary responses, i.e., gives all correct responses, yet continues to perform the task with increasingly faster reaction times, examining only the binary responses would suggest that learning has ceased, whereas the

Address for reprint requests and other correspondence: E. N. Brown, Neuroscience Statistics Research Laboratory, Dept. of Anesthesia and Critical Care, Massachusetts General Hospital, 55 Fruit St, Clinics 3, Boston, MA 02114-2696 (E-mail: brown@neurostat.mgh.harvard.edu).

faster reaction times would suggest that the subject is continuing to refine his/her understanding of the task.

We recently introduced a state-space approach to characterizing learning in behavioral experiments (Smith et al. 2004, 2005, 2007; Wirth et al. 2003). To develop a dynamic approach to analyzing data from learning experiments in which continuous and binary responses are simultaneously recorded, we extend our previously developed state-space model of learning to include both a lognormal probability model for the continuous measurements and a Bernoulli probability model for the binary measurements. We estimate the model using an approximate EM algorithm (Dempster et al. 1977) to compute the reaction-time curve, the learning curve, the ideal observer curve, the learning trial, and between-trial comparisons of performance. We illustrate our approach in the analysis of a simulated learning experiment and two actual learning experiments: one in which a monkey rapidly learns new associations within a single session and another in which a mouse slowly learns a T-maze task over many days.

METHODS

State-space model of learning

We assume that learning is a dynamic process that can be studied with the state-space framework used in engineering, statistics and computer science (Doucet et al. 2001; Durbin and Koopman 2001; Fahrmeir and Tutz 2001; Kitagawa and Gersch 1996; Mendel 1995; Smith and Brown 2003). The state-space model consists of two equations: a state equation and an observation equation. The state equation defines the temporal evolution of an unobservable process. Such state models with unobservable processes are often referred to as hidden Markov or latent process models (Fahrmeir and Tutz 2001; Smith and Brown 2003; Smith et al. 2004, 2005). In our analysis, the state equation defines an unobservable cognitive state that represents the subject's understanding of the task. The evolution of this cognitive state is tracked across the trials in the experiment. We formulate the model so that the state increases as learning occurs and decreases when learning does not occur. The observation equation defines how the observed data relate to the cognitive state process. The data we observe in the learning experiment are the series of continuous and binary responses. The objective of the analysis is to characterize learning by estimating the cognitive state process using simultaneously both sets of data.

We consider a learning experiment consisting of  $k$  trials in which on each trial a continuous and a binary response measurement of performance are recorded. Let  $r_k$  and  $n_k$  be, respectively, the values of the continuous and binary measurements on trial  $k$  for  $k = 1, \dots, K$ . We assume that the cognitive state model is the first-order autoregressive [AR(1)] process

$$x_k = \rho x_{k-1} + v_k \tag{1}$$

where  $0 < \rho < 1$  represents a forgetting factor, the  $v_k$  are independent, zero mean Gaussian random variables with variance  $\sigma_v^2$ . Let  $x = [x_0, x_1, \dots, x_K]$  be the unobserved cognitive state process for the entire experiment.

For the purpose of exposition, we assume that the continuous measurements are reaction times and that the observation model for the reaction times is

$$z_k = \log r_k = \alpha + \beta x_k + \varepsilon_k \tag{2}$$

where the  $\varepsilon_k$  are independent zero mean Gaussian random variables with variance  $\sigma_\varepsilon^2$ . Formulation of the log of the reaction time  $r_k$  as a linear function of the cognitive state variable ensures that at each trial

the reaction time is positive. We assume that  $\beta < 0$  to ensure that on average, as the cognitive state  $x_k$  increases with learning, the reaction time decreases. We let  $Z = [z_1, \dots, z_K]$  be the log of the reaction times on all  $\kappa$  trials.

We assume that the observation model for the binary responses is the Bernoulli probability model

$$\Pr(n_k | x_k) = p_k^{n_k} (1 - p_k)^{1-n_k} \tag{3}$$

where  $p_k$ , the probability of a correct response on trial  $k$ , is defined in terms of the unobserved cognitive state process as

$$p_k = \exp(\mu + x_k) [1 + \exp(\mu + x_k)]^{-1} \tag{4}$$

Formulation of  $p_k$  as a logistic function of the cognitive state process (Eq. 4) ensures that the probability of a correct response on each trial is constrained to lie between 0 and 1 and that as the cognitive state increases, the probability of a correct responses approaches 1. We estimate the parameter  $\mu$  from  $p_{\text{chance}}$ , the probability of a correct response by chance at the outset of the experiment. To do so, we assume that there is no bias in the subject's responses at the outset of the experiment, and we set  $p_0$  the probability of a correct response at the outset of the experiment equal to  $p_{\text{chance}}$ . For example, if the subject is executing a binary choice task, then the probability of a correct response by chance is  $p_{\text{chance}} = 0.5$ . Therefore using Eq. 4, with  $k = 0$ ,  $p_0 = p_{\text{chance}}$ , and  $x_0 = 0$ , we solve for the estimate of  $\mu$  as  $\hat{\mu} = \log[(1 - p_{\text{chance}})^{-1} p_{\text{chance}}]$ . After estimating the cognitive state process and the model parameters from the data across the entire learning experiment (described in the following text), the estimate of  $x_0, \hat{x}_0$ , may take on a value other than 0. If this is the case, then from Eq. 4, the final estimate of the probability of a correct response at the outset of the experiment,  $\hat{p}_0 = [1 + \exp(\hat{\mu} + \hat{x}_0)]^{-1} \exp(\hat{\mu} + \hat{x}_0)$ , will differ from  $p_{\text{chance}}$ . Any difference between  $\hat{p}_0$  and  $p_{\text{chance}}$  is due to  $\hat{x}_0$  not equaling 0 and suggests a potential learning bias for the subject. We let  $N = [n_1, \dots, n_K]$  be the sequence of correct and incorrect responses on all  $\kappa$  trials.

The state model (Eq. 1) provides a stochastic continuity constraint (Kitagawa and Gersch 1996) so that the current cognitive state, the reaction time (Eq. 2), and the probability of a correct response (Eq. 4) depend on the prior cognitive state. In this way, the state-space model provides a simple, plausible framework for relating performance on successive trials of the experiment.

Because  $x$  is unobservable and because  $\theta = (\rho, \alpha, \beta, \sigma_v^2, \sigma_\varepsilon^2, x_0)$  is a set of unknown parameters, we use the expectation-maximization (EM) algorithm to estimate them by maximum likelihood (Dempster et al. 1977). The EM algorithm is a well-known procedure for performing maximum likelihood estimation when there is an unobservable process or missing observations. The EM algorithm has been used to estimate state-space models from point process and binary observations with linear Gaussian state processes (Smith and Brown 2003; Smith et al. 2004, 2005). The current EM algorithm combines features of the ones in Shumway and Stoffer (1982) and Smith and Brown (2003). The key technical point that allows implementation of this algorithm is the combined filter algorithm in Eqs. A4–A8 of APPENDIX A. Its derivation is given in Prerau et al. (2008).

Estimation of the reaction-time and learning curves

Because we compute the maximum likelihood estimate of  $\theta = (\rho, \alpha, \beta, \sigma_v^2, \sigma_\varepsilon^2, x_0)$  using the EM algorithm, we compute the maximum likelihood estimate of each  $x_k$  for  $k = 1, \dots, K$  from the fixed-interval smoothing algorithm at the last iteration of the EM algorithm (Eqs. A9–A11). That is, the smoothing algorithm estimate at trial  $k$  is the estimate of  $x_k$  given  $Z$  and  $N$  with  $\theta$  replaced by its maximum likelihood estimate. Hence because the smoothing algorithm is based on all the data, it gives the estimate of the cognitive state from the perspective of the ideal observer. This estimate is both the maximum likelihood and empirical Bayes' estimate of  $x_k$  (Fahrmeir and Tutz

2001). The smoothing algorithm estimates the state as the Gaussian random variable with mean  $x_{k|K}$  (Eq. A9) and variance,  $\sigma_{k|K}^2$  (Eq. A11), where the notation  $x_{k|K}$  denotes the estimate of the cognitive state process at trial  $k$  given all of the data in the experiment. The state estimates have Gaussian distributions because we use Gaussian approximations to compute them in the filter (Eqs. A4–A8) and smoothing algorithms (Eqs. A9–A11).

Using the smoothing algorithm estimate of  $x_k$ , we can compute the probability density of the estimate of the log reaction time at trial  $k$ ,  $r_k$ , by using Eq. 2 and the standard change-of-variables formula from elementary probability theory (Rice 2007; Smith et al. 2004). Applying the change-of-variable formula to the relation  $r_k = \exp(\alpha + \beta x_k)$  and the fact that, conditional on the data  $Z$  and  $N$ ,  $x_k$  has an approximate Gaussian probability density with mean  $x_{k|K}$  (Eq. A9) and variance  $\sigma_{k|K}^2$  (Eq. A11), yields as the probability density of  $r_k$  given the data  $Z$  and  $N$

$$f(r | \mu, x_{k|K}, \sigma_{k|K}^2) = \frac{1}{r\beta} (2\pi\sigma_{k|K}^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma_{k|K}^2} \left(\frac{\log[r - \alpha]}{\beta} - x_{k|K}\right)^2\right\} \quad (5)$$

Eq. 5 is the probability density of a log normal random variable (Rice 2007). We define the reaction-time curve as the plot of  $r_{k|K}$  versus  $k$  for  $k = 1, \dots, K$ , where  $r_{k|K}$  is the mode of Eq. 5.

Similarly, if we apply the change-of-variables formula to the relation  $p_k = \exp(\mu + x_k)[1 + \exp(\mu + x_k)]^{-1}$  and use again the fact that conditional on the data,  $Z$  and  $N$ ,  $x_k$  has an approximate Gaussian probability density with mean  $x_{k|K}$  and variance  $\sigma_{k|K}^2$ , we obtain the probability density of  $p_k$  given the data  $Z$  and  $N$

$$f(p | \mu, x_{k|K}, \sigma_{k|K}^2) = \frac{1}{p(1-p)} (2\pi\sigma_{k|K}^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma_{k|K}^2} \left(\log\left[\frac{p}{(1-p)\exp(\mu)}\right] - x_{k|K}\right)^2\right\} \quad (6)$$

We define the learning curve as the plot of  $p_{k|K}$  versus  $k$  for  $k = 1, \dots, K$ , where  $p_{k|K}$  is the mode of Eq. 6. The derivations of Eqs. 5 and 6 are given in APPENDIX B.

### Ideal observer curve

To characterize learning, we need a systematic way to compare performance on each trial with performance on any other trial. Of particular importance is comparing performance on each trial with performance at the start of the experiment. If performance becomes consistently better than performance at the outset of the experiment and remains so for the balance of the experiment, it is plausible to conclude that learning has occurred. If in the learning experiment we only observed the subject's reaction times on each trial, then we can use the probability density for the learning curve in Eq. 5 to compute the ideal observer curve as

$$\Pr(r_{k|K} < r_0) \quad (7)$$

for  $k = 1, \dots, K$ . The reaction-time ideal observer curve computes at each trial  $k$  the probability that the reaction time at trial  $k$  is less than the reaction time at the outset of the experiment. Similarly, if only the binary responses are observed, then we use Eq. 6 to compute the ideal observer curve as

$$\Pr(p_{k|K} > p_0) \quad (8)$$

for  $k = 1, \dots, K$ . The binary response ideal observer curve describes for each trial  $k$  the probability that  $p_k$  is greater than the probability of a correct response at the outset of the experiment.

To generalize these definitions so that we can use simultaneously the reaction-time and binary behavioral observations, we note that for both the reaction-time and binary observations, the ideal observer curve can be expressed in terms of the cognitive state variables as

$$\Pr(x_{k|K} > x_0) = \Pr(r_{k|K} < r_0) = \Pr(p_{k|K} > p_0) \quad (9)$$

because  $r_{k|K}$  and  $p_{k|K}$  are monotonic functions of  $x_{k|K}$ . Therefore if the reaction-time and binary observations are both used to estimate the cognitive state process, then the ideal observer curve for the joint observations is defined by the left side of Eq. 9. Moreover, if only reaction times or the binary responses are observed, then Eq. 9 agrees with the individual ideal observer curves for these measurements. Therefore we define Eq. 9 as the ideal observer curve for the combined model.

### Identification of the learning trial

Although learning is a dynamic process, it is helpful to have a single summary that indicates where in the experiment performance is better than at the outset. We identify the learning trial from the ideal observer curve. We define the learning trial as the first trial for which the ideal observer can state with reasonable certainty that the subject performs better on that trial and all subsequent trials than at the outset of the experiment. For our analyses we define a level of reasonable certainty as 0.975 and term this trial the ideal observer learning trial with level of certainty 0.975 [IO (0.975)]. The IO(0.975) is the first trial  $j$  such that for all trials  $k \geq j$

$$\Pr(x_k > x_0) \geq 0.975 \quad (10)$$

### Between-trial comparisons of performance

We make between-trial comparisons of learning that are not restricted to comparisons between a given trial and performance at the outset of the experiment. These comparisons can be performed in a straightforward way in our paradigm because we estimate the joint probability distribution of  $x$  from the smoothing algorithm at the final iteration of the EM algorithm (Eqs. A9–A11). Therefore given any two trials, we can compute the probability that the subject's performance on one trial is better than the performance on the other trial. The transformation between the state variable and the probability of a correct response and the transformation between the state variable and the reaction time are monotonic. Therefore it suffices to compute the probability that the cognitive state at trial  $k$  is greater than the cognitive state on trial  $j$  for all distinct pairs  $k$  and  $j$ . To compute the probability  $\Pr(x_k > x_j)$  for all pairs  $j = 1, \dots, K - 1$  and  $k = 2, \dots, K$ , we use a Monte Carlo algorithm, which combines the fixed-interval smoothing and the state-space covariance algorithm at the last step of the EM algorithm to compute the covariance between the states at two different trials (Smith et al. 2004, 2005). This algorithm compares pairs of random variables from the  $K$ -dimensional joint probability density of the cognitive state process.

### Experimental protocols for learning

CONDITIONAL T-MAZE TASK. To illustrate our method applied to an actual learning experiment, we analyzed the responses of a mouse performing a T-maze task, similar to those described in Jog et al. (1999) and in Barnes et al. (2005). In this task, the mouse was required to use tactile cues (rough or smooth surface of the maze floor) to learn which one of two arms of a T-maze to enter to receive a reward. On each day of this 13-day experiment, the mouse performed 28–40 trials. The total number of trials was 461. In each trial, the mouse had to acquire the cue-turn association to

succeed. For this experiment, the probability of making a correct response by chance was 0.5. To characterize learning, we recorded at each trial the run time (the time from the animal's onset of locomotion to its arrival at the goal), and whether the animal's choice of goal arm was correct or incorrect. The objective of this study was to relate changes in learning across days to concurrent changes in neural activity in the striatum (Barnes et al. 2005; Jog et al. 1999).

**LOCATION-SCENE ASSOCIATION TASK.** As a second illustration of our methods applied to an actual learning experiment, we analyzed the responses of a macaque monkey in a location-scene association task, described in detail in Wirth et al. (2003). In this task, each trial started when the monkey fixated on a centrally presented cue on a computer screen. The animal was then presented with four identical targets (north, south, east, and west) superimposed on a novel visual scene. After the scene disappeared, the targets remained on the screen during a delay period. At the end of the delay period, the fixation point disappeared cueing the animal to make an eye-movement to one of the three targets. For each scene, only one target was rewarded and the positions of rewarded locations were counterbalanced across all new scenes. Between two to three novel scenes were typically learned simultaneously and trials of novel scenes were interspersed with trials in which three well-learned scenes were presented. Because there were three locations the monkey could choose as a response, the probability of a correct response occurring by chance was 0.33. To characterize learning we recorded for all trials the reaction times (time from to go-cue to the response), and the correct and incorrect responses. The objective of the study was to track learning as a function of trial number and to relate it to changes in the activity of simultaneously recorded hippocampal neurons (Wirth et al. 2003).

Experimental procedures employed for the experiments were in accordance with the National Institutes of Health, Massachusetts Institute of Technology, and New York University guidelines for use of laboratory animals.

RESULTS

Analysis of a simulated learning experiment

To illustrate our analysis paradigm, we simulated a learning experiment with a cognitive state process observed simultaneously through reaction times and binary responses. We modeled the cognitive state as the first-order autoregressive process as in Eq. 1 with the inclusion of a learning rate parameter  $\rho_0$ . That is, the cognitive state model used to simulate the data was

$$x_k = \rho_0 + \rho x_{k-1} + v_k \tag{12}$$

Choosing  $\rho_0 > 0$  ensured that on average the subject's performance as measured by the learning curve and the reaction-time curve would improve on successive trials. Equation 12 is an elementary version of a learning model developed by Usher and McClelland (2001). We chose as the parameters in our simulation  $\rho = 0.99$ ,  $\rho_0 = 0.01$ , and  $x_0 = -3$ . The  $v_k$  were independent, identically distributed Gaussian errors with  $E(v_k)$  and  $\text{Var}(v_k) = \sigma_v^2 = 0.09$ .

For the reaction-time process in Eq. 2, we chose the parameters  $\alpha = 2$ ,  $\beta = -0.2$ , and  $\sigma_\epsilon^2 = 1.3$ , and for the binary observation process, we modified Eq. 4 as

$$p_k = \exp(\mu + \gamma x_k) [1 + \exp(\mu + \gamma x_k)]^{-1} \tag{13}$$

where we chose  $\mu = -0.3$  and  $\gamma = 0.7$ . Using Eq. 12, we first simulated  $K = 100$  observations from the state process (Fig. 3, dashed curves). For each value of  $x_k$ , we used Eq. 2 to simulate the reaction times (Fig. 1A, ●). For each value of  $x_k$ , we computed  $p_k$  from Eq. 13, and we simulated the sequence of correct and incorrect responses using Eq. 3 by drawing a Bernoulli observation  $n_k$  with the probability of a correct response defined by  $p_k$  (Fig. 1C, ■ and □). To analyze the

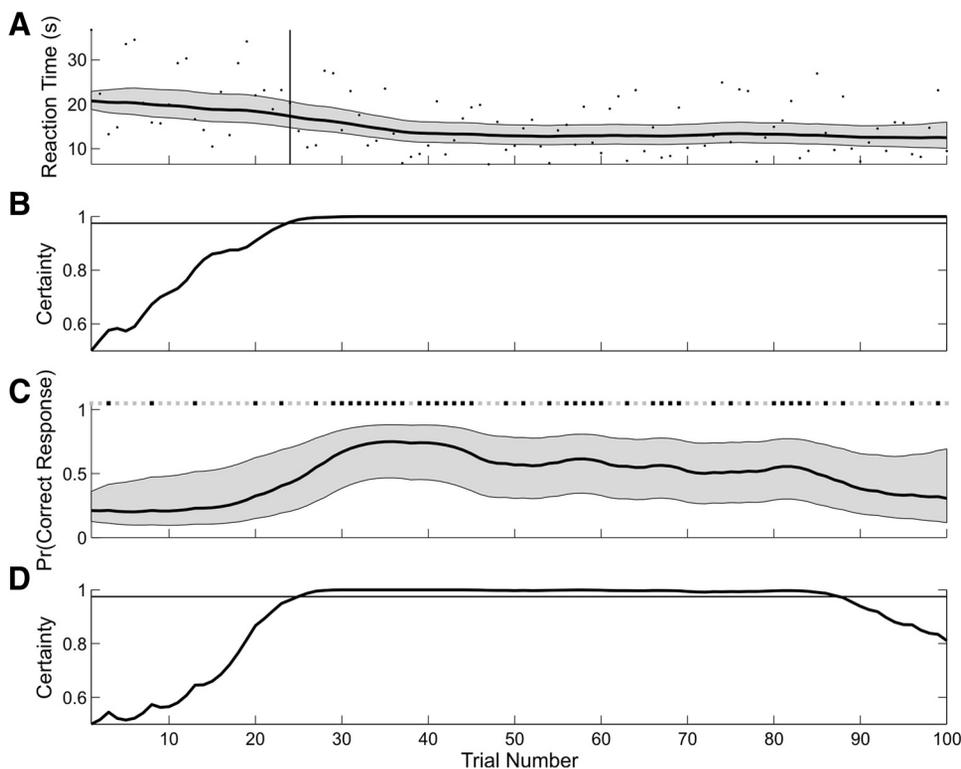


FIG. 1. A: black dots are the reaction times from the simulated experiment. Expectation maximization (EM) algorithm estimates of the reaction-time curve (thick black curve) and its 95% confidence intervals (gray area) computed from the reaction-time data. The vertical line marks trial 24, the IO(0.975) learning trial based only on the reaction-time data. B: the ideal observer curve computed from the analysis of the reaction-time data only crosses 0.975 (horizontal line) at trial 24. The horizontal line is 0.975. C: the simulated correct (black squares) and incorrect (gray squares) trial responses (top). EM algorithm estimate of the learning curve (thick black curve) and its 95% confidence interval (gray area) based on only the binary data. D: ideal observer curve computed from the analysis of the binary data only. The horizontal line is 0.975. There is no IO(0.975) learning trial from the analysis of the binary data only because the ideal observer curve drops <0.975 at trial 88.

simulated data, we fit the state-space model in *Eqs. 1–4* using the EM algorithm in APPENDIX A. We assumed in the model estimation that the parameter  $\rho_0 = 0$  and  $\gamma = 1$ . By so doing, the model used to simulate the data differed from the model used to analyze the data.

We first analyzed the simulated data assuming that only the reaction times were observed (Fig. 1A). We used this version of the state-space model (*Eqs. 1* and *2*) in the EM algorithm analysis to estimate the reaction-time curve and its 95% confidence intervals (Fig. 1A), the learning trial (A) and the ideal observer curve (B). The simulated observations were highly variable relative to the estimated reaction-time curve. The reaction times varied between 10 and 35 s from trials 1 to 25 and thereafter, between 5 and 20 s. The ideal observer curve (Fig. 1B) showed the probability that the reaction on each trial  $k$  was less than the reaction time at the outset of the experiment, i.e.,  $\Pr(r_k < r_0)$ , for  $k = 1, \dots, 100$ . This probability increased steadily from trials 1 to 25, and was  $>0.975$  from trial 24 on to the end of the experiment. For this reason, the IO(0.975) learning trial was estimated as trial 24.

We next analyzed the simulated data assuming that only the binary responses were observed (Fig. 1C). We used this version of the state-space model (*Eqs. 3* and *4*) in the EM algorithm analysis to estimate the learning curve, its 95% confidence intervals (Fig. 1C) and its ideal observer curve (D). This ideal observer curve (Fig. 1D) rose almost monotonically from trials 1 to 27 where it crossed 0.975 and remained above this level of certainty until trial 85. From trial 86 on to trial 100, the ideal observer curve decreased almost monotonically to 0.81. Because of this last decrease in performance, the criteria for the IO(9.75) learning trial were not satisfied. These results were consistent with the large number of correct responses from trial 29 to 45 and with the fact that there were only five correct responses in trials 85 to 100. We concluded that there is no learning trial based on the analysis of the binary data alone.

Last we analyzed the simulated data using both the reaction times and binary responses (Fig. 2). The combined algorithm estimates of the reaction-time curve (Fig. 2A) and learning curve (B) resembled approximately their respective continuous and binary estimates. At each trial, the combined algorithm 95% confidence intervals for the reaction-time curves (Fig. 2A)

were wider than their reaction-time-only model estimates (Fig. 1A). The greater variability in the reaction-time observations induced more variability in the learning curve estimate and wider 95% confidence intervals (Fig. 2B) at all trials compared with the binary data only estimates (Fig. 1C). The IO(0.975) learning trial was trial 23 for this analysis. This was consistent with trial 24 suggested by the reaction-time-only analysis rather than the estimate of no-learning trial suggested by the analysis of the binary data only.

Because the data were simulated, we also compared how well each of the algorithms estimated the unobservable cognitive state (Fig. 3) as a way of assessing goodness of fit. The estimated cognitive state for the reaction-time-only model (Fig. 3A, —) agreed closely with the true cognitive state (Fig. 3A, - - -) from trials 1 to 40. The estimated cognitive state overestimated the true cognitive state from trials 41 to 100. These 95% confidence intervals did not cover the true cognitive state at trial 60 and between trials 83 and 87. The cognitive state estimated from the binary observations alone (Fig. 3B, —) overestimated the true cognitive state (Fig. 3B, - - -) from trials 1 to 39 and underestimated the true cognitive state from trial 87 to 100. Except for trials 1 to 7 and trial 93, these 95% confidence intervals covered the true cognitive state at all trials.

The combined model estimate of the cognitive state (Fig. 3C, —) agreed closely with the true cognitive state (C, - - -). Except for trial 1, these 95% confidence intervals covered the true cognitive state for all trials. The performance of the combined model over the first 40 trials was greatly influenced by the reaction-time data (Fig. 1A). This is the trial range over which we observed the largest change in the reaction-time recordings. The overall shape of the combined model cognitive state estimate was very close to the binary cognitive state estimate between trials 29 and 45, where there was a long string of correct responses (Fig. 1C). Similarly, the downward trend in the combined model learning curve from trials 86 to 100 occurred because in this range there were only 10 of 15 correct responses (Fig. 2C). Overall, the combined model cognitive state estimate (Fig. 3C) was closer to the true cognitive state process than either the reaction-time estimate (A) or the binary state estimate (B).

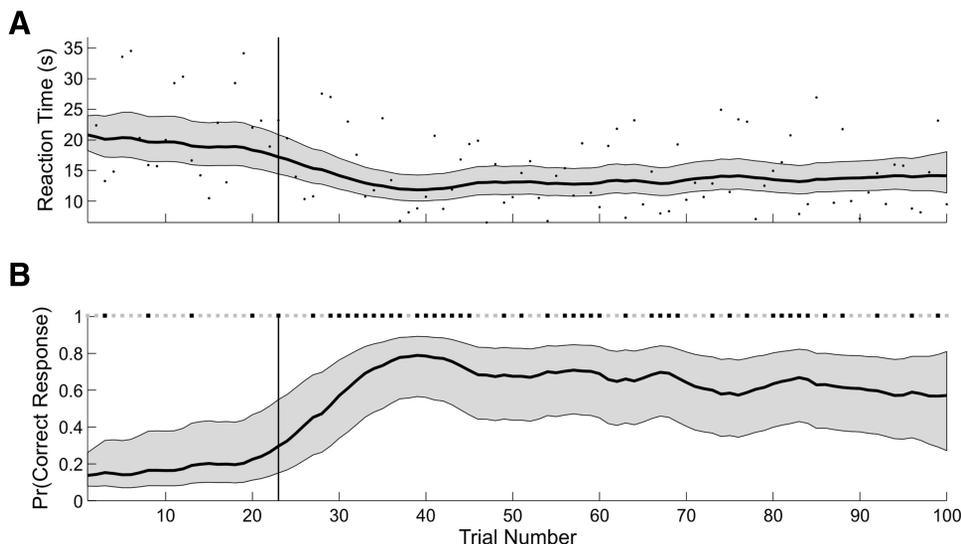


FIG. 2. A: black dots are the reaction times from the simulated learning experiment. EM algorithm estimates of the reaction-time curve (black curve) and its 95% confidence intervals (gray area) computed from the combined analysis of the reaction-time and binary data. B: the simulated correct (black squares) and incorrect (gray squares) trial responses are shown at the top of this panel. EM algorithm estimate of the combined analysis learning curve (black curve) and its 95% confidence intervals (gray area). The vertical lines mark trial 23, the IO(0.975) learning trial.

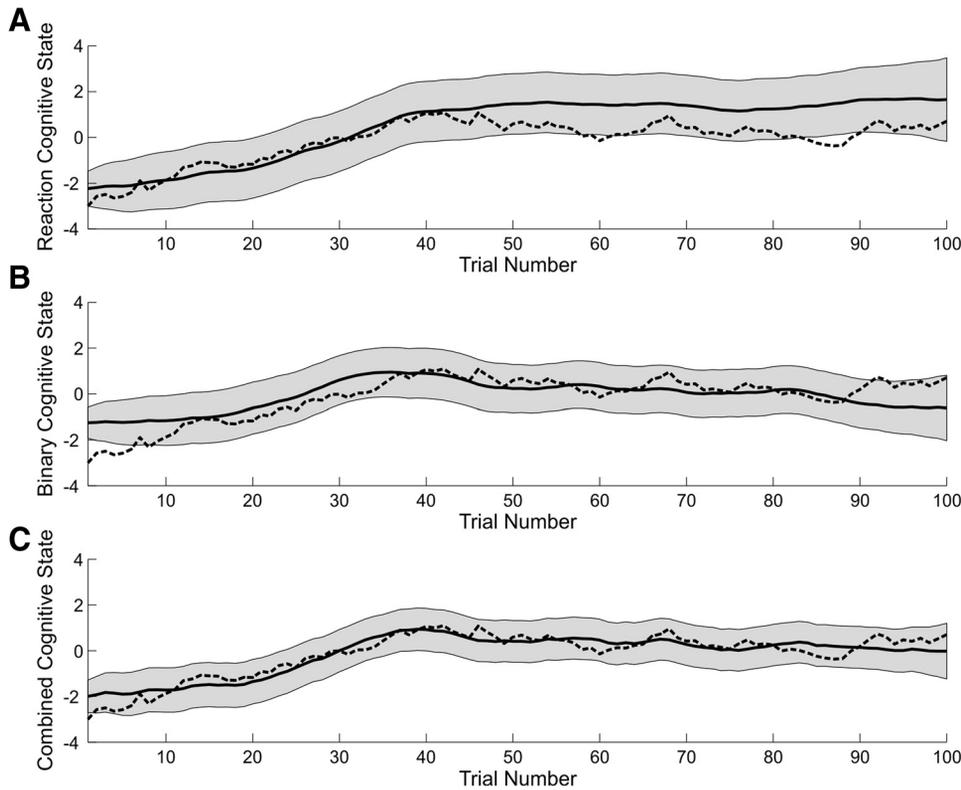


FIG. 3. EM algorithm estimate of the cognitive state (solid black curve) and its 95% confidence interval (gray area) based on the reaction-time data only (A), the binary data only (B), and the combined reaction and binary data (C). The true cognitive state process is the dashed black curve in all panels.

Consistent with this qualitative description, the combined model cognitive state estimate also had the smallest mean-squared error of 0.0001 compared with 1.3451 and 0.3957 for the reaction-time-only and binary-data-only models, respectively. As a further assessment of the accuracy and precision of the three algorithms, we computed for each the coverage probability and the median width of the 95% confidence intervals. The coverage probability is the fraction of times the 95% confidence intervals for a given algorithm cover the true cognitive state. This fraction should be 0.95. The coverage probabilities were 0.99 for the combined model, 0.88 for the reaction-time-only model, and 0.43 for the binary-data-only model. The median widths of the 95% confidence intervals for the cognitive state were 1.83 for the combined model, 2.64 for the reaction-time-only model, and 2.13 for the binary data only model. The performance of the combined model state estimates was consistent with the performance we recently reported for the filter algorithm based on the combined model relative to filter algorithms based only on continuous or only on binary behavioral measures (Prerau et al. 2008).

As reported in the preceding text, in the learning curve analyses (Fig. 2), the ideal observer curve based on the combined model (Fig. 4A, —) identified the IO(0.975) learning trial as trial 23. Because we estimated the joint probability density of the cognitive states, we also computed the probability that the cognitive state at a given trial  $k$  ( $x$  axis in Fig. 4B) was greater than the cognitive state for all trials  $j < k$  ( $y$  axis in B) and plotted this two-dimensional comparison for all pairs of trials (B). The red area shows the trial comparisons for which  $\Pr(x_k < x_j) \geq 0.95$ . Performance from trial 23 through the remainder of the experiment was better than performance at the outset of the experiment. This specific observation is simply a restatement of the findings from the ideal observer curve

analysis (Fig. 4A). However, this analysis goes further and shows that performance at trial 40 was better than performance at trials 35 and lower, whereas performance on trials 50–100 was better than performance on all trials 25 and lower (Fig. 4B, red area).

In summary, the combined model had the greatest accuracy and precision in estimating the cognitive state. The combined model recovered more information from the reaction-time and binary observations than state estimation with either of these sets of observations analyzed separately.

#### Combined model analysis of a mouse learning experiment

As an illustration of our methods applied to an actual learning experiment, we analyzed the behavioral data (Fig. 5) of a mouse performing a T-maze task.

The experiment consisted of 461 trials for this mouse. To analyze learning, we used the series of run times (Fig. 5A) and the series of correct or incorrect responses across all 461 trials. We estimated the parameters of the combined model using the EM algorithm. We computed the run-time curve (Fig. 5A, solid black curve), the learning curve (C, solid black curve), and the cognitive state estimate for the combined model, and their associated 95% confidence intervals (A, C, and D, gray areas) using Eqs. 5, 6, A9, and A11.

The average run time of the mouse was around 5 s for the first 60 trials and declined almost linearly to  $\sim 1.5$  s by trial 100 (Fig. 5, A and B). In the first 60 trials, there was large variability in the run times (Fig. 5B). Consistent with the large fluctuations in run times, the mouse gave many more incorrect responses during the first 60 trials than during any other period in the experiment (Fig. 5C). This variability decreased over the duration of the experiment. However, there were long run

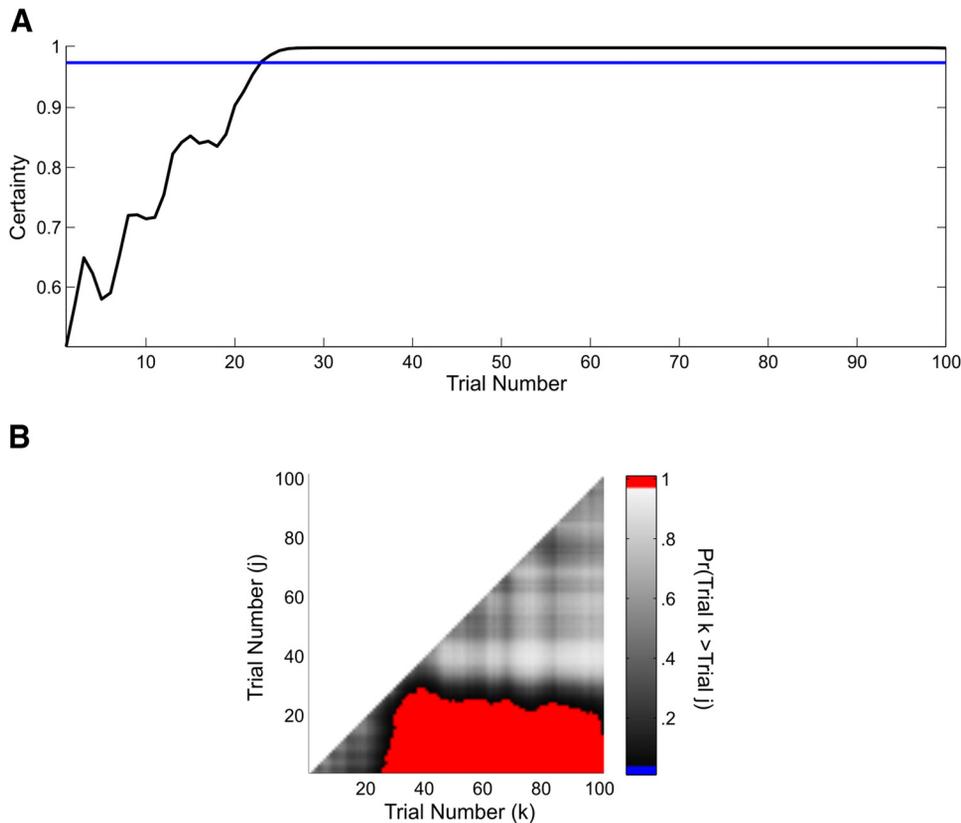


FIG. 4. A: ideal observer curve computed from the combined analysis of the simulated reaction-time and binary data. The horizontal line is 0.975. B: trial-by-trial comparison of the probability that the cognitive state at trial  $k$  ( $x$  axis) is greater than the cognitive state at trial  $j$  ( $y$  axis)  $<k$ , computed from the combined reaction-time and binary-data analysis, where  $k = 2, \dots, 100$  and  $j = 1, \dots, 99$ . Comparisons for which this probability is  $\geq 0.95$  are shown in red.

times at trials 139, 198, 257, 342, and 446. This variability affected the combined model estimate of the cognitive state and its 95% confidence intervals (Fig. 5D). The cognitive state curve and its 95% confidence intervals showed small downward dips at these long run times (Fig. 5D). The 95% confidence intervals for the learning curve was wider prior to trial 60, narrowed from trials 60 to 100 and maintained an approximately constant width from trials 100 to 461 (Fig. 5C). The 95% confidence intervals for the learning curve showed small downward dips at trials 198 and 267, where the run times were large relative to the run times immediately before and after those trials (Fig. 5A).

The fluctuations in the level of certainty of the ideal observer curve during the first 60 trials (Fig. 6A) were also consistent with high variability in the run times and the relatively larger number of incorrect responses during this period (Fig. 5B). By trial 61, the level of certainty of the ideal observer curve crossed and stayed above 0.975 for the balance of the experiment, making trial 61 the IO(0.975) learning trial (Figs. 5 and 6A). For comparison, we also included the learning trials for the run-time-only analysis (Fig. 5A, dashed gray vertical line) and the analysis using only the binary responses (Fig. 5C, dotted gray vertical line). Whereas the run-time-only analysis learning trial of 59 is consistent with the combined model learning trial of 61, the binary learning trial of 210 is substantially later. This is because for the binary analysis the estimated initial performance is 0.75 compared with 0.62 for the combined model. In addition, between trials 30 and 60 the animal had 18 incorrect responses (Fig. 5B). There was substantial variability in the run times during the first 80 trials. Moreover, the run times in this part of the experiment were not necessarily

fast when the animal responded correctly or slow when the animal responded incorrectly.

Instead of simply reporting the learning trial, we also computed the probability that the cognitive state at a given trial  $k$  ( $x$  axis in Fig. 6B) is greater than the cognitive state for all trials  $j < k$  ( $y$  axis in B) and plotted this two-dimensional comparison for all pairs of trials (B). The red areas show the trial comparisons for which  $Pr(x_k > x_j) \geq 0.95$ , whereas the blue areas show the trial comparisons for which this probability is  $< 0.05$ . The red area suggests that performance on trials 100–461 was better than performance on trials 98 and lower. Between trial 350 to approximately trial 425, performance was also better than on many of the trials between 100 and 300. In addition, the blue area suggests that performance between approximately trials 440 and 461 was worse than performance on trials 300–375.

These analyses demonstrate a clear change in the mouse's performance consistent with the conclusion that the animal learned the association between tactile cues and rewarded turning direction.

#### Combined model analysis of a monkey learning experiment

As a second illustration of the performance of our methods in the analysis of an actual learning experiment, we analyzed the reaction-times and binary responses of a macaque monkey executing 45 trials of a location-scene association task. Because there were four locations the monkey could choose as a response, the probability of a correct response occurring by chance was 0.25. The monkey's reaction times were  $\sim 4.5$  s for the first 21 trials and then declined linearly to  $\sim 2.5$  s by trial 45 (Fig. 7A). The monkey gave only three correct responses in

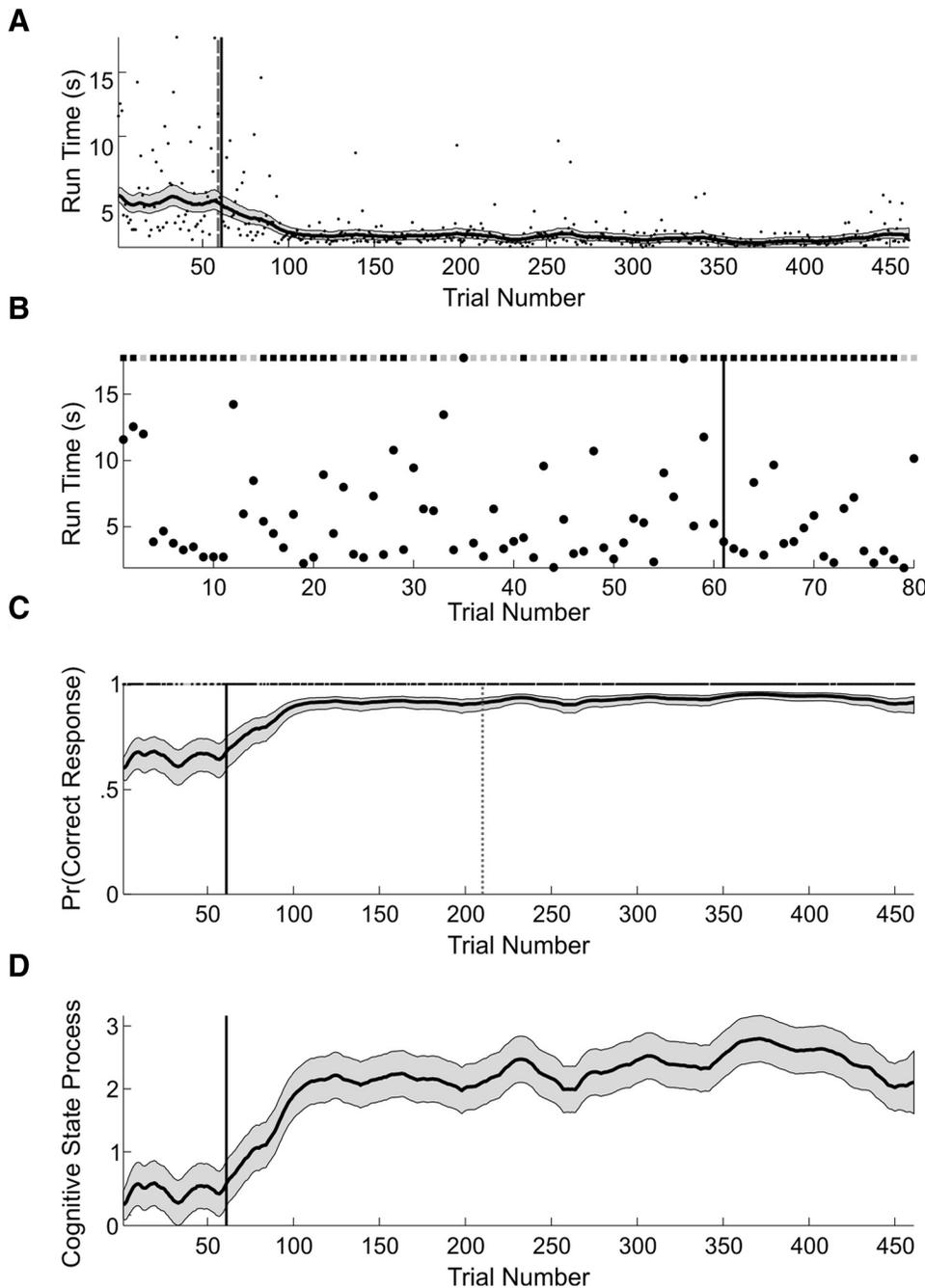


FIG. 5. A: black dots are the run times of the mouse executing the T-maze learning task. EM algorithm estimates of the run time curve (solid black curve) and its 95% confidence intervals (gray area) computed from the combined analysis of the run time and binary data. The solid black (dashed gray) vertical line is the IO(0.975) learning trial of 61 (59) computed using the combined run time and binary data (run time only). B: run time and binary data from trials 1 to 80 and IO(0.975) learning trial (black vertical line) for the combined model. C: the correct (black squares) and incorrect (gray squares) trial responses are shown at the top of this panel. EM algorithm estimate of the learning curve (solid black curve) and its 95% confidence interval (gray area) computed from the combined analysis. The black (gray dotted) vertical line marks the IO(0.975) learning trial of 61 (210) for the combined (binary data only) analysis. D: EM algorithm cognitive state estimate computed from the combined analysis of the run time and binary data. The gray areas are the 95% confidence intervals. The black vertical line marks the IO(0.975) learning trial of 61 for the combined model.

the first 25 trials then gave only correct responses for the remainder of the experiment (Fig. 7C).

As for the previous two examples, the combined model estimate of the reaction-time curve (Fig. 7A), the cognitive state process (B), and the learning curve (C) were estimated from the reaction-time and binary data using the model in Eqs. 1–4 and the EM algorithm. The combined model reaction-time curve estimate (Fig. 7A) increased slightly across trials 1–20 and declined monotonically from trials 21 to 45. The reaction times were approximately constant over these trials, and therefore the 19 consecutive incorrect responses strongly affected the cognitive state estimation. Similarly, the decline in reaction times beginning at trial 20 led to an increase in the combined model estimate of the cognitive state beginning at that trial. For this reason, the learning curve also increased at this trial even

though the animal had three of the five incorrect responses from trials 20 to 24. Because the monkey gave all correct responses from trials 26 to 45 and the reaction times decreased linearly over these trials, the combined model cognitive state and learning curve estimates increased and the reaction-time curve estimate decreased.

The outliers in the reaction times affected the learning curve estimates. The short reaction time at trial 6 and the long reaction time at trial 33 (Fig. 7A) caused a noticeable increase in the width of the learning curve confidence intervals on those trials (C). This is because these reaction times were inconsistent with the binary responses immediately preceding and following these trials. The declining reaction times between trials 20 and 24 (Fig. 7A) were responsible for the increase in the learning curve over these trials even though the animal

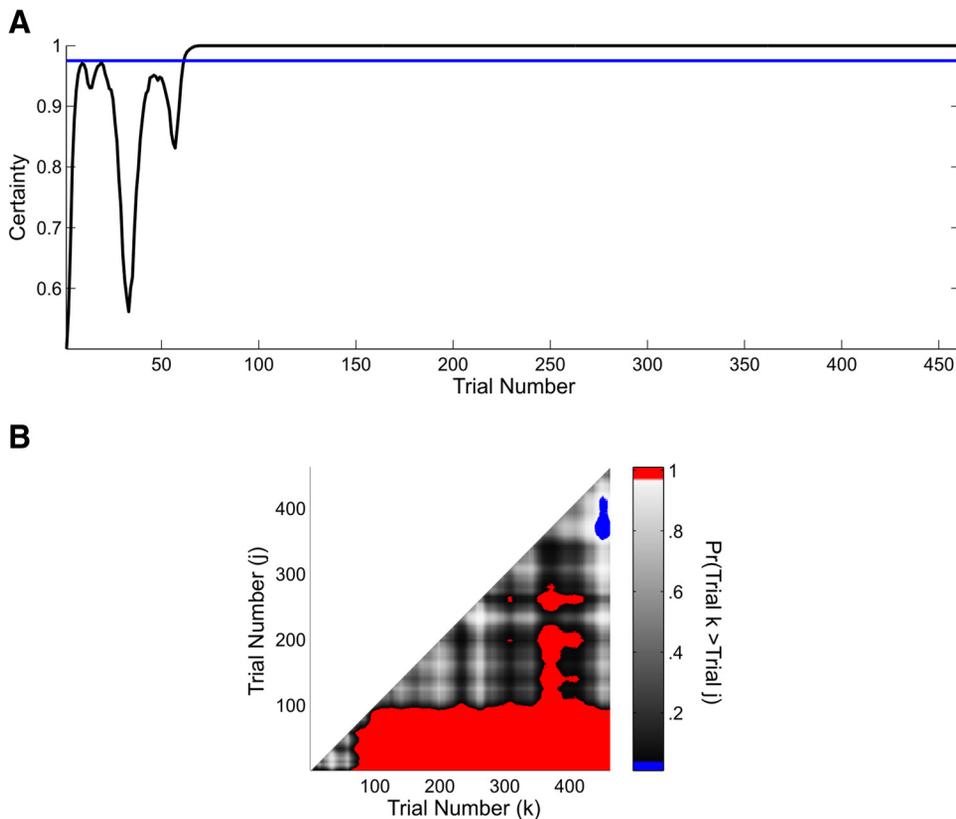


FIG. 6. A: ideal observer curve computed from the combined analysis of the run time and binary data of the mouse executing the T-maze task (Fig. 5). The horizontal line is 0.975. B: trial-by-trial comparison of the probability that the cognitive state at trial  $k$  ( $x$  axis) is greater than the cognitive state at trial  $j$  ( $y$  axis)  $< k$ , where  $k = 2, \dots, 461$  and  $j = 1, \dots, 460$ . Comparisons for which this probability is  $\geq 0.95$  are shown in red and comparisons for which this probability is  $\leq 0.05$  are shown in blue.

gave mostly incorrect responses (C). This brief inconsistency between the reaction time and the binary responses was reflected in the increased width of the learning curve confidence intervals over these trials. Similarly, the short reaction time at trial 30 in an interval in which all of the animal's trial

responses were correct, led to a sharp decrease in the width of the learning curve confidence intervals. For comparison we also included the learning trials for the reaction-time-only analysis (Fig. 7A, vertical line) and the analysis using only the binary responses (C, vertical line). The reaction-time-only and

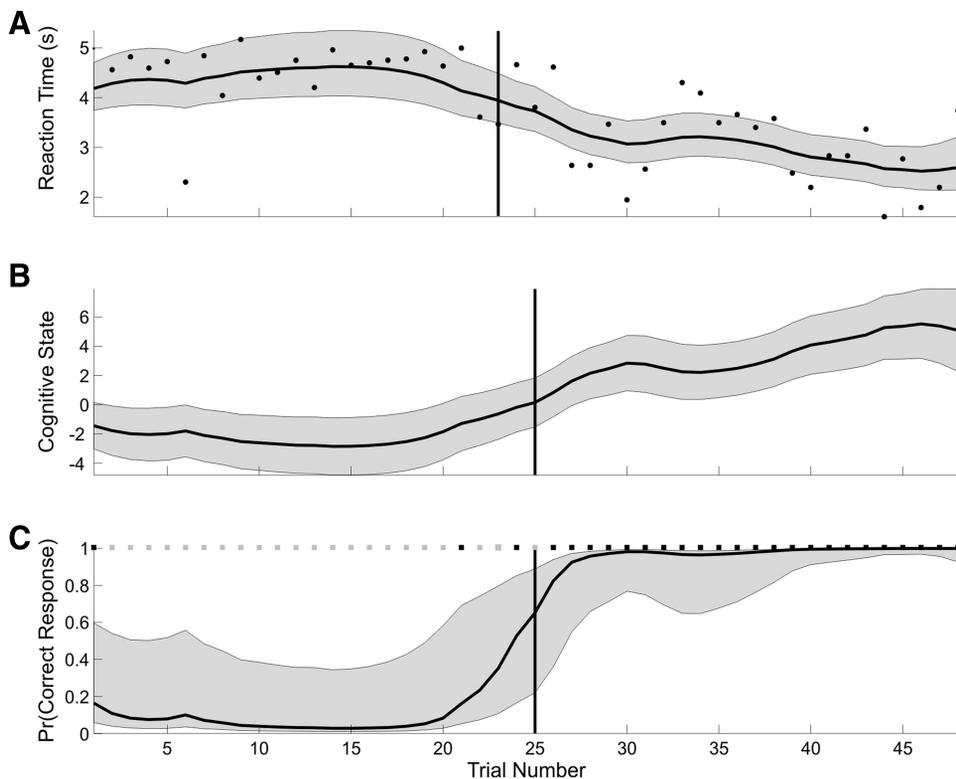


FIG. 7. A: black dots are the reaction times of the monkey executing a location-scene association task. EM algorithm estimates of the reaction-time curve (solid black curve) and its 95% confidence intervals (gray area) computed from the combined analysis of the reaction-time and binary data. The vertical line marks the IO(0.975) learning trial of 23 for the reaction-time-only analysis. B: EM algorithm cognitive state estimate computed from the combined analysis of the reaction-time and binary data. The gray areas are the 95% confidence intervals. The vertical line marks trial 25, the IO(0.975) learning trial computed using both the reaction-time and the binary data. C: the correct (black squares) and incorrect (gray squares) trial responses are shown at the top of this panel. EM algorithm estimate of the learning curve (solid black curve) and its 95% confidence intervals (gray area) computed from the combined reaction-time and binary data. The vertical line is the IO(0.975) learning trial of 25 for the binary-data-only analysis.

the binary-only learning-trial estimates of trials 23 and 25, respectively, were in close agreement with the combined model estimate of 25.

The ideal observer curve (Fig. 8A) mirrored the learning curve (Fig. 7C). The decline in certainty in the ideal observer curve between trials 1 and 5 reflected the constant reaction times and the fact that all of the binary responses on these trials except one were incorrect. The increase in certainty at trial 6 was due to the anomalously short reaction time of 2.1 s on that trial despite an incorrect answer. From trials 6 to 15, the ideal observer curve decreased as a result of the initial string of incorrect binary responses, with a sharp upward peak at trial 6 due to the single short reaction time at that trial. From trial 16 to 27, the ideal observer curve increased, and then remained essentially constant close to 1 for the remainder of the experiment. The IO(0.975) learning trial was trial 25. The between trial comparison showed that performance from trial 28 to the balance of the experiment was better than performance for all trials before trial 26 (Fig. 8B). These findings demonstrated a change in performance across the experiment that was consistent with the conclusion that the animal had learned this location-scene association.

## DISCUSSION

Continuous observations, such as reaction times and run times, and binary observations, such as correct/incorrect responses, are routinely recorded simultaneously in behavioral learning experiments. However, the two types of performance measures are not analyzed together to evaluate learning. In this work, we introduced a state-space model in which the observation model makes use of simultaneously recorded continuous and binary measures of performance to characterize learning.

We estimated the model from these simultaneously recorded performance measures by maximum likelihood using an EM algorithm. In addition to reformulating our previous definitions of the learning curve, the ideal observer curve, and the learning trial (Smith et al. 2004) in terms of the new model, we introduced the concept of the reaction- and run-time curves. We illustrated the new paradigm in the analysis of simulated data and data from actual learning experiments. In both the simulated and actual data analyses, the combined measures of performance provided highly informative measures of learning.

The technical innovation that enabled this combined model analysis is the recursive filter algorithm for mixed observation processes (i.e., continuous and binary observations) (Eqs. A4–A8), the fixed-interval smoothing algorithm, and an approximate EM algorithm for combined cognitive state and model parameter estimation. The mixed recursive filter algorithm (Prerau et al. 2008) combines the well-known Kalman filter with a recently developed binary filter (Smith and Brown 2003; Smith et al. 2004, 2005, 2007) and has both the Kalman and binary filters as special cases. The mixed recursive filter and the fixed-interval smoothing algorithm are used in the EM algorithm to estimate the model parameters, the combined model cognitive states, the reaction- or run-time curve, the learning curve, and ideal observer curve. In this way, the mixed filter makes possible a dynamic, multivariate analysis of behavioral performance data.

Unlike analyses that presuppose improvement during a learning experiment and therefore use learning models that change monotonically as a function of time or trial number, the dynamic nature of this state-space framework allows it to capture a range of behaviors. In the current work, this flexi-

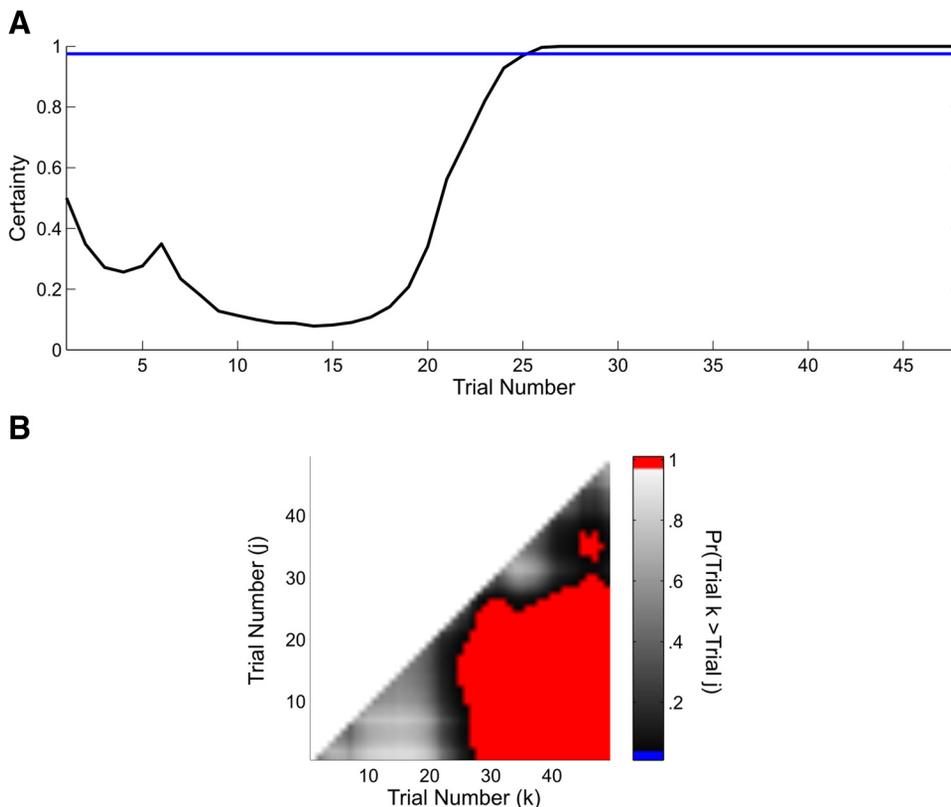


FIG. 8. *A*: ideal observer curve computed from the combined analysis of the reaction-time and binary data of the monkey executing the location-scene association task (Fig. 7). The horizontal line is 0.975. *B*: trial-by-trial comparison of the probability that the cognitive state at trial  $k$  ( $x$  axis) is greater than the cognitive state at trial  $j$  ( $y$  axis)  $< k$ , where  $k = 2, \dots, 45$  and  $j = 1, \dots, 44$ . Comparisons for which this probability is  $\geq 0.95$  are shown in red.

bility is illustrated by estimated performance indices that initially declined in the first half of the experiment, then improved in the latter half (Fig. 7).

*How well does the model work?*

In situations where the estimated cognitive state can be compared with a true cognitive state, standard goodness-of-fit procedures can be applied. However, in the actual data examples we considered, the true cognitive state was unknown. Therefore a standard goodness-of-fit analysis was not possible and alternative approaches had to be considered. For this reason, to assess performance, we conducted a simulation in which the model used to simulate the data differed from the model used in the estimation. We demonstrated that despite this difference, there was a highly accurate recovery of the underlying cognitive state process. While this result alone is not definitive, it suggests that the state-space model used in the analysis can differ from the dynamical model of the underlying cognitive process. Our approach provides a plausible way to fuse information from two different sources related by a common state process. As another performance measure, we could also use standard model selection approaches such as AIC and BIC to choose among a class of models that are appropriate for the data. In this case, we could compare the data description given by the models that had the smallest AIC and BIC. Despite use of these performance measures and model selection criteria, the interpretation of the model analysis in the context of the particular problem must also be plausible to infer that the particular model provides a reasonable data summary. A practical test of this is the extent to which the model can be used to design more efficient experiments and to conduct statistical power analyses. By making more efficient use of the data, we may be able to reduce the amount of recording required to estimate specific features of the learning process. Although we did not consider these experimental design issues here, results along these lines using state-space models for behavioral data have been reported in Smith et al. (2005).

*Future directions*

Several extensions of the current work are possible including developing models that can represent binary and continuous measures of performance along with simultaneously recorded spiking activity from multiple neurons, representing the cognitive state as a multidimensional rather than one-dimensional (Smith et al. 2006; Usher and McClelland 2001), and formulating the models in continuous rather than in discrete time. In addition, developing more complex, potentially nonmonotonic descriptions of the relationships between the cognitive states and the observations could yield models that offer insights into important phenomena such as over-training, behavioral strategies, and strategy switching (Smith et al. 2007). In this way, these more in-depth models would provide an opportunity to link our state-space methods to current work on learning strategies (Chater and Oaksford 2008; Dayan and Yu 2002; Gallistel 2008; Kakade and Dayan 2002; Smith et al. 2006; Usher and McClelland 2001).

In principle, the generalized observation models and the state models can be fit to experimental data using the EM

algorithm paradigm described here. To carry out these computations, the E-step would be most efficiently computed using Monte Carlo methods such as those described in Chan and Ledolter (1995). However, models such as these are now readily analyzed with sequential Monte Carlo (Doucet et al. 2001; Ergun et al. 2007) and Bayesian Markov Chain Monte Carlo methods (Gelman 2004). In the sequential Monte Carlo approach, the mixed filter algorithm could serve as a proposal density (Ergun et al. 2007).

Our approach provides a solution to the important problem of combining information from different measurement types to improve state estimation. Therefore it may be applied in the analysis of other problems in neuroscience in which different measurement types are being simultaneously recorded to study a neural system. For example, behavioral performance measures are frequently recorded with neural spike trains in learning experiments to establish neural correlates of behavior (Barnes et al. 2005; Wirth et al. 2003). As suggested by the model generalizations discussed in the preceding text, our paradigm offers an approach to establishing such correlates based on a joint model of the two measurement types (Smith et al. 2007). Second, neurophysiological experiments commonly record multiple neural spike trains and local field potentials simultaneously. Our paradigm may be used to develop dynamic approaches to analyzing the properties of a particular neural system by conducting a combined analysis of its measurements. Such algorithms could also make it possible to use simultaneously recorded ensemble neural spiking activity and local field potentials to improve control of neural prosthetic devices and brain machine interfaces (Hochberg et al. 2006; Srinivasan et al. 2007). Finally, a central focus of current functional neuroimaging research is the development of tools for multimodal imaging using functional magnetic resonance imaging, electroencephalography, magnetoencephalography, and diffuse optical tomography. This means conducting experiments in which imaging is performed with two or more of these modalities, simultaneously or in sequence, so that the information from the different sources can be combined (Liu et al. 1998; Molins et al. 2008). The concepts underlying our approach could be applied to this problem as well. These theoretical and applied extensions will be the topics of future investigations.

APPENDIX A. DERIVATION OF THE EM ALGORITHM

Use of the EM algorithm to compute the maximum likelihood estimates of  $\theta$  requires us to maximize the expectation of the complete data log-likelihood. The complete data likelihood is the joint probability density of  $Z, N$ , and  $x$ , which for our model is

$$\begin{aligned}
 p(Z, N, x|\theta) = & \prod_{k=1}^K p_k^{n_k} (1 - p_k)^{1-n_k} \\
 & \times \prod_{k=1}^K (2\pi\sigma_z^2)^{-\frac{1}{2}} \exp\{-2\sigma_z^2^{-1}(z_k - \alpha - \beta x_k)^2\} \\
 & \times \prod_{k=1}^K (2\pi\sigma_v^2)^{-\frac{1}{2}} \exp\{-2\sigma_v^2^{-1}(x_k - \rho x_{k-1})^2\} \tag{A1}
 \end{aligned}$$

where the first term on the right of Eq. A1 is defined by the Bernoulli probability mass function in Eq. 3, the second term is defined by the Gaussian probability density in Eq. 2, and the third term is the joint probability density of the state process defined by the Gaussian model

in Eq. 13. At iteration  $(\ell + 1)$  of the algorithm, we compute in the E-step the expectation of the complete data log likelihood given the responses  $N$  and  $Z$  across the  $K$  trials and  $\theta^{(\ell)} = (\rho^{(\ell)}, \sigma_e^{2(\ell)}, \alpha^{(\ell)}, \beta^{(\ell)}, \sigma_v^{(\ell)}, x_0^{(\ell)})$ , the parameter estimates from iteration  $\ell$ , which are defined as

E-step

$$\begin{aligned}
 & E\{\log[p(Z, N, x | \theta)] \mid Z, N, \theta^{(\ell)}\} \\
 &= E\left\{\sum_{k=1}^K \left[ n_k \log[p_k(1 - p_k)^{-1}] + \log(1 - p_k) \right] \mid Z, N, \theta^{(\ell)}\right\} \\
 &+ E\left[-\frac{1}{2}K \log(2\pi\sigma_e^2) - (2\sigma_e^2)^{-1} \sum_{k=1}^K (z_k - \alpha - \beta x_k)^2 \mid Z, N, \theta^{(\ell)}\right] \\
 &+ E\left[-\frac{1}{2}K \log(2\pi\sigma_v^2) - (2\sigma_v^2)^{-1} \sum_{k=1}^K (x_k - \rho x_{k-1})^2 \mid Z, N, \theta^{(\ell)}\right]
 \end{aligned}
 \tag{A2}$$

To evaluate the E-step we have to consider terms such as

$$\begin{aligned}
 x_{k|K} &\equiv E[x_k \mid Z, N, \theta^{(\ell)}] \\
 W_{k|K} &\equiv E[x_k^2 \mid Z, N, \theta^{(\ell)}] \\
 W_{k-1,k|K} &\equiv E[x_k x_{k-1} \mid Z, N, \theta^{(\ell)}]
 \end{aligned}
 \tag{A3}$$

for  $k = 1, \dots, K$  where the notation  $k|j$  denotes the expectation of the state variable at  $k$  given the responses up to time  $j$ . To compute these quantities efficiently, we follow a strategy first suggested by Shumway and Stoffer (1982) for linear Gaussian state and observation models and decompose the E-step into three parts: a nonlinear recursive filter algorithm to compute  $x_{k|k}$ , a fixed interval smoothing algorithm to estimate  $x_{k|K}$ , and a state-space covariance algorithm to estimate  $W_{k|K}$  and  $W_{k,k-1|K}$ .

Filter algorithm

Given  $\theta^{(\ell)}$  we can first compute recursively the state variable,  $x_{k|k}$  and its variance,  $\sigma_{k|k}^2$ . We accomplish this using the following nonlinear filter algorithm for combined continuous and binary processes as derived in Prerau et al. (2008)

$$x_{k|k-1} = \rho^{(\ell)} x_{k-1|k-1} \tag{A4}$$

$$\sigma_{k|k-1}^2 = \rho^{(\ell)2} \sigma_{k-1|k-1}^2 + \sigma_v^{2(\ell)} \tag{A5}$$

$$C_k = (\beta^{(\ell)2} \sigma_{k|k-1}^2 + \sigma_e^{2(\ell)})^{-1} \sigma_{k|k-1}^2 \tag{A6}$$

$$x_{k|k} = x_{k|k-1} + C_k [\beta^{(\ell)}(z_k - \alpha^{(\ell)} - \beta^{(\ell)} x_{k|k-1}) + \sigma_e^{2(\ell)}(n_k - p_{k|k})] \tag{A7}$$

$$\sigma_{k|k}^2 = [(\sigma_{k|k-1}^2)^{-1} + p_{k|k}(1 - p_{k|k}) + (\sigma_e^{2(\ell)})^{-1} \beta^{(\ell)2}]^{-1} \tag{A8}$$

for  $k = 1, \dots, K$ .

Fixed interval smoothing algorithm

Given the sequence of posterior mode estimates  $x_{k|k}$  (Eq. A7) and the variance  $\sigma_{k|k}^2$  (Eq. A8), we use the fixed-interval smoothing algorithm (Smith et al. 2004) to compute  $x_{k|K}$  and  $\sigma_{k|K}$ . This smoothing algorithm is

$$x_{k|K} = x_{k|k} + A_k(x_{k+1|K} - x_{k+1|k}) \tag{A9}$$

$$A_k = \sigma_{k|k}^2 (\sigma_{k+1|k}^2)^{-1} \tag{A10}$$

$$\sigma_{k|K}^2 = \sigma_{k|k}^2 + A_k^2 (\sigma_{k+1|k}^2 - \sigma_{k+1|K}^2) \tag{A11}$$

for  $k = K - 1, \dots, 1$  and initial conditions  $x_{k|K}$  and  $\sigma_{k|K}^2$ .

State-space covariance algorithm

The covariance estimate,  $\sigma_{k,u|K}$ , can be computed from the state-space covariance algorithm (Dejong and Mackinnon 1988) and is given as

$$\sigma_{k,u|K} = A_k \sigma_{k+1,u|K} \tag{A12}$$

for  $1 \leq k \leq u \leq K$ .

The variance and covariance terms required for the E-step are

$$W_{k|K} = \sigma_{k|K}^2 + x_{k|K}^2 \tag{A13}$$

$$W_{k-1,k|K} = \sigma_{k-1,k|K} + x_{k-1|K} x_{k|K} \tag{A14}$$

In the M-step, we maximize the expected value of the complete data log likelihood in Eq. A2 with respect to  $\theta^{(\ell+1)}$  giving

M-step

$$\rho^{(\ell+1)} = \sum_{k=1}^K W_{k-1,k|K} \left[ \sum_{k=1}^K W_{k-1|K} \right]^{-1} \tag{A15}$$

$$x_0^{(\ell+1)} = \rho x_{1|K} \tag{A16}$$

$$\begin{aligned}
 \sigma_e^{2(\ell+1)} &= K^{-1} \sum_{k=1}^K z_k^2 + K \alpha^{2(\ell+1)} \\
 &+ \beta^{2(\ell+1)} \sum_{k=1}^K W_{k|K} - 2\alpha^{(\ell+1)} \sum_{k=1}^K z_k \\
 &- 2\beta^{(\ell+1)} \sum_{k=1}^K x_{k|K} z_k + 2\alpha^{(\ell+1)} \beta^{(\ell+1)} \sum_{k=1}^K x_{k|K}
 \end{aligned}
 \tag{A17}$$

$$\begin{bmatrix} \alpha^{(\ell+1)} \\ \beta^{(\ell+1)} \end{bmatrix} = \begin{bmatrix} K & \sum_{k=1}^K x_{k|K} \\ \sum_{k=1}^K x_{k|K} & \sum_{k=1}^K W_{k|K} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{k=1}^K z_k \\ \sum_{k=1}^K x_{k|K} z_k \end{bmatrix} \tag{A18}$$

$$\sigma_v^2 = K^{-1} \sum_{k=1}^K [W_{k|K} - 2\rho W_{k-1,k|K} + \rho^2 W_{k-1|K}] \tag{A19}$$

The algorithm is iterated between the E-step (Eq. A2) and the M-step (Eqs. A16 and A17) using the filter algorithm, the fixed interval smoothing algorithm and the state covariance algorithm to evaluate the E-step. The maximum likelihood estimate of  $\hat{\theta} = \theta^{(\infty)}$ . The convergence criteria for the algorithm are those used in Smith and Brown (2003). The fixed-interval-smoothing algorithm evaluated at the maximum likelihood estimate of  $\theta$  (Eqs. A10–A12) together with Eqs. 5 and 6 give the maximum likelihood or empirical Bayes' estimates of the reaction-time curves and the learning curve.

APPENDIX B. PROBABILITY DENSITIES OF THE REACTION-TIME AND THE LEARNING CURVES

To compute the probability density of the reaction time  $r_k$  from the posterior probability density of the state process  $x_k$ , we use the standard change-of-variables formula (Rice 2007)

$$p(r_k \mid Z, N, \hat{\theta}) = p(x_k \mid Z, \hat{\theta}) \left| \frac{dx_k}{dr_k} \right| \tag{B1}$$

where  $p(x_k \mid Z, N, \hat{\theta})$  is the Gaussian probability density mean  $x_{k|K}$  and variance  $\sigma_{k|K}^2$ . Applying (Eq. A19) by substituting  $x_k = \beta^{-1}(\log r_k - \alpha)$  into  $p(x_k \mid Z, N, \hat{\theta})$  and multiplying by  $dx_k/dr_k = (\beta r_k)^{-1}$  yields the lognormal probability density for the reaction-time curve in Eq. 5. From the approximate posterior probability density of  $x_k$ , we can also compute the probability density of  $p_k$  by using Eq. 4 and the change-of-variable formula

$$p(p_k|Z, N, \hat{\theta}) = p(x_k|Z, N, \hat{\theta}) \left| \frac{dx_k}{dp_k} \right| \quad (B2)$$

Hence setting  $x_k = \{\log[p_k(1 - p_k)^{-1}] - \mu\}$  in  $p(x_k|Z, N, \hat{\theta})$  and computing  $dx_k/dp_k = [p_k(1 - p_k)^{-1}]^{-1}$ , we obtain Eq. 6.

GRANTS

Support was provided by National Institutes of Health Grants DA-015644 to E. N. Brown and W. Suzuki, DPIO003646, MH-59733 and MH-071847 to E. N. Brown, and P50 NS-38372 to A. M. Graybiel.

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.

**Chan KS, Ledolter J.** Monte-Carlo Em estimation for time-series models involving counts. *J Am Stat Assoc* 90: 242–252, 1995.

**Chater N, Oaksford M.** *The Probabilistic Mind: Prospects for Bayesian Cognitive Science.* Oxford, UK: Oxford Univ. Press, 2008.

**Dayan P, Yu A.** Expected and unexpected uncertainty: Ach and NE in the neocortex. In: *Advances in Neural Information Processing Systems*, edited by Dietterich T, Becker S, Ghahramani Z. Cambridge, MA: MIT Press, 2002, p. 189–196.

**Dejong P, Mackinnon MJ.** Covariances for smoothed estimates in state-space models. *Biometrika* 75: 601–602, 1988.

**Dempster AP, Laird NM, Rubin DB.** Maximum likelihood from incomplete data via Em algorithm. *J Roy Stat Soc B Met* 39: 1–38, 1977.

**Dias R, Robbins TW, Roberts AC.** Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from “on-line” processing. *J Neurosci* 17: 9285–9297, 1997.

**Doucet A, De Freitas N, Gordon N.** *Sequential Monte Carlo Methods in Practice.* New York: Springer, 2001.

**Durbin J, Koopman SJ.** *Time Series Analysis by State Space Methods.* Oxford, UK: Oxford Univ. Press, 2001.

**Dusek JA, Eichenbaum H.** The hippocampus and memory for orderly stimulus relations. *Proc Natl Acad Sci USA* 94: 7109–7114, 1997.

**Ergun A, Barbieri R, Eden UT, Wilson MA, Brown EN.** Construction of point process adaptive filter algorithms for neural systems using sequential Monte Carlo methods. *IEEE Trans Biomed Eng* 54: 419–428, 2007.

**Fahrmeir L, Tutz G.** *Multivariate Statistical Modelling Based on Generalized Linear Models.* New York: Springer, 2001.

**Fox MT, Barense MD, Baxter MG.** Perceptual attentional set-shifting is impaired in rats with neurotoxic lesions of posterior parietal cortex. *J Neurosci* 23: 676–681, 2003.

**Gallistel CR.** Learning and representation. In: *Learning Theory and Behavior. Learning and Memory- A Comprehensive Reference*, edited by Byrne J. Oxford: Elsevier, 2008, vol. 1, p. 227–242.

**Gallistel CR, Fairhurst S, Balsam P.** The learning curve: implications of a quantitative analysis. *Proc Natl Acad Sci USA* 101: 13124–13131, 2004.

**Gelman A.** *Bayesian Data Analysis.* Boca Raton, FL: Chapman and Hall/CRC, 2004.

**Hochberg LR, Serruya MD, Friehs GM, Mukand JA, Saleh M, Caplan AH, Branner A, Chen D, Penn RD, Donoghue JP.** Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* 442: 164–171, 2006.

**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 autoshaping. *Psychol Rev* 109: 533–544, 2002.

**Kitagawa G, Gersch W.** *Smoothness Priors Analysis of Time Series.* New York: Springer, 1996.

**Law JR, Flanery MA, Wirth S, Yanike M, Smith AC, Frank LM, Suzuki WA, Brown EN, Stark CEL.** Functional magnetic resonance imaging activity during the gradual acquisition and expression of paired-associate memory. *J Neurosci* 25: 5720–5729, 2005.

**Liu AK, Belliveau JW, Dale AM.** Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. *Proc Natl Acad Sci USA* 95: 8945–8950, 1998.

**Mendel JM.** *Lessons in Estimation Theory for Signal Processing, Communications, and Control.* Englewood Cliffs, NJ: Prentice Hall PTR, 1995.

**Molins A, Stufflebeam SM, Brown EN, Hamalainen MS.** Quantification of the benefit from integrating MEG and EEG data in minimum l(2)-norm estimation. *Neuroimage* 42: 1069–1077, 2008.

**Prerau MJ, Smith AC, Eden UT, Yanike M, Suzuki WA, Brown EN.** A mixed filter algorithm for cognitive state estimation from simultaneously recorded continuous and binary measures of performance. *Biol Cybern* 99: 1–14, 2008.

**Rice JA.** *Mathematical Statistics and Data Analysis.* Belmont, CA: Thomson/Brooks/Cole, 2007.

**Roman FS, Simonetto I, Soumireuourat B.** Learning and memory of odor-reward associatio—selective impairment following horizontal diagonal band lesions. *Behav Neurosci* 107: 72–81, 1993.

**Rondi-Reig L, Libbey M, Eichenbaum H, Tonegawa S.** CA1-specific N-methyl-d-aspartate receptor knockout mice are deficient in solving a nonspatial transverse patterning task. *Proc Natl Acad Sci USA* 98: 3543–3548, 2001.

**Shumway R, Stoffer D.** An approach to time series smoothing and forecasting using the EM algorithm. *J Time Ser Anal* 3: 253–264, 1982.

**Siegel S, Castellan NJ.** *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill, 1988.

**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 WA, 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 AC, Wirth S, Suzuki WA, Brown EN.** Bayesian analysis of interleaved learning and response bias in behavioral experiments. *J Neurophysiol* 97: 2516–2524, 2007.

**Smith MA, Ghazizadeh A, Shadmehr R.** Interacting adaptive processes with different time scales underlie short-term motor learning. *Plos Biol* 4: 1035–1043, 2006.

**Srinivasan L, Eden UT, Mitter SK, Brown EN.** General-purpose filter design for neural prosthetic devices. *J Neurophysiol* 98: 2456–2475, 2007.

**Stefani MR, Groth K, Moghaddam B.** Glutamate receptors in the rat medial prefrontal cortex regulate set-shifting ability. *Behav Neurosci* 117: 728–737, 2003.

**Usher M, McClelland JL.** The time course of perceptual choice: the leaky, competing accumulator model. *Psychol Rev* 108: 550–592, 2001.

**Whishaw IQ, Tomie JA.** Acquisition and retention by hippocampal rats of simple, conditional, and configural tasks using tactile and olfactory cues—implications for hippocampal function. *Behav Neurosci* 105: 787–797, 1991.

**Williams ZM, Eskandar EN.** Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat Neurosci* 9: 562–568, 2006.

**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: 1578–1581, 2003.

**Wise SP, Murray EA.** Role of the hippocampal system in conditional motor learning: mapping antecedents to action. *Hippocampus* 9: 101–117, 1999.