# Analysis and Design of Behavioral Experiments to Characterize Population Learning

**Anne C. Smith, Mark R. Stefani, Bita Moghaddam and Emery N. Brown**

---

**You might find this additional information useful...**

This article cites 23 articles, 10 of which you can access free at:
http://jn.physiology.org/cgi/content/full/93/3/1776#BIBL

Updated information and services including high-resolution figures, can be found at:
http://jn.physiology.org/cgi/content/full/93/3/1776

Additional material and information about *Journal of Neurophysiology* can be found at:
http://www.the-aps.org/publications/jn

---

This information is current as of March 7, 2005 .

# Analysis and Design of Behavioral Experiments to Characterize Population Learning

**Anne C. Smith,[1] Mark R. Stefani,[3] Bita Moghaddam,[3] and Emery N. Brown[1,2]**

[1]*Neuroscience Statistics Research Laboratory, Department of Anesthesia and Critical Care, Massachusetts General Hospital, Boston;*
[2]*Division of Health Sciences and Technology, Harvard Medical School/Massachusetts Institute of Technology, Cambridge, Massachusetts;*
*and* [3]*Department of Neuroscience, University of Pittsburgh, Pittsburgh, Pennsylvania*

**Smith, Anne C., Mark R. Stefani, Bita Moghaddam, and Emery N. Brown.** Analysis and design of behavioral experiments to characterize population learning. *J Neurophysiol* 93: 1776–1792, 2005; doi:10.1152/jn.00765.2004. In population learning studies, between-subject response differences are an important source of variance that must be characterized to identify accurately the features of the learning process common to the population. Although learning is a dynamic process, current population analyses do not use dynamic estimation methods, do not compute both population and individual learning curves, and use learning criteria that are less than optimal. We develop a state-space random effects (SSRE) model to estimate population and individual learning curves, ideal observer curves, and learning trials, and to make dynamic assessments of learning between two populations and within the same population that avoid multiple hypothesis tests. In an 80-trial study of an NMDA antagonist's effect on the ability of rats to execute a set-shift task, our dynamic assessments of learning demonstrated that both the treatment and control groups learned, yet, by trial 35, the treatment group learning was significantly impaired relative to control. We used our SSRE model in a theoretical study to evaluate the design efficiency of learning experiments in terms of the number of animals per group and number of trials per animal required to characterize learning differences between two populations. Our results demonstrated that a maximum difference in the probability of a correct response between the treatment and control group learning curves of 0.07 (0.20) would require 15 to 20 (5 to 7) animals per group in an 80-trial (60-trial) experiment. The SSRE model offers a practical approach to dynamic analysis of population learning and a theoretical framework for optimal design of learning experiments.

## INTRODUCTION

Two important challenges are at the center of research to accurately characterize learning as a dynamic process from performance measures recorded in behavioral experiments. First, the most common performance measure is the subject's binary sequence of correct and incorrect responses recorded across the trials in the experiment. Learning is established by using the sequence of trial responses to show that the subject can perform the previously unfamiliar task with a greater reliability than would be expected by chance. Developing optimal algorithms to characterize learning from binary responses is an active research question (Paton et al. 2003; Smith et al. 2004; Wirth et al. 2003).

Second, significant between-subject variation in responses is typical in learning studies. As a consequence, learning experiments often require multiple subjects to execute the same task to characterize the features of the learning process common to the population (Dias et al. 1997; Eichenbaum et al. 1986; Fox et al. 2003; Jonasson et al. 2004; Maclean et al. 2001; Roman et al. 1993; Rondi-Rieg et al. 2001; Stefani et al. 2003; Whishaw and Tomie 1991). Instead of formally characterizing between- and within-subject variation, current analyses of population learning compute only simple proportions of correct responses within a fixed number of trials, across multiple subjects. Furthermore, these analyses use definitions of learning that have been shown to be suboptimal (Smith et al. 2004). These shortcomings of current population analyses of learning have not been addressed.

Use of random effect models to estimate population and individual characteristics from the time series measurements of multiple subjects executing the same protocol is an established paradigm in statistics (Fahrmeir and Tutz 2001; Jones 1993; Laird and Ware 1982; Stiratelli et al. 1984). For learning studies, the random effects approach offers an efficient way to estimate the population curve, as well the individual learning curve for each subject. Although random-effects models have been widely applied in medical, epidemiological, and sample survey research, they have not been used to analyze population learning in behavioral experiments.

We introduced a state-space framework for conducting dynamic analyses of learning in behavioral experiments from time series of binary responses (Smith et al. 2004). The framework provided an estimate of the learning curve and its confidence intervals, gave a precise definition of the learning trial, and characterized learning more accurately and reliably in simulated and actual learning experiments than several currently accepted methods. To develop a dynamic approach to characterize simultaneously population and individual learning performance from time series of binary responses, we extend this framework by defining a state-space model with random effects. We present definitions of the learning curve, learning trial and the ideal observer curve for the population and individuals, and dynamic estimates of between- and within-group differences in learning. We illustrate the new approach by analyzing learning in a group of control rats and a group of rats treated with an NMDA (*N*-methyl-D-aspartate) receptor antagonist in a set-shift task. We also show how the paradigm may be used to design learning experiments optimally.

Address for reprint requests and other correspondence: E. 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).

METHODS

### State-space random effects model of population and individual learning

We assume that learning is a dynamic process that can be studied with the state-space framework (Kitagawa and Gersh 1996; Smith and Brown 2003). The state-space model consists of 2 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 (Fahrmeir and Tutz 2001; Roweis and Ghahramani 1999; Smith and Brown 2003; Smith et al. 2004). Because our objective is to characterize learning for the population and the individual subjects in our study, we formulate a state-space random effects (SSRE) model. That is, we assume that there is a population state learning process and that the learning processes for the individual subjects are drawn from a probability distribution, which has the population learning process as its mean.

We formulate the population and individual state learning processes so that they increase as learning occurs and decrease when it does not occur. From the learning state processes we compute population and individual learning curves that define the probability of a correct response as a function of trial number. The observation equations complete the state-space model setup and define how the observed data relate to the unobservable learning state processes. The data we observe in the learning experiment are the series of correct and incorrect responses for each subject as a function of trial number. Therefore, the objective of the analysis is to estimate the population and individual learning state processes and thus the population and individual learning curves from the observed data.

We conduct our analysis of the experiment from the perspective of an ideal observer. That is, given the state and observation models, we estimate the learning state processes at each trial after seeing the outcomes of all the trials of each subject 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 2-step process. In the first step, we estimate from the observed data the learning state process and thus, the learning curve. In the second step, we estimate when learning occurs by computing the confidence intervals for the population and individual leaning curves or, equivalently, by computing for each trial the ideal observer's assessment of the probability that each subject and the population perform better than chance.

To define the SSRE model, we assume that $J$ subjects participate in a learning experiment with $K$ trials, where we index the trials by $k$ for $k = 1, \ldots, K$ and the subjects by $j$ for $j = 1, \ldots, J$. To define the observation equation we let $n_j^k$ denote the response on trial $k$, from subject $j$ where $n_k^j = 1$ is a correct response and $n_k^j = 0$ is an incorrect response. We let $p_k^j$ denote the probability of a correct response $k$ from subject $j$. We assume that the probability of a correct response on trial $k$ from subject $j$ is governed by an unobservable learning state process $x_k$, which characterizes the dynamics of learning as a function of trial number. At trial $k$, for subject $j$, the observation model defines the probability of observing $n_k^j$ (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

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

where $p_k^j$ is defined by the logistic function

$$p_k^j = \exp(\mu + \beta^j x_k)[1 + \exp(\mu + \beta^j x_k)]^{-1} \qquad (2.2)$$

The parameter $\mu$ in *Eq. 2.1* is determined by the probability of a correct response by chance in the absence of learning or experience

and $\beta^j$ is the learning modulation parameter for subject $j$. We define the random effect component of our state-space model by assuming that the modulation parameters $\beta^j$ are independent Gaussian random variables with mean $\beta_0$ and variance $\sigma_\beta^2 I_{J \times J}$ where $I_{J \times J}$ is a $J \times J$ identity matrix. Therefore, we define the probability of a correct response for the population as

$$p_k = \exp(\mu + \beta_0 x_k)[1 + \exp(\mu + \beta_0 x_k)]^{-1} \qquad (2.3)$$

We define the unobservable learning state process as a random walk

$$x_k = x_{k-1} + \varepsilon_k \qquad (2.4)$$

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

An important concept that underlies all SSRE analyses is *exchangeability*, which means that the response data from each subject in a cohort provide information about the performance of every other subject in the cohort (Gelman et al. 1995). Therefore, the response data from each subject can be used to estimate the population learning curve and to estimate the learning curves for every subject in that cohort. To use the SSRE model optimally, it is key to define subgroups in the experiment for which exchangeability is a reasonable assumption. We illustrate this point in our analyses in the RESULTS.

In the learning experiment, we set the number of trials $K$ and we observe $N_{1:K} = \{n_1, \ldots, n_K\}$, the responses for each of the $K$ trials, where $n_k = \{n_k^1, \ldots, n_k^J\}$ is the set of responses from the $J$ subjects on trial $k$. The objective of our analysis is to estimate $x = \{x_1, \ldots, x_K\}$ $\beta = \{\beta^1, \ldots, \beta^J\}$ and $\theta = (\beta_0, \sigma_\beta^2, \sigma_\varepsilon^2)$ from these, data to estimate $p_k^j$, the probability of a correct response for subject $j$ and $p_k$, the probability of a correct response for the population for $j = 1, \ldots, J$ and $k = 1, \ldots, K$. If we can estimate $x$, $\beta$, and $\theta$ then, by *Eq. 2.2*, we can compute the probability of a correct response as a function of trial number given the data for each of the $J$ subjects and the population. Because $x$ and $\beta$ are unobservable and $\theta$ is an unknown parameter, we use the Expectation–Maximization (EM) algorithm to estimate them by maximum likelihood (Dempster et al. 1977). The EM algorithm is a well-known procedure for performing maximum-likelihood estimation when there is an unobservable process or missing observations. 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 an extension of the algorithm is Smith et al. (2004), and its derivation is given in APPENDIX A. We denote the maximum-likelihood estimate of $\theta$ as $\hat{\theta} = (\hat{\beta}_0, \hat{\sigma}_\beta^2, \hat{\sigma}_\varepsilon^2)$.

### Estimating individual and population learning curves

Given the maximum-likelihood estimates of the $x$ and $\theta$, we can compute for each $x_k x_{k|K}$, the smoothing algorithm (*Eqs. A16–A18*) estimate of the population learning state at trial $k$. It is the estimate of $x_k$ given $N_{1:K}$, all the data in the experiment with the parameter $\theta$ replaced by its maximum-likelihood estimate, where the notation $x_{k|K}$ means the learning state process estimate at trial $k$ given the data up through trial $K$. Similarly, the smoothing algorithm estimate of the individual learning modulation parameters is the estimate of $\beta$ given $N_{1:K}$ with the parameter $\theta$ replaced by its maximum-likelihood estimate. We denote the estimate of the learning modulation parameters as $\beta_{K|K} = (\beta_{K|K}^1, \ldots, \beta_{K|K}^J)$ given in *Eq. A16* of APPENDIX A. The smoothing algorithm gives the ideal observer estimate of the population learning states and the individual modulation parameters.

The smoothing algorithm estimate of the learning state at each trial $k$ is the Gaussian random variable with mean $x_{k|K}$ (*Eq. A16*) and variance, $\sigma_{k|K}^2$ (*Eq. A18*). The smoothing algorithm estimate of $\beta$ is the Gaussian random variable with mean $\beta_{K|K}$ and covariance matrix computed from $W_{K|K}$ defined in *Eq. A18* of APPENDIX A. The individual learning curve for subject $j$ is computed by *Eq. 2.2* at the maximum-likelihood estimates of $x_k$, $\beta^j$, and $\theta$ and is defined as

*Individual learning curve estimate*

$$p^j_{k|K} = \exp(\mu + \beta^j_{K|K}x_{k|k})[1 + \exp(\mu + \beta^j_{K|K}x_{k|k})]^{-1} \quad (2.5)$$

for $k = 1, \ldots, K$. Similarly, the population learning curve estimate is defined from *Eq. 2.4.*

*Population learning curve estimate*

$$p_{k|K} = \exp(\mu + \hat{\beta}_0 x_{k|k})[1 + \exp(\mu + \hat{\beta}_0 x_{k|k})]^{-1} \quad (2.6)$$

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

As *Eqs. 2.5* and *2.6* show, our approach to estimating population learning curves does not simply compute the average of the state-space estimates of the individual learning curves. Instead, using the exchangeability assumption, we estimate the population and individual learning curves simultaneously by extending the EM algorithm we previously developed to estimate individual learning curves (Smith et al. 2004). The key technical point that makes possible this extension is the augmented state-space model in *Eq. A8* (Eden et al. 2004; Jones 1993). This model represents the common learning state process and the individual learning modulation parameters in a single $J+1$-dimensional state equation so that the probability density of the current learning state depends on the value of the previous state (*Eq. 2.4*), whereas the modulation parameters have the same probability density (see text below *Eq. 2.2*) for the entire experiment. In other words, each learning modulation parameter is a random effect (variable) specific to each subject and each has the same probability density at each trial in the experiment. The probability density of the learning state variable changes from trial to trial depending on the value of the previous learning state variable. By using the augmented state-space to represent the properties of our model, we compute in the E-step of the EM algorithm (*Eqs. A16–A18*) both the best estimates of the state variable at each trial and the subject-specific modulation parameters given all the responses of the cohort recorded in the experiment (Jones 1993).

Estimating a common population learning state from the binary responses of subjects belonging to the same cohort is analogous to decoding a biological signal from the spiking activity of an ensemble of neurons using a state-space model to characterize the signal and point process models to represent the spiking activity. For this reason, the filter algorithm (*Eqs. A9–A11*) and smoothing algorithm (A16–A18) used in the E-step of our EM algorithm are respectively the analogs of the Bayes' filter and the Bayes' smoother used in Brown et al. (1998) to decode the position of a rat in its environment from the ensemble spiking activity of place cell neurons in the CA1 region of the animal's hippocampus.

To construct confidence intervals for the learning curves, we must obtain their probability densities. For the population learning curve we can compute the probability density of any $p^j_{k|K}$ using *Eq. 2.2* and the standard change of variables formula from elementary probability theory. That is, the smoothing algorithm estimates the state as the Gaussian random variable with mean $x_{k|K}$ (*Eq. A16*) and variance, $\sigma^2_{k|K}$ (*Eq. A18*). Because the population learning curve estimate is a function of this random variable, we can compute its probability density by standard change of variable formula from elementary probability theory. Applying the change of variable formula to the Gaussian probability density with mean $x_{k|K}$ and variance $\sigma^2_{k|K}$ yields (Smith et al. 2004)

$$f(p|\mu, \beta_0, x_{k|K}, \sigma^2_{k|K}) = [(2\pi\sigma^2_{k|K})^{1/2}\beta_0 p(1-p)]^{-1}$$
$$\times \exp\{-(2\beta_0^2\sigma^2_{k|K})^{-1}[\log[p[(1-p)\exp(\mu)]^{-1}] - x_{k|K}]^2\} \quad (2.7)$$

The individual learning curves are functions of two random variables $\beta^j$ and $x_k$, and the joint distributions of these 2 random variables, given the data $N_{1:K}$, is given by the smoothing algorithm. Because this learning curve is a function of 2 random variables it is more difficult to derive its probability density in closed form. Therefore, we compute it by the Monte Carlo algorithm in APPENDIX B.

*The ideal observer curve and the ideal observer learning trial*

Having estimated the learning curve, we compute for each trial the ideal observer's assessment of the probability that the subject or the population performs better than chance. We term this function the *ideal observer curve.* The ideal observer curve for individual subject $j$ is $\Pr(p^j_{k|K} > p_0)$, where $p^j_{k|K}$ is defined in *Eq. 2.5,* $p_0$ is the probability of a correct response by chance in the experiment and $k = 1, \ldots, K$. We compute this curve for each of the $J$ subjects. The ideal observer curve for the population is $\Pr(p_{k|K} > p_0)$, which is the probability that the population performs better than chance for trials $k = 1, \ldots, K$. The probability that the population performs better than chance on trial $k$ is computed using the smoothing algorithm and *Eq. 2.7*, where the ideal observer curve for each individual is computed using the Monte Carlo algorithm in APPENDIX C. An important advantage of the ideal observer curve is that it provides, together with the learning curve, a dynamic assessment of learning in terms of how sure an ideal observer is that learning has occurred on each trial in the experiment.

Contrary to the approach taken by the current hypothesis-testing methods for analyzing learning, this analysis makes explicit the fact that learning is not a yes–no process (Smith et al. 2004). Nevertheless, for the purpose of making comparisons with these and other methods, it is important to define a learning trial. We define the population (individual) learning trial as the earliest trial in the experiment such that the ideal observer is reasonably certain that the performance of the population (individual) is better than chance from that trial through the balance of the experiment. Because we define learning as performance that is better than chance, identifying a learning trial indicates that learning has occurred. 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)].

In terms of the ideal observer learning curve, we define the learning trial as follows. Given a level of certainty of 0.95, the learning trial of subject $j$ is the earliest trial $r$ such that $\Pr(p^j_{k|K} > p_0) \geq 0.95$ for all trials $k \geq r$. Given a level of certainty of 0.95, the population learning trial is the earliest trial number $r$ such that $\Pr(p_{k|K} > p_0) \geq 0.95$ for all trials $k \geq r$.

For either an individual or the population learning curves, the ideal observer learning trial can be computed from the lower confidence bounds for $p^j_k$ and $p_k$, respectively. The ideal observer learning trial for the individual (population) is the first trial on which the lower 95% confidence bound for the probability of a correct response, $p^j_k$ $(p_k)$ is greater than chance $p_0$ and remains above $p_0$ for the balance of the experiment.

*Comparing learning between and within groups*

An objective of population learning studies is to compare learning between 2 or more groups. This comparison can be carried out in a straightforward way in our paradigm because we have the probability distribution associated with each learning curve (*Eq. 2.7*). Therefore, given any 2 learning curves we can compute at each trial, the probability that curve one is greater than curve two, or vice versa, and plot this probability as a function of trial number. Therefore, we can state for each trial how sure we are that one curve is greater than the other, and test hypotheses about differences in learning between the 2 groups. We explain in APPENDIX C how we compute these comparison probabilities by Monte Carlo from the probability models for learning curves of 2 different groups.

Another objective of population and individual learning studies is to compare learning within a group. This comparison can also be carried out in a straightforward way in our paradigm because we estimate the joint probability distribution associated with each learning curve (*Eq. 2.7*). Therefore, given any 2 trials we can compute the probability that the population (individual) performance at one trial is greater than the performance at any other trial. A plot of this 2-di-

Journal of Neurophysiology

mensional comparison for all trial pairs illustrates how sure we are that performance on one trial is greater than performance on any other trial. We explain in APPENDIX D how we compute these comparison probabilities by Monte Carlo from the joint probability distribution of the learning states for a given group.

The Matlab (MathWorks, Natick, MA) code for the algorithms we present here can be downloaded from our website: https://neurostat. mgh.harvard.edu/BehavioralLearning/Matlabcode.

### Learning analysis using the 8-trial blocks and the 8 consecutive correct response methods

Stefani et al. (2003) estimated population learning curves from group responses in learning experiments by computing the fraction of correct responses across all animals in nonoverlapping blocks of 8 trials. We termed this method the 8-trial blocks (8TB) method. This gave a 10-point estimated learning curve for each group. Stefani and colleagues considered an animal to have learned the task when it gave 8 consecutive correct responses. We termed this method the 8 consecutive correct responses (8CCR) method. We compared the 8TB method with our state-space random-effects method for estimating the population learning curve and the 8CCR method with our IO(0.95) method for identifying the learning trial.

### Experimental protocol for a set-shift task

To illustrate the performance of our method on actual experimental data, we analyzed the responses from 2 groups of rats performing a set-shift task. In the set-shift task the animal learned one task during the first phase (Set 1) then during a second phase (Set 2) had to shift and learn a second task with the confound of the response options of the first task were present as the animal learned the second task (Stefani et al. 2003). The task consisted of 2 discriminations, performed on consecutive days in the same 4-arm maze. The arms of the maze differed along 2 stimulus dimensions: texture and brightness. Texture was either rough or smooth and brightness was either light or dark. For each trial, one arm was blocked so that the maze was in a T-configuration. Thus, from each start arm a rat had a choice of a left or right turn, and simultaneously by design, a choice between rough and smooth, and a choice between light and dark (Fig. 1). Each trial began from a different start arm, chosen pseudo-randomly so that in each block of 8 consecutive trials there were 2 starts from each of the 4 arms.

On the first day (Set 1), rats were trained to discriminate maze arms on the basis of one of the 2 stimulus dimensions. For a given dimension, a rat was rewarded only for making entries into arms with a particular stimulus attribute. For example, if the dimension was texture and reward was associated with rough texture, then the rat would be rewarded only if it chose the arm with the rough texture regardless of brightness. On the second day (Set 2), rats were trained on the other dimension. That is, a rat trained to discriminate arm texture on Set 1, was trained to discriminate arm brightness on Set 2. To avoid overtraining, each rat was trained to a criterion of 8 consecutive correct arm entries on Set 1. The choice of 8 consecutive correct responses had been chosen by Stefani et al. (2003) as the training criterion because, under an assumption of trial independence, this event gave an approximate significant level of less than 0.005. Pseudorandomization of the start arm order ensured that using a criterion of 8 consecutive correct responses, a rat would have to make correct choices from each of 4 start arms at least once, and usually twice, thereby strengthening the determination that learning had occurred. On Set 2, each rat was trained for 80 consecutive trials regardless of performance. The 80-trial maximum for Set 2 was adopted because it was the number of ten 8-trial blocks that pilot data collected before the experiments in Stefani et al. (2003) indicated was sufficient to demonstrate stable performance in the control rats and to
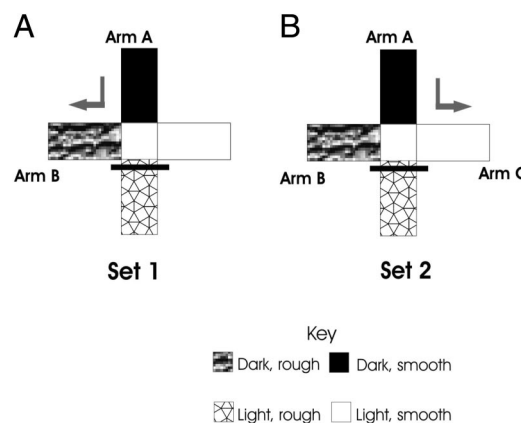


FIG. 1. Design of the set-shift task. There are 2 dimensions, brightness and texture, and 2 categories within brightness (light and dark) and 2 categories within texture (smooth and rough). At any given trial the maze was in a T-configuration (i.e., the opposing arm was blocked, shown as the black horizontal bar) with all 4 arms used as start arms. *A*: in Set 1, each rat was trained to discriminate one of the categories of brightness to receive a food reward. Each rat was trained to seek reward until it achieved 8 consecutive correct responses on Set 1. In the example shown, the rat was rewarded for choosing the dark arm, regardless of the arm's texture. Thus, if it started in arm A and entered arm B, it would be rewarded. *B*: in Set 2, the rat was trained over 80 trials to seek reward according to one of the categories of texture. In the example shown, the rat was rewarded for choosing the smooth arm. Thus, if it again started in arm A, it would have to enter arm C to get reward, ignoring its previous training in Set 1 that the dark arm (arm B) contained a reward.

distinguish performance differences between the control and drug-treated groups.

Twenty minutes before beginning training on Set 2, each rat received a bilateral microinjection into the medial prefrontal cortices of either a vehicle solution (145 mM NaCl, 2.7 mM KCl, 1.0 mM $MgCl_2$, and 1.2 mM $CaCl_2$) or the vehicle solution plus the NMDA-receptor antagonist MK801 at a dose of 3 $\mu g$ per hemisphere. The hypothesis tested by Stefani and colleagues was that treatment with MK801 should alter the ability of the rats to execute the set-shift compared to the animals receiving the vehicle. We termed animals receiving only the vehicle solution the Vehicle group and animals receiving the vehicle solution with MK801 the Treatment group.

### RESULTS

#### Learning in a set-shift task

To illustrate application of our methods, we analyzed the learning behavior of the Vehicle and Treatment groups from Set 2 from the Brightness–Texture part of the set-shift experiment. The trial responses are shown in Fig. 2 as blue and red marks corresponding respectively to correct and incorrect responses. Figure 2*A* and 2*B* (2*C* and 2*D*) are the responses from the Vehicle (Treatment) group. We subdivided each group according to the rewarded arm in Set 1. Figure 2*A* (2*C*) are the Vehicle (Treatment) animals rewarded for the light reward arm in Set 1 and Fig. 2*B* (2*D*) are the Vehicle (Treatment) animals rewarded for the dark arm Set 1. We denote the subgroups Vehicle light, Vehicle dark, Treatment light, and Treatment dark.

The 13 animals in the Vehicle group performed at or above chance from the outset of this experiment (Fig. 2*A* and *B*). In the first 4 trials, the numbers of correct responses were 8 of 13, 7 of 13, 9 of 13, and 9 of 13. For the Vehicle animals, there was a noticeable improvement in performance in the second half of the experiment. For example, animals 2, 3, 4, and 5 in the
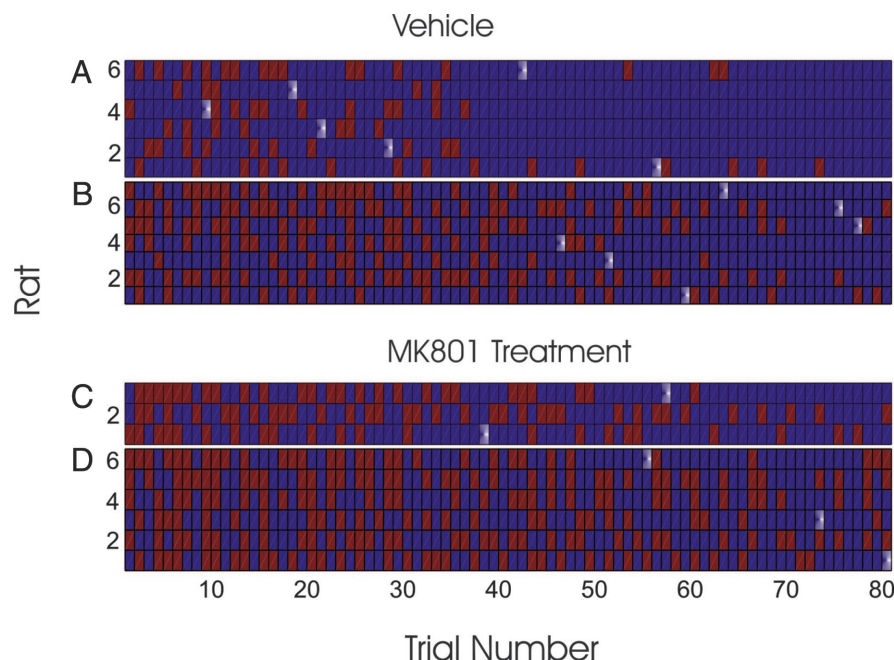
FIG. 2. Responses of the 13 rats in the Vehicle group (A, B) and 9 rats in the MK801 Treatment group (C, D) during the 80 trials of Set 2. Blue and red squares indicate correct and incorrect responses, respectively. Each row gives the responses of a different animal and each column is a different trial. Both groups were trained to discriminate one of the categories of the brightness dimension in Set 1. *A* (*B*) are the responses of the 6 (7) Vehicle rats that were rewarded for entering the lighter (darker) arm in Set 1. *C* (*D*) are the responses of the 3 (6) Treatment rats that were rewarded for entering the lighter (darker) arm in Set 1. Light blue squares indicate the 8CCR (consecutive correct responses) learning trial estimate of the learning trial for each individual. Rats not achieving the 8CCR criterion were assigned a learning trial of 80.

Vehicle light subgroup (Fig. 2*A*) had all correct responses from trial 37 to trial 80. Similarly, animals 3, 4, and 7 in the Vehicle dark subgroup had all correct responses from trials 61, 50, and 55 respectively to trial 80 (Fig. 2*B*).

The performance of the 9 Treatment animals began close to or slightly below chance with 4 of 9, 3 of 9, 4 of 9, and 3 of 9 correct responses in the first 4 trials (Fig. 2, *C* and *D*). At the end of the 80-trial experiment, the performance of the Treatment group was greater than chance, but with many more incorrect responses than the Vehicle group. Only animal 3 (Fig. 2*C*) in the Treatment light subgroup had an uninterrupted sequence of correct responses at the end of the experiment. This sequence began at trial 60, much later than those of animals 2, 3, 4, and 5 in the Vehicle light subgroup (Fig. 2*A*).

We performed 3 analyses using the state-space paradigm: *1*) a state-space (SS) analysis in the Vehicle and the Treatment groups in which the response data are pooled across all the animals in each group (Smith et al. 2004); *2*) a state-space analysis of the response data of each individual animal; and *3*) a state-space random effects (SSRE) analysis on the response data within each of the 4 subgroups: Vehicle light, Vehicle dark, Treatment light, and Treatment dark. The state-space analysis of the pooled response data illustrated population learning curve estimation under the assumption that there was no between-subject variation. The state-space analysis of individual responses illustrated the estimation of individual learning curves under the assumption that there was no common or population feature shared by the members of any of the subgroups. The state-space random effects analysis illustrated characterization of between-subject variation in learning by estimating simultaneously population and individual learning curves within a subgroup.

We compared the learning curves estimated from our state-space methods with the learning curve estimated by the 8-trial nonoverlapping block method (8TB) (Stefani et al. 2003) and we compared our IO(0.95) learning trial estimates from the state-space analyses with those computed by the 8 consecutive correct responses criterion (8CCR) (Stefani et al. 2003).

*Analysis of learning from the pooled responses within the vehicle and the treatment groups*

We first analyzed the response data without taking into account the reward arm during Set 1. That is, we combined the responses across the Vehicle light (Fig. 2*A*) and the Vehicle dark (Fig. 2*B*) subgroups and analyzed the experimental data as the number of correct responses from the 13 animals by trial across the 80 trials. For the Treatment group, we combined the responses across the Treatment light (Fig. 2*C*) and the Treatment dark (Fig. 2*D*) subgroups and analyzed the experimental data as the number of correct responses from the 9 animals by trial across the 80 trials. This analysis thereby assumes that there is no between-subject variation within either the Vehicle or the Treatment group.

To do this, we replaced the Bernoulli model in *Eq. 2.1* with the binomial observation model

$$\Pr(n_k|p_k, x_k, m) = \binom{m}{n_k} p_k^{n_k}(1 - p_k)^{m-n_k} \qquad (3.1)$$

where *k* indexes the trial, *m* is 13 (9) animals for the Vehicle (Treatment) group, and $n_k$ is now the number of correct responses from the *m* Vehicle (Treatment) animals on trial *k*. As in Smith et al. (2004), we defined $p_k$ as

$$p_k = \exp(\mu + x_k)[1 + \exp(\mu + x_k)]^{-1} \qquad (3.2)$$

and fit the Gaussian state-space model in *Eq. 2.4* to the pooled response data by an EM algorithm using *Eqs. 3.1* and *3.2* as the observation model.

The SS learning curve estimated from the pooled responses of the Vehicle group provided a trial-by-trial estimate of the probability of a correct response that increased monotonically from 0.58 on trial 1 to 0.94 on trial 80 (Fig. 3*A*). This learning curve was >0.5, the probability of a correct response by chance (Fig. 3*A*, horizontal dashed line), for the entire experiment. The behavior of this learning curve is consistent with the performance apparent from the pattern of correct and incorrect responses seen in the Vehicle group (Fig. 2, *A* and *B*).
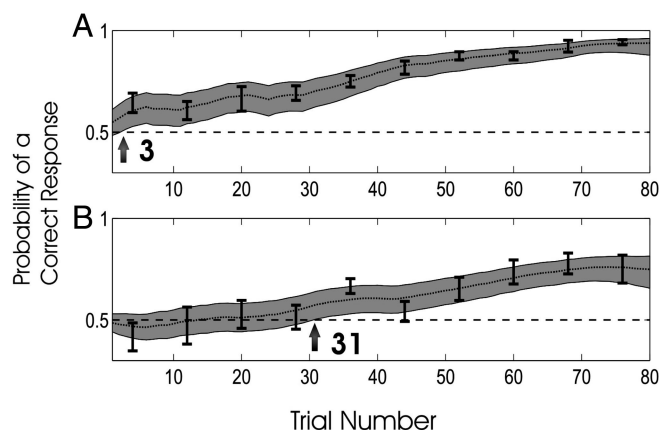
FIG. 3. State-space analysis of the pooled responses within the Vehicle group (*A*) and within the Treatment group (*B*). Probability of a correct response by chance is 0.5 (dashed horizontal line). Dotted lines are the learning curves estimated for the Vehicle (*A*) and Treatment groups (*B*) using the state-space approach and the binomial observation model (*Eq. 3.1*). Gray-shaded area is the 90% confidence region. Arrow indicates the ideal observer [IO(0.95)] learning trials of 3 (*A*) and 31 (*B*) for the Vehicle and Treatment groups, respectively; 10 black error bars are the learning curve estimates (with SE) computed by the 8TB method.

The SS learning curve estimated from the pooled responses of the Treatment group began at 0.5, decreased slightly to 0.46 at trial 5, increased almost monotonically to a maximum of 0.79 at trial 70, and decreased to 0.77 at trial 80 (Fig. 3*B*). The behavior of this learning curve is also strongly consistent with the performance apparent from the pattern of correct and incorrect responses of the Treatment group (Fig. 2, *C* and *D*). In particular, the large number of incorrect responses in the early trials was the reason this learning curve initially fell to <0.5. Similarly, the several incorrect responses on trials 77 to 80, particularly in the Treatment dark subgroup, were responsible for the decline in the learning curve at the end of the experiment.

The learning curve for the Vehicle group lies above the learning curve for the Treatment group at each trial. The IO(0.95) learning trial for the Vehicle group was trial 3 because this is where the lower 95% confidence bound of the learning curve crossed 0.5 (Fig. 3*A*). For the Treatment group the IO(0.95) learning trial was trial 31 (Fig. 3*B*). These results show the strong effect of the MK801 on the learning process.

To compare our state-space model analysis of the pooled responses with the approach taken in Stefani et al. (2003), we estimated for both the Vehicle and the Treatment groups the learning curves using the 8TB method and we identified the learning trials for both groups using the 8CCR method. The population learning curve computed using the 8TB method provided only 10 estimates for the 80 trials for each of the 2 groups. For the Vehicle group, this curve (Fig. 3*A,* black SE error bars) increased from 0.64 in the first block to 0.92 in the 10th block. This learning curve was in close agreement with the SS learning curve for this group. Neither this curve nor any of its lower SE bars—defining an approximate 67% confidence interval in each block—dropped below 0.5 (dashed horizontal line), the probability of a correct response by chance. For the Treatment group (black SE error bars, Fig. 3*B*), the population learning curve began at 0.41 in the first block, increased to a maximum of 0.69 in the ninth block, and decreased slightly to 0.65 by the last block. This learning curve was also in close

agreement with the corresponding SS learning curve. As was true for the SS learning curves for the Vehicle and Treatment groups, the 8TB learning curve for the Treatment was below the 8TB learning curve for the Vehicle group at each of the 10 trial estimates.

The 67% confidence intervals for the 8TB learning curve for the Vehicle group were wider through trial 24 and became smaller for the trials beyond trial 24 (Fig. 3*A*). The 90% confidence intervals for the SS learning curve showed a similar change in width. Based on the width of the 67% confidence intervals for the 8TB learning curve the corresponding 90% confidence intervals for the 8TB learning curve would be larger than the 90% confidence intervals for the SS learning curve. The larger intervals occurred because for the vehicle group each 8TB confidence interval was based on 8 trials $\times$ 13 = 104 observations, whereas the SS confidence intervals were based on 80 trials $\times$ 13 = 1,040 observations. The 67% confidence intervals for the 8TB learning curve were slightly wider than 90% confidence intervals for the SS Treatment group learning curve because the SS confidence intervals were based on all the 80 trials $\times$ 9 animals = 720 observations, whereas each 8TB interval was based on only 8 trials $\times$ 9 animals = 72 observations (Fig. 3*B*).

The population learning trial in the analysis of Stefoni et al. (2003) was computed by using the 8CCR method to compute the learning trial for each animal in the Vehicle (Treatment) group (Fig. 2, light blue squares) and then taking the population learning trial to be the mean of the individual Vehicle (Treatment) learning trial estimates. The mean (median) of the individual learning trials for the Vehicle group was trial 48.1 (51). As predicted by the analyses of actual and simulated data in Smith et al. (2004), this learning trial is much later than the IO(0.95) learning trial estimate of trial 3 for this group. The mean (median) of the individual learning trials for the Treatment group was trial 70.0 (80) under the assumption used by Stefani and colleagues that trial 80 was assigned as the learning trial to an animal that did not reach the criterion of 8 consecutive correct responses. This differed from the IO(0.95) learning trial for this group of trial 31. For both the Vehicle and the Treatment groups the population learning trial was later than what might have been expected by analyzing the 8TB learning curve. This discrepancy between the 8TB learning curve estimate and the 8CCR estimate of the learning trial arises because the two methods, unlike the SS learning curve estimate and the IO(0.95) learning trial, are not related. In particular, it is possible to have many more correct than incorrect responses yet not have 8 consecutive correct responses. The 8TB learning curves agree closely with the SS learning curves in this pooled analysis and showed clearly the difference in learning between the Vehicle and Treatment groups. By construction the IO(0.95) learning trials gave estimates of the learning trial consistent with the SS learning curves. The 8CCR learning trials suggest that learning occurred much later in the Vehicle group and perhaps not at all in the Treatment group.

*State-space analysis of individual learning within the vehicle and treatment groups*

The pooled analysis treated all the responses within each group as if there was no subject-specific effect. To analyze the learning on a subject-specific basis, we estimated the SS learning curve for each animal using the state-space model for
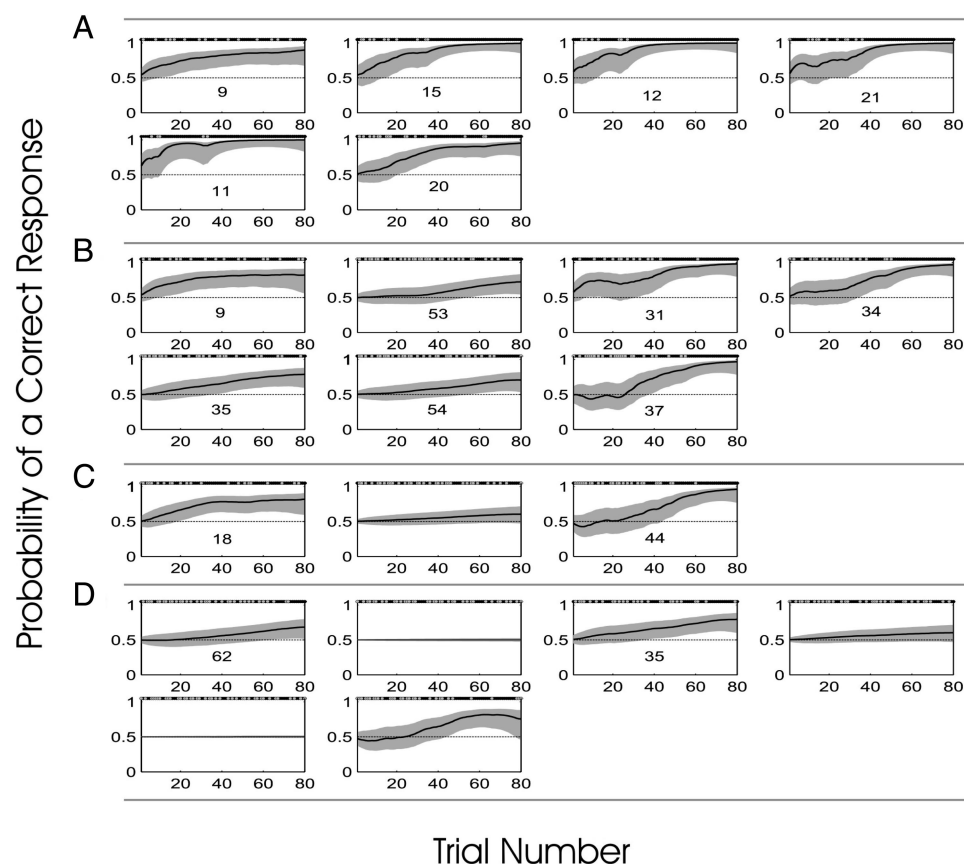
FIG. 4. State-space analysis of the individual responses for each animal. Probability of a correct response occurring by chance is 0.5 (horizontal dashed lines). *A*: solid black curves are the learning curves for the 6 rats in the Vehicle light group and the gray-shaded areas are the associated 90% confidence regions. Correct/incorrect responses for each animal are shown as black/gray squares above each panel (as in Fig. 2*A*). The IO(0.95) learning trial for each animal is marked on each panel. *B*: learning curves for the 7 rats in the Vehicle dark group. Learning curves for the 3 rats in the Treatment light group (*C*) and the 6 rats in the Treatment dark group (*D*). Five of the 9 rats in the Treatment group did not meet the criterion for an IO(0.95) learning trial.

a single individual defined in Smith et al. (2004) (Fig. 4). This corresponded to using the state-space model in *Eq. 2.4* and observation model in *Eqs. 3.1* and *3.2* with $m = 1$. For the Vehicle group all the SS learning curves increased. In agreement with the responses (Fig. 2), the individual SS learning curves for the Vehicle light subgroup (Fig. 4*A*) increased more rapidly than the individual SS learning curves for the Vehicle dark subgroup (Fig. 4*B*). The IO(0.95) learning trials for the Vehicle light (dark) subgroup ranged from trial 9 (9) to 21 (54) with a median of trial 13.5 (35).

For the Treatment group, the SS learning curves had a wider range of shapes (Fig. 4, *C* and *D*). Four of the animals—one in the Treatment light group and 3 in the Treatment dark group—had essentially flat learning curves. For these animals, no IO(0.95) learning trial could be identified. For animal 7 in the Treatment dark group (Fig. 4*D*), the SS learning curve mirrored the SS learning curve for the Treatment group in the pooled analysis (Fig. 3*B*). It began below 0.5, decreased further until trial 5, increases monotonically to 0.75 at trial 62 then decreases to 0.70 at trial 80. Because the lower 95% confidence bound dropped below 0.5 on trial 80, this animal did not meet the strict IO(0.95) definition of learning. The remaining 4 animals, 2 in the Vehicle light group and 2 in the Vehicle dark group, all had monotonically increasing learning curves with IO(0.95) learning trials 18, 44, 62, and 35. The width of the 90% confidence intervals for the individual analyses are wider than the confidence intervals in the pooled analyses because each of the former is based on only 80 observations, whereas the latter were based on 1,040 (720) observations for the Vehicle (Treatment) groups.

These analyses confirm the finding from the pooled analysis that learning in the Treatment group was impaired relative to the Vehicle group. They also show that learning differed as well between the Vehicle light and the Vehicle dark subgroups) and, even though the numbers were small, there was a difference in learning between the Treatment light and the Treatment dark subgroups. These analyses further suggest that because the learning behavior was similar within the Vehicle and Treatment subgroups, the SSRE analysis should be carried out within these subgroups.

*State-space random effects analysis of population and individual learning*

We applied the SSRE analysis to the Vehicle subgroups (Fig. 5, *A* and *B*) and Treatment subgroups (Fig. 5, *C* and *D*). For the Vehicle light subgroup, the SSRE population learning curve (Fig. 5*A*, red line) increased monotonically from 0.5 at trial 1 to 0.99 by trial 45 and remained constant at this level for the balance of the experiment. The individual SSRE learning curves (Fig. 5*A*, green lines) were distributed evenly about this population learning curve. The 90% confidence intervals for the population learning curve (Fig. 5*A*, gray shaded region) were wide through trial 30 and began to decrease as the learning curve began to climb monotonically. The lowest individual learning curve, which was slightly below the lower 95% confidence bound for the population learning curve toward the end of the experiment, corresponded to animal 1. This animal continued to make errors throughout the experiment (Fig. 2*A*). The IO(0.95) learning trial identified from the SSRE population learning curve is trial 11 (Fig. 5*A*). The ideal
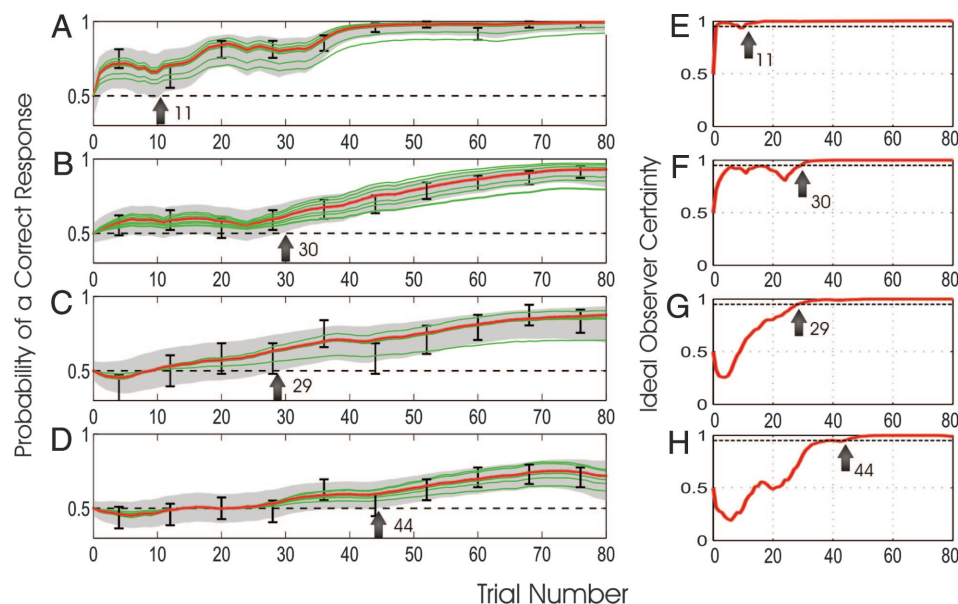
FIG. 5. State-space random effects (SSRE) estimation of population and individual learning. Population learning curves (red curves) for the Vehicle light (*A*), Vehicle dark (*B*), Treatment light (*C*), and Treatment dark (*D*) subgroups and the associated 90% confidence intervals (gray-shaded area). Green curves are the individual SSRE learning curves in each subgroup. Black error bars are the learning curve estimates from the 8TB method ± SE. Arrows indicate the respective IO(0.95) learning trials. Ideal observer curves (red curves) for the Vehicle light (*E*), Vehicle dark (*F*), Treatment light (*G*), and Treatment dark subgroups (*H*). All the ideal observer curves (red curves) start at 0.5, the probability of a correct response by chance. Trial where an ideal observer curve crosses and stays above 0.95 (dashed horizontal line) is the IO(0.95) learning trial for that group.

observer curve (Fig. 5*E*) showed that the IO(0.95) learning trial would have occurred earlier were it not for a series of incorrect responses on trials 9 and 10 (Fig. 2*A*).

For the Vehicle dark subgroup, the SSRE population learning curve (Fig. 5*B*, red line) increased slightly from 0.5 at trial 1, to 0.60 at trial 5, remained constant and slightly decreased to 0.57 at trial 24, and then increased monotonically to 0.94 at trial 80. The individual SSRE learning curves (Fig. 5*B*, green lines) were distributed evenly about this population learning curve. The individual learning curves (green lines) were much closer to the population learning curve in the first half of the experiment, especially at trials 11 and 25, where nearly all the animals in this subgroup made incorrect responses. The 90% confidence intervals for the population learning curve (Fig. 5*B*, gray shaded region) were approximately constant in their width. The IO(0.95) learning trial identified from the SSRE population learning curve was trial 30 (Fig. 5*C*), a trial shortly after the learning curve began its monotonic increase. The ideal observer curve for the Vehicle dark group (Fig. 5*F*) had an initially sharp increase, and remained just below the 0.95 level of certainty before crossing this level on trial 30. These results support the suggestion from the individual analysis in the previous section that the learning behavior differs between the 2 Vehicle subgroups.

The Treatment light subgroup had only 3 animals (Fig. 2*C*). For this subgroup, the SSRE population learning curve (Fig. 5*C*, red line) decreased slightly from 0.5 at trial 1, to 0.45 at trial 5, and then increased monotonically to 0.75 at trial 80. The 90% confidence intervals for this population learning curve were broad across the entire experiment because the responses were pooled across only 3 animals. One of the individual SSRE learning curves (Fig. 5*C*, green lines) was above the population learning curve and one was almost indistinguishable from the population curve. The third learning curve, which was well below the population learning curve, corresponded to animal 2. This animal's individual learning curve increased only slightly above 0.50 (Fig. 4*C*) and its analysis did not identify an IO(0.95) learning trial. In this case, the good performance of the other 2 animals in this subgroup (with individual IO(0.95) learning trials of 18 and 44) pulled up the learning curve of this

animal. The population ideal observer curve (Fig. 5*G*) mimicked the behavior of the population learning curve and identified the population IO(0.95) as trial 29.

For the Treatment dark subgroup, the SSRE population learning curve (Fig. 5*D*, red line) decreased slightly from 0.5 at trial 1, to 0.45 at trial 5, increased slightly and remained constant at trial at 0.5 from trial 11 to trial 27. From this trial, the population learning curve increased monotonically to 0.70 at trial 70 and decreased to 0.65 at trial 80. The 90% confidence intervals for this population learning curve had a constant width across the entire experiment that was narrower than the widths of the 90% confidence intervals for the learning curve of the Treatment light subgroup. Between trials 11 to 27 the population and individual learning curves were indistinguishable because nearly all 6 of the animals in this subgroup made many incorrect responses in this interval. The population IO(0.95) learning trial for this group was trial 44 (Fig. 5*D*, 5*H*).

From the individual learning curve analyses (Fig. 4*D*), we concluded that 4 of the 6 animals did not learn by the IO(0.95) learning criterion, whereas the remaining 2 animals learned at trials 62 and 35. From the individual learning curves computed as part of the SSRE analysis, we found 5 of the 6 animals had learning trials that ranged from trial 44 to 49. As was true for the 3 animals in the Treatment light subgroup, by pooling the data to estimate the population and individual learning curves for the Treatment dark group, more animals showed learning than would be indicated by the individual analyses. Moreover, the population analysis showed that although the individual animals in this subgroup performed poorly at the outset of the experiment, this subgroup showed population learning. The 15-trial difference in the learning trial between the Treatment light and the Treatment dark group suggests that learning in these 2 subgroups was different.

For each subgroup, the 8TB learning curve agreed with the SSRE population learning curves (Fig. 5). For each subgroup, the 67% confidence intervals for the 8TB learning curves were close to the width of the 90% confidence intervals for the SSRE learning curves because the former were computed from only the points in the given 8-trial block, whereas all the SSRE confidence intervals are based on all the responses in each

group. Whereas the IO(0.95) learning trials for the SSRE analysis for the Vehicle light, Vehicle dark, Treatment light, and the Treatment dark were trials 11, 30, 29, and 44, respectively, the 8CCR learning trial estimates for these groups were identified much later at trials 25, 63, 58, and 80, respectively.

The SSRE analysis computed, for each subgroup, the population learning and an individual learning curve for each member of the subgroup. In this way, the data from each group member contributed to the individual learning curve estimate of every other group member. For each animal, we also computed its SS individual learning curve (Fig. 4) (Smith et al. 2004). To show the difference between an individual learning curve estimated from the SSRE analysis and the individual SS learning curve we compared these 2 learning curves for animal 2 from the Vehicle light group and animal 4 from the Treatment dark group. Animal 2 clearly performed better than chance in the task (Fig. 6*A*), with no incorrect responses from trial 36 to the end of the experiment. Based on the responses it was less apparent whether animal 4 performed better than chance. It had 47/80 correct responses with 9/11 correct responses in the last 11 trials of the experiment (Fig. 6*B*).

For both animals the SSRE and SS learning curves resemble each other, but with some noticeable differences. The SSRE learning curves (Fig. 6, black dotted lines) were more variable than their SS counterparts (Fig. 6, gray solid lines), reflecting the variability in the responses across the entire subgroup. The SSRE individual learning curves resembled more closely the population learning curves (Fig. 5*A*) for their respective subgroups than the corresponding individual SS learning curve. The confidence intervals for the SSRE learning curves are narrower than those for the SS learning curves because the former used all the data in the subgroup in their estimation. The SS IO(0.95) learning trial for animal 2 was trial 15 (Fig. 6*A*), whereas the SSRE IO(0.95) learning trial was trial 11, which in this case, was also the population learning trial estimate for the Vehicle light subgroup. For animal 4, no learning trial was
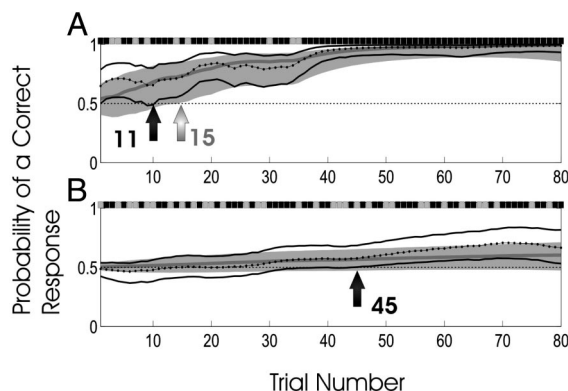


FIG. 6. Comparison between the individual learning curves computed using SS and SSRE analyses for (*A*) rat 2 from the Vehicle light group and (*B*) rat 4 from the Treatment dark group showing the SS learning curve estimates (gray curves) and the SSRE learning curve estimates (dotted black curves). Correct/incorrect responses for each animal are shown respectively as black/gray bars above each panel. Gray-shaded regions and black lines are, respectively, the SS and SSRE 90% confidence intervals computed by Monte Carlo (APPENDIX B). The SSRE 90% confidence intervals are narrower than the SS intervals, particularly in the latter trials, reflecting more certainty that the animal has learned based on the subgroup's performance (*A*). SS and SSRE IO(0.95) learning trials are 15 and 11 for animal 2, respectively. For animal 4 the SSRE IO(0.95) learning trial was trial 45 and no SS IO(0.95) learning trial was identified (*B*).

identified by the SS analysis; however, the individual SSRE IO(0.95) learning trial for this animal based as well on the performance of the other 5 animals in this subgroup was trial 45, one trial after this subgroup's IO(0.95) learning trial of 44.

*Comparing population learning between the vehicle and treatment subgroups*

The aim of the experiment was to test the effect of MK801 on the ability of the rats to shift a learned strategy. As a result, we were interested in whether the learning curves for the treatment animals were different from the learning curves for the vehicle animals. Because we predicted that MK801 would impair learning, we estimated the trial-by-trial probability that the population performance in the Vehicle light (dark) subgroup was greater than the population performance in the Treatment light (dark) subgroup (Fig. 7*A*). That is, using the Monte Carlo algorithm in APPENDIX C, we computed $\Pr(p_k^{\text{Vehicle light}} > p_k^{\text{Treatment light}})$ and $\Pr(p_k^{\text{Vehicle dark}} > p_k^{\text{Treatment dark}})$ for trials $k = 0, \ldots, K$. We considered the performance in the Vehicle subgroup to be greater than the performance in corresponding Treatment subgroup on trial $k$ if this probability was $\geq 0.95$.

The performance in the Vehicle light subgroup was better than the performance in the Treatment light subgroup from trials 2 to 6, trials 16 to 27, and from trial 35 to the end of the experiment at trial 80 (Fig. 7*A,* blue line). Similarly, the performance in the Vehicle dark subgroup was better than the performance in the Treatment dark subgroup from trials 6 to 7 and from trial 42 to the end of the experiment (Fig. 7*A,* red line). We also found that the performance in the Vehicle light subgroup was better than the performance in the Vehicle dark subgroup from trials 17 to 31, and from trial 35 to trial 80 (Fig. 7*B,* blue line). On the other hand, the level of certainty that the performance in the Treatment light subgroup was better than the performance in the Treatment dark subgroup for any trial in the experiment was never >0.90 (Fig. 7*B,* red line).

We concluded from this analysis that the rats injected with the NMDA-receptor antagonist MK801 were significantly impaired in their ability to learn compared to those injected only with the vehicle for most of the later half of the experiment (trial 42 to trial 80). We also concluded that although performance in the Vehicle light subgroup was better than performance in the Vehicle dark subgroup, a difference in performance between the Treatment light and the Treatment dark subgroups was less apparent.

*Comparing learning within the vehicle and treatment subgroups*

The learning trial identifies the trial on which the ideal observer is 0.95 certain that the animal is performing better than chance from that trial through the balance of the experiment. This analysis compares the performance on trial 0 with performance on each of the 80 trials. Another frequently asked question is whether learning performance differs between trials within a group. In these analyses, learning in the later trials of the experiment is frequently compared with learning in the earlier trials. Using the Monte Carlo algorithm in APPENDIX D, we computed $\Pr(p_{k_2} > p_{k_1})$, the probability that the learning curve at trial $k_2$ was greater than the learning curve at trial $k_1$ for all $k_1 < k_2$. These results consist of $K(K + 1)/2$ within-
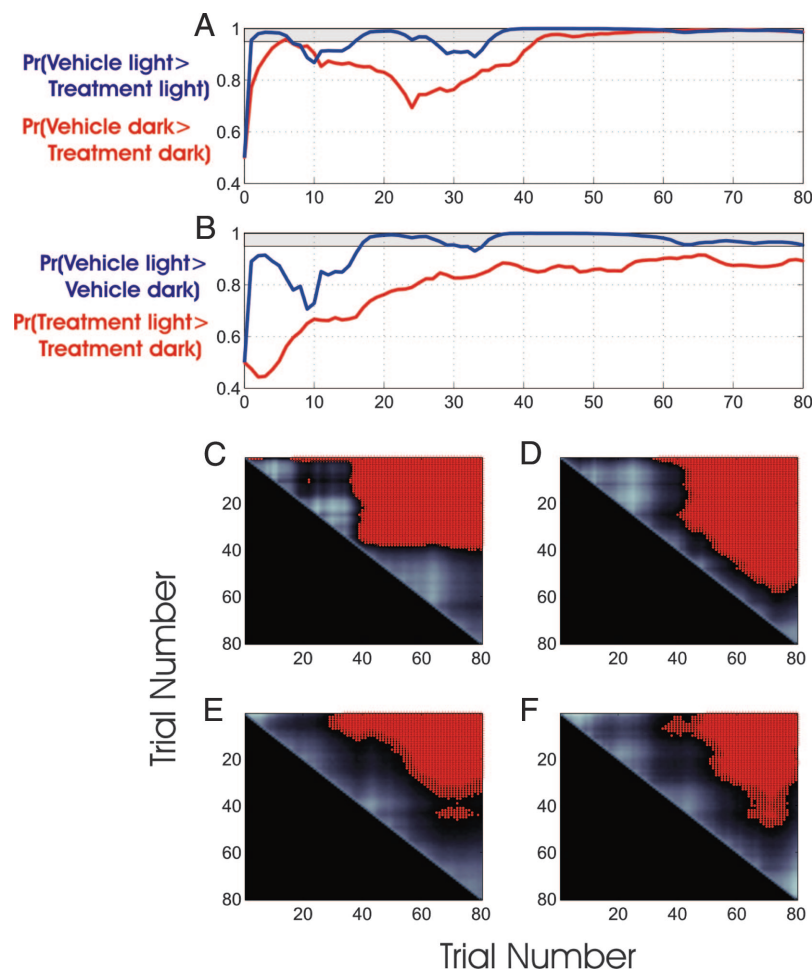
FIG. 7. Comparison between and within Vehicle and Treatment group population learning curves. *A*: blue curve (red curve) shows the trial-by-trial estimates of the probability that performance in the Vehicle light group (Vehicle dark group) is greater than the performance in the Treatment light group (Treatment dark group) computed using Monte Carlo (APPENDIX C). For the light group comparisons, this probability is $\geq 0.95$ from trials 2–6, trials 16–27, and from trial 35 to the end of the experiment (gray-shaded area). For the dark group comparisons, this probability is $\geq 0.95$ from trials 6–7 and from trial 42 to the end of the experiment. *B*: blue curve (red curve) shows the trial-by-trial estimates of the probability that performance in the Vehicle light group (Treatment light group) is greater than the performance in the Vehicle dark group (Treatment dark group). For the Vehicle group comparisons, this probability is $\geq 0.95$ from trials 17–32 and from trial 35 to the end of the experiment. For the Treatment group comparisons, this probability is never $>0.90$. *C–F*: trial-by-trial within-group comparisons for the Vehicle light (*C*), Vehicle dark (*D*), Treatment light (*E*), and Treatment dark (*F*) subgroups, computed as the probability that performance on trial k (*x*-axis) is greater than performance on trial j (*y*-axis) using Monte Carlo (APPENDIX D). Color of the probability surface changes from lighter to darker as the probability values increase to 1. Red areas denote trial combinations for which the probability is $\geq 0.95$ that performance on trial $k$ (*x*-axis) is greater than performance on trial $j$ (*y*-axis).

subgroup comparisons (probabilities) for the Vehicle light subgroup (Fig. 7*C*), Vehicle dark subgroup (Fig. 7*D*), Treatment light subgroup (Fig. 7*E*), and Treatment dark subgroup (Fig. 7*F*). Comparisons on which $\Pr(p_{k_2} > p_{k_1})$ was $\geq 0.95$ are shown in red. The algorithm in APPENDIX D shows that the computations involve evaluating comparisons between pairs of random variables from the *K*-dimensional joint probability density of the learning state process. This probability density was estimated by our model-fitting analysis in APPENDIX A. For this reason, there is no problem with multiple hypothesis tests in this analysis.

For the Vehicle light subgroup (Fig. 7*C*), the learning curve at trial 40 onward was significantly greater than the learning curve from trials 1 to 37. The steplike structure in the probability surface resulted from the steplike increase in the learning curve around trial 40 (Figs. 7*C* and 5*A*). Because of the large increase in the learning curve at the start of the experiment, where the probability of a correct response is 0.5, there is a line of red along the top of the probability surface, indicating the animals' performances were significantly above chance early in the experiment. The Vehicle dark subgroup (Fig. 7*D*) also showed a significant increase around trial 40 but in this case, the improvement continued through the length of the experiment. For this subgroup, the learning curve for any trial greater than trial 40 was consistently larger than the learning curve 10 trials earlier or more.

Beginning at trial 30 for the Treatment light group (Fig. 7*E*) performance was better on this trial than that on all trials 20 trials earlier or more. This level of difference in between-trial performance was maintained for the balance of the experiment. A similar pattern held for the Treatment dark group (Fig. 7*F*). These analyses show that within each of the 4 subgroups there is substantial improvement in performance consistent with learning within each group.

*Optimal design of a learning experiment*

Two important questions that arise in the design of behavioral experiments that compare population learning are how many animals per group and how many trials per experiment are required to detect accurately between-group differences in learning. To study these question, we used our SSRE model to conduct a theoretical study of how well we can distinguish differences in learning between 2 populations as a function of the true differences in their learning propensity, *J* the number of animals per group and *K* the number trials in the experiment. We assumed that learning in both groups (denoted by Control group and Treatment group) was dependent on the same unobservable learning process defined at trial *k* by the logistic equation

$$x_k = 3.5[1 + \exp(-0.05(k - 80)]^{-1} + \eta_k \qquad (3.3)$$

where $\eta_k$ is a zero mean Gaussian random variable with

variance $\sigma_\eta^2 = 0.04$ for $k = 1, \ldots K$. In the analyses we compared a Control group with 3 different Treatment groups. For each group we assumed that, given its learning modulation parameter $\beta_0$, its population probability of a correct response was given by evaluating the expected value of the state model (*Eq. 3.3*) in *Eq. 2.2*. We assumed that the Control and Treatment groups differ only in their learning modulation parameters. This analysis simulated the situation in which the ability to learn was modulated by treatment or previous experience.

We assumed that each group consists of $J$ individuals and that $\beta_j$, the learning parameter for individual $j$, is drawn from a Gaussian probability distribution with mean $\beta_0$ and variance $\sigma_\beta^2$ for $j = 1, \ldots, J$. Therefore, for each individual in each group, we assumed that given its learning modulation parameter $\beta_j$, the individual's probability of a correct response was given by evaluating the state model in *Eq. 3.3* using *Eq. 2.2*. For the Control group, we set $\beta_0 = 2.6$ and $\sigma_\beta^2 = 0.04$. To simulate a treatment effect that induced impaired learning propensity, we chose $\sigma_\beta^2 = 0.04$ and 3 different values of $\beta_0$: 1.8, 1.4, and 1. That is, the differences between the population learning parameters of the Control and Treatment groups in these cases were given by $\delta$, where $\delta = -0.8, -1.2,$ and $-1.6$.

The resulting population learning curves are shown in Fig. 8*A*. We chose this model because the resulting Control and Treatment group learning curves resembled, respectively, smoothed versions of the Vehicle and MK801 Treatment group learning curves that we estimated in our real data example. In addition, the parameter values are similar to those estimated from the analysis of the true data. As we did for the analysis shown in Fig. 6*A,* we computed by Monte Carlo the probability that the population learning curve for the Control group differed from the population learning curve of the Treatment group for each of the 3 values of the Treatment group parameters, assuming that there was a sample of 10,000 individuals per group (Fig. 8*B*) and 120 trials in the experiment. The differences between the learning curves of these groups are the between-group difference curves we would like to detect in our SSRE model analysis.

As the differences between the Control and the Treatment group population learning curves increased (Fig. 8*A*), the trial on which we are able to identify, with a certainty of at least 0.95, that the groups were different moved earlier in the experiment (Fig. 8*B*). Thus, as $\delta$ decreased from $-0.8$ to $-1.2$ to $-1.6$, the maximum difference between the Control and Treatment learning curves increased from 0.07, to 0.13, to 0.20 (Fig. 8*B*) and the earliest detectable learning trial decreased from trial 58, to 43, to 34 (Fig. 8*C*). This feature of the simulation was important because it indicated that a longer experiment might be needed to detect smaller differences between the learning curves.

For each of the 3 differences in population learning curves between the Control and Treatment groups, we tested 6 different numbers of subjects per group $J = 3, 5, 7, 11, 15,$ and 20, and 7 different numbers of trials per experiment $K = 40, 50, 60, 70, 80, 100, 120,$ and 140. This represents a reasonable range of number of subjects and number of trials per experiment that might be used in a population learning study. For each of the $3 \times 6 \times 8 = 144$ triplets of parameter values, we simulated a learning curve for each of the $J$ subjects in each group and from each subject's learning curve we simulated
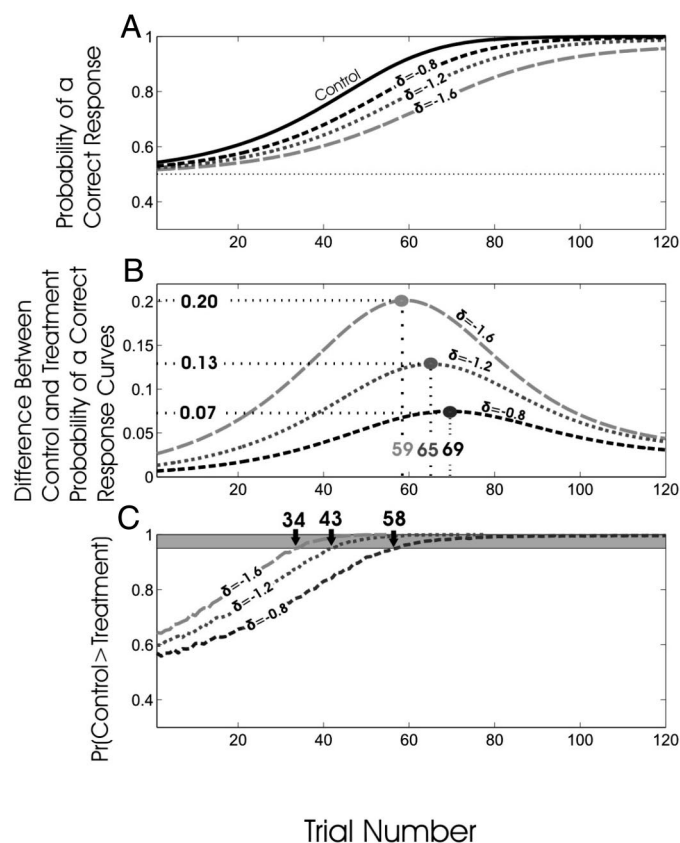


**Trial Number**

FIG. 8. *A*: population learning curves used to generate binary response data for the experimental design study. Mean learning curve for Control (solid line) remains fixed and the learning curve for Treatment group (dashed lines) takes 3 values representing 3 different group separations. Control and Treatment group learning curves are generated from the same noisy latent process (*Eq. 3.3*) with different modulation parameters. Change in modulation parameter $\delta$ required to generate the Treatment group learning curve from the Control learning curve is shown on each Treatment group curve. *B*: distance between the mean of the Control and the 3 Treatment group learning curves. Positions of the maxima are shown on each curve along with their *x, y* values. *C*: dashed lines represent the probability that the Control group learning curve is larger than the Treatment group learning curve for the 3 Treatment group learning curves in *A,* computed using Monte Carlo (APPENDIX C). Earliest trial, the difference between groups that can be detected with probability $\geq 0.95$, is marked with an arrow above the panel.

experimental data that constituted a sequence of correct and incorrect responses of length $K$.

We used our SSRE model to estimate from the sample of simulated binary response data the population and individual learning curves in each group. As we did for the analysis shown in Fig. 7*A*, we computed for each of the 144 triplets of parameters a trial-by-trial estimate of whether the population learning curve of the Control group was greater than the population learning curve for the Treatment group. We performed 100 simulation experiments for each of the 144 pairs of population curves. For each of the 144 pairs of population curves, we computed the earliest trial on which the between-group difference could be detected with a probability of at least 0.95. We reported the earliest detectable trial (detection trial) from the average of the 100 simulation experiments for each of the 144 pairs (Fig. 9) for comparison with the theoretical earliest trials of 58, 43, and 34 for $\delta$ values of $-0.8, -1.2,$ and $-1.6$ (Fig. 8*B*), respectively, computed from 10,000 Monte Carlo individuals per trial.
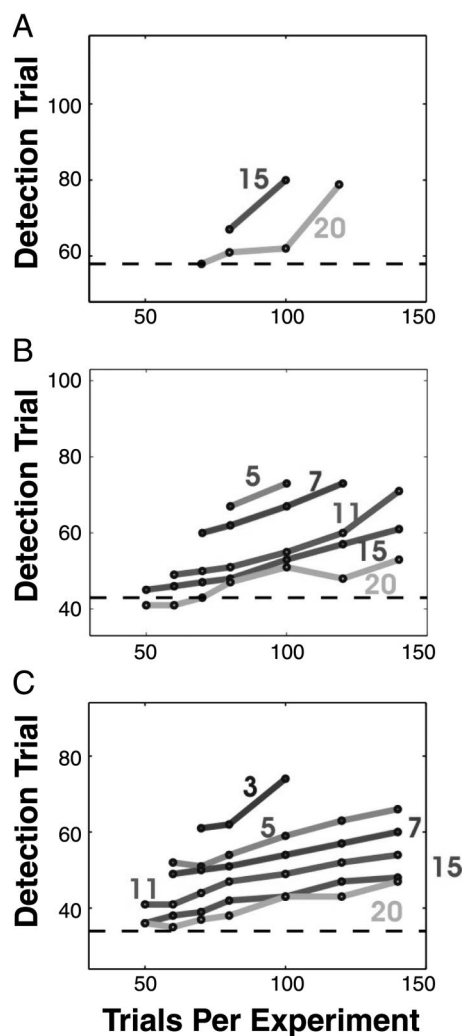
FIG. 9. Results of theoretical study to estimate how efficiently simulated Control and Treatment groups can be distinguished for different group sizes and trials per experiment. Three values of theoretical group separation were considered (*A*, $\delta = -0.8$; *B*, $\delta = -1.2$; *C*, $\delta = -1.6$). Each panel shows the estimate from our SSRE simulation study of the earliest detection trial against the number of trials per experiment for different group sizes. Group size J is indicated next to each line. Horizontal dashed lines indicate the theoretical detection trial for each of the 3 separations A, 58; B, 43; C, 34. Each point on the graph is estimated from the average of 100 simulations. For many parameter combinations, a between group difference could not be detected with certainty $\geq 0.95$. For example, *A* has data only for group sizes of $J = 15$ and 20 because for smaller group sizes the groups were not distinguishable.

The smallest difference of $\delta = -0.8$ between the population modulation parameters of the Control and Treatment groups corresponds to a maximum difference of only 0.07 between their respective population learning curves (Fig. 8*B*). Because this treatment effect is so small, it could be detected with a level of certainty of only $\geq 0.95$ (Fig. 9*A*) if there were either 15 or 20 animals per group. For the study with 15 animals per group, experiments with 80 trials could detect the difference between the Control and the Treatment group learning curves. That difference was detected at trial 70 compared with the theoretically earliest detectable trial of trial 58, shown as the horizontal dashed line (Fig. 9*A*). An experiment with 15 animals per group and 100 trials did not detect this difference until trial 79. Therefore, for 15 animals per group, an experi-

ment with 80 trials was best. For the studies with 20 animals per group, an experiment with 70 trials could detect this small difference between the Control and Treatment learning curves at trial 58 (Fig. 9*A*). An experiment with 20 animals per group and 80 trials or more would detect the difference between the Control and Treatment learning curves at a later trial. Hence, for a study with 20 animals per group an experiment with 70 trials was best. Because it is most likely to be less costly to perform 10 more trials per animal on 15 animals than to prepare 5 additional animals for a learning study, these results suggest that, for this difference between the Control and Treatment group population learning curves, 15 animals per group with 80 trials per experiment would be the best design.

The difference of $\delta = -0.12$ between the population modulation parameters of the Control and Treatment group corresponds to a maximum difference of 0.13 between their respective population learning curves (Fig. 8*B*). It was possible to detect this difference with a level of certainty of $\geq 0.95$ with all of the experiments except those with 3 animals per group (Fig. 9*B*). For this difference we clearly saw a pattern in the analysis that was hinted at with the difference of $\delta = -0.8$. It is that, beyond a certain point, for any number of animals per group, the number of trials per experiment required to achieve the detection trial increased. This is because to study realistic structure in learning experiments, the true population learning curves approach each other (Fig. 8*A*)—that is, the animals in all the groups learn. As a consequence, beyond a certain point, more trials per experiment makes distinguishing the between-group differences more challenging.

For a given number of animals per group, the earliest detectable trial occurred at the smallest number of trials per experiment for which the learning curve could be estimated for that number of animals. For example, with 7 animals per group, the smallest number of trials per group for which the population learning curve could be estimated was 70 trials and the detection trial was achieved at trial 62 (Fig. 9*B*). Similarly, for 11 animals per group the smallest number of trials per group for which the population learning curve could be estimated was 60 and the detection trial was achieved at trial 48 (Fig. 9*B*). For the difference of $\delta = -1.2$ between the Control and Treatment Group learning curves the theoretical detection trial was trial 43 (Figs. 8*B* and 9*B*, horizontal dashed line). The experiments with 15 (Fig. 9*B*) and 20 (Fig. 9*B*) animals per group achieved this detection trial with 50 trials per experiment. Although for this difference between the learning curves of the Control and Treatment group the theoretical detection trial can be achieved with 15 animals per group, this analysis shows that a study with 7 animals per group and 70 trials per experiment can detect this difference at trial 62 and thus, may be a more cost-effective design.

The difference of $\delta = -0.16$ between the population modulation parameters of the Control and Treatment groups corresponds to a maximum difference of 0.20 between their respective population learning curves (Fig. 8*B*). For this difference, the detectable trial could be identified for at least 3 choices of number of trials per experiment for any number of animals per group (Fig. 9*C*). The results for this group of simulations resembled closely those from the $\delta = -0.12$ analysis (Fig. 9*B*). For any number of animals per group,

beyond a certain point, increasing the number of trials resulted in the detection trial being identified later in the experiment. Again, the detection trial moved to a later trial in the experiment because the difference between the Control and Treatment group learning curves decreased at later trials. For a given number of trials per experiment, increasing the number of animals per group decreased the trial number on which the detection trial was identified. The theoretical detection trial for this set of experiments was trial 34 (Fig. 9*C,* horizontal dashed line). Both the studies with 15 and 20 animals per group nearly achieved the theoretical detection trial with either 50 or 60 trials per experiment. An experiment with 60 trials and 5 animals per group (Fig. 9*C*) had a detection trial of 51, whereas an experiment with 60 trials and 7 animals per group (Fig. 9*C*) had a detection trial of trial 49. Although it is possible to design for this between-group comparison an experiment with either 15 or 20 animals per group and 50 trials per experiment that can identify the theoretical lower limit of the detection trial, experiments with either 5 or 7 animals per group can reliably distinguish this difference with 60 trials per experiment. Given the trade-off between conducting a longer experiment and training more than double the number of animals to execute a task with almost as many trials, the design with 5 or 7 animals per group and 60 trials per experiment offers an efficient solution for detecting this difference between the Control and Treatment groups.

### DISCUSSION

In a population learning study, the between-subject differences in responses are an important source of variance that must be characterized to identify accurately the features of the learning process common to the population. To address this problem, we have developed an SSRE model of learning from which we defined for both the population and each subject studied on an experimental protocol, the learning curve, the ideal observer curve and the IO(0.95) learning trial. We presented dynamic comparisons of learning both between and within subgroups. When used to analyze actual learning experiments, the SSRE model gave a more informative characterization of population performance, between-subject variation in performance, and both population and individual subject learning trials than current non-model-based methods. Furthermore, we illustrated how our analysis paradigm may be used theoretically to assess the design efficiency of a learning experiment and to plan these studies prospectively.

### *A state-space random effects model of population and individual learning*

To address the several conceptual and technical challenges for characterizing between-subject variation in population learning studies, we performed 3 different state-space analyses of the set-shift experiment. First, in the pooled population analysis we treated all the response data within the Vehicle group or the Treatment group as identical. This approach estimated the population learning curves for the Vehicle and Treatment groups with high precision (i.e., with narrow confidence intervals), but at the expense of assuming that that there was no between-subject variation. The assumption of no between-subject variation was also made by the non-model-based

8TB and 8CCR methods. In a second state-space analysis, we estimated a separate learning curve for each individual in the study. Although this approach illustrated the maximum possible between-subject variation, it lost precision because each individual learning curve was estimated from only the data for that subject and no population learning curve was estimated.

In our third analysis, we used the SSRE model to estimate simultaneously population and individual learning curves within subgroups of both the Vehicle and Treatment groups. The SSRE model is an extension of the state-space model for dynamically analyzing learning in an individual subject (Smith et al. 2004). In the SSRE model each subject had a subject-specific learning modulation parameter (Jones 1993) distributed as a Gaussian random variable about the mean population learning parameter. This combined state-space and random effects structure in our model made it possible to estimate simultaneously population and individual learning curves by using the augmented state-space model in *Eq. A3* (Eden et al. 2004; Jones 1993) to extend the EM algorithm in Smith et al. (2004). Because each individual SSRE learning curve is estimated using the responses from all the subjects in the subgroup, each individual SSRE learning curve is a compromise between the population learning curve for that subgroup and the corresponding SS individual learning curve (Fig. 6). For this reason, the individual SSRE learning curves also have greater precision than that of the individual SS learning curves.

Exchangeability is an important concept that underlies all SSRE analyses. In the set-shift experiment, whether the Set 1 training dimension was light or dark affected appreciably the Set 2 response pattern in both the Vehicle and Treatment groups (Fig. 1). Therefore, we used the Set 1 training dimension to define the subgroups for the SSRE analysis and we performed the SSRE analysis within the light and dark subgroups in both the Vehicle and Treatment groups. As is standard for use of random effects models, we recommend using block covariates, such as the Set 1 training dimension, individual and pooled population learning analyses with the state-space models, to help identify the largest possible subgroups within which the SSRE analysis can be applied.

An alternative approach to estimating the between-subgroup differences would have treated the Vehicle (Treatment) cohort as one group and estimate a fixed effect, i.e., a specific coefficient to distinguish the light and dark subgroups within the Vehicle (Treatment) cohort (Fahrmeir and Tutz 2001; Jones 1993). We found that the structure in the between-group differences was not accurately described by this mixed model.

### *Non-model-based analyses of population learning*

The non-model-based 8TB and 8CCR methods used by Stefani et al. (2003) to analyze these data had several shortcomings. In particular, the 8TB method treated each block of trials as independent, gave only a 10-point learning curve estimate in an 80-trial experiment, provided no learning curve estimate for an individual animal, and gave error estimates based only on the responses in the blocks. Because the choice of block length is arbitrary, this method suffers from the bias variance trade-off problem; estimates in longer blocks have smaller bias and larger variance, whereas estimates in shorter blocks have smaller bias but more variance. The 8CCR method ignored between-subject variability and estimated the popula-

tion learning trial as the simple average across the individual learning trials. This learning trial estimation method was unrelated to the 8TB method. Moreover, to use a consecutive correct response method to identify the learning trial at a significance level of 0.05, the number of consecutive correct responses in an 80-trial experiment should be 10 rather than 8 (Smith et al. 2004). Finally, the analysis of Stefani and colleagues required a third, unrelated technique, the Wilcoxon sign rank test to assess trial-by-trial learning differences within subgroups.

### Analysis of population learning and optimal design of population learning experiments

Beyond estimating population and individual learning curves, our SSRE model paradigm has 4 features that make possible a comprehensive, dynamic characterization of population learning experiments. First, population ideal observer curves (Fig. 5) provided a trial-by-trial assessment of the probability that a population was performing better than chance. Although we prefer the dynamic assessment of learning given by the ideal observer curve, the IO(0.95) criterion offered more credible assessments of the learning trial than the 8CCR method. Second, our analysis allowed a direct comparison of learning between subgroups. That is, because we estimated a probability model for each subgroup, we compared the subgroups by computing trial-by-trial the probability that the performance in the Vehicle light (dark) was greater than performance in the Treatment light (dark) subgroup (APPENDIX C). Learning in each MK801 Treatment subgroup was impaired relative to the corresponding Vehicle subgroup (Fig. 7A). Moreover, learning in the Vehicle dark subgroup was significantly impaired relative to learning in the Vehicle light subgroup (Fig. 7B). Although there was a suggestion that learning in the Treatment dark subgroup was impaired relative to the Treatment light subgroup, this impairment was not as significant as that seen between the two Vehicle subgroups.

Third, our analysis allowed us to make a direct comparison of learning within subgroups. For this, we used our SSRE model to estimate the $K$-dimensional joint probability density of all the learning states within a subgroup (APPENDIX A, *Eqs. A16–A19*). We assessed learning within each subgroup by computing the probability that performance on a give trial was greater than performance on any previous trial (APPENDIX D). Together, the within- and between-subgroup analyses demonstrated a strong difference in learning between the respective Vehicle and Treatment subgroups even though learning occurred within each subgroup (Fig. 7, *C–F*). Because the between- and within-subgroup comparisons of performance are computed from the estimated $K$-dimensional joint probability densities of the state variables and involved no null hypotheses, we obviated the problem of multiple hypothesis tests. Although Stefani et al. (2003) reported similar findings, our SSRE model gave a more detailed dynamic assessment of learning that allowed us to understand in one analysis the effects on learning of both the NMDA receptor antagonist and the Set 1 response dimension (i.e, light vs. dark).

Fourth, the current analysis uses an elementary state-space model to impose the constraint that performance on adjacent trials is related and a random effects model to relate formally individual and population performance. By using a more de-

tailed state-space model, the current methods can be extended from simple data analysis tools to ones that are rudimentary models of learning (Luce et al. 1965; Suppes 1959, 1990; Usher and McClelland 2001; Kakade and Dayan 2002). For example, a more detailed form of the state-space model in *Eq. 2.4* might be

$$x_k = \mu + \rho x_{k-1} + \gamma I_k + \varepsilon_k \qquad (4.1)$$

where $\mu$ is a drift or learning rate, $\exp(-\alpha \Delta_k) = \rho$, $\alpha$ defines a forgetting time constant, $\Delta_k$ is the time between trials $(k - 1)$ and $k$, $I_k$ a possible external covariate and now the $\varepsilon_k$ are zero mean, Gaussian random variables whose variance $\sigma_k^2$ is a function of the learning state variable (Kadake and Dayan 2002). Similarly, it is possible to extend the random effects model to represent between-subject variation beyond the current formulation in terms of the modulation parameters (Fahrmeir and Tutz 2001). These extensions will be the topic of a future report.

Finally, an important feature of our SSRE paradigm is its use to design experiments and assess the efficiency of a given design. Our experimental design study showed how to predict the extent to which a control and treatment group could be distinguished as a function of the magnitude of the underlying differences in the learning modulation parameter (maximum difference in the population learning curves), the number of subjects per group and the number of trials per experiment. For the magnitude of the learning effect identified in the current study, our results showed that for a maximum difference in the probability of correct response between the Control and Treatment group learning curves of 0.20 (0.07), 5 to 7 (15 to 20) animals per group and 60 (80) trials per experiment allowed discrimination between the 2 groups that approached the theoretical limit of what would be possible with an unlimited number of subjects per group. We foresee preliminary data from learning experiments being used in design simulations to predict across a reasonable range of outcomes how many animals per group and trials per experiment will be needed to characterize learning reliably. As is standard practice in medical studies and clinical trials, we recommend use of a design analysis in the early stages of a learning experiment to increase the likelihood of accurately characterizing learning behavior and to make efficient use of valuable experimental resources.

## APPENDIX

### A. Derivation of the EM algorithm

Use of the EM algorithm to compute the maximum likelihood estimates of $\theta = (\beta_0, \sigma_\varepsilon^2, \sigma_\beta^2)$ requires us to maximize the expectation of the complete data log-likelihood. The complete data likelihood is the joint probability density of $N_{1:K}$ $x$ and $\beta$, which for our model is

$$p(N_{1:K}, x, \beta | \theta) = p(N_{1:K}|x, \beta, \theta)p(x, \beta|\theta) = \prod_{k=1}^{K} \prod_{j=1}^{J} (p_k^j)^{n_k^j}(1 - p_k)^{1 - n_k^j}$$

$$\times \prod_{k=1}^{K} (2\pi\sigma_\varepsilon^2)^{-1/2} \exp\{-(2\sigma_\varepsilon^2)^{-1}(x_k - x_{k-1})^2\}$$

$$\times \prod_{j=1}^{J} (2\pi\sigma_\varepsilon^2)^{-1/2} \exp\{-(2\sigma_\beta^2)^{-1}(\beta^j - \beta_0)^2\} \quad (A1)$$

where the first term on the right of *Eq. A1* is defined by the Bernoulli

probability mass function in *Eq. 2.1* and second term is the joint probability density of the learning state process defined by the Gaussian model in *Eq. 2.4* and the random effects model for the $\beta^j$. At iteration $(l + 1)$ 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 $\theta^{(l)}$, the parameter estimates from iteration $l$, which is defined as

E-STEP

$$Q(\theta|\theta^{(l)}) = E\left[\log p(N_{1:K}, x, \beta|\theta^{(l)})\|N_{1:K}, \theta^{(l)}, x_0\right]$$

$$= E\left[\sum_{k=1}^{K}\sum_{j=1}^{J}[n_k^j \log p_k^j + (1 - n_k^j)\log(1 - p_k^j)]\|N_{1:K}, \theta^{(l)}, x_0\right]$$

$$+ E\left[-\frac{K}{2}\log(2\pi\sigma_\varepsilon^{2(l)}) - (2\sigma_\varepsilon^{2(l)})^{-1}\sum_{k=0}^{K}(x_k - x_{k-1})^2\|N_{1:K}, \theta^{(l)}, x_0\right]$$

$$+ E\left[-\frac{J}{2}\log(2\pi\sigma_\beta^{2(l)}) - (2\sigma_\beta^{2(l)})^{-1}\sum_{j=1}^{J}(\beta^j - \beta_0)^2\|N_{1:K}, \theta^{(l)}, x_0\right]$$

$$(A2)$$

where $E[\cdot\|N_{1:K}, \theta^{(l)}, x_0]$ denotes the expectation of the indicated quantity taken with respect to the probability density of x and $\beta$ given $N_{1:K}$, $\theta^{(l)}$, and $x_0$. Upon expanding the right side of *Eq. A2* we see that calculating the expected value of the complete data log-likelihood requires computing the expected value of the state variables, which we denote as

$$x_{k|K} \equiv E[x_k\|N_{1:K}, \theta^{(l)}, x_0] \tag{A3}$$

$$R_{k|K} \equiv E[x_k^2\|N_{1:K}, \theta^{(l)}, x_0] \tag{A4}$$

$$R_{k-1,k|K} \equiv E[x_{k-1}, x_k\|N_{1:K}, \theta^{(l)}, x_0] \tag{A5}$$

for $k = 1, \ldots, K$ and

$$\beta_{K|K}^j \equiv E[\beta^j\|N_{1:K}, \theta^{(l)}, x_0] \tag{A6}$$

$$R_{K|K}^j \equiv E[(\beta^j)^2\|N_{1:K}, \theta^{(l)}, x_0] \tag{A7}$$

$j = 1, \ldots, J,$ where the notation $k|j$ denotes the state variable at $k$ given the responses up to time $j$. We construct a nonlinear recursive filtering algorithm, a fixed-interval smoothing and a covariance smoothing algorithm to evaluate these expectations as in Smith and Brown (2003) and Smith et al. (2004). To do so, we first construct the augmented state-space model (Eden et al. 2004, Jones 1993) to include the random-effects component of the model in the state equation. The augmented state-space model is

$$\beta_k^* = \beta_{k-1}^* + \varepsilon_k^* \tag{A8}$$

where $\beta_k^* = (x_k, \beta_k^1, \beta_k^2, \beta_k^3, \ldots, \beta_k^J)$ and $\varepsilon_k^* = (\varepsilon_k, 0, \ldots, 0)$. The stochastic properties of $x_k$ are defined by *Eq. 2.4*, whereas the stochastic properties of $\beta$ come from the assumption that the modulation parameters are Gaussian with mean $\beta_0$ and covariance $\sigma_\beta^2 I_{J \times J}$. Our representation of the random effects in the state-space model ensures that the stochastic properties of these parameters remain constant as the filter and smoothing algorithms evolve across the trials (Jones 1993). The algorithms are

FILTER ALGORITHM.    Given $\theta^{(l)}$ we can first compute recursively the state variable $\beta_{k|k}^*$ and its variance $W_{k|k}$. We accomplish this by using the following vector-valued nonlinear filter algorithm for the augmented state-space model in this problem (Eden et al. 2004), which gives

$$\beta_{k|k-1}^* = \beta_{k-1|k-1}^* \tag{A9}$$

$$W_{k|k-1} = W_{k-1|k-1} + W_{\beta^*} \tag{A10}$$

$$\beta_{k|k}^* = \beta_{k|k-1}^* + W_{k|k-1}F_k \tag{A11}$$

$$W_{k|k} = [W_{k|k-1}^{-1} - G_k]^{-1} \tag{A12}$$

$k = 1, \ldots, K,$ where $W_{\beta^*}$ is the $(J + 1) \times (J + 1)$ diagonal covariance matrix whose 1,1 element is $\sigma_\varepsilon^{2(l)}$ and whose remaining elements are zero and where $F_k$ is the $(J + 1) \times 1$ vector whose elements are

$$F_k = \begin{bmatrix} \sum_{j=1}^{J}\beta_k^j(n_k^j - p_k^j) \\ x_k(n_k^1 - p_k^1) \\ \vdots \\ x_k(n_k^J - p_k^J) \end{bmatrix} \tag{A13}$$

and $G_k$ is the $(J + 1) \times (J + 1)$ matrix whose elements are

$$G_{k,i,m} = \begin{cases} -\sum_{j=1}^{J}(\beta_k^j)^2 p_k^j(1 - p_k^j) & i = m = 1 \\ -x_k\beta_k^j p_k^j(1 - p_k^j) + (n_k^j - p_k^j) & i = 1, m = 2, \ldots, J + 1; \\ & i = 2, \ldots, J + 1, m = 1 \\ -x_k^2 p_k^j(1 - p_k^j) & i = m = 2, \ldots, J \\ 0 & \text{otherwise} \end{cases}$$

$$(A14)$$

The initial conditions are $\beta^{*(l)} = (x_0^{(l)}, \beta_0^{(l)})$ and

$$W_0 = \begin{bmatrix} \sigma_\varepsilon^{2(l)} & 0 \\ 0 & \sigma_\beta^{2(l)} I_{J \times J} \end{bmatrix} \tag{A15}$$

In these analyses we take $x_0^J = 0$.

FIXED-INTERVAL SMOOTHING ALGORITHM.    Given the sequence of posterior mode estimates $\beta_{k|k}^*$ (*Eq.* A9) and the variance $W_{k|k}$ (*Eq.* A12) we use the fixed-interval smoothing algorithm (Shumway and Stoffer 1982; Brown et al. 1998) to compute $\beta_{k|k}^*$ and $W_{k|k}$. This smoothing algorithm is

$$\beta_{k|K}^* = \beta_{k|k}^* + A_k(\beta_{k+1|K}^* - \beta_{k+1|k}^*) \tag{A16}$$

$$A_k = W_{k|k}W_{k+1|k}^{-1} \tag{A17}$$

$$W_{k|K} = W_{k|k} + A_k(W_{k+1|K} - W_{k+1|k})A_k' \tag{A18}$$

for $k = K - 1, \ldots, 1$ and the initial conditions $\beta_{K|K}^*$ and $W_{K|K}$.

STATE-SPACE COVARIANCE ALGORITHM.    The covariance estimate $W_{k,u|K}$ can be computed from the state-space covariance algorithm (De Jong and MacKinnon 1988; Smith and Brown 2003; Smith et al. 2004) and is given as

$$W_{k,u|K} = A_k W_{k+1,u|K} \tag{A19}$$

for $1 \leq k \leq u \leq K$. The covariance terms required for the E-step are

$$R_{k|K} = W_{k|K}^{(1,1)} + x_{k|K}^2 \tag{A20}$$

$$R_{K|K}^j = W_{K|K}^{(j+1,j+1)} + (\beta_{K|K}^j)^2 \tag{A21}$$

$$R_{k-1,k|K} = W_{k-1,k|K}^{(1,1)} + x_{k-1|K}x_{k|K} \tag{A22}$$

for $j = 2, \ldots, J$ and $k = 2, \ldots, K$, where the superscripts on the covariance matrix indicate the element in the $(J + 1) \times (J + 1)$ matrix.

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

M-STEP

$$\sigma_{\varepsilon}^{2(l+1)} = K^{-1}\left[2\left(\sum_{k=2}^{K} R_{k|K} - \sum_{k=2}^{K} R_{k-1,k|K}\right) + 2R_{1|K} - R_{K|K}\right] \qquad (A23)$$

$$\beta_0^{(l+1)} = J^{-1}\sum_{j=1}^{J} \beta_{K|K}^{j(l)} \qquad (A24)$$

$$\sigma_{\beta}^{2(l+1)} = J^{-1}\left[\sum_{j=1}^{J} R_{K|K}^{j(l)} - 2\beta_0^{(l+1)}\sum_{j=1}^{J} \beta_{K|K}^{j(l)} + J(\beta_0^{(l+1)})^2\right] \qquad (A25)$$

The algorithm iterates between the E-step (*Eq. A2*) and the M-step (*Eqs. A23* to A25), and gives the maximum likelihood estimate of $\theta$ as $\theta^{(\infty)}$, using the same convergence criteria in Smith and Brown (2003). By the invariance property of maximum likelihood estimates (Pawitan 2001), the learning curves, ideal observer curves, the learning trials, and the between- and within-trial comparisons of performance are now computed by evaluating their respective formulae using the maximum likelihood estimates of $\theta$, $x$, and $\beta$.

## B. Computing confidence intervals for individual learning curves by Monte Carlo

The learning curve for each subject is

$$p_k^j = \exp(\mu + \beta^j x_k)[1 + \exp(\mu + \beta^j x_k)]^{-1} \qquad (B1)$$

Under our state-space model assumption $\beta^j$ and $x_k$ are Gaussian random variables. By fitting the model to the data we estimated the joint distribution of $\beta_k^*$. The distribution of this random variable is defined by the smoothing algorithm at any trial $k$ as the Gaussian distribution with mean $\beta_{k|K}^*$ and covariance matrix $W_{k|K}^*$ for $k = 1, \ldots, K$. Let $\beta_{k|K}^{*(j)}$ denote the $2 \times 1$ subvector from $\beta_{k|K}^*$ and $W_{k|K}^{*(j)}$ denote the $2 \times 2$ submatrix from $W_{k|K}^*$, which defines the joint Gaussian distribution between $\beta^j$ and $x_k$. Because $\mu$ is fixed according to $p_0$, the probability distribution of $p_k^j$ can be computed from the joint distribution of $\beta^j$ and $x_k$ by Monte Carlo. Confidence limits as well as any other function of this distribution can be computed from this simulated distribution. The algorithm is as follows.

*1)* For $i = 1, \ldots, M_c$, draw $(\beta^j)^i$ and $x_k^i$ from the Gaussian distribution with mean $\beta_{k|K}^{*(j)}$ and covariance matrix $W_{k|K}^{*(j)}$.

*2)* For each draw compute $(p_k^j)^i = \exp[\mu + (\beta^j)^i x_k^i][1 + \exp(\mu + (\beta^j)^i x_k^i)]^{-1}$.

*3)* Order the estimates of $(p_k^j)^i$ from smallest to the largest and denote the ordered estimates as $(p_k^j)^{(i)}$.

*4)* For $\alpha \epsilon(0, 1)$ the level $\alpha/2$ $(1 - \alpha/2)$ lower (upper) confidence bound is $(p_k^j)^{(i')}$ $[(p_k^j)^{(i'')}]$ such that $\alpha/2 = i' M_c^{-1}$ $(1 - \alpha/2 = i'' M_c^{-1})$. Hence $(p_k^j)^{(i')}$ and $(p_k^j)^{(i'')}$ define respectively the lower and upper bound of the $100\% (1 - \alpha)$ confidence interval. For example, choosing $\alpha = 0.05$ ($\alpha = 0.10$) yields 95% (90%) confidence intervals. In all our analyses we take the number of Monte Carlo samples $M_c = 10,000$.

## C. Comparing between-group learning by Monte Carlo

Let $p_k^C$ and $p_k^T$ denote respectively the probability of a correct response at trial $k$ for the control and treatment populations. From our analysis we estimate the probability distribution of $p_k^C$ and $p_k^T$. Therefore, we can compare the learning curves by computing for each trial the probability that $p_k^C$ is greater than $p_k^T$, or vice versa and plotting the resulting probability as a function of trial number $k$. An easy way to compute this curve is by Monte Carlo using the following algorithm.

For a given trial $k$, pick $M_c$.

*1)* Set $i = 1$; $S_{Mc} = 0$.
*2)* Draw $p_k^{C,i}$ from $f_k^C(p)$ (*Eq. 2.7*).
*3)* Draw $p_k^{T,i}$ from $f_k^T(p)$ (*Eq. 2.7*).

*4)* If $p_k^{C,i} > p_k^{T,i}$ then $S_{Mc} = S_{Mc} + 1$.
*5)* $i = i + 1$.
*6)* If $i > M_c$ stop; else go to 2.
We compute $\Pr(p_k^{C,i} > p_k^{T,i}) = M_c^{-1} S_{Mc}$ for each trial $k = 1, \ldots, K$. In our analyses we chose $M_c = 10,000$.

## D. Comparing within-group learning by Monte Carlo

Because the transformation between the state variable and the probability of a correct response is monotonic, to compute the probability that the probability of a correct response at trial $k_1$ is greater than the probability of a correct response at trial $k_2$ it suffices to compute the probability that the learning state at trial $k_1$ is greater than the learning state at trial $k_2$. To find the probability that the learning state at trial $k_1$ is greater than the learning state at trial $k_2$, we present a Monte Carlo algorithm similar to those in APPENDICES B and C. Combining the fixed interval smoother (A17) and the state-space covariance algorithm (A19), we can compute the covariance between the augmented state-space trials $k_1$ and $k_2$ where $k_2 > k_1$ from

$$W_{k_1,k_2|K} = \prod_{j=k_1}^{k_2-1} A_j W_{k_2,k_2|K} \qquad (D1)$$

The algorithm is as follows.

*1)* Set $i = 1$; $S_{Mc} = 0$.
*2)* Draw $x_{k_1}^i$ and $x_{k_2}^i$ from a Gaussian distribution with mean

$$\begin{pmatrix} x_{k_1|K} \\ x_{k_2|K} \end{pmatrix}$$

and covariance matrix

$$\begin{pmatrix} W_{k_1,k_1|K}^{(1,1)} & W_{k_1,k_2|K}^{(1,1)} \\ W_{k_2,k_1|K}^{(1,1)} & W_{k_2,k_2|K}^{(1,1)} \end{pmatrix}$$

where the (1,1) superscript indicates the element (1,1) from the indicated $(J + 1) \times (J + 1)$ matrix.

*3)* If $x_{k_1}^i > x_{k_2}^i$ then $S_{Mc} = S_{Mc} + 1$.
*4)* $i = i + 1$.
*5)* If $i > M_c$ stop; else go to 2.
We compute $\Pr(x_{k_1}^i > x_{k_2}^i) = M_c^{-1} S_{Mc}$ for each trial $k = 1, \ldots, K$. In our analyses we chose $M_c = 10,000$.

REFERENCES

**Brown EN, Frank LM, Tang D, Quirk MC, and Wilson MA.** 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, 1998.

**De Jong P and Mackinnon MJ.** Covariance for smoothed estimates in state-space models. *Biometrika* 75: 601–602, 1988.

**Dempster AP, Laird NM, and Rubin DB.** Maximum likelihood from incomplete data via EM algorithm. *J Royal Stat Soc Series B-Methodol* 39: 1–38, 1977.

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

**Eden UT, Frank LM, Barbieri R, Solo V, and Brown EN.** Dynamic analyses of neural encoding by point process adaptive filtering. *Neural Comput* 16: 971–998, 2004.

**Eichenbaum H, Fagan A, and Cohen NJ.** 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, 1986.

**Fahrmeir L and Tutz D.** *Multivariate Statistical Modeling Based on Generalized Linear Models.* New York: Springer-Verlag, 2001.

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

**Gelman A, Carlin JB, Stern HS, and Rubin DB.** *Bayesian Data Analysis.* London: Chapman & Hall, 1995.

**Jonasson Z, Ballantyne JK, and Baxter MG.** Preserved anterograde and retrograde memory of rapidly acquired discriminations after neurotoxic hippocampal lesions. *Hippocampus* 14: 28–39, 2004.

**Jones RH.** *Longitudinal Data with Serial Correlation: A State-Space Approach.* New York: Chapman & Hall, 1993.

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

**Kitagawa G and Gersh W.** *Smoothness Priors Analysis of Time Series.* New York: Springer-Verlag, 1996.

**Laird NM and Ware JH.** Random-effects models for longitudinal data. *Biometrics* 38: 963–974, 1982.

**Luce RD, Bush RR, and Galanter E.** *Handbook of Mathematical Psychology.* New York: Wiley, 1965.

**Maclean CJ, Gaffan D, Baker HF, and Ridley RM.** Visual discrimination learning impairments produced by combined transactions of the anterior temporal stem, amygdala and fornix in marmoset monkeys. *Brain Res* 888: 34–50, 2001.

**Paton JJ, Belova MA, and Salzman CD.** Emotional learning in the rhesus monkey. Program No. 718.5. 2003 Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience. Online, 2003.

**Pawitan Y.** *In All Likelihood: Statistical Modeling and Inference Using Likelihood.* New York: Oxford Univ. Press, 2001.

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

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

**Roweis S and Ghahramani Z.** A unifying review of linear Gaussian models. *Neural Comput* 11: 305–345, 1999.

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

**Smith AC and Brown EN.** Estimating a state-space model from point process observations.

**Smith AC, Frank LM, Wirth S, Yanike M, Hu D, Kubota Y, Graybiel AM, Suzuki WA, and Brown EN.** Dynamic analysis of learning in behavioral experiments. *J Neurosci* 24: 447–461, 2004.

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

**Stiratelli R, Laird N, and Ware JH.** Random-effects models for serial observations with binary response. *Biometrics* 40: 961–971, 1984.

**Suppes P.** A linear model for a continuum of responses. In: *Studies in mathematical learning theory,* edited by Bush RR and Estes KK. Stanford, CA: Stanford University Press, 1959, p. 400–414.

**Suppes P.** On deriving models in the social-sciences. *Math Comput Model* 14: 21–28, 1990.

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

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

**Wirth S, Yanike M, Frank LM, Smith AC, Brown EN, and Suzuki WA.** Single neurons in the monkey hippocampus and learning of new associations. *Science* 300: 1578–1584, 2003.

**Yu A and Dayan P.** Expected and unexpected uncertainty: ACH and NE in the neocortex. *News Physiol Sci* 2002, 2003.

Journal of Neurophysiology