

# Dynamic Analysis of Learning in Behavioral Experiments

Anne C. Smith,<sup>1,2</sup> Loren M. Frank,<sup>1,2</sup> Sylvia Wirth,<sup>3</sup> Marianna Yanike,<sup>3</sup> Dan Hu,<sup>4</sup> Yasuo Kubota,<sup>4</sup> Ann M. Graybiel,<sup>4</sup> Wendy A. Suzuki,<sup>3</sup> and Emery N. Brown<sup>1,2</sup>

<sup>1</sup>Neuroscience Statistics Research Laboratory, Department of Anesthesia and Critical Care, Massachusetts General Hospital, Boston, Massachusetts 02114, <sup>2</sup>Division of Health Sciences and Technology, Harvard Medical School–Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, <sup>3</sup>Center for Neural Science, New York University, New York, New York 10003, and <sup>4</sup>Department of Brain and Cognitive Sciences and McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Understanding how an animal's ability to learn relates to neural activity or is altered by lesions, different attentional states, pharmacological interventions, or genetic manipulations are central questions in neuroscience. Although learning is a dynamic process, current analyses do not use dynamic estimation methods, require many trials across many animals to establish the occurrence of learning, and provide no consensus as how best to identify when learning has occurred. We develop a state–space model paradigm to characterize learning as the probability of a correct response as a function of trial number (learning curve). We compute the learning curve and its confidence intervals using a state–space smoothing algorithm and define the learning trial as the first trial on which there is reasonable certainty ( $>0.95$ ) that a subject performs better than chance for the balance of the experiment. For a range of simulated learning experiments, the smoothing algorithm estimated learning curves with smaller mean integrated squared error and identified the learning trials with greater reliability than commonly used methods. The smoothing algorithm tracked easily the rapid learning of a monkey during a single session of an association learning experiment and identified learning 2 to 4 d earlier than accepted criteria for a rat in a 47 d procedural learning experiment. Our state–space paradigm estimates learning curves for single animals, gives a precise definition of learning, and suggests a coherent statistical framework for the design and analysis of learning experiments that could reduce the number of animals and trials per animal that these studies require.

**Key words:** learning; behavior; state–space model; hidden Markov model; change-point test; association task; EM algorithm

Learning is a dynamic process that generally can be defined as a change in behavior as a result of experience. Learning is usually investigated by showing that an animal can perform a previously unfamiliar task with reliability greater than what would be expected by chance. The ability of an animal to learn a new task is commonly tested to study how brain lesions (Whishaw and Tomie, 1991; Roman et al., 1993; Dias et al., 1997; Dusek and Eichenbaum, 1997; Wise and Murray, 1999; Fox et al., 2003), attentional modulation (Cook and Maunsell, 2002), genetic manipulations (Rondi-Reig et al., 2001), or pharmacological interventions (Stefani et al., 2003) alter learning. Characterizations of the learning process are also important to relate an animal's behavioral changes to changes in neural activity in target brain regions (Jog et al., 1999; Wirth et al., 2003).

In a learning experiment, behavioral performance can be analyzed 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 suggesting learning can be documented using a statistical criterion (Siegel and Castellan, 1988; Jog et al., 1999; Wirth et al., 2003). Methods for estimating the learning curve typically require multiple trials measured in multiple animals (Wise and Murray, 1999; Stefani et al., 2003), whereas learning curve estimates do not provide confidence intervals. Among the currently used methods, there is no consensus as to which identifies the learning trial most accurately and reliably. In many, if not most, experiments, the subject's trial responses are binary, i.e., correct or incorrect. Although dynamic modeling has been used to study learning with continuous-valued responses, such as reaction times (Gallistel et al., 2001; Kakade and Dayan, 2002; Yu and Dayan, 2003), they have not been applied to learning studies with binary responses.

To develop a dynamic approach to analyzing learning experiments with binary responses, we introduce a state–space model of learning in which a Bernoulli probability model describes behavioral task responses and a Gaussian state equation describes the unobservable learning state process (Kitagawa and Gersh, 1996; Kakade and Dayan, 2002). The model defines the learning curve as the probability of a correct response as a function of the state process. We estimate the model by maximum likelihood using the expectation maximization (EM) algorithm (Dempster et al., 1977), compute both filter algorithm and smoothing algorithm estimates of the learning curve, and give a precise statistical

Received June 11, 2003; revised Oct. 8, 2003; accepted Oct. 20, 2003.

This work was supported by National Institute on Drug Abuse Grant DA015644 (E.N.B., W.A.S.), National Institute of Mental Health Grants MH59733 and MH61637 (E.N.B.), MH58847 (W.A.S.), MH60379 (A.M.G.), and MH65108 (L.M.F.), and a McKnight Foundation grant (W.A.S.). We are grateful to Peter Dayan and Howard Eichenbaum for helpful suggestions that significantly improved the work in this article.

Correspondence should be addressed to Emery N. Brown, Neuroscience Statistics Research Laboratory, Department of Anesthesia and Critical Care, Massachusetts General Hospital, 55 Fruit Street, Clinics 3, Boston, MA 02114-2696. E-mail: brown@neurostat.mgh.harvard.edu.

L. M. Frank's present address: Keck Center for Integrative Neuroscience, Department of Physiology, University of California, San Francisco, San Francisco, CA 94143.

DOI:10.1523/JNEUROSCI.2908-03.2004

Copyright © 2004 Society for Neuroscience 0270-6474/04/240447-15\$15.00/0

definition of the learning trial of the experiment. We compare our methods with learning defined by a moving average method, the change-point test, and a specified number of consecutive correct responses method in a simulation study designed to reflect a range of realistic experimental learning scenarios. We illustrate our methods in the analysis of a rapid learning experiment in which a monkey learns new associations during a single session and in a slow learning experiment in which a rat learns a T-maze task over many days.

**Materials and Methods**

*A 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 (Kitagawa and Gersh, 1996; Smith and Brown, 2003). The state–space model consists of two equations: a state equation and an observation equation. The state equation defines an unobservable learning process whose evolution is tracked across the trials in the experiments. Such state models with unobservable processes are often referred to as hidden Markov or latent process models (Roweis and Gharamani, 1999; Fahrmeir and Tutz, 2001; Smith and Brown, 2003). We formulated the learning state process so that it increases as learning occurs and decreases when it does not occur. From the learning state process, we compute a 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 process increases the probability of a correct response, and a decrease in the learning process decreases the probability of a correct response. The observation equation completes the state–space model setup and defines how the observed data relate to the unobservable learning state process. The data we observe in the learning experiment are the series of correct and incorrect responses as a function of trial number. Therefore, the objective of the analysis is to estimate the learning state process and, hence, the learning curve from the observed data.

We conduct our analysis of the experiment from the perspective of an ideal observer. That is, we estimate the learning state process at each trial after seeing the outcomes of all of the trials in the experiment. This approach is different from estimating learning from the perspective of the subject executing the task, in which case, the inference about when learning occurs is based on the data up to the current trial (Kakade and Dayan, 2002; Yu and Dayan, 2003). Identifying when learning occurs is therefore a two-step process. In the first step, we estimate from the observed data the learning state process and, hence, the learning curve. In the second step, we estimate when learning occurs by computing the confidence intervals for the learning curve or, equivalently, by computing for each trial the ideal observer’s assessment of the probability that the subject performs better than chance.

To define the state–space model, we assume that there are  $K$  trials in a behavioral experiment, and we index the trials by  $k$  for  $k = 1, \dots, K$ . To define the observation equation, we let  $n_k$  denote the response on trial  $k$ , where  $n_k = 1$  is a correct response, and  $n_k = 0$  is an incorrect response. We let  $p_k$  denote the probability of a correct response  $k$ . We assume that the probability of a correct response on trial  $k$  is governed by an unobservable learning state process  $x_k$ , which characterizes the dynamics of learning as a function of trial number. At trial  $k$ , the observation model defines the probability of observing  $n_k$ , i.e., either a correct or incorrect response, given the value of the state process  $x_k$ . The observation model can be expressed as the Bernoulli probability mass function:

$$Pr(n_k | p_k, x_k) = p_k^{n_k} (1 - p_k)^{1 - n_k}, \tag{2.1}$$

where  $p_k$  is defined by the logistic equation:

$$p_k = \frac{\exp(\mu + x_k)}{1 + \exp(\mu + x_k)}, \tag{2.2}$$

and  $\mu$  is determined by the probability of a correct response by chance in the absence of learning or experience. We define the unobservable learning state process as a random walk:

$$x_k = x_{k-1} + \epsilon_k, \tag{2.3}$$

where the  $\epsilon_k$  are independent Gaussian random variables with mean 0 and variance  $\sigma_\epsilon^2$ .

Formulation of the probability of a correct response on each trial as a logistic function of the learning state variable (Eq. 2.3) ensures that, at each trial, the probability is constrained between 0 and 1. The state model (Eq. 2.3) provides a continuity constraint (Kitagawa and Gersh, 1996) so that the current state of learning and, hence, the probability of a correct response in the current trial depend on the previous state of learning or experience. Under the random walk model, the expected value of  $x_k$  given  $x_{k-1}$  is  $x_{k-1}$ . Therefore, in the absence of learning, the expected probability of a correct response at trial  $k$  is  $p_{k-1}$ . In other words, the Gaussian random walk model enforces the plausible assumption that immediately before trial  $k$ , the probability of a correct response on trial  $k$  is simply the probability from the previous trial  $k - 1$ . We compute the parameter  $\mu$  before each experiment from  $p_0$ , the probability of a correct response occurring by chance at the outset of the experiment. To do so, we note that the parameter  $x_0$  describes the subject’s learning state before the first trial in the experiment. We set  $x_0 = 0$ , and then by Equation 2.2,  $\mu = \log[p_0(1 - p_0)^{-1}]$ . For example, given a particular visual cue, if a subject has five possible response choices, then there is 0.2 probability of a correct response by chance at the start of the experiment. In this case, we have  $\mu = \log(0.2(0.8)^{-1}) = -1.3863$ . Choosing  $\mu$  this way ensures that  $x_0 = 0$  means that the subject uses a random strategy at the outset of the experiment. In each analysis, we estimate  $x_0$  because the subject may have a response bias or may be using a specific nonrandom strategy. The parameter  $\sigma_\epsilon^2$  governs how rapidly changes can occur from trial to trial in the unobservable learning state process and in the probability of a correct response. As we describe next, the value of  $\sigma_\epsilon^2$  is estimated from the set of trial responses in an experiment.

In the learning experiment, we set the number of trials  $K$ , and we observe  $N_{1:K} = \{n_1, \dots, n_K\}$ , the responses for each of the  $K$  trials. The objective of our analysis is to estimate  $x = \{x_0, x_1, \dots, x_K\}$  and  $\sigma_\epsilon^2$  from these data to estimate  $p_k$  for  $k = 1, \dots, K$ . That is, if we can estimate  $x$  and  $\sigma_\epsilon^2$ , then by Equation 2.2, we can compute the probability of a correct response as a function of trial number given the data. Because  $x$  is unobservable and  $\sigma_\epsilon^2$  is a parameter, we use the 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. We used the EM algorithm to estimate state–space models from point process observations with linear Gaussian state processes (Smith and Brown, 2003). Our EM algorithm is a special case of the one by Smith and Brown (2003), and its derivation is given in Appendix A.

*Estimation of the learning curves*

Because of how we compute the maximum likelihood estimate of  $\sigma_\epsilon^2$  using the EM algorithm, we derive two estimates of each  $x_k$  for  $k = 1, \dots, K$ . The first is  $x_{k|k}$  and comes from the filter algorithm in Appendix A (Eqs. A.6–A.9). The second,  $x_{k|K}$ , comes from the fixed-interval smoothing algorithm (Eqs. A.10–A.12) and is both the maximum likelihood and empirical Bayes estimate (Fahrmeir and Tutz, 2001). The notation  $x_{k|j}$  means the learning state process estimate at trial  $k$  given the data up through trial  $j$ . The filter algorithm estimate is the estimate of  $x_k$  at trial  $k$ , given  $N_{1:k}$ , the data up through trial  $k$  with the true parameter  $\sigma_\epsilon^2$  replaced by its maximum likelihood estimate. The smoothing algorithm estimate at trial  $k$  is the estimate of  $x_k$  given  $N_{1:K}$ , all of the data in the experiment with the true parameter  $\sigma_\epsilon^2$  replaced by its maximum likelihood estimate. Hence, the filter algorithm (Kakade and Dayan, 2002) gives the state estimate of the subject, whereas the smoothing algorithm gives the estimate of the ideal observer.

The filter algorithm estimates the learning state at time  $k$  as the Gaussian random variable with mean  $x_{k|k}$  (Eq. A.8) and variance  $\sigma_{k|k}^2$  (Eq. A.9), whereas the smoothing algorithm estimates the state as the Gaussian random variable with mean  $x_{k|K}$  (Eq. A.10) and variance,  $\sigma_{k|K}^2$  (Eq. A.12). Because our analysis gives two estimates of  $x_k$ , by using Equation 2.2, we can obtain two estimates of  $p_k$ , namely, the filter algorithm estimate  $p_{k|k}$  and the smoothing algorithm estimate  $p_{k|K}$ . Similarly,  $p_{k|k}$  defines the

probability of a correct response at trial  $k$  given the data  $N_{1:k} = \{n_1, \dots, n_k\}$  up through trial  $k$ , and  $p_{k|K}$  defines the probability of a correct response at trial  $k$  given all of the data  $N_{1:K} = \{n_1, \dots, n_K\}$  in the experiment. We can therefore compute the probability density of any  $p_{k|j}$  using Equation 2.2 and the standard change of variables formula from elementary probability theory, where  $j = k$  denotes the filter algorithm estimate, and  $j = K$  is the smoothing algorithm estimate. Applying the change of variable formula to the Gaussian probability density with mean  $x_{k|j}$  and variance  $\sigma_{k|j}^2$  yields the following:

$$f(p|\mu, x_{k|j}, \sigma_{k|j}^2) = [(2\pi\sigma_{k|j}^2)^{-1/2} p(1-p)]^{-1} \exp\left(-\frac{1}{2\sigma_{k|j}^2} [\log[p/(1-p)\exp(\mu)]^{-1}] - x_{k|j}\right)^2. \quad (2.4)$$

Equation 2.4 is the probability density for the correct response probability at trial  $k$  using either the filter algorithm ( $j = k$ ) or the smoothing algorithm ( $j = K$ ) and is derived in Appendix B. Therefore, we define the learning curve on the basis of the filter (smoothing) algorithm as the sequence of trial estimates  $p_{k|k}$  ( $p_{k|K}$ ), where  $p_{k|j}$  is the mode (most likely value) of the probability density in Equation 2.4 for  $k = 1, \dots, K$  and  $j = k$  or  $j = K$ .

#### Identification of the learning trial

Having completed the first step of estimating the learning curve, we identify the learning trial by computing for each trial the ideal observer's assessment of the probability that the subject performs better than chance or, equivalently, by computing the confidence intervals for the learning curve. We define the trial on which learning occurs as the first trial for which the ideal observer can state with reasonable certainty that the subject performs better than chance from that trial to the end of the experiment. For our analyses, we define a level of reasonable certainty as 0.95 and term this trial the ideal observer learning trial with level of certainty 0.95 [IO (0.95)].

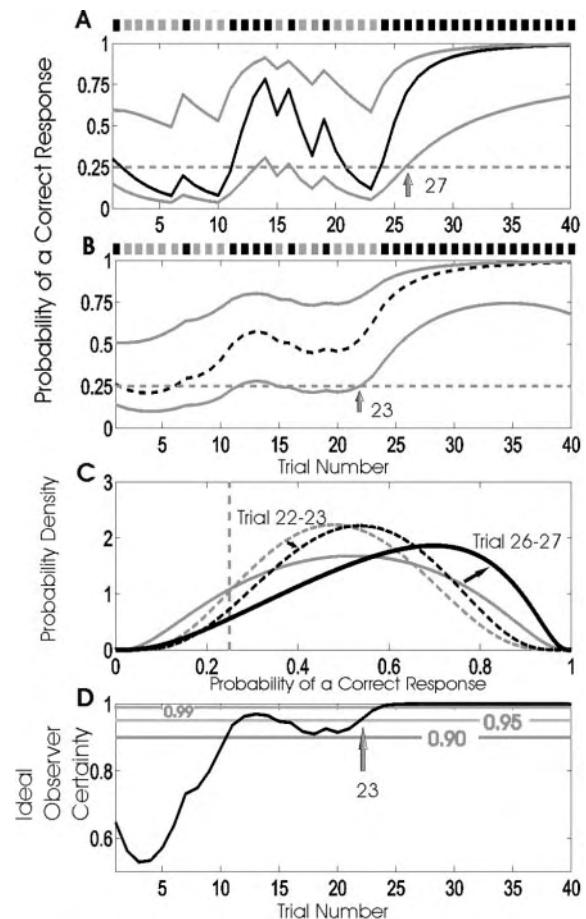
To identify the ideal observer learning trial, we first construct confidence intervals for  $p_k$ . The ideal observer learning trial is the first trial on which the lower 95% confidence bound for the probability of a correct response is greater than chance  $p_0$  and remains above  $p_0$  for the balance of the experiment. This definition takes account of the fact that the probability of a correct response on a trial is estimated and that there is uncertainty in that estimation. That is, the ideal observer (or the smoothing algorithm) estimates the probability of a correct response on each trial  $k$  with error. Therefore, we ask, what is the smallest the true probability of a correct response can be on trial  $k$ ? If the smallest value the ideal observer is 95% sure of (the lower 95% confidence bound) is greater than  $p_0$ , then we conclude that the performance on that trial is better than chance. Because the ideal observer can observe the outcomes of the entire experiment, he/she can make certain that the lower 95% confidence bound exceeds  $p_0$  from a given trial through the balance of the experiment. If the smallest value we are 95% sure of is less than  $p_0$ , then the ideal observer cannot distinguish the subject's performance from what can be expected by chance, and he/she cannot conclude that the subject has learned.

Because our analysis also provides  $p_{k|k}$ , the filter algorithm estimate of  $p_k$ , we can construct a definition of the learning trial using this estimate and its associated confidence intervals as well.

The Matlab software (MathWorks, Natick, MA) used to implement the methods presented here is available at our website (<http://neurostat.mgh.harvard.edu/BehavioralLearning/Matlabcode>).

#### An illustration of learning curve estimation and learning trial identification

Figure 1 illustrates use of the filter algorithm (Fig. 1A) and the smoothing algorithm (Fig. 1B) to estimate the learning curve in a simulated learning experiment consisting of 40 trials, in which the probability of a correct response occurring by chance is 0.25 (Fig. 1A,B, horizontal dashed lines). The trial responses are shown above the figures as gray and black marks, corresponding, respectively, to incorrect and correct responses. In the first 10 trials, there are two correct responses, followed by a sequence of four correct responses beginning at trial 11. Beginning at trial



**Figure 1.** Example of the filter algorithm (A) and the smoothing algorithm (B) applied in the analysis of a simulated learning experiment. The correct and incorrect responses are shown, respectively, by black and gray marks above the panels. The probability of a correct response occurring by chance is 0.25 (dashed horizontal line). Black lines are the learning curve estimates, and the gray lines are the associated 90% confidence intervals. The 90% confidence intervals are defined by the upper and lower 95% confidence bounds. The learning trial is defined as the trial on which the lower 95% confidence bound exceeds 0.25 and remains above 0.25 for the balance of the experiment. The filter algorithm identified trial 27 as the learning trial (arrow in A), whereas the smoothing algorithm, which used all of the data, identified trial 23 as the ideal observer learning trial with level of certainty 0.95 (arrow in B). The confidence limits at a given trial were constructed from the probability density of a correct response at that trial using Equations 2.4 and B.4. The probability densities of the probability of a correct response at the learning trial and the trial immediately preceding the learning trial are shown in C for both the filter (solid lines) and smoothing (dashed lines) algorithms. For the filter algorithm, the learning trial was 27 (C, solid black line) and the preceding trial was 26 (solid gray line), whereas for the smoothing algorithm, the IO (0.95) learning trial was trial 23 (C, dashed black line) and the preceding trial was 22 (dashed gray line). D shows the level of certainty the ideal observer has that the animal's performance is better than chance at each trial. From trial 23 on, the ideal observer is 0.95 certain that the performance is better than chance, whereas this observer can be 0.90 certain of performance better than chance from trial 12 on.

15, there are two correct responses until trial 23, after which all of the responses are correct. The smoothing algorithm learning curve estimate (Fig. 1B, dashed black line) is smoother than that of the filter algorithm learning curve (Fig. 1A, solid black line), and its 90% confidence intervals (Fig. 1B, gray lines) are narrower than those of the filter algorithm (Fig. 1A, gray lines) because the smoothing algorithm intervals are based on all of the data in the experiment.

The learning trials for the filter algorithm and IO (0.95) (Fig. 1A,B, arrows) are trials 27 and 23, respectively. The lower confidence bounds of the filter and smoothing algorithms first exceed the probability of a correct response by chance at trials 14 and 12, respectively. However, because the lower confidence bounds for the two estimates do not remain

above 0.25 until trials 23 and 27, these latter trials are, respectively, the filter algorithm and IO (0.95) learning trial estimates. Given either the known or estimated value of  $\sigma_e^2$ , the filter algorithm estimates the learning trial as later in the experiment than the IO (0.95) because the filter algorithm estimate at trial  $k$  uses only the data collected from the start of the experiment up through trial  $k$ . The learning curve estimate at the last trial  $K$  and the associated confidence intervals are the same for both algorithms because, by design (Eqs. A.10–A.12), the smoothing algorithm performs its state estimation in reverse from trial  $K - 1$  to 1 starting with the filter estimate at trial  $K$ .

To illustrate how the confidence bounds and the learning trial are computed in Figure 1, *A* and *B*, Figure 1*C* shows the probability density of the probability of a correct response at the learning trial and the trial immediately preceding the learning trial for both the filter (solid lines) and the smoothing (dashed lines) algorithms. For the filter algorithm, the learning trial is 27 (Fig. 1*C*, solid black line) and the preceding trial is 26 (solid gray line), whereas for the IO (0.95), learning trial is 23 (Fig. 1*C*, dashed black line) and the preceding trial is 22 (dashed gray line). That is, the areas under both gray curves from 0 to 0.25 are  $>0.05$ , whereas the areas under the corresponding black curves are  $<0.05$ . The learning state process estimates are Gaussian densities by Equations A.6 and A.9 and Equations A.10 and A.12. The probability density of the probability of a correct response is non-Gaussian by Equation 2.4. The closer the mode of Equation 2.4 is to either 0 or 1, the more non-Gaussian this probability density is. This is why the probability density for trial 27 is more skewed than the probability density for trial 23.

Just as the learning curve is a measure of the probability of a correct response as a function of trial number, we compute from the probability density in Equation 2.4 the probability that the ideal observer views the animal's performance as better than chance as a function of trial number (Eq. B.4) (Fig. 1*D*). This curve can be displayed in each analysis or it can be inferred by the distance at each trial between the lower confidence bound of the learning curve and the constant line denoting the probability of a correct response by chance (Fig. 1*B*). We have chosen the trial on which this probability first crosses 0.95 and remains above this probability for the balance of the experiment as the IO (0.95) learning trial. At trial 12, the probability that the animal is performing better than chance first crosses 0.95. Immediately after this trial, it dips below 0.95 until trial 23, at which point it passes and remains greater than 0.95 for the balance of the experiment. This is why, in this experiment, we define trial 23 as the learning trial. Between trials 12 and 22, the ideal observer has a high level of confidence ( $>0.90$ ) that the animal's performance is better than chance but not greater than 0.95 for this entire interval. To apply our definition, the investigator must specify the desired level of certainty. In all of our analyses, we use 0.95. Choosing a higher level of certainty, such as 0.99 when we wish to ensure overlearning of a task, will tend to move the learning trial to a later trial in the experiment. For this experiment, a 0.99 level of certainty would identify the learning trial at trial 24. Choosing a lower level of certainty, such as 0.90, will tend to move the learning trial to an earlier trial. In this experiment, that would be trial 12.

#### Alternative methods for estimating learning

We compared our algorithm with three methods commonly used to estimate learning: the moving average method, the change-point test for binary observations, and the fixed number of consecutive correct responses method. Although all three methods estimate learning by conducting a hypothesis test, only the moving average method can be used to estimate a learning curve.

**Moving average method.** This technique, based on using the binomial test in a sliding window, has been used frequently to identify a learning trial (Eichenbaum et al., 1986). The moving average method estimates the learning curve by computing in a series of overlapping windows of length  $2w + 1$  the probability of a correct response at trial  $k$  as follows:

$$p_k = (2w + 1)^{-1} \sum_{i=k-w}^{k+w} n_i. \quad (2.5)$$

Because the response at trial  $k$  is an average of the responses in the trials on both sides, the resulting learning curve can only be estimated from trial  $w + 1$  to trial  $K - w$ . To estimate the learning trial, this method uses the binomial probability distribution to compute in each window the probability of seeing the observed number of correct responses in the window under the null hypothesis of no learning with the probability of a correct response being  $p_0$ . Using the moving average estimation formula in Equation 2.5, the learning trials are identified as the middle trial in the windows for which the probability of seeing the observed number of correct responses is 0.05 or less. To illustrate, we chose  $w = 4$ , giving a window length of nine trials. Hence, for  $p_0 = 0.125, 0.25$ , and  $0.5$ , we require four, five, and eight correct responses within a nine-trial window to identify the middle trial of that window as the learning trial.

The simplicity of the moving average method makes it highly appealing. However, it uses multiple statistical tests that do not take account of the number of trials in the experiment. This method is therefore likely to yield an unacceptable proportion of false-positive results. To reduce the likelihood of false positives and to maintain consistency in the comparison with our ideal observer definition of learning, we compute the probability of a correct response for trials  $w + 1$  to  $K - w$  and define the learning trial as the first trial such that all subsequent trials have  $p < 0.05$  for the number of observed correct trial responses in the window.

**The change-point test for binary observations.** The change-point test is based on a null hypothesis that, during the  $K$  trials, there is a constant probability of a correct response (Siegel and Castellan, 1988). This constant probability is not  $p_0$ , but rather, it is estimated at the end of the experiment as the proportion of correct responses across all of the trials. If the null hypothesis is rejected, the change-point statistic is used to identify the trial on which learning occurred. If we let

$$S_k = \sum_{j=1}^k n_j, \quad k = 1, \dots, K$$

be the total number of correct responses up through trial  $k$ , then the change-point statistic computes as follows:

$$D(k) = \left| \frac{K}{(K - S_k)S_k} \left( S_k - \frac{kS_k}{K} \right) \right|, \quad (2.6)$$

for  $k = 1, \dots, K - 1$ . We compare the maximum value of  $D(k)$  with the tabulated distribution of the Kolmogorov–Smirnov statistic (Siegel and Castellan, 1988) to decide whether there has been a change in the probability of a correct response. If the null hypothesis is rejected, then the trial on which learning occurred is the one with the maximum value of the statistic  $D(k)$ .

**Fixed-number of consecutive correct responses method.** For a learning experiment with  $K$  trials, a standard criterion for establishing learning is to require that a fixed number of consecutive correct responses be observed (Fox et al., 2003; Stefani et al., 2003). We let  $j$  denote this observed number of consecutive responses. Like the change-point test, the fixed number of consecutive correct responses method is based on a null hypothesis that, during the  $K$  trials, there is a constant probability of a correct response. Unlike the change-point test, the probability of a correct response is  $p_0$ . If  $K$  is large relative to  $j$ , then  $j$  consecutive correct responses are more likely to occur by chance. Hence, this approach is predicated on showing that, for  $j$  appropriately chosen relative to  $K$ , the probability of  $j$  consecutive correct responses occurring by chance is small.

The probability of observing a sequence of  $j$  consecutive correct responses for several combinations of  $K$  and  $j$  and two levels of significance are tabulated in Table 1. The number of consecutive correct responses required to establish learning increases with increases in the probability of a correct response, increases in the number of trials per experiment and decreases in the desired significance level. For example, from column 1 in Table 1, if there are  $K = 20$  trials in an experiment, and the probability of a correct response by chance is  $p_0 = 0.125$ , then only  $j = 3$  consecutive correct responses are required to reject a null hypothesis of no learning with  $p < 0.05$ . On the other hand, if  $p_0 = 0.5$ , then  $j = 8$  consecutive responses are required to reject a null hypothesis of no learn-

**Table 1. Tabulation of the probability of  $j$  consecutive correct responses in  $K$  trials**

$p_0$	Total number of trials, $K$								
	20	30	40	50	60	70	80	90	100
0.500	8 (10)	8 (11)	9 (11)	9 (11)	10 (12)	10 (12)	10 (12)	10 (12)	10 (13)
0.250	5 (6)	5 (6)	5 (6)	5 (6)	5 (7)	5 (7)	6 (7)	6 (7)	6 (7)
0.125	3 (4)	3 (4)	4 (4)	4 (5)	4 (5)	4 (5)	4 (5)	4 (5)	4 (5)

For a given number of trials  $K$  (top row) and given probability of a correct response by chance  $p_0$  (first column), the table entries are the smallest number of consecutive correct responses required to reject a null hypothesis of no learning with significance level 0.05 (or 0.01, in parentheses). The observed number of consecutive correct responses required to establish learning increases with increases in the probability of a correct response, increases in the number of trials per experiment, and decreases in the desired significance level. The table entries were computed using the algorithm in Appendix C.

ing with  $p < 0.05$ . In Appendix C, we give an algorithm to compute the significance of observing  $j$  correct responses in  $K$  trials.

For the simulated learning experiment in Figure 1, the learning trial identified by the moving average method was trial 25, the change-point test estimate was trial 23, and the one identified using the criterion of five consecutive correct responses in 40 trials (Table 1) was trial 28.

### Experimental protocols for learning

**A location–scene association task.** To illustrate the performance of our methods in the analysis of an actual learning experiment, we analyzed the responses of a macaque monkey in a location–scene association task, described in detail by Wirth et al. (2003). In this task, the monkey fixated on a point on a computer screen for a specified period and then was presented with a novel scene. A delay period followed, and, to receive a reward, the monkey had to associate the scene with the correct one of four target locations: north, south, east, and west. Once the delay period ended, the monkey indicated its choice by making a saccadic eye movement to the chosen location. Typically, between two and four novel scenes were learned simultaneously, and trials of novel scenes were interspersed with trials in which four well learned scenes were presented. 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 objective of the study was to track learning as a function of trial number and relate the learning curve to the activity of simultaneously recorded hippocampal neurons (Wirth et al., 2003).

**A T-maze task.** As a second illustration of our methods applied to an actual learning experiment, we analyzed the responses of a rat performing a T-maze task, described in detail by Jog et al. (1999) and Hu et al. (2001). In this task, the rat used auditory cues to learn which one of two arms of a T-maze to enter to receive a reward. On each day of this 47 d experiment, the rat performed 40 trials, except on days 1 and 46, on which it performed 20 and 15 trials, respectively. The total number of trials was 1835. For this experiment, the probability of making a correct response by chance was 0.5. The objective of this study was to relate changes in learning across days to concurrent changes in neural activity in the striatum (Jog et al., 1999; Hu et al., 2001).

Experimental procedures used for both tasks were in accordance with the National Institutes of Health guidelines for the use of laboratory animals.

## Results

### Simulation study of learning curve estimation

We designed two simulation studies to investigate the performance of our algorithms. In the first, we tested the accuracy of the algorithms in estimating a broad family of learning curves. In the second, we tested their performance in estimating three specific learning curves seen in actual experiments.

In the first study, we compared the performance of the filter algorithm, the smoothing algorithm, and the moving average method with a window width of nine ( $w = 4$ ) in the analysis of simulated learning experiments from a family of sigmoid curves (Fig. 2A). Each learning curve was defined by  $p_k$ , the probability of a correct response at trial  $k$ , and was defined using the following logistic equation:

$$p_k = p_0 + \frac{(p_f - p_0)}{1 + \exp(-\gamma(k - \delta))}, \quad (3.1)$$

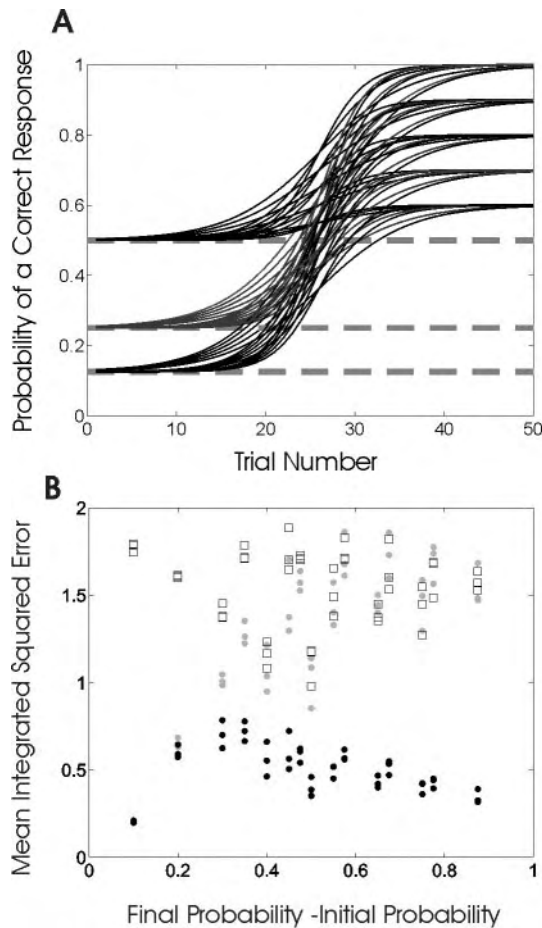
where for  $k = 1, \dots, K$ ,  $p_0$ , the initial probability, is the probability of a correct response by chance,  $p_f$  is the final probability of a correct response,  $\gamma$  is a constant governing the rate of rise of the learning curve, i.e., the learning rate, and  $\delta = 25$  is the inflection point of the curve.

In these simulated learning experiments, we chose three values of the initial probability  $p_0$  (0.125, 0.25, and 0.5), five values of the final probability  $p_f$  (0.6, 0.7, 0.8, 0.9, and 1), and three values of  $\gamma$  (0.2,

0.3, and 0.4) (Fig. 2A). With this family of curves, we tested systematically how the three methods performed as a function of the probability of a correct response by chance, the learning rate, and the final probability of a correct response the animal achieved. For each of the  $3 \times 5 \times 3 = 45$  learning curves, we simulated 100 50-trial experiments. For example, for a given triplet of the parameters  $p_0$ ,  $p_f$ , and  $\gamma$ , we simulated 50 trials of experimental data by using Equation 3.1 to draw Bernoulli random variables with probability  $p_k$  of a correct response for  $k = 1, \dots, K$ . That is, on trial  $k$  a coin is flipped with the probability of heads, a correct response, being  $p_k$  and the probability of tails, an incorrect response being  $1 - p_k$ . The result recorded from each trial was a one if there was a correct response and a zero if the response was incorrect. This procedure was repeated 100 times for each of the 45-parameter triplets for Equation 3.1. We compared the filter algorithm, smoothing algorithm, and the moving average estimates of the true learning curves using mean integrated squared error (MISE) (Rustagi, 1994).

We compared the MISE across the three estimation methods for each of the 45 triplets of parameters  $p_0$ ,  $p_f$  and  $\gamma$ . The MISE for the smoothing algorithm was smaller than the MISE for the moving average method for each of the 45 triplet combinations (Fig. 2B). The MISE for the smoothing algorithm was smaller than the filter algorithm in 44 of the 45 triplet combinations. For the one exception, the difference in the two MISEs was  $< 10^{-3}$  and occurred with the difficult to estimate learning curve with  $p_0 = 0.5$ ,  $p_f = 0.6$ , and  $\gamma = 0.3$ . To examine learning as a function of the difference between the initial and final probabilities of a correct response, we plotted MISE against  $p_f - p_0$  (Fig. 2B). The MISEs for the smoothing algorithm estimates were similar across all values of  $p_f - p_0$  (Fig. 2B, black dots). The moving average method performed poorly in estimating the learning curve for all values of  $p_f - p_0$  (Fig. 2B, squares). The filter algorithm MISE estimates increased (Fig. 2B, gray dots) with  $p_f - p_0$  because these estimates lag behind the smoothing algorithm estimates (Fig. 1) and the true learning curve (Fig. 3A, D, G). As a consequence, the MISE between the filter estimate and the true learning curve paradoxically increases as  $p_f - p_0$  becomes larger.

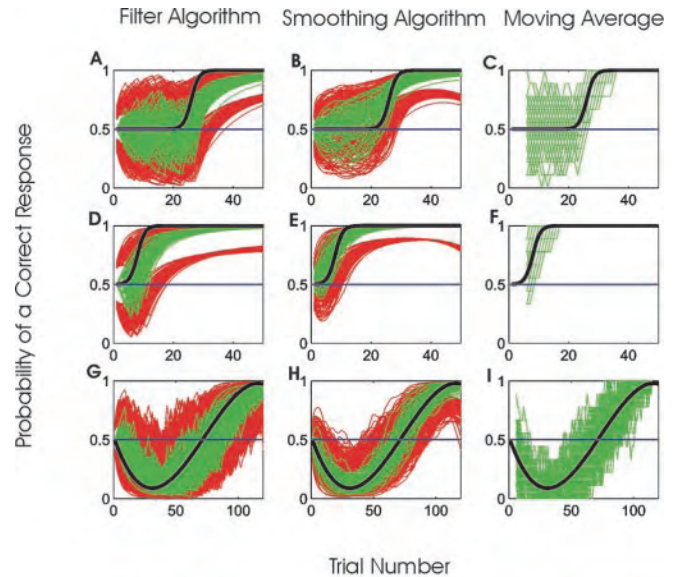
Having established that the smoothing algorithm outperformed the other techniques for a broad family of sigmoidal learning curves, we next analyzed the performance of this algorithm in estimating three types of learning curves: delayed rapid learning (Fig. 3A, black line), immediate rapid learning (Fig. 3D, black line), and learning after initially declining performance (Fig. 3G, black line). The first learning curve (Fig. 3, top row) was from the sigmoid family in Equation 3.1 with  $\gamma = 0.8$ ,  $p_0 = 0.5$ , and  $p_f = 1$ . This learning curve simulates rapid learning because  $\gamma = 0.8$  is twice the largest learning rate of  $\gamma = 0.4$  used in the first simulation study (Fig. 2A). The second curve (Fig. 3, middle row) had the same parameters as the first learning curve, except with  $\delta = 7$  instead of  $\delta = 25$  to simulate rapid learning at the outset of



**Figure 2.** *A*, Family of sigmoidal curves (Eq. 3.1) used to simulate the learning experiments. Learning curves were constructed using three values of the initial probability of a correct response  $p_0$  (0.125, 0.25, and 0.5), five values of the final probability of correct response  $p_f$  (0.6, 0.7, 0.8, 0.9, and 1), and three values of  $\gamma$  (0.2, 0.3, and 0.4), which governs the rate of rise or learning rate of the curves. For each of the  $3 \times 5 \times 3 = 45$  learning curves, we simulated 100 learning experiments for a total of 4500. All of the curves increase monotonically from  $p_0$ , indicating that performance is better than chance on all trials, i.e., learning starts immediately. *B*, MISE for the filter algorithm (gray dots), the smoothing algorithm (black dots), and the moving average method (squares) for each of the 45 learning curves in *A* plotted as a function of  $p_f - p_0$ . The smoothing algorithm MISE is smaller than those of the filter algorithm and the moving average method for all values of  $p_f - p_0$  above 0.1.

the experiment. The third learning curve (Fig. 3, bottom row) was a cubic equation that like the first two curves had  $p_0 = 0.5$  and  $p_f = 1$ . However, for this learning curve the probability of a correct response decreased and then increased. This type of learning profile is seen when an animal has a response bias and performs poorly at the outset of an experiment (Stefani et al., 2003). The simulated experiments based on the first two learning curves had 50 trials each, whereas those based on the third curve had 120 trials each.

For each curve, we simulated 100 learning experiments, estimated learning curves with each of the three methods, and compared them with true learning curves. The moving average method performed the least well of the three methods, because it was unable to estimate the learning curves reliably for any of the three curves (Fig. 3*C, F, I*). Because the moving average method is a two-sided filter, it could not estimate learning for the initial four and final four trials. As in Figure 1, the filter algorithm estimated the learning curves with a noticeable delay for each of these three learning curves (Fig. 3*A, D, G*). For each of the three learning



**Figure 3.** Analysis of three simulated learning experiments by the filter algorithm, the smoothing algorithm, and the moving average method. *A–C*, Delayed rapid learning. *D–F*, Immediate rapid learning. *G–I*, Learning after initially declining performance. We compared the 100 estimated learning curves (green), the true learning curve (black), and the 90% confidence intervals (red) using the filter algorithm (first column, *A, D*, and *G*), the smoothing algorithm (second column, *B, E*, and *H*), and the moving average method (third column, *C, F*, and *I*). The moving average method estimates fluctuate more, do not provide confidence intervals, and do not track well the true learning curves. The filter algorithm learning curve estimates consistently lag behind the true learning curves. The smoothing algorithm gives the best overall estimates of the learning curves with the narrowest confidence intervals.

**Table 2. MISE for the filter algorithm, smoothing algorithm, and moving average method in estimating the three learning curves in Figure 3**

	MISE		
	Delayed rapid learning	Early rapid learning	Delayed learning
Filter	1.2262	0.5789	2.3894
Smoothing	0.5025	0.1445	0.6240
Moving average	0.7963	0.2213	1.6276

For each of the three learning curves, the MISE was computed for each method using 100 simulated learning experiments. The number of trials per experiment was 50 for the two rapid learning experiments and 120 for the delayed learning experiment.

curves, the smoothing algorithm followed the true learning curve most closely (Fig. 3*B, E, H*) and tracked especially well the trials in which the performance was worse than chance (Fig. 3*H*). The MISE for the smoothing algorithm was 0.5025 for the delayed rapid learning, 0.1445 for the immediate rapid learning curve, and 0.6240 for the learning after a decline in performance. The MISEs for the smoothing algorithm were smaller than those for the filter algorithm and moving average method by a factor of 1.5 to 4 (Table 2), suggesting again that the smoothing algorithm provided the best estimate of the true learning curve.

**Simulation study of learning trial identification**

We used the two simulation studies of learning curve estimation to compare the smoothing algorithm with the change-point test, the consecutive correct responses method, and the moving average method in identifying the learning trial. In the analysis of the simulated data and in the actual data analyses, we only compare the smoothing algorithm with the alternative methods because the smoothing algorithm is the maximum likelihood (most probable) estimate of the learning curve given the data and because its

performance was superior to the filter algorithm in the two previous simulation studies. For the smoothing algorithm, we estimated learning from Equation B.4 using the IO (0.95) criterion. For the change-point, consecutive correct responses method, and moving average method, we set the significance level at 0.05.

For the first simulation study, there were 50 trials in each simulated experiment and three different probabilities,  $p_0 = 0.125, 0.25, \text{ and } 0.5$ , so the minimal numbers of consecutive responses required to identify learning by the consecutive responses method were, respectively, four, five, and nine (Table 1, column 4) for a significance level of 0.05. For each of these curves, the true performance increased monotonically following the logistic model (Fig. 2) in Equation 3.1 so that the probability of a correct response was greater than chance from the outset of the experiment. Therefore, the method that performed best was the one that identified the earliest learning trial.

In the 4500 simulated experiments of the first simulation study, the smoothing algorithm identified a learning trial in 3752 (83%) simulated experiments, the change-point test in 3225 (72%) experiments, the consecutive correct responses method in 3439 (76%) experiments, and the moving average method in 3579 (80%) experiments. Estimating a learning trial for some of the learning curves was more challenging than for others (Fig. 4A). The shorter the distance between the initial and final probability, the more difficulty each procedure had in identifying the learning trial. On the other hand, as the distance between the initial and final probability increased, it was easier for each procedure to identify the learning trial. The IO (0.95) criterion identified a learning trial in a higher proportion of experiments than the other techniques for any difference between the initial and final probability of a correct response. The only exception was when the difference between the initial and final probabilities was 0.1, in which case, all three methods performed poorly (Fig. 4A).

We compared the estimates of learning trial in 2892 of the 4500 (64%) experiments in the first simulation study in which all four methods identified a learning trial. The IO (0.95) identified learning in advance of the change-point test in 2237 of these 2892 (77%), at the same trial as the change-point test in 233 (8%), and after the change-point test in 422 (15%) of the simulated experiments (Fig. 4B). The IO (0.95) identified learning in advance of the consecutive correct responses method in 2875 of these 2892 (>99%), at the same trial as the consecutive correct responses method in 0, and after the consecutive correct responses method in 17 (<1%) of the simulated experiments (Fig. 4C). Similarly, the IO (0.95) identified learning in advance of the moving average method in 2810 (97%), at the same trial as the moving average in 70 (<3%), and after in 12 (<1%) of the simulated experiments (Fig. 4D). For the simulated experiments in which the IO (0.95) identified the learning trial before the change-point test (consecutive correct responses method/moving average method), the median difference in the learning trial estimates was 5 (10/5) trials with a range from 1 to 27 (1 to 35/1 to 32) trials.

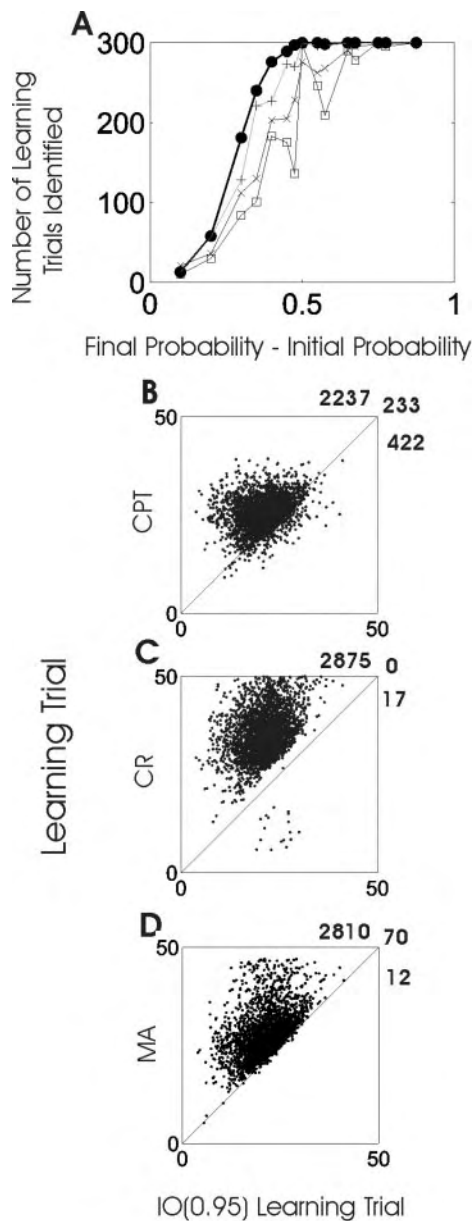
To investigate further the performance of the IO (0.95) relative to the change-point test, the consecutive correct responses method, and the moving average method, we used these three methods to identify the learning trials from the second set of simulated learning experiments in Fig. 3 (black lines). Each method identified a learning trial in at least 94 of the 100 simulated experiments for each of the three learning curves, except for the change-point test, which identified a learning trial in only 81 of the 100 experiments for the second rapid learning curve (Fig. 3D). The IO (0.95) and moving average method identified

learning trial in more of the simulated experiments (299/300) than either the change-point test (279/300) or the consecutive correct responses method (295/300).

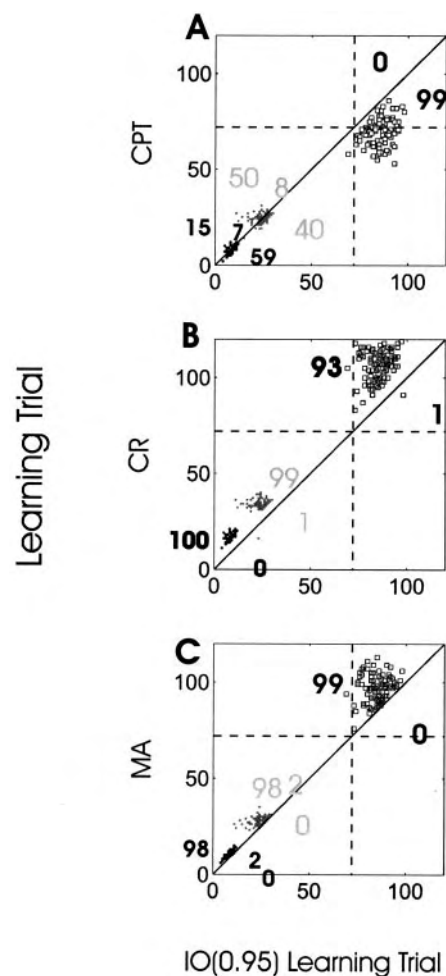
For the two rapid learning curves (Fig. 3A–F), the probability of a correct response exceeded chance from the outset, so that, as in the previous analysis, the method that identified the earliest learning trial was the best. For the early rapid learning experiments (Figs. 3D, black curve, 5, black dots), the change-point test identified more of the learning trials earlier than the IO (0.95) (Fig. 5A, black dots), whereas the consecutive correct responses method identified all of them later than the IO (0.95) (Fig. 5B, black dots). For the delayed rapid learning experiments (Figs. 3A, black curve, 5, gray dots), the IO (0.95) identified the majority of the learning trials earlier than the change-point test (Fig. 5A, gray dots), the consecutive correct responses method (Fig. 5B, gray dots), and the moving average method (Fig. 5C, gray dots). For the experiments from both the early and delayed rapid learning curves in which the change-point test identified the learning trial earlier than the IO (0.95), the median differences were one and two trials, respectively, and the maximum difference for both was five trials. The slightly better performance of the change-point test is attributable to the fact that its null hypothesis is that there is no change in the overall proportion of correct responses. In these experiments, there were many correct responses in the latter trials, making the null hypothesis probability close to one. As a result, the change-point test detected learning near the start of the experiments where the probability of a correct response was 0.5, the farthest value from its null hypothesis probability.

For the analysis of the simulated experiments based on the learning curve that involved learning after declining performance, the true learning trial was trial 72. This was the trial on which the true learning curve first exceeded the line  $p_0 = 0.5$ , defining the probability of a correct response by chance (Fig. 3G–I). Therefore, the best method for identifying the learning trials in these experiments is the one which identified them at the earliest trials on or after trial 72. The IO (0.95) identified 98 of its 99 learning trials for this experiment on or after trial 72 (Fig. 5A, open squares and vertical dashed line). The change-point test identified only 38 of its 99 learning trials on or after trial 72 (Fig. 5A, open squares and horizontal dashed line). Of the 38 identified after trial 72, all were earlier than the corresponding trials identified by the IO (0.95) (Fig. 5B, diagonal line). The consecutive correct responses method identified all 94 of its learning trials after trial 72; however, 93 of those 94 were later than the learning trials identified by the IO (0.95) (Fig. 5B, open squares). Similarly, the moving average method identified all of its 99 learning trials after trial 72, but each was later than the corresponding one estimated by the IO (0.95) (Fig. 5C, open squares).

The bias of the change-point test toward identifying early learning trials can be explained by the way its null hypothesis was formulated. The change-point test null hypothesis probability is the total proportion of correct responses observed in an experiment (see Material and Methods). For this learning curve, the null hypothesis probability for the change-point test was on average 0.47. This was the average proportion of correct responses over the 100 simulated experiments for this learning curve (Fig. 3G). Because the change-point test identified the earliest trial in which the observed proportion of correct responses up to that trial differed from that predicted by 0.47, it detected consistently the increase from the nadir in the probability of a correct response of 0.10 near trial 30 as significant. This increase was ap-



**Figure 4.** Analysis of learning trial estimation for the family of sigmoidal learning curves in Figure 2. *A* shows the number of learning trials identified by the smoothing algorithm [IO (0.95); black dots], the change-point test (CPT; crosses), the consecutive correct response method (CR; squares), and the moving average method (MA; plus signs) as a function of  $p_f - p_0$ , the difference between the final and the initial probabilities of a correct response. Each point is the number of learning trials identified by a given method for one of the 15 combinations of  $p_f - p_0$  summed over the three values of  $\gamma$  (Fig. 2). The maximum value of each point would be  $3 \times 100 = 300$ , if the method identified a learning trial for each simulated experiment. The IO (0.95) identified a learning trial in more experiments than the other three methods, for all values of  $p_f - p_0$  except when  $p_f - p_0$  was 0.1. Scatterplots of the learning trial estimates of the change-point test versus the IO (0.95) (*B*), the consecutive correct responses method versus the IO (0.95) (*C*), and the moving average method versus the IO (0.95) (*D*), for the 2892 of the 4500 simulated learning experiments (Fig. 2) in which all three methods identified a learning trial. Because in all of the simulated experiments true performance was greater than chance from the outset, the method that identified the earliest learning trial performed the best. The 45° line indicates that the two methods being compared identified the same learning trial. The change-point learning trial estimate was later than the IO (0.95) estimate in 2237 of 2892 experiments (77%), at the same trial in 233 of 2892 (8%), and earlier in 422 of 2892 (15%). The consecutive correct response method estimate of the learning trial was later than the IO (0.95) estimate in 2875 of 2892 experiments (>99%), at the same trial as in 0 of 2970, and earlier in 17 of 2970 (<1%). The moving average method estimated the learning trial after the IO (0.95) criterion in 2810 of 2892 experiments (97%), at the same trial as the moving average in 70 of 2892 (<3%), and earlier in 12 of 2892 (<1%) of the simulated experiments.

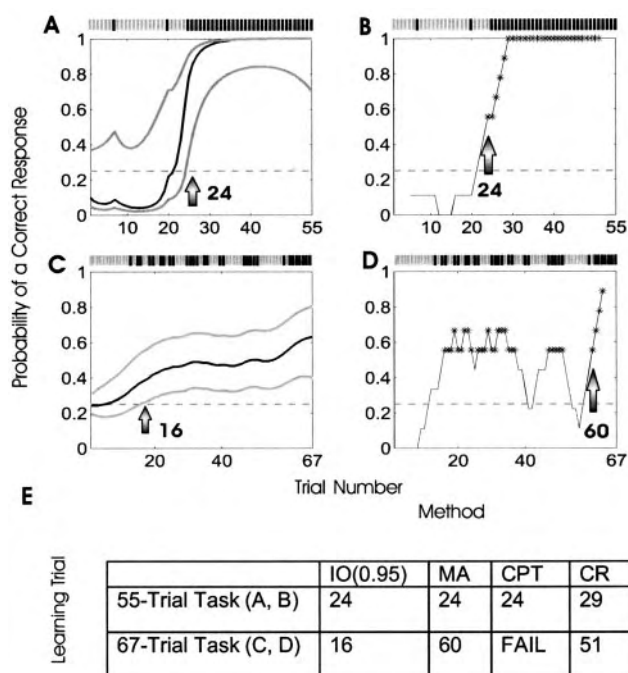


**Figure 5.** Comparison of learning trials estimated by the IO (0.95), change-point test, the consecutive correct response method, and the moving average method for the three simulated learning experiments in Figure 3. *A* compares learning trial estimates from the change-point test (CPT) and the IO (0.95) for delayed rapid learning (gray dots), immediate rapid learning (black dots), and learning after declining performance (squares). The number of learning trial estimates above and below the 45° line are marked in the panel. Trial 72 (dashed lines) is the trial on which the learning curve in the learning after declining performance simulations crossed  $p_0 = 0.5$ , the line for the probability of a correct response by chance. For the delayed rapid learning curve (gray dots, *A*), there were 50 simulated experiments in which the change-point test estimated learning later than the IO (0.95), 40 experiments in which it estimated learning earlier, and eight experiments in which the change-point test and IO (0.95) estimated identical learning trials. For the immediate rapid learning (black dots, *A*), the change-point test estimated learning earlier than the IO (0.95) for the majority of simulated experiments (59 of 81) because of the observation-dependent null hypothesis in the change-point test (see Results). For the learning after declining performance curve (open squares, *A*), the change-point test estimated learning earlier than the IO (0.95) in all of the simulated experiments. However, the majority of the estimates (61 of 99) occurred before trial 72, indicating that the change-point test incorrectly identified learning before the true learning curve was above chance. *B* compares the learning trial estimated from the consecutive correct responses method (CR) and the IO (0.95) for the same three simulated learning curves using the same symbol definition as in *A*. The IO (0.95) estimated learning earlier than the consecutive responses method in nearly all (292 of 294) simulated experiments. For the learning after declining performance curve (open squares, *B*), the IO (0.95) only estimated learning in advance of trial 72 in one experiment as trial 69. The consecutive correct responses method never detected learning before trial 72. *C* compares learning trials estimated with the moving average method (MA) and the IO (0.95) criterion. In most cases, the moving average estimates of the learning trial were later (295 of 299) than the IO (0.95) learning trial estimates and were identical to the IO (0.95) estimates in 4 of 299.

parent but, before trial 72, it did not indicate performance that was better than chance.

The results of these two simulation studies demonstrate that, when learning is defined as performance better than predicted





**Figure 6.** Performance of the smoothing algorithm compared with the moving average method in analyzing association learning by a macaque monkey. The monkey performed a 55-trial location–scene association task (A, B) and a 67-trial location–scene association task (C, D). The incorrect and correct responses of the monkey in each experiment are shown, respectively, as gray and black tick marks above A–D. The solid black curves are the learning curve estimates for the smoothing algorithm, and the gray curves are the associated 90% confidence intervals in A and C. In B and D, the gray line shows the moving average estimate of the probability of a correct response. The asterisks denote the nine-trial windows in which the number of correct responses is significantly more than would be predicted by chance if there were no learning based on a local binomial probability distribution function. In both experiments, the probability of a correct response occurring by chance was 0.25 (horizontal dashed line). In the 55-trial experiment, the monkey gave only two correct responses in the first 24 trials (A, B). The IO (0.95) criterion identified learning at trial 24 (arrow in A). The moving average method (B) identified the learning trial as occurring in the window centered at trial 24 (arrow in B). In the 67-trial experiment, the monkey gave its first correct response at trial 13 and subsequently performed better than chance but with a smaller total proportion of correct responses compared with the 55-trial experiment. The IO (0.95) identified the learning trial as trial 16 (arrow in C), whereas the moving average method identified the learning trial as occurring near trial 60 (arrow in D). E summarizes learning trial estimates for the two experiments using the IO (0.95) criterion, the moving average method (MA), the consecutive correct responses method (CR), and change-point test (CPT).

bychance, the IO (0.95) identifies the learning trial more reliably and more accurately than the change-point test, the consecutive correct responses method, or the moving average method across a broad range of learning curves.

### Analysis of learning in an association task

To illustrate the performance of our methods in the analysis of an actual learning experiment, we used the IO (0.95) to characterize the behavioral performance of a macaque monkey in a location–scene association task (Wirth et al., 2003) described in Materials and Methods. We analyzed the monkey's responses in two different learning experiments: one with 55 trials and the second with 67 trials. The monkey's observed correct (black) and incorrect (gray) responses are shown in Fig. 6.

In the 55-trial experiment, the monkey responded incorrectly in all of the first 24 trials, except in trials 7 and 20 (Figs. 6A, B). The smoothing algorithm (Fig. 6A) and the moving average method (Fig. 6B) produced learning curve estimates with similar shapes. Because the smoothing algorithm estimates the probabil-

ity of a correct response at each trial, it estimated the IO (0.95) learning trial as trial 24 (Fig. 6A, arrow). The moving average method identified the learning trial as occurring within the nine-trial window centered at trial 24 (Fig. 6B, arrow), containing at least five correct responses. The change-point test identified the learning trial as trial 24, and the consecutive responses method identified it as trial 29 (Fig. 6E). The smoothing algorithm, however, had the advantage of providing an easily interpretable estimate of the probability of a correct response as a function of trial number. For example, we can use the confidence limits to analyze performance worse than or better than chance. That is, the monkey appears to perform below chance at the outset of the experiment because it has only two correct responses in 24 trials. Its performance, however, was not worse than what could be expected by chance, because the upper confidence bounds were never below 0.25 (Fig. 6A).

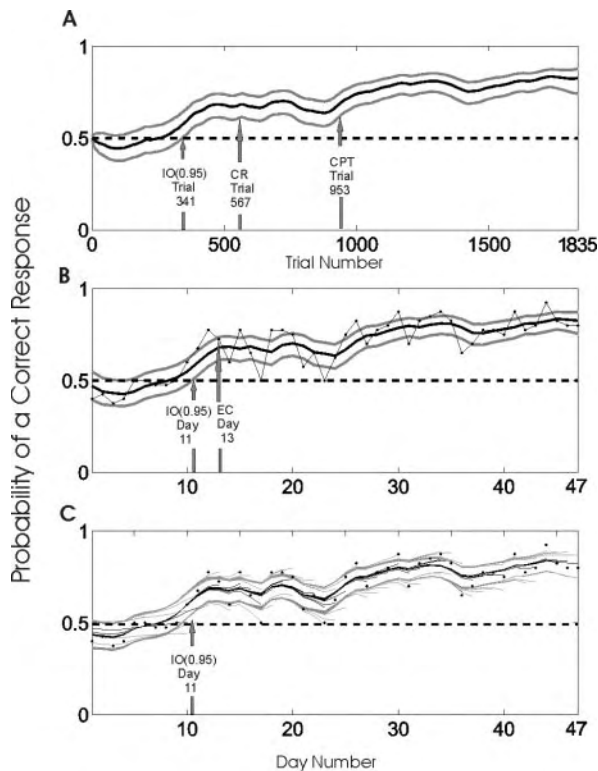
An additional analysis of the 55-trial experiment comparing the current state–space model with a state–space model with state-dependent variance based on Kakade and Dayan (2002) is presented in the supplemental data to this article ([www.jneurosci.org](http://www.jneurosci.org)). Although the results did not differ from those presented here, it illustrates how our paradigm may be extended to analyze whether the learning rate declines as learning progresses.

For the 67-trial location–scene association task, the IO (0.95) showed that the animal learned the correct association at trial 16 (Fig. 6C, arrow). The moving average method identified the learning trial at trial 60 (Fig. 6D). The trials at the center of a nine-trial window with at least five correct responses are marked by an asterisk. On the basis of the results in Appendix C, the consecutive correct responses method required a minimum of five consecutive correct responses in 67 trials to identify the learning trial, and identified it as trial 51 (Fig. 6E). The change-point test could not identify a learning trial in this experiment (Fig. 6E).

On the basis of the proportion of correct responses for the entire experiment, 30/67, it is apparent that the monkey has learned. The likelihood of observing this proportion of correct responses if the monkey were performing with a probability of a correct response by chance on each trial of 0.25 would be  $<10^{-3}$ . The moving average method first identified learning at trial 16, but, because there were windows in which no learning was detected after trial 16, an ideal observer would not be able to conclude learning had definitely occurred until trial 60. The change-point test failed to reject the null hypothesis of no learning because its null hypothesis is that the probability of a correct response by chance is  $p_0 = 30/67$  and not 0.25. The monkey performed consistently from the start of the experiment so that the proportion of correct responses up to any trial was well predicted by the proportion of correct responses for the entire experiment. When the proportion of correct responses for the entire experiment predicts well the performance up to any trial, then by Equation 2.6, the change-point test does not reject the null hypothesis and, hence, does not identify a learning trial. Again, the advantage of the smoothing algorithm is that, in addition to identifying the ideal observer learning trial, it provides a learning curve estimate for the entire experiment.

### Analysis of learning in a T-maze task

We also analyzed the performance of a rat in a procedural learning task requiring the animal to use auditory cues to learn which of two arms in a T-maze to enter to receive a reward (see Materials and Methods). The observed proportions of correct responses on each day are shown in Figure 7B (dotted line).



**Figure 7.** Performance of a rat during a 47 d T-maze procedural learning task. The rat performed 40 trials per day, except on day 1 (20 trials) and day 46 (15 trials), for a total of 1835 trials. The data were analyzed using the smoothing algorithm and the IO (0.95) criterion in a trial-by-trial analysis (*A*) and in a day-by-day analysis using the smoothing algorithm with the binomial observation model (Eq. 3.2) and the IO (0.95) criterion and the number of correct responses on each day as the observed data (*B*, *C*). The learning curve estimates for each analysis (solid black lines) and the associated 90% confidence intervals (gray lines) are shown. In *B*, the black dots connected by the gray lines are the proportions of correct responses on each day. The probability that correct response occurred by chance was 0.5 in this experiment (dashed horizontal lines in all panels). The learning trial estimates for the trial-by-trial analysis in *A* were as follows: IO (0.95), trial 341 on day 9; the change-point test (CPT), trial 953 on day 24; and the fixed criterion of 15 consecutive correct responses (CR), trial 567 on day 14. The estimates of the learning day for the day-by-day analysis in *B* were as follows: IO (0.95), day 11; and the empirical criterion (EC) of Jog et al. (1999), day 13. *C*, The data were also analyzed on a day-by-day basis as the experiment progressed by computing the learning curve from the start to the end of the data series as the end was moved from day 3 to day 47. The 45 learning curves (black lines) and their associated 90% confidence bounds (gray lines) are plotted in *C* with the proportion of correct responses on each day (black dots). The “whisper” appearance of the plot is attributable to the fact that each learning curve and its associated confidence intervals were estimated with a different number of days of data, and, hence, each has different end points. In each analysis from day 12 to 47, the IO (0.95) criterion identified the learning day as day 11. The fact that the upper 95% confidence bound fell below 0.5 at trial 4 (*B*, *C*) suggests that the animal might have had a response bias at the start of the experiment.

We analyzed this experiment using our state–space model paradigm in three ways. First, we concatenated all of the responses across all the days and analyzed learning as a function of trial number across the 1835 trials (Fig. 7*A*). The smoothing algorithm showed a near monotonic increase in the learning curve beginning almost at the outset of the experiment. The IO (0.95) criterion identified the learning trial as trial 341, which occurred on day 9 [Fig. 7*A*, IO (0.95)]. The change-point test identified trial 953 on day 24 as the learning trial (Fig. 7*A*, CPT). Given  $K = 1835$ , the consecutive response method required a minimum of 15 consecutive correct responses to establish learning. This occurred at trial 567 on day 14 (Fig. 7*A*, CR). In a similar experiment, Jog et al. (1999) defined learning as the point at which there were at

least 72.5% correct responses on 2 consecutive days. This corresponds to 29 correct responses or more of 40 per day. By this criterion, learning occurred in this experiment on day 13 (Fig. 7*B*, EC).

In the second analysis, we assumed that only the number of trials per day and the number of correct responses per day were recorded instead of all of the sequences of individual trial outcomes within each day. Therefore, we replaced the Bernoulli model in Equation 2.1 with the following binomial observation model:

$$Pr(n_k | p_k, x_k) = \binom{m_k}{n_k} p_k^{n_k} (1 - p_k)^{m_k - n_k}, \quad (3.2)$$

where  $k$  indexes the day,  $m_k$  is the number of trials on day  $k$ , and  $n_k$  is now the number of correct responses on day  $k$ . Following the arguments in Appendix A, we derived a smoothing algorithm using the same Gaussian state model with the observation model in Equation 3.2. We used the reformulated state–space model to analyze this coarser data series. The day-by-day learning curve for this analysis (Fig. 7*B*) resembles closely that of the trial-by-trial learning curve (Fig. 7*A*), with the exception that the latter had narrower confidence limits. The day-by-day learning curve was smoother than the plot of the proportion of correct responses in each day because the latter does not relate performance on one day to that on any other day (Fig. 7*B*). This analysis shows only a minimal loss of information relative to the trial-by-trial analysis as the IO (0.95) criterion identified the learning day as day 11 of the experiment instead of on day 9. In either case, the IO (0.95) identified the learning day 2 to 4 d before the day estimated by the criterion of Jog et al. (1999).

In the third analysis of this experiment, we estimated learning as the experiment progressed. That is, we used the day-by-day algorithm starting at day 1 and progressively computed the learning curve as the end point of the data series was moved back from day 3 to day 47 (Fig. 7*C*). Once the data series length reached 12 d, the learning day was identified as day 11, and this estimate remained unaltered for all series lengths from days 13 to 47 (Fig. 7*C*). Thus, the day-by-day smoothing algorithm and the IO (0.95) criterion predicted on day 12 that learning occurred on day 11, suggesting that the smoothing algorithm and learning trial definition could be used in real time as data accrue to identify the learning trial and, possibly, to shorten the length of experiments or to study “overlearning” on-line.

In the trial-by-trial and the day-by-day analyses (Fig. 7*A*, *B*), the smoothing algorithm learning curve estimates are less than  $p_0 = 0.5$  until trial 255 and day 8, respectively. In the day-by-day analyses (Fig. 7*B*, *C*), the smoothing algorithm estimated that the upper 95% confidence bound of the probability of a correct response as below  $p_0 = 0.5$  at trial 4, suggesting that the animal might have had a response bias at the start of the experiment.

## Discussion

At present, there is no consensus as to the most reliable and accurate way to analyze sequences of trial responses in learning experiments. To address this question, we developed a state–space model that characterizes learning as a dynamic process. The smoothing algorithm derived from the model estimates a learning curve from the trial responses of a single animal in a learning experiment. The algorithm provides at each trial the ideal observer’s assessment of the probability that a subject is performing better than chance and, as a consequence, the first trial on which an ideal observer is 0.95 certain that the subject is performing better than chance for the balance of the experiment. We tested the state–space model on a representative set of simulated learning experi-

ments and applied it in the analysis actual learning experiments. The superior performance of both the smoothing algorithm and ideal observer learning trial definition in the analysis of the simulated and actual learning experiments relative to current methods suggests that the state–space model paradigm offers a coherent statistical framework for analyzing learning experiments.

### State–space modeling of learning versus testing hypotheses

The superior performance of our state–space model paradigm relative to the hypothesis testing methods was expected because building a statistical model of a process, in general, gives a more informative analysis than simply designing a procedure to test a hypothesis (Box, 1980; Casella and Berger, 1990). Whereas the hypothesis testing methods test a yes–no hypothesis, the state–space model characterizes the relationship between the hidden learning process and the animal's responses across trials. From a properly constructed statistical model, the hypothesis of interest can be tested, confidence intervals can be constructed, and model goodness-of-fit assessments can be made (Box, 1980). The state–space model represents the trials as dependent, whereas the hypothesis testing methods assume them to be independent.

The change-point test, the consecutive response method, the moving average method, and the learning criterion by Jog et al. (1999) test the null hypothesis of no learning defined by a constant probability of a correct response. The probability of a correct response under the null hypothesis should be defined by how likely the animal is to respond correctly to the task by chance before attempting it. For the association task with four choices on any trial, this probability was 0.25, whereas for the T-maze task with two choices on any trial, this probability was 0.50. In contrast, the null hypothesis probability of a correct response for the change-point test is independent of the task and is defined after the experiment has been conducted as the proportion of correct responses recorded during the experiment. The change-point test (Eq. 2.6) compares by trial the difference between the observed cumulative number of correct responses and the cumulative number of correct responses predicted by the total proportion of correct responses. This test identifies any large difference as evidence to reject the null hypothesis. Our analysis demonstrates that the null hypothesis of the change-point test is not relevant for a learning experiment and that, with this method, every experiment has a different null hypothesis formulated after the experimental observations have been recorded.

In previous applications of the consecutive correct responses method, no specific justification has been given for the number of consecutive correct responses needed to reject the null hypothesis (Fox et al., 2003; Stefani et al., 2003). Table 1 provides what we believe is the first demonstration of how the number of correct responses depends on the number of trials in the experiment, the probability of a correct response by chance, and the desired significance level. This method tests a null hypothesis of no learning in which  $p_0$  is chosen correctly as the preexperiment probability of a correct response. However, the analysis of the simulated learning experiments showed that learning could occur, and this procedure could fail to detect it or would detect it consistently later than the IO (0.95). This is because the animal's performance can be better than chance, and the required number of consecutive correct responses may not occur. If they do occur, they may not occur early in the experiment.

The moving average method identifies the learning trial by testing a null hypothesis of no learning using a number of trials equal to the window width of the local estimation filter (Eq. 2.5). We required that the criterion for this method for rejecting the

null hypothesis hold from an identified trial to the end of the experiment for that trial to be the learning trial. Otherwise, the moving average  $p$  value is not correct because it does not consider the total number of trials in the experiment.

### The smoothing and filter algorithms

Under the state–space model, the smoothing algorithm uses all of the experimental data and gives the maximum likelihood (optimal) estimate of the learning curve (Fahrmeir and Tutz, 2001). Given either the known or estimated value of the model parameter  $\sigma_e^2$ , the smoothing algorithm represents the ideal observer of the experiment. The filter algorithm estimate at trial  $k$  is causal (Kakade and Dayan, 2002) because it does not depend on the animal's responses beyond that trial, and, therefore, its learning curves (Figs. 1, 3) lag behind those of the smoothing algorithm. This lag contributed to the paradoxical increase in the MISE of this procedure as the difference between the initial and final probabilities of a correct response increased (Fig. 2B). The filter algorithm remains important because, as shown in Appendix A, the smoothing algorithm is computed from the filter algorithm. Our simulation studies along with those by Smith and Brown (2003) show that the Gaussian approximations applied in the filter and smoothing algorithms are sufficiently accurate to allow reliable learning curve estimation across a broad range of shapes. The filter and smoothing algorithms are, respectively, the analogs for the current problem of the Bayes' filter and the Bayes' smoother neural spike train decoding algorithms developed by Brown et al. (1998).

### State–space models of learning: advantages and future directions

The state–space model paradigm offers important advantages over current techniques. First, individual learning curve estimates provide a direct characterization of between animal differences in learning propensity for the same task. Second, our analysis can estimate the probability that the subject has learned (performs better than chance) on each trial and allows the investigator to choose a minimal level of certainty he/she wishes to accept to identify the learning trial. Our IO (0.95) learning trial definition identified the learning trial more reliably and accurately than all of the methods we tested in the analyses of both simulated and actual learning experiments. In this regard, the smoothing algorithm analysis of the T-maze task was especially instructive because the IO (0.95) criterion identified learning 2 to 4 d earlier than the criterion used by Jog et al. (1999) and Hu et al. (2001) (Fig. 7A) in similar experiments. Third, because the smoothing algorithm can estimate learning on a trial-by-trial basis for a single animal, our analyses can be used to correlate behavioral changes associated with learning with changes in neural activity (Jog et al., 1999; Wirth et al., 2003).

Finally, sequential use of the smoothing algorithm to analyze the T-maze experiment showed that, by day 12 of this 47 d experiment, the IO (0.95) criterion estimated that learning had occurred by day 11. This finding suggests that the smoothing algorithm can be used to track learning in real time. This is analogous to a sequential analysis of a medical clinical trial in which the outcomes of study participants are continuously monitored to determine as early as possible whether a new therapy is effective (Everitt and Pickles, 2000). For learning studies, obtaining reliable real-time estimates of the learning trial opens the possibility of changing a protocol in real time to test an animal's performance under new conditions. These advantages of our approach over current methods can therefore reduce the number of animals and the number of trials per animal required to execute a learning study.

Our state–space model expands easily to take account of other features of learning experiments. These include a forgetting time constant or autoregressive coefficient for the state model (Smith and Brown, 2003) to capture the decline in performance that can be seen when an animal executes multiple trials over many days (Hu et al., 2001) and use of covariates (Smith and Brown, 2003) to help disambiguate learning from response biases and nonrandom strategies. Likewise, the decreasing learning rate frequently seen during learning may be represented by using a state-dependent model for the noise variance in Equation 2.3 (Kakade and Dayan, 2002) (see supplemental data at [www.jneurosci.org](http://www.jneurosci.org)). The model can also be extended to estimate dependence among simultaneously learned tasks such as in our first example in which the monkey learned two to four novel location–scene associations per session (Wirth et al., 2003). Finally, our state–space framework can be combined with hierarchical Bayesian (Gilks et al., 1996) and longitudinal data (Jones, 1993; Fahrmeir and Tutz, 2001) methods to pool information properly across multiple animals executing the same task to estimate simultaneously a population learning curve and a learning curve for each individual animal.

**Appendix**

**A: derivation of the EM algorithm**

Use of the EM algorithm to compute the maximum likelihood estimates of  $\sigma_\epsilon^2$  and  $x_0$  requires us to maximize the expectation of the complete data log likelihood. We treat  $x_0$  as a parameter like  $\sigma_\epsilon^2$ . The complete data likelihood is the joint probability density of  $N_{1:K}$  and  $x$ , which for our model is as follows:

$$p(N_{1:K}, x | \sigma_\epsilon^2, x_0) = \prod_{k=1}^K p_k^{n_k} (1 - p_k)^{1-n_k} p(x | \sigma_\epsilon^2, x_0) = p(N_{1:K} | x) p(x | \sigma_\epsilon^2, x_0), \tag{A.1}$$

where the first term on the right of Equation A.1 is defined by the Bernoulli probability mass function in Equation 2.1, and second term is the joint probability density of the learning state process defined by the Gaussian model in Equation 2.3. At iteration  $(\ell + 1)$  of the algorithm, we compute in the E-step the expectation of the complete data log likelihood given the responses  $N_{1:K}$  across the  $K$  trials and  $\sigma_\epsilon^{2(\ell)}$  and  $x_0^{(\ell)}$ , the parameter estimates from iteration  $\ell$ , which is defined as follows:

$$Q(\sigma_\epsilon^2, x_0 | \sigma_\epsilon^{2(\ell)}, x_0^{(\ell)}) = E[\log[p(N_{1:K}, x | \sigma_\epsilon^2)]] | N_{1:K}, \sigma_\epsilon^{2(\ell)}, x_0^{(\ell)} = E \left[ \sum_{k=1}^K n_k (\mu + x_k) - \log(1 + \exp(\mu + x_k)) \right] | N_{1:K}, \sigma_\epsilon^{2(\ell)}, x_0^{(\ell)} + E \left[ \sum_{k=1}^K -\frac{1}{2} \frac{(x_k - x_{k-1})^2}{\sigma_\epsilon^2} - \frac{K+1}{2} \log 2\pi - \frac{K+1}{2} \log \sigma_\epsilon^2 - \frac{x_0^2}{2\sigma_\epsilon^2} \right] | N_{1:K}, \sigma_\epsilon^{2(\ell)}, x_0^{(\ell)}. \tag{A.2}$$

On expanding the right side of Equation A.2, we see that calculating the expected value of the complete data log likelihood re-

quires computing the expected value of the state variable  $E[x_k | N_{1:K}, \sigma_\epsilon^{2(\ell)}, x_0^{(\ell)}]$  and the covariances  $E[x_k^2 | N_{1:K}, \sigma_\epsilon^{2(\ell)}, x_0^{(\ell)}]$  and  $E[x_k x_{k-1} | N_{1:K}, \sigma_\epsilon^{2(\ell)}, x_0^{(\ell)}]$ . We denote them as follows:

$$x_{k|K} \equiv E[x_k | N_{1:K}, \sigma_\epsilon^{2(\ell)}, x_0^{(\ell)}] \tag{A.3}$$

$$W_{k|K} \equiv E[x_k^2 | N_{1:K}, \sigma_\epsilon^{2(\ell)}, x_0^{(\ell)}] \tag{A.4}$$

$$W_{k,k-1|K} \equiv E[x_k x_{k-1} | N_{1:K}, \sigma_\epsilon^{2(\ell)}, x_0^{(\ell)}] \tag{A.5}$$

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 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  $W_{k,k-1|K}$ . These algorithms for our model are as follows.

*Filter algorithm*

Given  $\sigma_\epsilon^{2(\ell)}$  and  $x_0^{(\ell)}$ , 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 (Brown et al., 1998; Smith and Brown, 2003):

(One-step prediction)

$$x_{k|k-1} = x_{k-1|k-1} \tag{A.6}$$

(One-step prediction variance)

$$\sigma_{k|k-1}^2 = \sigma_{k-1|k-1}^2 + \sigma_\epsilon^{2(\ell)} \tag{A.7}$$

(Posterior mode)

$$x_{k|k} = x_{k|k-1} + \sigma_{k|k-1}^2 [n_k - p_{k|k}] \tag{A.8}$$

(Posterior variance)

$$\sigma_{k|k}^2 = [(\sigma_{k|k-1}^2)^{-1} + p_{k|k}(1 - p_{k|k})]^{-1}, \tag{A.9}$$

where  $p_{k|k}$  is the mode of Equation 2.4 for  $k = 1, \dots, K$  and  $j = k$ . The initial condition is  $x_0 = x_0^{(\ell)}$  and  $\sigma_{0|0}^2 = \sigma_\epsilon^{2(\ell)}$ . The algorithm is nonlinear because  $x_{k|k}$  appears on the left and right of Equation A.8. The derivation of this algorithm for arbitrary point process observation model and linear state–space model is given by Smith and Brown (2003).

*Fixed-interval smoothing algorithm*

Given the sequence of posterior mode estimates  $x_{k|k}$  (Eq. A.8) and the variance estimates  $\sigma_{k|k}^2$  (Eq. A.9), we use the fixed-interval smoothing algorithm (Shumway and Stoffer, 1982; Mendel, 1995; Brown et al., 1998; Smith and Brown, 2003) to compute  $x_{k|K}$  and  $\sigma_{k|K}^2$ . This smoothing algorithm is as follows:

$$x_{k|K} = x_{k|k} + A_k(x_{k+1|K} - x_{k+1|k}) \tag{A.10}$$

$$A_k = \sigma_{k|k}^2 (\sigma_{k+1|k}^2)^{-1} \tag{A.11}$$

$$\sigma_{k|K}^2 = \sigma_{k|k}^2 + A_k^2 (\sigma_{k+1|K}^2 - \sigma_{k+1|k}^2), \tag{A.12}$$

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 (De Jong and Mackinnon, 1988) and is given as follows:

$$\sigma_{k,u|K} = A_k \sigma_{k+1,u|K}, \quad (\text{A.13})$$

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

It follows that the expectations required in Equation A.3 is given by Equation A.10 and that the covariance terms required for Equations A.4 and A.5 of the E-step in the EM algorithm are as follows:

$$W_{k,k-1|K} = \sigma_{k-1,k|K} + x_{k|K}x_{k-1|K} \quad (\text{A.14})$$

and

$$W_{k|K} = \sigma_{k|K}^2 + x_{k|K}^2. \quad (\text{A.15})$$

In the M-step, we maximize the expected value of the complete data log likelihood in Equation A.2 with respect to  $\sigma_\epsilon^{2(\ell+1)}$  and  $x_0^{(\ell+1)}$  giving:

$$\sigma_\epsilon^{2(\ell+1)} = (K+1)^{-1} \left( 2 \left( \sum_{k=2}^K W_{k|K} - \sum_{k=2}^K W_{k-1,k|K} \right) + \frac{3}{2} W_{1|K} - W_{K|K} \right) \quad (\text{A.16})$$

and

$$x_0^{(\ell+1)} = \frac{1}{2} x_{1|K}. \quad (\text{A.17})$$

The algorithm iterates between the E-step (Eq. A.2) and the M-step (Eqs. A.16, A.17) using the filter algorithm, the fixed-interval smoothing algorithm, and the state–space covariance algorithm to evaluate the E-step. The maximum likelihood estimates of  $\sigma_\epsilon^2$  and  $x_0$  are  $\sigma_\epsilon^{2(\infty)}$  and  $x_0^{(\infty)}$ . The convergence criteria for the algorithm are those used by Smith and Brown (2003). The filter algorithm (Eqs. A.6–A.9) evaluated at the maximum likelihood estimates of  $\sigma_\epsilon^2$  and  $x_0$  together with Equation 2.4 give the filter algorithm estimate of the learning curve, whereas the fixed-interval smoothing algorithm (Eqs. A.10–A.12) evaluated at maximum likelihood estimates of  $\sigma_\epsilon^2$  and  $x_0$  together with Equation 2.4 give the smoothing algorithm estimate of the learning curve.

**B: derivation of the probability density of the probability of a correct response at trial  $k$** 

The filter (smoothing) algorithm produces at trial  $k$  the estimate of the state process described as a Gaussian probability density with mean  $x_{k|j}$  and variance  $\sigma_{k|j}^2$  with  $j = k$  ( $j = K$ ). We use the standard change of variable formula from elementary probability theory (Ross, 1993) and Equation 2.2 to derive the probability density of  $p_{k|j}$  as follows:

$$f(p|\mu, x_{k|j}, \sigma_{k|j}^2) = f(x|x_{k|j}, \sigma_{k|j}^2) \left| \frac{dx}{dp} \right| \quad (\text{B.1})$$

$$= (2\pi\sigma_{k|j}^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma_{k|j}^2}(x - x_{k|j})^2\right\} [p(1-p)]^{-1} \quad (\text{B.2})$$

$$= [(2\pi\sigma_{k|j}^2)^{\frac{1}{2}}p(1-p)]^{-1} \times \exp\left(-\frac{1}{2\sigma_{k|j}^2}[\log[p[(1-p)\exp(\mu)]^{-1}] - x_{k|j}]^2\right), \quad (\text{B.3})$$

where Equation B.2 follows from Equation B.1 by the definition of a Gaussian probability density and because by Equation 2.2  $dx/dp = [p(1-p)]^{-1}$ . Equation B.3 follows from Equation B.2 because by Equation 2.2  $x = \log[p[(1-p)\exp(\mu)]^{-1}]$ . The estimate of  $p_{k|j}$  is the mode of Equation B.3.

The 95% lower confidence bound for  $p_{k|j}$  is used to define the trial on which learning occurs. This lower confidence bound  $z(0.05)$  is defined as follows:

$$0.95 = \int_{z(0.05)}^1 f(p|\mu, x_{k|j}, \sigma_{k|j}^2) dp. \quad (\text{B.4})$$

The lower confidence bound is computed by evaluating Equation B.4 numerically using Equation B.3.

**C: probability of a fixed number of consecutive correct responses**

We derive the probability of  $j$  consecutive correct responses ( $CR$ ) in  $K$  independent trials in which the probability of a correct response on each trial is  $p_0$ . If we assume  $2 < j < K$  and let  $m$  denote the trial on which the  $j$   $CR$  start, then we have that  $m = 1, \dots, K - j + 1$ . If  $E$  denotes the event

$$E = \{j \text{ consecutive responses in } K \text{ independent trials}\}$$

and  $E_m$  denotes the  $K - j + 1$  disjoint events defined by

$$E_1 = \{a \text{ length } j \text{ } CR \text{ sequence beginning at trial } 1 \cap \text{any outcomes in trials } j + 1 \text{ to } K\}$$

$$E_2 = \{an \text{ incorrect response at trial } 1 \cap a \text{ length } j \text{ } CR \text{ sequence from trial } 2 \text{ to } j + 1 \cap \text{any outcomes in trials } j + 2 \text{ to } K\}$$

$$m = 3, \dots, j + 1$$

$$E_m = \{any \text{ outcomes in trials } 1 \text{ to } m - 2 \cap an \text{ incorrect response at trial } m - 1 \cap a \text{ length } j \text{ } CR \text{ from trial } m \text{ to } m + j - 1 \cap \text{any outcomes in trials } m + j \text{ to } K\}$$

$$m = j + 2, \dots, K - j + 1$$

$$E_m = \{no \text{ } j \text{ length } CR \text{ beginning in trials } 1 \text{ to } m - 2 \cap an \text{ incorrect response at trial } m - 1 \cap a \text{ length } j \text{ } CR \text{ from trial } m \text{ to } m + j - 1 \cap \text{any outcomes in trials } m + j \text{ to } K\}.$$

By construction, we have that

$$E = \bigcup_{m=1}^{K-j+1} E_m.$$

From the definition of the  $E_m$ s and the independence of the trials, we have

$$\Pr(E_1) = p_0^j \sum_{\ell=1}^{K-j} \binom{K-j}{\ell} p_0^j (1-p_0)^{K-j-\ell} = p_0^j \tag{C.1}$$

$$\begin{aligned} \Pr(E_2) &= (1-p_0)p_0^j \sum_{\ell=1}^{K-j-1} \binom{K-j-1}{\ell} p_0^j (1-p_0)^{K-j-\ell} \\ &= (1-p_0)p_0^j \end{aligned} \tag{C.2}$$

and for  $m = 3, \dots, j + 1$

$$\begin{aligned} \Pr(E_m) &= \sum_{\ell=1}^{m-2} \binom{m-2}{\ell} p_0^{\ell(1-p_0)^{m-2-\ell}} \\ &\cdot (1-p_0)p_0^j \sum_{\ell=1}^{K-j-m} \binom{K-j-m}{\ell} p_0^{\ell} (1-p_0)^{K-j-m-\ell} \\ &= (1-p_0)p_0^j. \end{aligned} \tag{C.3}$$

We now compute  $\Pr(E_m)$  for  $m = j + 2$  to  $K - j + 1$ , and, in so doing, we consider only those events not previously counted. In particular, because  $m \geq j + 2$ , there can be CR sequences of length  $j$  or greater beginning in trials  $E_m$  for  $m = 1, \dots, j + 1$ . These are the events whose probabilities have been computed in Equations C.1–C.3. For each of the current  $E_m$ , we must consider only those events in which there is no CR sequence of length  $j$  or greater beginning in trials 1 to  $m - 2$ . Using the independence of the trials and the definition of  $\Pr(E_m)$ , we have

$$\begin{aligned} \Pr(E_m) &= [1 - \Pr(\text{a CR response sequence of } j \text{ or greater in trials } \\ &\quad \text{1 to } m - 2)] \Pr(\text{of an incorrect response at trial } m - 1) \\ &\Pr(\text{a length } j \text{ CR from trial } m \text{ to trial } m + j - 1) \\ &\Pr(\text{any outcomes in trials } m + j \text{ to } K) \\ &= \left( 1 - \sum_{\ell=1}^{m-j-1} \Pr(E_\ell) \right) \\ &(1-p_0)p_0^j \cdot \sum_{\ell=1}^{K-j-m} \binom{K-j-m}{\ell} p_0^{\ell} (1-p_0)^{K-j-m-\ell} \\ &= \left( 1 - \sum_{\ell=1}^{m-j-1} \Pr(E_\ell) \right) (1-p_0)p_0^j. \end{aligned} \tag{C.4}$$

The  $E_m$  are disjoint events so that

$$\Pr(E) = \Pr\left(\bigcup_{m=1}^{K-j+1} E_m\right) = \sum_{m=1}^{K-j+1} \Pr(E_m). \tag{C.5}$$

For a given  $p$  value and given  $K$ , we pick  $j$  to be the smallest CR sequence such that  $\Pr(E) \leq p$  value. The values in Table 1 were computed for  $p$  values of 0.05 and 0.01.

### References

Box GEP (1980) Sampling and Bayes' inference in scientific modeling and robustness. *J Roy Statist Soc A* 143:383–430.

Brown EN, Frank LM, Tang D, Quirk MC, Wilson MA (1998) A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *J Neurosci* 18:7411–7425.

Casella G, Berger RL (1990) *Statistical inference*, p 373. Belmont, NY: Duxbury.

Cook EP, Maunsell JHR (2002) Attentional modulation of behavioral performance and neuronal responses in middle temporal and ventral intraparietal areas of macaque monkey. *J Neurosci* 22:1994–2004.

De Jong P, Mackinnon MJ (1988) Covariances for smoothed estimates in state space models. *Biometrika* 75:601–602.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J Roy Statist Soc B* 39:1–38.

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

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

Eichenbaum H, Fagan A, Cohen NJ (1986) Normal olfactory discrimination learning set and facilitation of reversal learning after medial-temporal damage to rats: implications for an account of preserved learning abilities in amnesia. *J Neurosci* 6:1876–1884.

Everitt BS, Pickles A (2000) *Statistical aspects of the design and analysis of clinical trials*, pp 255–279. London: Imperial College.

Fahrmeir L, Tutz D (2001) *Multivariate statistical modeling based on generalized linear models*. New York: Springer.

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

Gallistel CR, Mark TA, King A, Latham PE (2001) The rat approximates an ideal detector of changes in rates of reward: implications for the law of effect. *J Exp Psych: Anim Behav Proc* 27:354–372.

Gilks WR, Richardson S, Spiegelhalter DJ (1996) *Markov chain Monte Carlo in practice*. New York: Chapman and Hall.

Hu D, Kubota Y, Graybiel AM (2001) Successive resculpting of task-related activity patterns in the striatum during action-sequence procedural learning, extinction and relearning. *Soc Neurosci Abstr* 27:514.9.

Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM (1999) Building neural representations of habits. *Science* 286:1745–1749.

Jones RH (1993) *Longitudinal data with serial correlation: a state-space approach*. New York: Chapman and Hall.

Kakade S, Dayan P (2002) Acquisition and extinction in autoshaping. *Psych Rev* 109:533–544.

Kitagawa G, Gersh W (1996) *Smoothness priors analysis of time series*. New York: Springer.

Mendel JM (1995) Lessons in estimation theory for signal processing, communication, and control, pp 317–344. Upper Saddle River, NJ: Prentice Hall.

Roman FS, Simonetto I, Soumireu-Mourat B (1993) Learning and memory of odor-reward association: selective impairment following horizontal diagonal band lesions. *Behav Neurosci* 107:72–81.

- Rondi-Reig L, Libbey M, Eichenbaum H, Tonegawa S (2001) 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.
- Ross SM (1993) Introduction to probability models, p 55. London: Academic.
- Roweis S, Ghahramani Z (1999) A unifying review of linear Gaussian models. *Neural Comp* 11:305–345.
- Rustagi JS (1994) Optimization techniques in statistics, p 273. London: Academic.
- Shumway RH, Stoffer DS (1982) An approach to time series smoothing and forecasting using the EM algorithm. *J Time Series Analysis* 3:253–264.
- Siegel S, Castellan NJ (1988) Nonparametric statistics for the behavioral sciences, p 64. New York: McGraw-Hill.
- Smith AC, Brown EN (2003) Estimating a state-space model from point process observations. *Neural Comp* 15:965–991.
- Stefani MR, Groth K, Moghaddam B (2003) Glutamate receptors in the rat prefrontal cortex regulate set-shifting ability. *Behav Neurosci* 117:728–737.
- Whishaw IQ, Tomie J-A (1991) 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.
- Wirth S, Yanike M, Frank LM, Smith AC, Brown EN, Suzuki WA (2003) Single neurons in the monkey hippocampus and learning of new associations. *Science* 300:1578–1584.
- Wise SP, Murray EA (1999) Role of the hippocampal system in conditional motor learning: mapping antecedents to action. *Hippocampus* 9:101–117.
- Yu A, Dayan P (2003) Expected and unexpected uncertainty: ACH and NE in the neocortex. In: *Advances in neural information processing systems* 15. Cambridge, MA: MIT.