Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Improving accuracy and temporal resolution of learning curve estimation for within- and across-session analysis

Matthias Deliano[1], Karsten Tabelow[2], Reinhard König[1], Jörg Polzehl[2]

submitted: October 28, 2015

[1] Leibniz Institute for Neurobiology
Brenneckestr. 6
39118 Magdeburg
Germany
E-Mail: deliano@lin-magdeburg.de
rkoenig@lin-magdeburg.de

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: karsten.tabelow@wias-berlin.de
joerg.polzehl@wias-berlin.de

No. 2170
Berlin 2015

2010 *Mathematics Subject Classification.* 62P10, 62G05, 62G15.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Estimation of learning curves is ubiquitously based on proportions of correct responses within moving trial windows. In this approach, it is tacitly assumed that learning performance is constant within the moving windows, which, however, is often not the case. In the present study we demonstrate that violations of this assumption lead to systematic errors in the analysis of learning curves, and we explored the dependency of these errors on window size, different statistical models, and learning phase. To reduce these errors for single subjects as well as on the population level, we propose adequate statistical methods for the estimation of learning curves and the construction of confidence intervals, trial by trial. Applied to data from a shuttle-box avoidance experiment with Mongolian gerbils, our approach revealed performance changes occurring at multiple temporal scales within and across training sessions which were otherwise obscured in the conventional analysis. The proper assessment of the behavioral dynamics of learning at a high temporal resolution clarified and extended current descriptions of the process of avoidance learning. It further disambiguated the interpretation of neurophysiological signal changes recorded during training in relation to learning.

1 Introduction

Learning, the acquisition of knowledge through experience, manifests as behavioral changes in the course of training. Learning behavior relies on a multitude of neural and cognitive processes which act on different spatial and temporal scales (Pessoa, 2008; Balleine and O'Doherty, 2010; Rothe et al., 2009); however, many of these processes are not accessible experimentally. Therefore, any particular learning experiment is influenced by numerous uncontrolled variables. This entails a certain degree of unaccountable variability of behavior across time, for a subject as well as between subjects (Stark et al., 2008).

As a consequence, single behavioral responses of individual subjects are difficult to interpret with respect to learning (Stark et al., 2007). Correct behavioral responses are defined by the experimenter and are influenced by many aspects, conceivably even unrelated to learning. Correct responses may occur spontaneously by chance or even systematically with some spurious relations to task contingencies and pre-experience, for example due to response biases. Erroneous responses may occur, even in well-trained subjects, by chance due to attentional lapses, systematically due to fatigue, or due to a decrement of attention and motivation over time.

In order to relate behavioral performance changes to learning, statistical analyses across trials and subjects are indispensable. In conventional learning curve analyses behavioral performance is analyzed in moving trial windows. At any instant in time performance is quantified by the success probability, i.e. the probability of a correct response, estimated as the proportion of

correct trials among all n trials of a window centered over the current trial. Generally, the window is then moved along trials in fixed steps. Its size is adjusted to the particular interest of the experimenter and may range from a few trials to entire sessions. Learning is then identified from the time course of the resulting learning curve as an increase of the estimated success probability across trial steps (Suzuki and Brown, 2005).

This approach to the analysis of learning curves is ubiquitous, although it includes some serious statistical problems. First, the estimation of the success probability by the proportion of correct trials within a window relies on the assumption that this probability is constant within the window. The validity of this assumption, as well as the temporal resolution, the accuracy, and the precision of the estimate therefore depend on the chosen window size. However, the a-priori choice of an adequate window size is difficult, because it is unclear beforehand, when and how fast learning-related changes in the subject's responses occur, and how strong they will be. Furthermore, as many learning tasks extend over long periods of time, training is often carried out in several consecutive sessions with sometimes prolonged breaks between them. This introduces discontinuities in the learning process, e.g. by memory consolidation during inter-session sleep (Schicknick and Tischmeyer, 2006; Hennevin et al., 2007). By continuously moving the analysis window across trials in fixed steps, and by choosing the window size a-priori, neither discontinuities across session breaks nor the temporal scale of the learning process itself are accounted for properly in conventional learning curve analysis. To avoid these problems, learning is often analyzed only on the session level. Within-session changes of responding are thereby rather treated as a disturbance that has to be controlled experimentally than as a learning-related process of interest (McSweeney and Roll, 1993). However, Gallistel et al. (2004), for example, have demonstrated that performance changes within a session are not just a confounding factor, but are systematically related to learning.

Second, even with a proper choice of window size, learning curves display a high inter-subject variability (Bathellier et al., 2013). Statistical analysis of learning curves is therefore predominantly carried out on the population level by means of artihmetic averaging of learning curves across subjects (Brown and Heathcote, 2003). However, it is commonly neglected that the grand-mean learning curve is the mean of Bernoulli probabilities bound between 0 and 1 estimated across discrete trials, and that standard errors of the mean are not well suited to display the variability of such a quantity.

In this work, we frame the analysis of learning curves as the statistical problem of estimating the Bernoulli probability of a correct response in a trial based on the observed responses in a small (n<200) trial window centered over that trial. We apply different statistical methods designed for estimating these success probabilities from small samples, which also provide more appropriate confidence intervals than conventional analysis. These methods were evaluated in comparison to conventional analysis by applying them to simulated and experimental data derived from an exemplary learning experiment in which rodents were trained in a two-way active avoidance paradigm in a shuttle-box (Cain and LeDoux, 2003; Stark et al., 2008). This evaluation showed that moving window estimation of learning curves is prone to various random and systematic errors depending on learning phase, window size, and the statistical model employed.

To better account for *systematic* errors (bias) arising from a non-stationary success probability in the analysis window, we applied a generalized linear model (GLM) (McCullagh and Nelder,

1998), which allows the success probability to vary systematically across trials in the analysis window. Furthermore, we used a Firth's penalized likelihood approach (Firth, 1993) to estimate the parameters of the GLM. This adjusts for biases in case of small or large probabilities and guarantees that estimates will be finite, opposite to the conventional likelihood estimate. Optimal window sizes balancing variability and systematic errors were derived by minimizing cross-validated estimates of the mean squared error of prediction. To retain discontinuities in the learning process introduced by session breaks, learning curves were estimated separately for each training sessions. Finally, population learning curves and their confidence intervals were estimated with a fixed effects model, which is statistically more adequate than conventional population averaging. Taken together, these methods improved the accuracy and temporal resolution of learning curve analysis both for single subject and population analysis.

The application of this approach to experimental data revealed performance changes otherwise obscured by conventional analysis. These changes were learning-related, and covaried with other continuous behavioral and physiological signals recorded during training, namely reaction time and magnitude of prefrontal cortical potential in response to the conditioned stimulus. Our results indicate that shuttle-box avoidance learning is multi-phasic, and that it involves learning processes on various time scales.

2 Methods

2.1 Learning curve analysis

2.1.1 Estimation of individual learning curves

For each trial $i = 1, \ldots, n$ of a training session, the response of a subject in the learning process can be regarded as a Bernoulli variable $Y_i \sim B(1, p_i)$ describing a correct $(Y_i = 1)$ or an incorrect $(Y_i = 0)$ response. The probability p_i of correct responses is a measure of learning performance for each trial. The time course of this success probability p_i over trials is called a learning curve.

To estimate individual learning curves, we employed two different models, a *constant model* and a *generalized linear model.* In the constant model, the success probability is estimated by the binomial proportion of correct responses within a window of bandwidth h, i.e. a window centered at trial i preceded by up to h trials and succeeded by up to h trials. The constant model underlies the conventional moving window analysis (from now on referred to as constant model (conv)), in which the window is moved in steps of a single trial across all sessions irrespective of session breaks; hence, potential discontinuities at the session boundaries are ignored. In our approach, we accounted for discontinuities of the learning process between sessions, by performing the moving window analysis separately for each training session (termed constant model (sep) in this work), i.e. windows were *not* moved across session breaks. For equidistant trial times $t_1, \ldots t_n$ within a session, this window encompasses the subset of trials given by

$$W_i = \{j | \max(t_1, t_i - h) \leq t_j \leq \min(t_i + h, t_n)\}. \tag{2.1}$$

The size of the window, i.e. the number of trials within the window $\#W_i = \min(t_i + h, t_n) - \max(t_1, t_i - h) \leq 2 * h + 1$, depends on the bandwidth h. In case the assumption of a constant model is met, i.e. that the success probabilities in W_i are constant, the maximum likelihood success probability for trial i can be estimated by means of the binomial proportion of correct responses in W_i:

$$\hat{p}_i^{(c)} = \sum_{j \in W_i} Y_j / \#W_i. \tag{2.2}$$

Many of the commonly applied statistical inference methods approximate the binomial distribution of this estimated success probability by a Gaussian distribution. Thus, standard statistics textbooks (see, e.g., Collett (2003)) recommend the use of confidence intervals with nominal coverage probability $(1 - \alpha)$ (i.e. the probability that the confidence interval contains the true value) for the evaluation of the success probability, based on asymptotic normal approximations given as

$$CI(p_i) = \left(\hat{p}_i - \kappa \sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{\#W_i}}, \hat{p}_i + \kappa \sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{\#W_i}} \right), \tag{2.3}$$

where κ is the $(1 - \alpha/2)$-quantile of the standard Gaussian distribution, and the term $\sqrt{\hat{p}_i(1-\hat{p}_i)/\#W_i}$ is an estimate of the standard deviation of the estimated probability $\hat{p}_i$. The computation of confidence intervals based on this approximation entails, due to the discreteness and skewness of the binomial distribution, severe problems with respect to their coverage probability, even if $\#W_i$ and $\#W_i\hat{p}_i(1-\hat{p}_i)$ are moderate, see Brown et al. (2001) for a detailed discussion. Among the various alternatives given by Blyth and Still (1983) and Brown et al. (2001), we employ the Agresti-Coull interval (Agresti and Coull, 1998)

$$CI(p_i) = \left(\tilde{p}_i - \kappa \sqrt{\frac{\tilde{p}_i(1-\tilde{p}_i)}{(\#W_i + \kappa^2)}}, \tilde{p}_i + \kappa \sqrt{\frac{\tilde{p}_i(1-\tilde{p}_i)}{(\#W_i + \kappa^2)}} \right), \tag{2.4}$$

which is centered at $\tilde{p}_i = \frac{\sum_{j \in W_i} Y_j + \kappa^2/2}{\#W_i + \kappa^2}$ for single subject learning curves estimated using the constant model. The interval can be viewed as a standard interval obtained from a sample with $\kappa^2/2$ failures and successes added.

If the success probability varies significantly across the trials within window W_i, the constant model is no longer the appropriate model to use, since the resulting estimate will be biased (see Sect. 3.1.2). This bias can be severe, for example if the window W_i is not centered at trial i (border effect), i.e. when t_i is close to t_1 or t_n, or if the success probability is strongly and rapidly changing within W_i.

In situations when the success probabilities within the window are time dependent, the constant model can, for each window W_i, be replaced by a generalized linear model (McCullagh and Nelder, 1998) that allows the logit transform of the success probability to vary linearly over trials. The GLM relates, for trials j within a fixed window W_i, the success probability p_j to a linear predictor

$$\eta_j(\beta) = \beta_0 + \beta_1(t_j - t_i) \tag{2.5}$$

with local trial parameters $\beta = (\beta_0, \beta_1)$, and $p_j = h(\eta_j(\beta))$. The inverse relation $\eta_j(\beta) = h^{-1}(p_j) = g(p_j)$ is determined by the link function $g()$. Its canonical choice for the binomial family is the logit link $g(p) = \log\left(\frac{p}{(1-p)}\right)$ which provides $h(\eta) = \frac{\exp\eta}{1+\exp\eta}$.

Local parameters β are estimated from the trial outcomes within window W_i as

$$\hat{\beta} = \text{argmax}_\beta \sum_{j \in W_i} Y_j \log(h(\eta_j(\beta))) + (1 - Y_j) \log(1 - h(\eta_j(\beta))). \tag{2.6}$$

Estimates of the success probabilities are derived as

$$\hat{p}_i = h(\eta_i(\hat{\beta})) = h(\hat{\beta}_0). \tag{2.7}$$

Point-wise confidence intervals for $\hat{p}_i$ are obtained in the GLM assuming asymptotic normality of $\hat{\eta}_i$ sa

$$(h(\eta_i - \kappa s_i), h(\eta_i + \kappa s_i)) \tag{2.8}$$

where s_i denotes the standard deviation of the linear predictor $\eta_i(\hat{\beta})$ and κ is the $(1 - \alpha/2)$-quantile of the standard Gaussian distribution. Standard deviations s_i were determined from the asymptotic covariance matrix of the parameter estimate.

In the generalized linear model with logistic link (logistic regression), estimates are biased if the sample size is small or counts for a possible outcome are either very high or very low. To account for this problem we use Firth's bias adjusted estimates (Firth, 1993; Heinze and Schemper, 2002). In contrast to the standard maximum likelihood procedure used in standard GLM analyses, Firth's approach maximizes a penalized likelihood function and guarantees finite parameter estimates. All analyses were implemented in R (R Core Team, 2014).

2.1.2 Bandwidth selection of the moving window

Properties of the estimates strongly depend on the chosen bandwidth h (i.e. on the maximum window size $\#W = 2h + 1$). Large bandwidths lead to a potential bias of the estimated probabilities, while estimates employing small bandwidths suffer from high variability (see Sect. 3). As a criterion for bandwidth selection, we suggest to use the expected mean squared error *(MSEP)* of prediction given by

$$MSEP(h) = \mathbf{E}_Y(Y - \hat{p}(h))^2 \tag{2.9}$$

An estimate with minimal *MSEP* balances bias and variability of the estimates. The *MSEP* can be estimated by cross-validation as

$$M\hat{S}EP(h) = \sum_i (Y_i - \hat{p}_i^{(-i)}(h))^2 \tag{2.10}$$

where $\hat{p}_i^{(-i)}(h)$ is the estimate of p_i employing a window $W_i(h)/\{i\}$, i.e. with the observation from trial i removed.

2.1.3 Grand mean learning curves

The calculation of grand mean learning curves for Bernoulli trials was based on a GLM fixed effects model centered at time point i

$$\begin{aligned} \eta_{ji}^{(k)} =& \beta_{0,i} + \beta_{1,i}(t_j - t_i) + \beta_{0,i}^{(k)} + \beta_{1,i}^{(k)}(t_j - t_i) \\ \text{with} \sum_k \beta_{0,i}^{(k)} =& \sum_k \beta_{0,i}^{(k)} = 0 \end{aligned} \tag{2.11}$$

Estimated probabilities $\hat{p}_i^{(k)}$ for subject k at trial i can be represented by their logit transform $\hat{\eta}_{ii}^{(k)} = \log \frac{\hat{p}_i^{(k)}}{1-\hat{p}_i^{(k)}}$. In the GLM framework, standard deviations $s_i^{(k)}$ of the $\hat{\eta}_{ii}^{(k)}$ can be obtained from the asymptotic covariance matrix of the parameters $\beta_{.,i} + \beta_{.,i}^{(k)}$. In case of the constant model, the variance of $\hat{\eta}_{it}$ is given by

$$v_i^{(k)} = 1/(\#W_i \hat{p}^{(k)}(1 - \hat{p}^{(k)}))$$

The GLM assumes the logit transforms $\hat{\eta}_{ii}^{(k)}$ to be approximately Gaussian. Values for the grand mean curves are therefore computed as

$$\begin{aligned} \hat{\eta}_i^{GM} =& \left(\frac{1}{\#Rodents} \sum_{r \in Rodents} (\hat{\eta}_{ii}^{(r)}) \right) = \beta_{0,i} \\ \hat{p}_i^{GM} =& \frac{\exp \hat{\eta}_i^{GM}}{1 + \exp \hat{\eta}_i^{GM}}. \end{aligned} \tag{2.12}$$

with variance of the logit transform $\hat{\eta}_i^{GM}$ given as

$$v_i^{GM} = \frac{1}{\#Rodents} \sum_{r \in Rodents} v_i^{(r)}$$

Pointwise confidence intervals for the grand mean are obtained again according to (2.8) as

$$h(\eta_i - \kappa\sqrt{v_i}), h(\eta_i + \kappa\sqrt{v_i})]. \tag{2.13}$$

2.1.4 Analysis of continuous behavioral and physiological covariates of learning

In addition to the learning curve itself, we further analysed the temporal dynamics of reaction times for the conditioned responses as an exemplary behavioral, and amplitudes of CS-evoked prefrontal, cortical potentials as an exemplary neurophysiological covariate. Along with the estimation of the learning curve as outlined in Sect. 2.1.1, these covariates provide important additional information for the investigation of neural correlates of learning. As for the GLM in Eqn. (2.5), we used a local linear model (Fan and Gijbels, 1996) employing the estimate implemented in the package lpridge for R (Maechler, 2013). To determine significant changes, point-wise confidence intervals were computed employing a Gaussian assumption for the estimates. Grand means were calculated by population averaging with confidence intervals derived from standard errors under Gaussian assumptions.

2.2 Experimental data

2.2.1 Animal training and acquisition of behavioral data

We used data from a two-way active avoidance paradigm in which 20 Mongolian gerbils (Meriones unguiculatus) were trained to detect frequency-modulated tones (FM sweeps). The training took place in a shuttle-box consisting of two compartments separated by a small hurdle (Figure 1A). It was carried out in three consecutive sessions with maximum two sessions per day separated by a pause of at least two hours. The remaining session(s) was performed on the following day. Each session consisted of 60 trials and lasted about 30 minutes. Within each session, and after an initial three-minute phase of habituation, a new trial was presented every 30 seconds.

The structure of a trial is shown in Figure 1B. It started with the onset of the conditioned stimulus (CS) which consisted of a train of linear FM sweeps presented at an intensity of 60 dB sound-pressure level. Within a train, 0.25-s long FM sweeps were repeated with a stimulus-onset interval of 0.5 s for a maximum of 15 s. In each trial, either a train of rising (1–2 kHz) or falling (2–1 kHz) FM sweeps was presented. In order to avoid the unconditioned stimulus (US)—a mild electric footshock elicited by 600-μA current pulses delivered through the grid-floor of the shuttle-box—animals had to learn to change the compartment in response to the CS by crossing the hurdle within 5 s after CS-onset. If the animal did not change the compartment, the footshock was applied 5 s after CS-onset for a maximum duration of 10 s. A conditioned avoidance response (CR) was defined as a compartment change within 5 s after CS-onset, before US-onset. CS was immediately turned off after a CR of the animal. An unconditioned response (UCR) consisted of a compartment change in response to the footshock, i.e. an escape response after US-onset. Both CS and US were immediately switched off after an escape response. A response was counted as correct, if the animal changed the compartment within the 5-s interval after CS-onset thus avoiding the footshock.

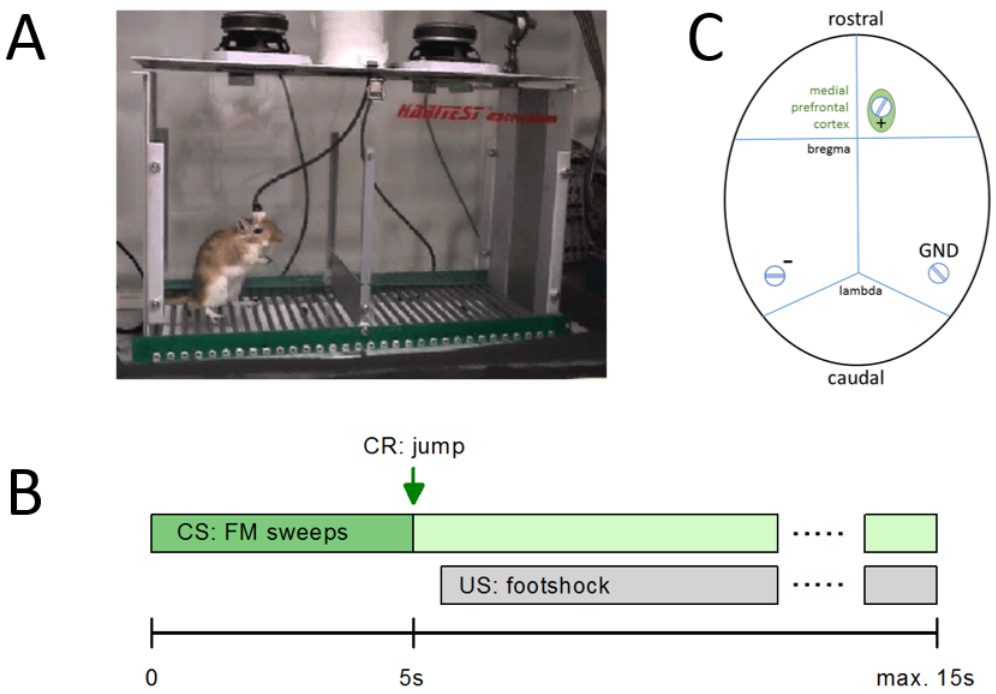

Figure 1 – *Experimental setup. (A) Snapshot of a gerbil in the shuttle-box. (B) Trial structure (time course) of the experiment. (C) Sketch of the setup for the recording of the electrocorticogram. The recording electrode (stainlees-steel screw) was located in the rostral part of the brain, above the medial prefrontal cortex.*

2.2.2 Electrocorticogram recording

Electrocorticograms (ECoGs) were recorded epidurally from the surface of the brain via a stainless-steel screw (1.5 mm diameter, Fig. 1C) implanted through a hole above the medial prefrontal cortex (mPFC), and connected to an amplifier via a cable. ECoG signals were recorded relative to two further skull screws which were fixed to the parietal bones and served as reference/ground electrodes. ECoG signals were amplified 10.000 times, low-pass filtered at 100 Hz, and digitized at a sampling rate of 500 Hz. Note that the timing of the US-onset is not very precise; it can occur up to 0.5 s before the actual onset.

3 Results

3.1 Estimation of individual learning curves and their statistical properties

3.1.1 Experimental data

As we have argued in the introduction, the quality of learning curve estimation relies on the chosen window size, the employed statistical model, the learning phase, and on discontinuities arising from breaks between training sessions. To systematically explore the influence of these factors, learning curves were estimated from exerimental data obtained from Mongolian gerbils trained with a two-way active avoidance paradigm in the shuttle-box (Figure 1). A correct avoidance response consisted of a compartment change (jump) within 5 s after CS-onset. Figure 2 shows correct ($Y = 1$) and incorrect ($Y = 0$) responses as a function of trial, exemplarily for

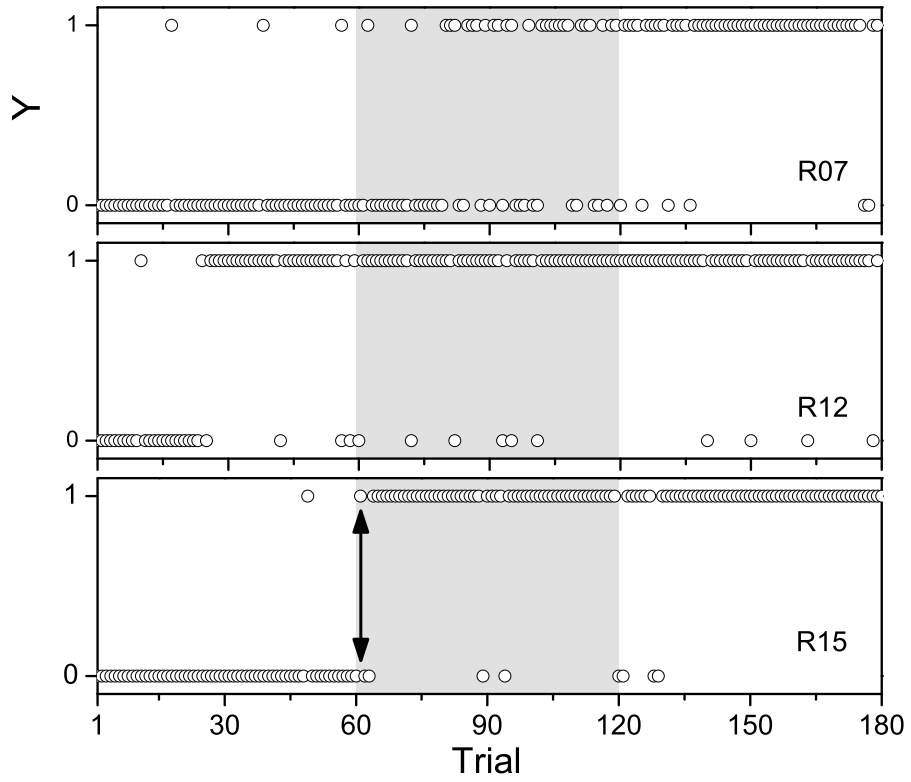

Figure 2 – *Time courses of correct ($Y = 1$) and incorrect ($Y = 0$) responses of three exemplary data sets recorded from gerbils R07, R12, and R15. In total 180 trials were presented in three different sessions of 60 trials each, with a break of at least two hours between two consecutive sessions. Black arrows indicate the approximate transitions from the naive state to the state where the task had been learned by the animal. The three experimental sessions are separated by different background colors.*

three rodents (R07, R12, R15). The figure emphasizes the individuality of the learning process for different subjects. For each animal, the learning transition occurs at a distinctive point in time,

within the second session (R07), within the first session (R12), or at the border between first and second session (R15).

In Figure 3 exemplary learning curves (black traces), estimated from the correct and incorrect responses (indicated by red symbols in upper left panel of Figure 3), are shown for a single rodent for four different window sizes (11, 19, 31, 51) and for two different statistical models (constant model and GLM, see Sect. 2.1.1).

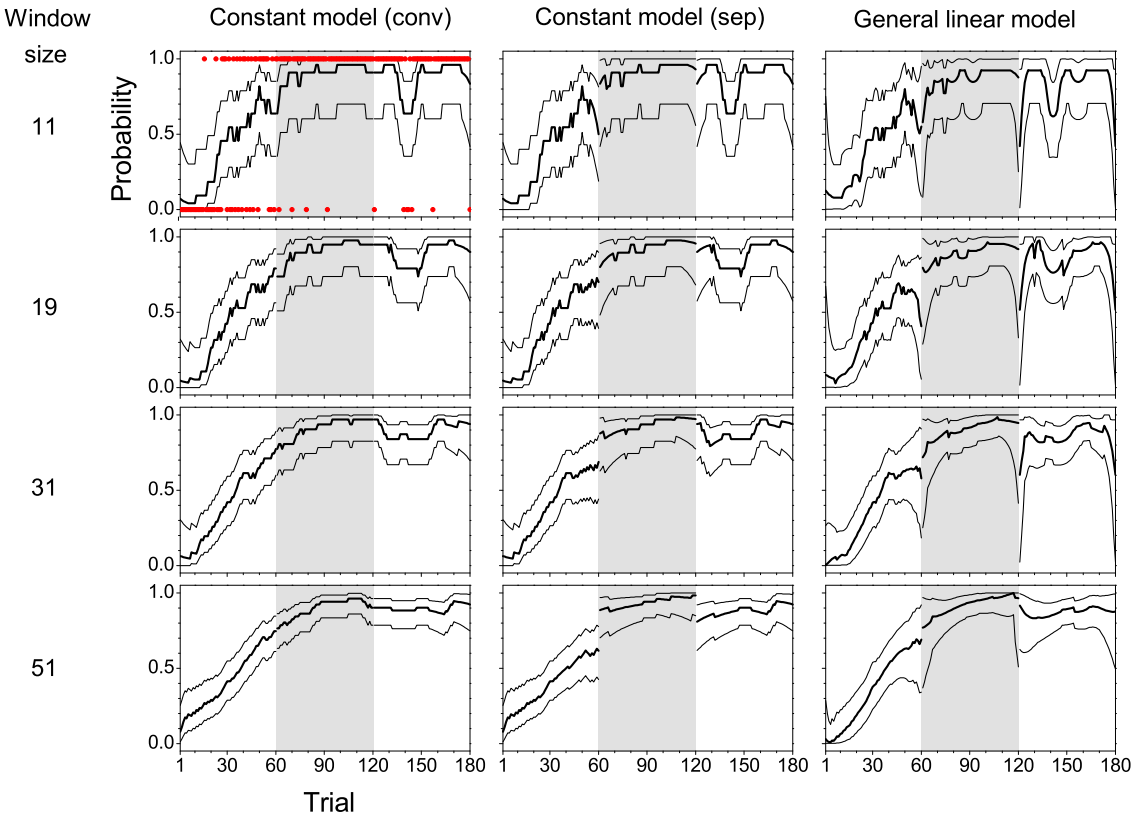

Figure 3 – *Exemplary learning curves (success probability, thick black lines) and confidence intervals (thin black lines) from a single animal estimated with three different approaches: constant model with moving window analysis across sessions ((conv), left column), constant model with moving window analysis within sessions ((sep), middle column), and general linear model with moving window analysis within sessions (right column). Results for four different window sizes are shown: 11, 19, 31, and 51. The underlying correct (1) and incorrect (0) single trial responses are displayed by red symbols in the top left panel.—The three experimental sessions are separated by different background colors.*

Conventional moving window analysis is displayed in the panels of the left column of Figure 3 (constant model (conv)), in which the window is moved trial by trial across all sessions irrespective of session breaks. The panels in the middle column of this figure show results from a moving window analysis likewise based on the constant model but now applied separately to each session (constant model (sep)). Here, approximate pointwise confidence intervals were computed according to the Agresti-Coull interval (Eqn. (2.4)), which avoids the problems of asymptotic normal approximation of binomial proportion intervals. The panels on the right col-

umn of Figure 3 show the sessionwise learning-curve estimates obtained with the GLM using Firth's bias-adjusted estimates.

The results of the three approaches have in common that the learning curves were getting smoother with increasing window size, and that the larger the window size, the smaller the width of the confidence intervals which reflects the decreasing variability of the estimate with increasing sample size within the window. Moreover, in all three approaches the slope of the learning curves decreased with increasing window size. However, compared to the learning curves estimated by means of constant model (conv) and constant model (sep) shown in the left and middle column of Figure 3, respectively, the slopes obtained with the GLM (right column of Figure 3) remained steeper at all window sizes. To obtain a more quantitative assertion of this observation, slopes of single-subject learning curves were calculated within sessions as differences in performance between consecutive trials, separately for the four window sizes, and for constant model (sep) and GLM. Maximum slopes were averaged across subjects. Although these average maximum slopes decreased with increasing window size for the two models, they always remained nearly twice as large with the GLM compared to the constant model (sep). For example, for the GLM, the average maximum slope of 0.172 ± 0.005 per trial was obtained for window size 11, and of 0.044 ± 0.004 per trial for window size 51. For the constant model (sep), the average maximum slope was only 0.089 ± 0.001 per trial for window size 11 and 0.021 ± 0.001 per trial for window size 51. The flatter slopes obtained with the constant model (sep) for large window sizes might be due to systematic errors that are expected to become substantial at large window sizes, because the assumption of a constant success probability within a window becomes then more and more violated. This bias is particularly prominent at points near session borders (asymmetry of windows) and in case of asymmetric changes of the success probability in the analysis window. Effects of this systematic error should, at least to a certain extent, be less prominent using the GLM which takes into account the variation of the success probability over time. Steeper slopes can be therefore interpreted such that the estimation of learning curves using the GLM is less biased than using the constant model. However, for a given window size, this reduction in error comes at the cost of loss in precision due to the larger flexibility of the GLM as indicated by the broader confidence intervals in the right column of Figure 3.

Whereas the moving window analysis across session borders yielded smooth, continuous learning curves (constant model (conv), left column of Figure 3), step-like increases and decreases (discontinuities) between sessions appeared, when moving window analysis was carried out separately for each session (constant model (sep), middle column of Figure 3). In contrast to constant model (sep), the session-wise GLM revealed fast changes of success probabilities at the session borders within less than 10 trials, for example steep decreases in performance at the end of sessions 1, and steep increases at the beginning of sessions 2 and 3 (right column of Figure 3, particularly for small window sizes). These changes could emphasize or de-emphasize the discontinuities in learning performance between the last trial of a session and the first trial of the following session. Both, fast performance changes at the session borders and abrupt performance changes across session breaks affect learning curve analysis across sessions.

3.1.2 Simulations

Learning curve estimation suffers from random and systematic errors. Random errors can be assessed from the experimental data by means of variance and confidence intervals. However, the evaluation of the systematic errors requires the knowledge of the true success probabilities underlying the observed responses. Systematic errors can be assessed by estimation bias, which is the difference between the expected estimated value and the true value. They can also be revealed by the coverage probability, which is the probability that the confidence interval contains the true value. If it is significantly below its nominal 95%-value, confidence intervals are systematically undersized or misplaced.

To assess systematic errors, we carried out a simulation study by generating simulated responses from known true success probabilities. In order to generate simulated data which closely resemble the measured data of the shuttle-box two-way active avoidance paradigm, we estimated a population learning curve using a fixed effects model as described in the methods section. The success probabilities of the population learning curve was then used to randomly generate 1000 different trial sequences of 180 correct or incorrect responses. For each of the simulated trial sequences, we estimated a learning curve using the constant model (Eqn. (2.2)) and the GLM (Eqn. (2.7)), for the same four window sizes (11, 19, 31, 51) as in Figure 3. Estimates were calculated separately for the three sessions (trials $1-60$, $61-120$, and $121-180$). We then assessed the effect of window size and model selection (constant model (sep) versus GLM) on different measures of random and systematic estimation errors.

Results of the simulation are summarized in Figure 4. In all panels of this figure, black curves result from the constant model (sep) and red curves from the GLM. The panels of the left column show the trial dependence of the bias, i.e. the differences between the grand mean across the 1000 simulated trial sequences of estimated success probabilities and the true population learning curve used to simulate the trial sequences. Positive bias indicates overestimation and negative bias underestimation of success probability. Generally, magnitude of bias increased with window size. Particularly large positive and negative bias even for small windows was found in the first five trials of sessions two and three. In these phases, a fast increase was seen in many individual learning curves. Larger negative bias was also observed in the last third of session 1 (trials 45 to 55) during a phase when most learning curves displayed a first increase. Lowest bias was observed in the last two thirds of session three, when most learning curves reached a stable upper plateau. Bias observed with GLM was generally lower than with the constant model (sep).

The second column of panels in Figure 4 shows, for each trial, the standard deviation (SD) of the estimates as measure of random errors. As expected, we observed an overall decrease of SD with increasing window size. The SDs of the two models did not differ much from each other within the sessions, but strongly increased towards the session borders where the effective window size decreased. This increase was much more pronounced for the SD of the GLM.

The mean absolute error, $MAE(i) = \sum_{s=1}^{1000} \mathrm{abs}(\hat{p}_i^s - p_i)$, for trial i is displayed in the third column of panels in Figure 4. *MAE* quantifies the combined effects of bias and variance and is a robust measure of the quality of the estimate. It shows a dominant effect of variability for smaller window sizes and a decrease of overall estimation quality especially at the session borders.

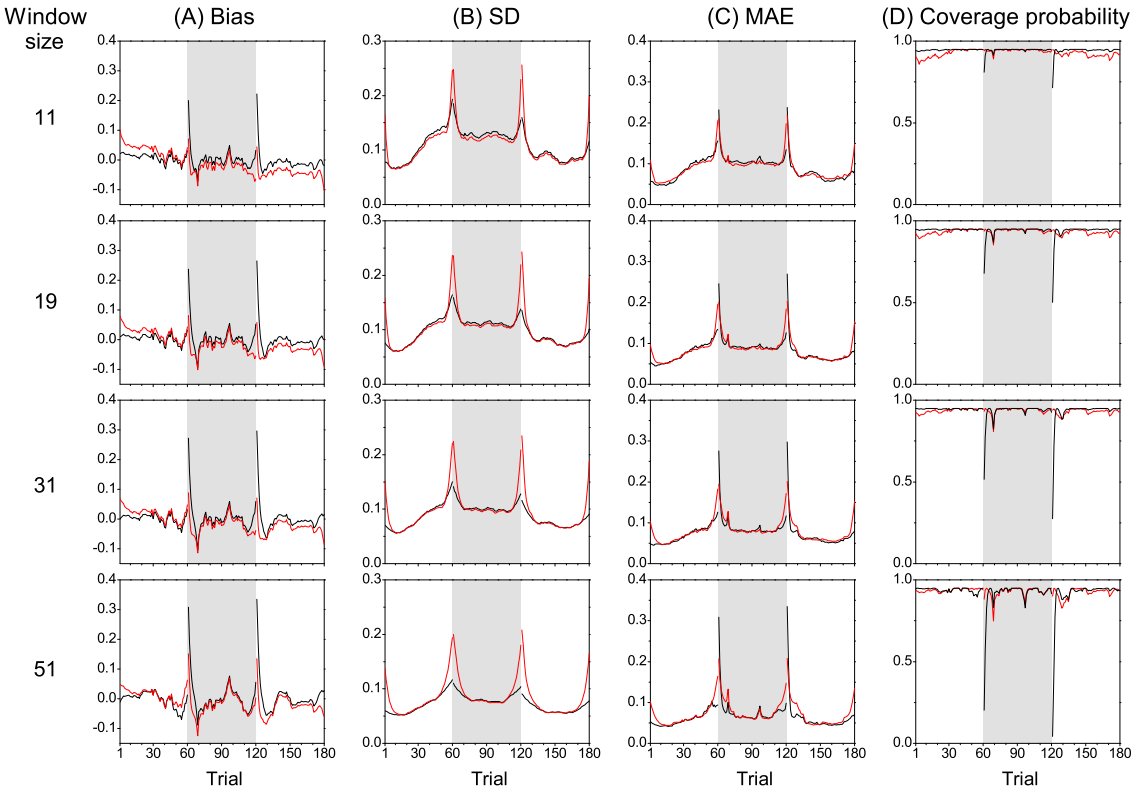

Figure 4 – *Simulation results derived from 1000 response sequences (180 trials) based on a population learning curve (n = 20) estimated from the data using a GLM and an optimal window size of 15. Individual learning curves (n=1000) were estimated from these data for four different window sizes (11, 19, 31, and 51) employing either the constant model (conv) (black lines) or the GLM (red lines). (A) Bias, i.e. the mean difference between original learning curve and curves estimated from each of the simulated response sequences. (B) Standard deviation (SD), (C) mean absolute error (MAE), and (D) coverage probabilities of confidence intervals of the estimated learning curves. Note the large values of all measures as well as the strong differences between the two models at the session borders.—The three experimental sessions are separated by different background colors.*

The right column of panels in Figure 4 shows the trial dependence of the coverage probability of the nominal 95%-confidence interval. Empirical coverage probability should be itself close to 95%, otherwise estimated confidence intervals are undersized or misplaced. Deviations of coverage probability from its 95% nominal value therefore indicate systematic errors. Similar to bias, deviation of coverage probability from 95% increased with window size. Coverage probabilities far below 95% were observed for the constant model (sep) at the beginning of sessions 2 and 3, particularly at trials 61 and 121, i.e. when bias was at its maximum. In these trials, interpretation of learning curves obtained with the constant model therefore becomes highly questionable. As with bias, deviations were much more pronounced for the constant model (sep) than for the GLM.

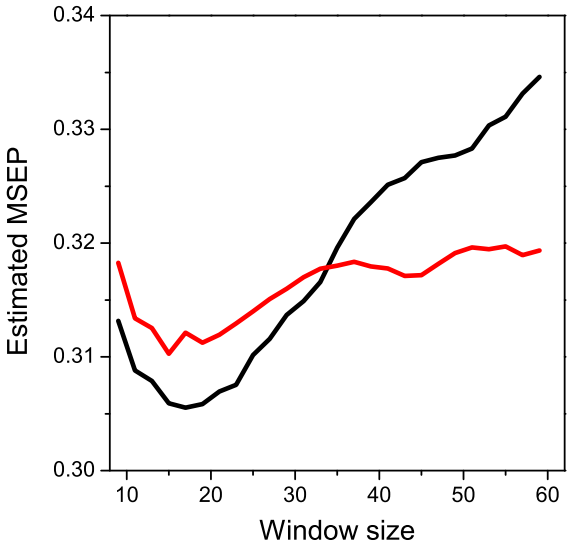

Figure 5 – *Window-size dependence of the mean squared error of prediction (MSEP) determined across trials and subjects in a cross-validation procedure for the constant model (sep) (black line) and the GLM (red line).*

3.2 Selection of window size

Our simulation shows how both random and systematic errors of learning curve estimation strongly depend on the window size. Window sizes therefore should be chosen with the aim of minimizing both, random and systematic errors. However, systematic errors assessed by measures of bias and coverage probability as in the simulation analysis, cannot be determined from the experimental data, for which the true success probabilities underlying the observed responses are not known. Though, a combined measure of estimation errors can be calculated from the experimental data through the expected mean squared error of prediction (Eqn. (2.9)) derived from a cross-validation procedure (Eqn. (2.10)). By means of *MSEP*, we quantified how well the observed response in each single trial could be predicted by the estimated success probability as a function of window size. Figure 5 shows the *MSEP* for estimates with the constant model (sep) (black trace) and with the GLM (red trace). *MSEP* was affected by an increase in random errors with shorter, as well as an increase in systematic errors with longer trial windows. To minimize and balance the influence of random and systematic errors, the use of a window size is recommended where *MSEP* displays a minimum. For the constant model (sep) and the GLM, this minimum was found at a window size of 17 and 15 trials, respectively. The slightly larger *MSEP* at its minimum obtained with the GLM compared to the respective value of the constant model (sep) reflects the larger variance observed with the GLM. Notably, for window sizes larger than 33 trials, the *MSEP* computed for the GLM remained below that for the constant model (sep). This is consistent with the finding from our simulation that the use of the GLM reduced systematic errors of learning curve estimation even for larger window sizes.

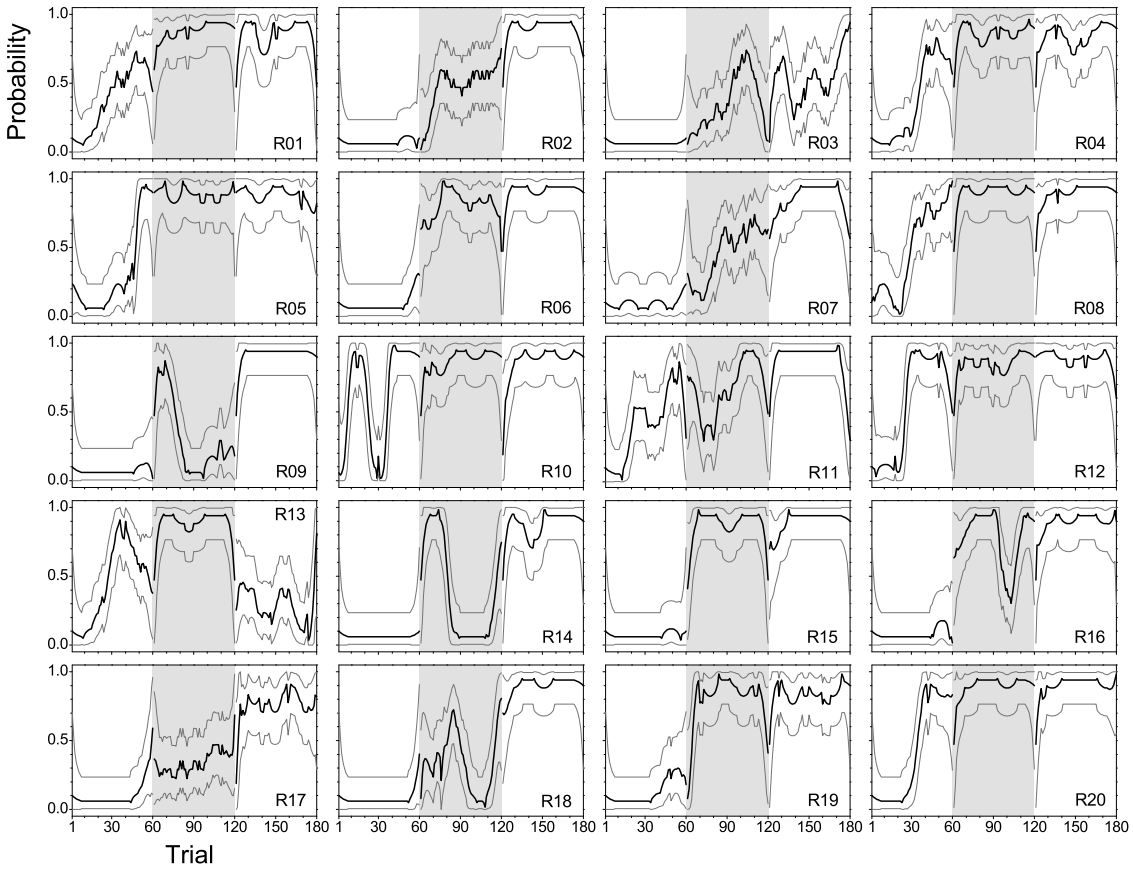

Figure 6 – *Learning curves (black lines) for the 20 individual rodents (R01 to R20) with 95%-confidence intervals (thin grey lines) derived from the constant term of a GLM maximum likelihood estimation within a moving window of optimal size of 15 trials.—The three experimental sessions are separated by different background colors.*

3.3 Intra- and intersubject variability of learning curves

Learning curves display a large variability between subjects. Figure 6 shows learning curves of all 20 rodents together with their 95%-confidence intervals estimated for each subject individually with the GLM and the optimum window size of 15 trials (see Sect. 3.2). Non-overlapping confidence intervals can be used to identify non-random changes of performance between consecutive trials or within a narrow range of neighboring trials as an indicator of learning. The prevailing pattern observed in a large number of animals ($n = 11$; R01, R04, R05, R08, R10–R13, R17, R20) was an initial increase of the success probability during the first session. For all other animals, the first non-random increase in task performance occurred in the second session. The transition between the different learning states occurred mainly within a rather narrow range of about 10 to 20 trials, but could also appear very rapidly, almost from one trial to the next. After an initial increase in performance, some animals also showed transient declines in performance (e.g. R10, R14, R16, R18) during the session.

The application of the optimal window size, and the use of a GLM instead of a constant model, further revealed fast changes in learning performance at the session boarders. Some animals

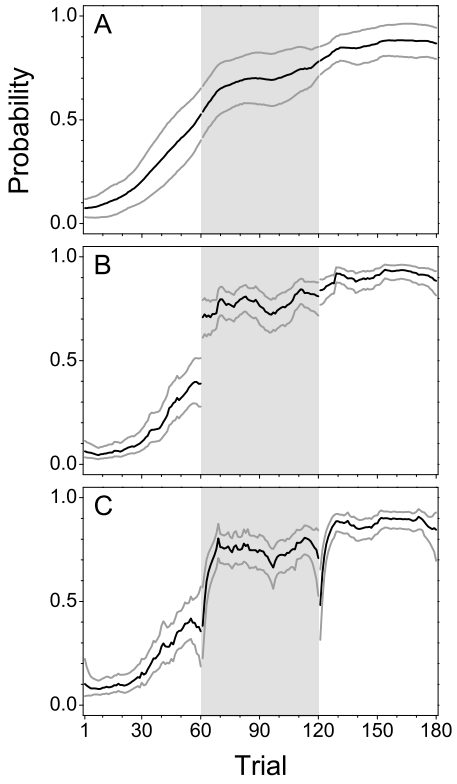

Figure 7 – *(A) Grand mean learning curve ($n = 20$, black line) obtained by conventional analysis employing the constant model (conv) in which the window is continously moved across session borders. Point-wise 95% confidence intervals (grey lines) are estimated by means of standard errors. (B) Population learning curve ($n = 20$, black line) estimated by the constant model (sep). Point-wise 95% confidence intervals (grey lines) were approximated by the Agresti-Coull interval. (C) Population learning curve ($n = 20$, black line) estimated by a general linear fixed effects model. Point-wise 95% confidence intervals (grey lines) were obtained from Firth's penalized likelihood approach for the GLM.—In all three cases, the window size was 15 trials. The three experimental sessions are separated by different background colors.*

showed a drop in performance towards the end of a session (e.g. R01, R03, R09, R11, R13, R19), and some others showed a rapid increase in learning performance within the first 10 trials in sessions 2 and 3 after initial acquisition (e.g. R01, R04, R14, R20). However, due to the broader confidence intervals at the session border, these changes were often not significant on the single subject level.

3.4 Population learning curve and its behavioral and physiological covariates

Figure 7 shows populations learning curves and their confidence intervals derived from the responses of 20 subjects obtained with the three different statistical approaches introduced in Figure 3. Figure 7A displays the conventional arithmetic grand mean learning curve obtained by

a continuous moving window analysis across session borders (constant model (conv)). Confidence intervals were based on the standard errors of the mean under Gaussian assumptions. This analysis resulted in a smooth, sigmoid population learning curve, without any discontinuities. Figures 7B and 8C show population learning curves and their confidence intervals estimated separately for each of the three sessions using the constant (B) and the generalized linear (C) fixed effects model, respectively (see Sect. 2.1.3). Confidence intervals of learning curves estimated by a fixed effects model were narrower than those obtained from standard errors of the mean (see Figure 7A). Fixed effects models therefore provided not only statistically more appropriate, but also more precise estimates. Similar to the conventional analysis, learning curves obtained from the fixed effects models showed a moderate increase of performance in the second half of the first session, although this increase was smaller and occurred somewhat later compared to the conventional model. With the constant fixed effects model (Figure 7B), a step-like increase from the last trial of session 1 to the first trial of session 2 appeared in the estimate. This step may reflect an improvement of performance during the session break without training, e.g. by memory consolidation. In the conventional analysis (Figure 7A), such discontinuities in performance were obscured. However, population learning curves based on the GLM additionally revealed fast and strong increases of performance within the first five trials of sessions 2 and 3, respectively (Figure 7C). Opposite to the constant fixed effects model, performance in the first trials of session 2 was the same as in the last trials of session 1, or even worse in the first trials of session 3 than in the last trials of session 2. After the rapid increase at the beginning of session 2 and 3, however, the performance level reached a plateau in less than 10 trials, and exceeded that of the preceding session. In the case of the generalized linear fixed effects model, the increase in performance from session 1 to session 2 would not occur unobserved in the session break as in the conventional approach, but will be expressed by a fast learning process at the beginning of session 2.

Additional insight into the learning process was gained from behavioral and physiological covariates recorded during training. Reaction time is the continuous behavioral criterion variable used to determine a correct avoidance response at a latency below 5 s relative to CS-onset. Figure 8A shows the population reaction time curve and its confidence interval obtained by local linear smoothing (see Sect. 2.1.4). Some features of the reaction-time curves are not reflected in the learning curve, e.g. the strong initial decreases of the reaction time at the beginning of each session. Even before the initial increase of the learning curve in the first session (see Figure 7), the reaction time already decayed in a fast initial phase by about 2 s within the first 10 trials. A slower decrease by about two further seconds followed within the subsequent 40 trials. Note that learning performance did not increase until the second half of the slow phase of the reaction time decay (see Figure 7). Other within-session changes of the reaction time were reflected by the population learning curves obtained by the GLM (Figure 7C). The fast drop of the reaction time at the beginning of sessions 2 and 3 was mirrored by the initial increase the population learning curve in these sessions revealed by the use of a GLM. The reaction time was larger than 5 s at the beginning of these two sessions, but only for a few trials. Then it quickly dropped below the 5-s threshold, reaching a plateau for the remaining session. This initial drop of the reaction time became steeper with each consecutive session, and the plateau decreased in steps across sessions. This mirrors the step-like increase in performance level across sessions seen in population learning curves estimated with the fixed effects models (Figures 7B, C).

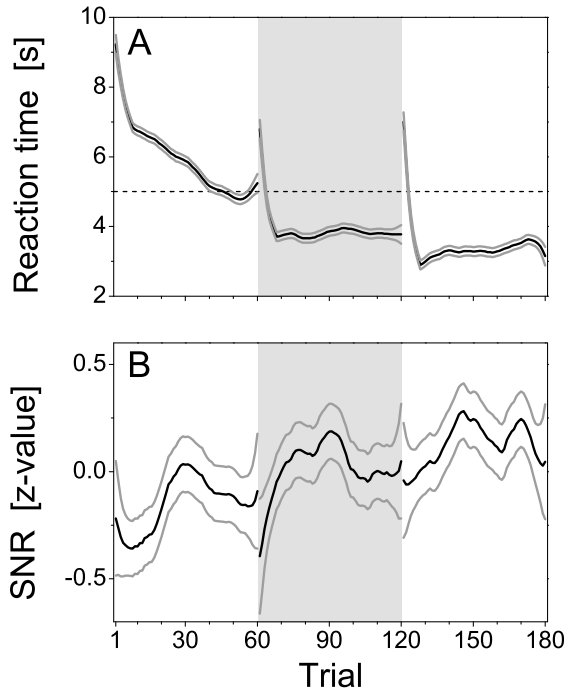

Figure 8 – *Results obtained from behavioral and physiological covariates. (A) Grand mean of reaction times ($n = 20$, black line) obtained by local linear smoothing over trials with window size 15. Point-wise 95% confidence intervals (grey lines) were determined from standard errors under Gaussian assumptions. The 5-s threshold for a correct response is indicated by the dashed horizontal line. (B) Grand mean of the z-standardized signal-to-noise ratio of the CS-evoked electric potential at the prefrontal electrode over trials after local linear smoothing with window size 15 ($n = 20$, black line). Point-wise 95% confidence intervals (grey lines) were determined from standard errors under Gaussian assumptions.—In both panels, the three experimental sessions are separated by different background colors.*

The behavioral changes displayed by learning performance and reaction time were also reflected by neurophysiological changes. During the training, CS-evoked cortical potentials were obtained from ECoG recordings over prefrontal cortex as a neurophysiological covariate of learning. In each trial, the magnitude of this CS-evoked potential was quantified by its signal-to-noise ratio (SNR) relative to pre-stimulus baseline. SNR values were calculated for each trial as the ratio of the logarithm of the root-mean-square (RMS) amplitude of the ECoG from 0 to 0.5 s after and during a 0.5-s long baseline before CS-onset. For each subject, SNR values were z-standardized across trials. Figure 8B shows the population SNR curve along with its confidence interval obtained by local linear smoothing. The SNR varied within and across sessions. SNR increased at the beginning of each session, reached a maximum after about 30 trials, and then decayed again. In the first session, SNR increased to its session maximum during the slow phase of reaction-time decay, just before the initial increase of the population learning curves. Also, SNR showed a steep increase during the first 10 trials of session 2, which paralleled the fast decay of reaction time and the fast increase in performance observed in the population learning curves estimated with a GLM. Besides within-session variation of SNR, the mean SNR

in a session increased from session to session, which might be related to the step-like changes of reaction time and performance levels across sessions. Apparently, changes in prefrontal, CS-evoked cortical potential within and across sessions were highly correlated with reaction time and behavioral performance.

4 Discussion

The aim of the present work was to improve the precision and the accuracy of the analysis of learning curves, and to allow for a characterization of the learning dynamics at high temporal resolution within a single session as well as across sessions. To achieve this—both for single subjects and on the population level—we introduced appropriate statistical methods for estimating Bernoulli success probabilities and their confidence intervals in small trial windows. Properly determined confidence intervals yield criteria for non-random changes in performance in relation to learning, which are less prone to errors than often used ad hoc criteria (e.g. three consecutive correct responses), predefined thresholds of the success probability (e.g. 0.5), or statistical tests based on Gaussian assumptions (Suzuki and Brown, 2005). Using established statistical methods, we aimed to minimize the assumptions about the learning process itself, still staying as close as possible to the most frequently used, conventional moving window analysis.

An approach for the trial-by-trial analysis of learning curves which has some similarity with this work has been developed by Smith et al. (2004). This highly sophisticated approach is based on the estimation of a state-space random-effects model of the learning process. Although this approach is only based on model assumptions that are very general and apply to most behavioral data, it is more complex and deviates from conventional learning curve analysis since it explicitly models the underlying learning process. Other trial-by-trial analyses of learning processes are based on strong assumptions about the learning process itself, as it is, for example, the case with reinforcement-learning models. Although these models can explain several aspects of avoidance learning (Maia, 2010; Moutoussis et al., 2008; Myers et al., 2014), they have been rarely fit to experimental data, because it is difficult to obtain stable paramter estimates.

4.1 Reducing errors in the estimation of learning curves by appropriate statistical methods

By applying different learning curve estimators to experimental and simulated data, we first systematically assessed the effects of window size, statistical model, learning phase, and session breaks on random and systematic estimation errors. As revealed by our simulations, learning curve estimation suffers from systematic errors. Particularly, estimation based on the constant model was severly biased in numerous trials. Moreover, estimated confidence intervals in the same trials were largely undersized or misplaced, as revealed by the coverage probability being significantly below its nominal 95%-value. Such systematic errors were found during phases of large and fast performance changes, e.g. at the beginning of sessions 2 and/or 3. These estimation errors led to a decrease of the slope of single-subject and population learning curves, and obscured fast, learning-related changes in responding (Gallistel et al., 2004). Similar systematic

errors occurred when windows were moved across session breaks obscuring fast within-session changes at session borders as well as step-like changes in performance levels across sessions. By moving windows across session breaks, such changes got lost due to smoothing.

As a consequence, we propose not to continuously step the moving window across the entire trial sequence, but to carry out the analysis for each session separately. Notably, with experimentally obtained learning curves, systematic errors might be even larger. Compared to the population learning curve used in the simulation, individual learning curves are expected to be less smooth and success probabilities therefore less constant.

When a model of constant success probability in the analysis window is applied, systematic errors are expected when the model assumptions are violated, e.g. when the distribution of the success probability in the window is varying across trials. Because changes in success probability over trials is a hallmark of learning, the constant model will never be able to describe the learning process exactly. To reduce systematic errors, we therefore replaced the constant model by a generalized linear model to account for changes of success probability within the analysis windows. Using maximum likelihood methods to estimate Bernoulli probabilities in small trial windows by a GLM, as we did in our study, can lead to bias and possible infinite parameter estimates. To avoid such problems, we employed Firth's penalized likelihood estimates. In accordance with the reduction of systematic errors at the beginning of sessions 2 and 3, the GLM revealed fast and strong increases in population performance within the first five trials of these session, which were not detected by the constant model. Corresponding fast and strong behavioral and physiological changes were found in the continuous reaction times and the magnitude of evoked prefrontal cortical potentials, respectively. Moreover, similar changes in responding at the beginning of post-acquisition sessions have been repeatedly reported in avoidance learning, a phenomenon called 'warm up' (see Sect. 4.2). All these findings strongly suggest that the fast changes detected by the GLM approach were true changes, although confidence intervals at session borders were larger for the GLM than for the constant model.

Both, the observed random and systematic errors were not only dependent on the chosen model, but also on the window size and the learning phase. Systematic errors (bias) generally increased with increasing window size, whereas random errors of the estimation assessed by variance and confidence intervals decreased. Smaller random errors can be explained by the increase of sample size with window size. Systematic errors were mainly due to the non-stationarity of success probabilities within a trial window. However, systematic errors did not increase for all trials, but only during phases of fast and large performance change. The magnitude of systematic errors thereby depends on the temporal scales of the involved learning processes in relation the chosen window size. If changes within a given window are strong and relatively fast in comparison to window size, the proportion of trials in the window that differ in their success probability from the estimated trial will increase. Also, with growing window size, the estimate at a trial becomes increasingly influenced by distant trials from an other phase of the learning process potentially governed by different success probabilities.

Thus, window size is a major determinant of the quality of learning curve analysis. Optimal choice of window size thereby requires to minimize and balance random and systematic errors. This is achieved by calculating an overall measure of estimation error, the *MSEP*, for different window sizes in a cross-validation procedure, and, subsequently, by selecting the window size

with minimum *MSEP*. We found an optimal window size of 17 or 15 trials for our data, for the constant model (sep) and the GLM, respectively. This number fits well to the observed temporal scales of performance changes which were on the order 5 to 25 trials.

Employing a GLM with Firth logistic regression seemed to be advantageous for protecting estimation against systematic errors, particularly when the chosen window size was large relative to the temporal scale of the learning-related performance changes. Also, the use of a GLM improved the temporal resolution of learning curve analysis. With this approach, systematic errors still increased with increasing window size, although to a lesser extent. Changes of the success probability within the window over trials presumably became non-monotonic, violating the assumption of the GLM. In contrast to the systematic errors, however, random errors of the GLM were much larger, and decreased less with growing window size compared to the constant model (sep). This can be explained by the increased number of parameters that have to be estimated in the GLM, which increases the degrees of freedom, and therefore decreases the precision of the estimates. Hence, if the length of the window is at or below the expected temporal scale of learning, the choice of a constant model can be appropriate as well, and would then yield a higher precision.

4.2 Implications for the analysis and interpretation of experimental data

To reduce the errors in estimating learning curves from our experimental data, we combined the following methodological steps: (1) the use of a GLM and a fixed effects model fitted by Firth logistic regression for single subject and population analysis, (2) the selection of an optimal window size by cross-validation, and (3) session-wise estimation. The application of these measures yields results differing from conventional analysis, and provides an extension of the interpretation of the process of avoidance learning.

The major objective of reinforcement learning is maximization of reward and minimization of punishment. In avoidance learning, this is equivalent to a reduction in footshock exposure. The first significant change in behavior and physiology was a fast decay in reaction time during the first 10 trials of the first session (Figure 8A). Although this change was not accompanied by a change in avoidance performance, a decay in reaction time shortens the exposure to the footshock, which was turned off by an escape response. Therefore, this decay presumably reflects a first phase of instrumental learning to escape from the shock (Cain and LeDoux, 2003). This phase was followed by a second, slower phase of reaction time decay over about 40 trials, that might reflect a second instrumental escape learning phase, in which animals acquire the temporal relationship between CS and footshock (US) onset. The simultaneous increase of the magnitude of the CS-evoked prefrontal cortical potential might be related to perceptual, motor, emotional, or cognitive aspects of this learning process, but it could also more specifically reflect aspects of forming Pavlovian associations between CS and US known to be involved in shuttlebox avoidance learning (Cain and LeDoux, 2003; Stark et al., 2008; Choi et al., 2010). In the second half of this phase, reaction time further decreased reaching the avoidance criterion of 5 s reflected in a first increase in avoidance performance. This raises the question, whether early correct responses at reaction times below 5 s were genuine avoidance responses. In contrast to an avoidance response controlled by the CS predicting the footshock, these responses might be

due to a premature timing of an escape response controlled by the US, which is primed but not elicited by the CS. This could explain that animals only reached success rates of below 50% in this early acquisition phase: If animals timed their escape response by the CS-US interval, and if the error distribution of this interval estimation were be symmetric, correct responses would occur in maximally 50% of the cases. Fast, transient increases of performance, decay in reaction time, and increases in the magnitude of evoked prefrontal potentials were observed during the first ten trials in sessions 2 and 3. Similar changes in responding have been described at the beginning of post-acquisition sessions in various avoidance paradigms (Servatius et al., 2008; Spear et al., 1973; Kamin, 1963; McSweeney and Roll, 1993), and have been called 'warm up', which is interpreted as a reinstatement of memory, i.e. a fast relearning of instrumental responses weakened by memory interference due to inter-session activities (Myers et al., 2014). However, the fast performance increase in session 2 differed from the described 'warm up' phenomena, as the learning performance in the first trial did not fall below the performance level reached in the preceding session 1, and the fast performance increase largely exceeded the level reached in session 1.

Changes in performance were not only found within, but also across sessions. Thus, levels of behavioral performance increased from one session to the next, markedly between session 1 and 2. This was also reflected by decreasing levels of reaction time, and increasing session means of the magnticude of CS-evoked potentials across session. These across-session changes might be related to memory consolidation during session breaks, which can improve performance in the time between sessions, especially when animals had the opportunity to sleep (Schicknick and Tischmeyer, 2006; Hennevin et al., 2007). However, changes in performance level did not occur unobserved during the session break, but were instantiated within the first trials of a session. Seemingly, increases in performance levels were learned within a few trials.To become behaviorally effective, consolidation might therefore require fast learning in the succeeding session. Improvement of responding might thereby not only rely on learning new associations, but also on a better utilization of previously learned information. Such a view would extend to the current interpretation of "warm-up"phenomena as reinstatement of memory.

Finally, it should be noted that inter-individual variability of learning was high. Initial increase in avoidance performance occurred at different times in different animals, even not always within the first session. Also, the initial increase tended to be faster in single subjects. Moreover, in learning curves of some subjects, within session drops in performance were observed after a success rate of about 80% had been reached for several trials. Whereas possible fatigue over trials should be similarly observed in all sessions, behavioral and physiological changes towards the session end depended on the learning state, and might therefore reflect a decrement in attention or motivation. Alternatively, these changes might be due to a fall back to an effective escape strategy by responding to the US after priming by the CS. Further statistical methods have to be developed that cope with biases arising from trial-by-trial variance inhomogeneities across subjects (Gallistel et al., 2004; Bathellier et al., 2013).

In summary, our results show that even in simple tasks, learning can be multi-phasic, and involve multiple behavioral and physiological processes acting on different time scales (Cain and LeDoux, 2003; Pessoa, 2008). Such a dynamic view of learning is in accordance with studies characterizing learning as governed by the interplay of various conflicting and cooperating brain

systems (Balleine and O'Doherty, 2010; Choi et al., 2010). Whether and how these processes depend on each other could be investigated by our methods through a systematic variation of task parameters like CS-US interval, and shock strength, or by pseudo-learning experiments with randomized CS-US contingencies and joked US presentation. As noted by McSweeney and Roll (1993), describing such a rich learning dynamics only on a molar level, e.g. by session means of success rates, might be too gross for an adequate understanding of learning. Otherwise, a molecular description of learning by the responses of each single trial as by reinforcement learning models must also account for the described multi-scale learning dynamics (Myers et al., 2014). In any case, a detailed, quantitative assessment of the behavioral learning dynamics, as provided by our approach, is crucial for properly interpreting behavioral and neural changes observed at different time scales during a task with respect to learning.

Acknowledgments

We thank Thomas Rothe and late Holger Stark for providing the experimental data. We are also grateful for the discussions with André Brechmann at an early stage of the manuscript. Financial support as a LIN Special Project 2010 is gratefully acknowledged.

References

Agresti, A. and Coull, B. A. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. The American Statistician 52, 119–126.

Bathellier, B., Tee, S.P., Hrovat, C., Rumpel, S. (2013) A multiplicative reinforcement learning model capturing learning dynamics and interindividual variability in mice. Proceedings of the National Academy of Sciences of the United States of America, 110 (49), 19950-19955.

Balleine, B. W. and O'Doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology 35, 48–69.

Bickel, P. J. and Doksum, K. A. (1977). Mathematical statistics: Basic Ideas and Selected Topics. Holden-Day: San Francisco.

Blyth, C. R. and Still, H. (1983). Binomial Confidence Intervals. Journal of the American Statistical Association 78 (381), 108–116.

Brown, L. D., Cai, T. T., and DasGupta, A. (2001). Interval estimation for a binomial proportion. Statistical Science 16, 101–133.

Brown, S. and Heathcote, A. (2003). Averaging learning curves across and within participants. Behavior Research Methods, Instruments & Computers 35 (1), 11–21.

Cain, C.K. and LeDoux, J.E. (2008). Chapter 3.1 Brain mechanisms of Pavlovian and instrumental aversive conditioning. Handbook of Behavioral Neuroscience 17, 103-124.

Choi, J.-S., Cain, C.K., Ledoux, J.E. (2010). The role of amygdala nuclei in the expression of auditory signaled two-way active avoidance in rats. Learning and Memory, 17 (3), 139-147.

Collett, D. (2003). Modelling Binary Data. Chapman and Hall, London.

Fan, J. and Gijbels, I. (1996). Local polynomial modelling and its applications. Chapman and Hall, London.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. Biometrika 80 (1), 27-38.

Gallistel, C. R., Fairhurst, S., and Balsam, P. (2004). The learning curve: Implications of a quantitative analysis. Proceedings of the National Academy of Sciences of the USA 101 (36), 13124–13131.

Heinze, G. and Schemper, M. (2002). A solution to the problem of separation in logistic regression. Statistics in Medicine, 21 (16), 2409-2419.

Hennevin, E., Huetz, C., Edeline, J.-M. (2007). Neural representations during sleep: From sensory processing to memory traces. Neurobiology of Learning and Memory, 87 (3), pp. 416-440.

Kamin, L.J. (1963). Retention of an incompletely learned avoidance response: some further analyses. Journal of Comparative and Physiological Psychology 56, 713-718.

Maia, T.V. (2010). Two-factor theory, the actor-critic model, and conditioned avoidance. Learning and Behavior 38 (1), 50-67.

Maechler, M. (2013). lpridge: Local Polynomial (Ridge) Regression. R package version 1.0-7. URL `http://CRAN.R-project.org/package=lpridge`

McCullagh, P. and Nelder, J. A. (1998). Generalized linear models. In: Monographs on Statistics and Applied Probability. Chapman and Hall, London.

McSweeney, F.K. and Roll, J.M. (1993). Responding changes systematically within sessions during conditioning procedures. Journal of the Experimental Analysis of Behavior 60 (3), 621–640.

Moutoussis, M., Bentall, R. P., Williams, J., and Dayan, P. (2008). A temporal difference account of avoidance learning. Network 19, 137-160.

Myers, C.E., Smith, I.M., Servatius, R.J., and Beck, K.D. (2014). Absence of “warm-up” during active avoidance learning in a rat model of anxiety vulnerability: Insights from computational modeling. Frontiers in Behavioral Neuroscience, 8, art. no. 283.

R Core Team (2014). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL `http://www.R-project.org/`

Pessoa, L. (2008). On the relationship between emotion and cognition. Nature Reviews Neuroscience 9 (2), 148–158.

Rothe, T., Deliano, M., Scheich, H., and Stark, H. (2009). Segregation of task-relevant conditioned stimuli from background stimuli by associative learning. Brain Research 1297, 143–159.

Servatius, R. J., Jiao, X., Beck, K. D., Pang, K. C., and Minor, T. R. (2008). Rapid avoidance acquisition in Wistar-Kyoto rats. Behavioural Brain Research 192, 191-197.

Spear, N. E., Gordon, W. C., and Martin, P. A. (1973). Warm-up decrement as failure in memory retrieval in the rat. Nature Journal of Comparative and Physiological Psychology 85, 601-614.

Stark, H., Rothe, T., Deliano, M., and Scheich, H. (2007). Theta activity attenuation correlates with avoidance learning progress in gerbils. Neuroreport 18 (6), 549–552.

Stark, H., Rothe, T., Deliano, M., and Scheich, H. (2008). Dynamics of cortical theta activity correlates with stages of auditory avoidance strategy formation in a shuttle-box. Neuroscience 151 (2), 467–475.

Schicknick, H. and Tischmeyer, W. (2006). Consolidation of auditory cortex-dependent memory requires n-methyl-d-aspartate receptor activation. Neuropharmacology 50 (6), 671–676.

Smith, A.C., Frank, L.M., Wirth, S., Yanike, M., Hu, D., Kubota, Y., Graybiel, A.M., Suzuki, W.A., and Brown, E.N. (2004). Dynamic Analysis of Learning in Behavioral Experiments. Journal of Neuroscience 24 (2), 447–461.

Suzuki, W.A. and Brown, E.N. (2005). Behavioral and neurophysiological analyses of dynamic learning processes. Behavioral and Cognitive Neuroscience Reviews, 4 (2), 67–95.