bioRxiv preprint doi: https://doi.org/10.1101/072678; this version posted November 3, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Running title: PENUMBRA OF LEARNING

Behavioral tagging and the penumbra of learning

Samuel J. Gershman
Department of Psychology and Center for Brain Science
Harvard University

Address for correspondence:
Samuel Gershman
Department of Psychology
Harvard University
52 Oxford St., room 295.05
Cambridge, MA 02138
Phone: 773-607-9817
E-mail: gershman@fas.harvard.edu

bioRxiv preprint doi: https://doi.org/10.1101/072678; this version posted November 3, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

In noisy, dynamic environments, organisms must distinguish genuine change (e.g., the movement of prey) from noise (e.g., the rustling of leaves). Expectations should be updated only when the organism believes genuine change has occurred. Although individual variables can be highly unreliable, organisms can take advantage of the fact that changes tend to be correlated (e.g., movement of prey will tend to produce changes in both visual and olfactory modalities). Thus, observing a change in one variable provides information about the rate of change for other variables. We call this the *penumbra of learning*. At the neural level, the penumbra of learning may offer an explanation for why strong plasticity in one synapse can rescue weak plasticity at another (synaptic tagging and capture). At the behavioral level, it has been observed that weak learning of one task can be rescued by novelty exposure before or after the learning task. Here, using a simple number prediction task, we provide direct behavioral support for the penumbra of learning in humans, and show that it can be accounted for by a normative computational theory of learning.

bioRxiv preprint doi: https://doi.org/10.1101/072678; this version posted November 3, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Some of the most significant developments in synaptic physiology have sprung from the insight that no synapse is an island: changes in plasticity at one synapse can exert a variety of effects on plasticity at other synapses, such as synaptic scaling [1], heterosynaptic plasticity [2], and (the focus of this paper) synaptic tagging and capture [3, 4, 5]. The synaptic tagging and capture hypothesis holds that weak stimulation at a synapse, while insufficient to produce synthesis of plasticity-related proteins necessary for long-lasting potentiation, sets a molecular tag that allows the synapse to subsequently capture proteins synthesized at another convergent synapse undergoing strong stimulation. Capture is possible in both a "weak-before-strong" scenario (weak stimulation at synapse A followed by strong stimulation at synapse B) and a "strong-before-weak" scenario (strong stimulation at synapse B followed by weak stimulation at synapse A). Tagging and capture also occurs for synapses undergoing depression [6]. Furthermore, strong depression at one synapse can rescue weak potentiation at another synapse, and vice versa ("cross-capture" [6]).

In an effort to organize the functional principles underlying these phenomena, Gershman [7] proposed a normative computational model of synaptic tagging and capture. According to this model, a basic challenge facing any learning system is distinguishing between change in the environment (which should be learned) and noise (which should be ignored). Under the assumption that the rate of change is correlated across different variables, observing a large change for one variable is informative about the global rate of change, which in turn alters learning about other changes—the *penumbra of learning*. Thus, changes that appear insignificant under some circumstances (due to noise) may attain significance when the environment is perceived to be changing quickly. Applied to the synaptic level, this principle explains why strong learning at one synapse can enhance weak learning at another synapse.

One advantage of such an abstract model formulation is that it can be applied at multiple levels; for example, we should be able to observe the penumbra of learning behaviorally. Indeed, "behavioral tagging" experiments demonstrate that exposing an animal to novelty can enhance weak learning on a separate task [8, 9, 10, 11]. One limitation of these experiments is that they are not amenable to a systematic computational analysis, since it is unclear how to quantitatively model the novelty manipulation. In the present paper, we study a simple behavioral task that allows us to test the model's predictions in humans.

Participants performed a number prediction task [12] in which they had to guess the number a colored button would produce. Each button's output changed over time according to a Gaussian random walk, such that participants had to continually update their predictions. Trials alternated between red and green buttons (Figure 1), allowing us to measure the penumbra of learning in terms of whether a large or small change in prediction for one of the buttons was associated with a large or small change in the other button. By looking at the direction of change (i.e., positive or negative), we can also measure a behavioral analogue

bioRxiv preprint doi: https://doi.org/10.1101/072678; this version posted November 3, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

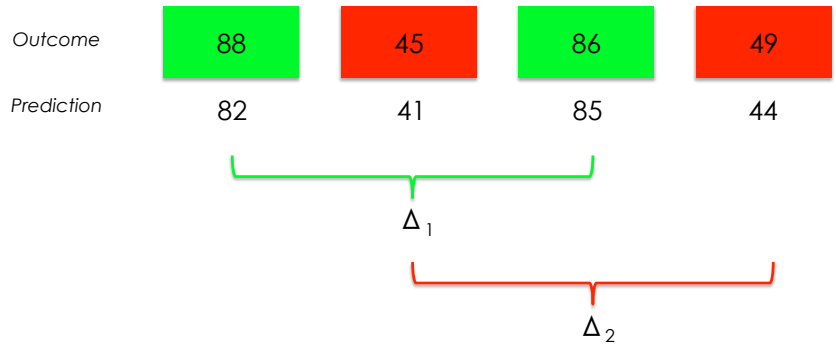

Figure 1: **Task and analysis schematic**. Participants alternated between making predictions for red and green buttons. After entering a predicted number, participants observed feedback about the true number, drawn from a Gaussian distribution whose mean fluctuated over time according to a Gaussian random walk. Two variables were computed from the sequence of predictions: the difference between contiguous green predictions (Δ_1) and the difference between contiguous red predictions (Δ_2). These variables were used to quantify the penumbra of learning, as explained in the main text.

of cross-capture: an association between large positive changes on one button with large negative changes in the other button, or vice versa.

Results

Experiment 1

The experimental results are summarized in Figure 2. First, we measured "forward tagging": the absolute change in prediction on trial t ($|\Delta_2|$) as a function of whether the absolute change in prediction on trial $t-1$ ($|\Delta_1|$) is large (higher than average) or small (lower than average). Consistent with the penumbra of learning hypothesis, we observed that $|\Delta_2|$ was larger when $|\Delta_1|$ was large compared to when $|\Delta_1|$ was small [$t(148) = 6.23, p < 0.0001$; Figure 2A]. We also observed "backward tagging": $|\Delta_1|$ was larger when $|\Delta_2|$ was large compared to when $|\Delta_2|$ was small [$t(148) = 5.47, p < 0.0001$; Figure 2B].

We next examined a behavioral analogue of "cross-capture," where a change in prediction for one button amplifies change in prediction for the other button, even when the directions of these changes go in opposite directions. Thus, a negative change on trial $t-1$ amplifies a positive change on trial t [$t(117) = 3.92, p < 0.001$; Figure 2C], and a positive change on trial $t-1$ amplifies a negative change on trial t [$t(120) = 3.44, p < 0.001$; Figure 2D]. These results support the claim that it is the magnitude, not direction, of change that determines the strength of the penumbra. Consistent with this claim, we found no significant correlation

bioRxiv preprint doi: https://doi.org/10.1101/072678; this version posted November 3, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

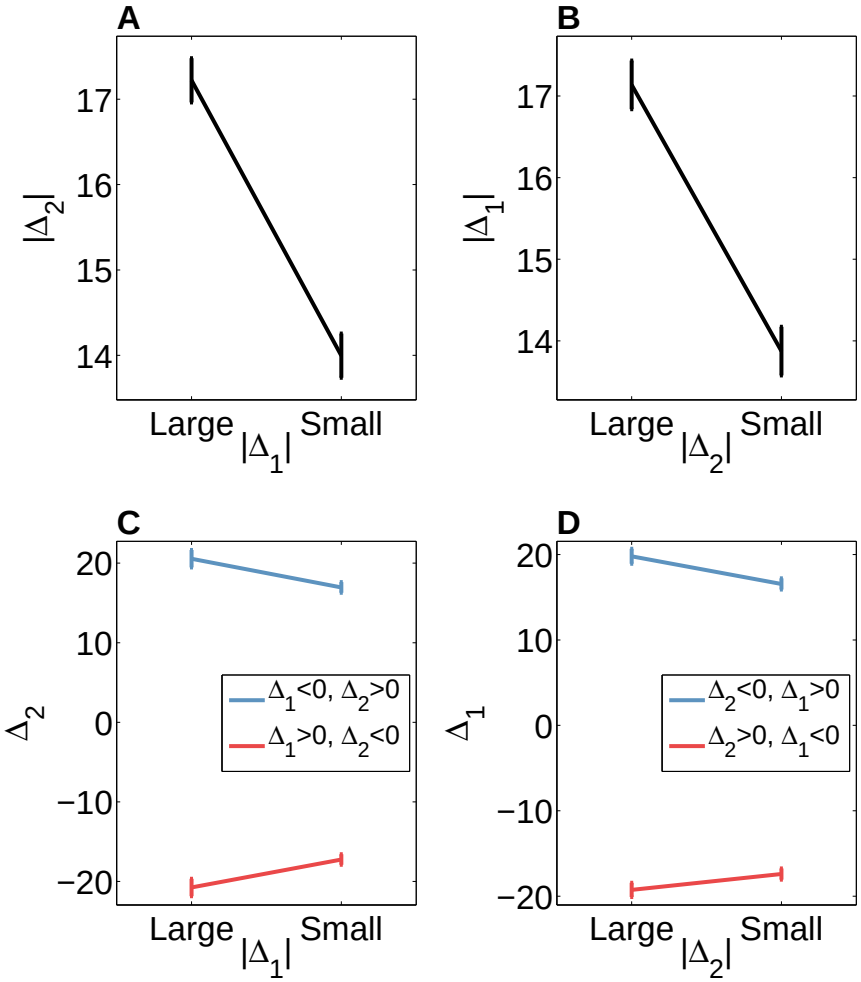

Figure 2: **Experiment 1 results**. (A) Forward tagging: absolute change in prediction on trial t as a function of the size of change in prediction on trial $t-1$. The "large" and "small" categories were determined on the basis of a mean split. (B) Backward tagging: absolute change in prediction on trial $t-1$ as a function of the size of change in prediction on trial t. (C) Forward cross-tagging: a negative change on trial $t-1$ amplifies a positive change on trial t. Similarly, a positive change on trial $t-1$ amplifies a negative change on trial t. (D) Backward cross-tagging: a negative change on trial t amplifies a positive change on trial $t-1$. Similarly, a positive change on trial t amplifies a negative change on trial $t-1$. Error bars show standard error of the mean.

between predictions on trial t and trial $t-1$ (median Pearson correlation coefficient: $0.06, p = 0.15$, signed rank test).

bioRxiv preprint doi: https://doi.org/10.1101/072678; this version posted November 3, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

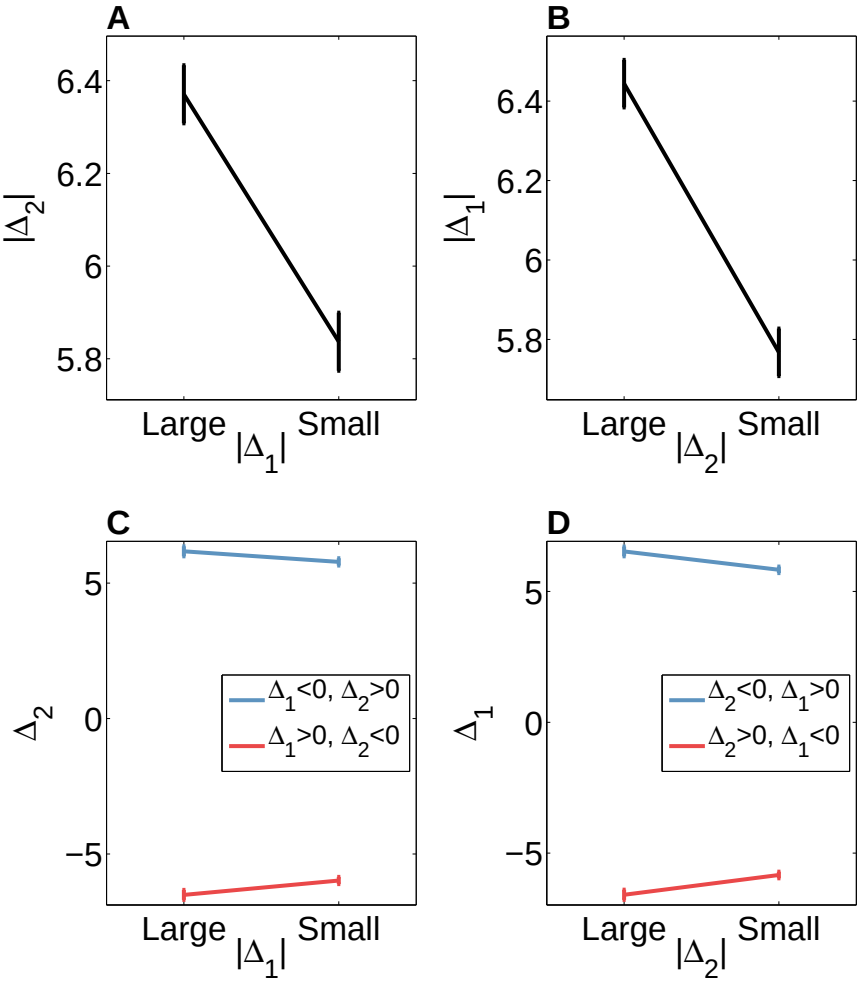

Figure 3: **Experiment 1 model simulations**. (A) Forward tagging. (B) Backward tagging. (C) Forward cross-tagging. (D) Backward cross-tagging. Error bars show standard error of the mean.

These empirical results are well captured by the normative computational model described in the Methods section (Figure 3), which was previously used to explain a variety of synaptic tagging and capture phenomena [7]. The key principle is that participants are simultaneously estimating both the number and the diffusion variance (or volatility) q. Change on one button is, according to the model, informative about the rate of change for all buttons. Formalized mathematically, without any parameter fitting, this principle is sufficient to reproduce forward/backward tagging and forward/backward cross-tagging.

bioRxiv preprint doi: https://doi.org/10.1101/072678; this version posted November 3, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

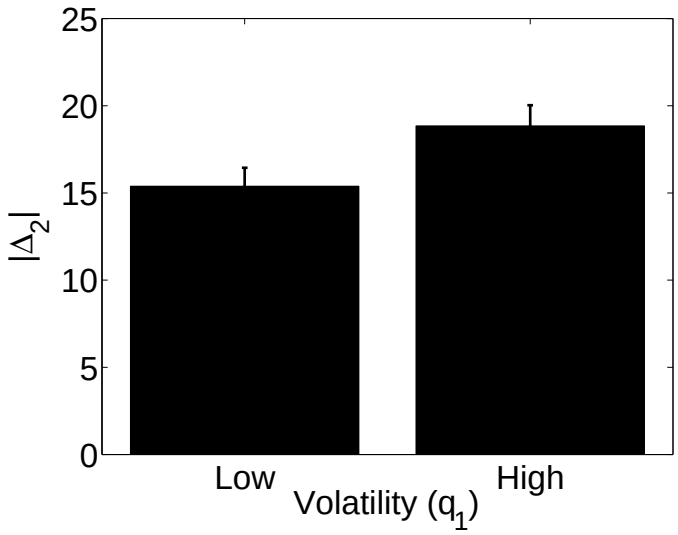

Figure 4: **Experiment 2 results**. Absolute change in prediction for button 2 ($|\Delta_2|$) as a function of volatility for button 1 (Low volatility: $q_1 = 25$; High volatility: $q_1 = 225$.). Error bars show standard error of the mean.

Experiment 2

In Experiment 1, the diffusion variance was the same across buttons, and we exploited natural variability in prediction errors to test the model's predictions. In Experiment 2, we sought to directly manipulate the size of prediction errors for button 1 while keeping the prediction errors for button 2 constant. A separate set of participants ($N = 46$) performed the same task as in Experiment 1, with diffusion variance for button 1 (q_1) set to 225 and the diffusion variance for button 2 (q_2) set to 25. We refer to this new condition as "high volatility" for button 1, compared to the "low volatility" condition studied in the subset of participants in Experiment 1 ($N = 53$) for whom both q_1 and q_2 were set to 25 (see Methods for details). Crucially, the diffusion variance for button 2 was the same across the high volatility and low volatility conditions.

As shown in Figure 4, we found that the absolute change in prediction for button 2 ($|\Delta_2|$) was significantly greater in the high volatility condition compared to the low volatility condition $[t(97) = 2.19, p < 0.05]$. This finding supports the model prediction that high volatility for one button will induce greater plasticity of learning for the other button (whose volatility was fixed). Experiment 2 thus complements the correlational results of Experiment 1 with a causal experimental manipulation.

bioRxiv preprint doi: https://doi.org/10.1101/072678; this version posted November 3, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discussion

We have shown a form of behavioral tagging in humans that recapitulates analogues of several important synaptic tagging and capture phenomena. First, we found that a large change in predictions for one stimulus amplified learning about another stimulus. This "penumbra of learning" operated both forward and backward in time. Second, we found that a large positive change for one stimulus amplified negative changes for the other stimulus, and vice versa, indicating that the penumbra depends on the magnitude (rather than direction) of change. Third, we found that increasing the volatility (and hence the average change in predictions) for one stimulus amplified learning about the other stimulus, indicating that the penumbra of learning can be exogenously manipulated.

Our paradigm made it possible to test a computational account of synaptic tagging and capture [7], applied with minimal modifications to the behavioral task. This model formalizes the penumbra of learning in terms of inference about the volatility of the environment. Under the assumption that stimulus dynamics are coupled via an unknown volatility parameter, perceiving a change in one stimulus increases the inferred volatility and thereby amplifies learning about other stimuli. The notion that humans and animals adapt their learning rate based on inferred volatility has become an important facet of modern learning theory [13, 14, 15, 16], with roots in earlier models of classical conditioning [17]. The penumbra of learning model extends this logic to multidmensional environments. Although originally applied to synaptic plasticity phenomena [7], we have shown that the model can accurately predict the behavioral tagging phenomena reported in this paper, lending credence to the idea that synaptic and behavioral learning reflect common principles. However, we still lack physiological evidence for a direct connection.

Our results pave the way for several future directions of inquiry. First, it has been demonstrated in animals that behavioral tagging is dopamine-dependent [8, 18, 10], suggesting that systemically depleting dopamine in humans should eliminate the penumbra of learning. Behavioral tagging is also dependent on β-adrenergic function [18], which could also be systemically manipulated. The role of β-adrenergic function is particularly intriguing given the role of noradrenergic neurons in adaptation of learning [19, 20]. Pupil diameter, tightly linked to firing of noradrenergic neurons [20], tracks changes in learning rate [15], and could be easily measured in our task.

Methods

Participants

For Experiment 1, 149 participants were recruited online through the Amazon Mechanical Turk web service and paid for their participation. For Experiment 2, 46 participants were

bioRxiv preprint doi: https://doi.org/10.1101/072678; this version posted November 3, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

recruited. The experiments were approved by the Harvard Internal Review Board.

Procedure

Participants completed 100 trials of a number prediction task. On each trial, participants were shown a red or green button (on alternating trials) and had to guess a number between 1 and 100. Participants then received feedback. The numbers generated by the buttons evolved over time according to a Gaussian random walk:

$$w_t \sim \mathcal{N}(f(w_{t-1}), q\mathbf{I}) \tag{1}$$

$$y_t \sim \mathcal{N}(w_t x_t, r) \tag{2}$$

where y_t is the number observed on trial t, $w_t \in \mathbb{R}^2$ is the vector of mean values, $\mathbf{I}$ is the 2×2 identity matrix, $x_t \in \mathbb{R}^2$ is a vector representing the button presented on trial t ($x_{tk} = 1$ if button k was present on trial t, and 0 otherwise), q is the diffusion variance of the random walk, and r is the noise variance of the emission distribution. The function $f(w)$ specifies the deterministic component of the dynamics; here we assume $f(w) = w$, but in the next section we discuss different assumptions about $f(w)$. To explore the parameter space in Experiment 1, 53 participants completed the task with $q = 25$ and $r = 25$, 50 with $q = 1$ and $r = 25$, and 46 with $q = 25$ and $r = 100$. However, the experimental results were nearly indistinguishable across conditions, so in the remainder of the paper we report results pooled across conditions. Experiment 2 used an identical procedure, with the only difference that the two buttons had different diffusion variances: $q_1 = 225$ and $q_2 = 25$.

Computational model

Here we summarize the computational model first presented in [7]; we refer the reader to that paper for details. Our approach is to impute to the learner a set of probabilistic assumptions about the environment (the *generative process*), and then formalize the learning problem as inference over the latent causes that gave rise to observed data under the generative process. Thus, we posit that the learner is executing a reasonable accurate approximation of probabilistic inference. In [7], the learner was conceptualized as a set of synapses; in this paper, we apply essentially the same model at the organismic level, only changing some of the parameters to match our experimental paradigm.

Formally, the computational problem facing the learner is to infer both the dynamic mean w_t and the diffusion variance q after observing data $(\mathbf{x}_{1:t}, \mathbf{y}_{1:t})$, where the subscript $1:t$ denotes the collection of variables on trials 1 to t. The optimal statistical inference is stipulated by Bayes' rule:

$$P(w_t, q|\mathbf{x}_{1:t}, \mathbf{y}_{1:t}) \propto P(y_t|w_t, x_t)P(w_t|\mathbf{x}_{1:t-1}, \mathbf{y}_{1:t-1}, q)P(q), \tag{3}$$

bioRxiv preprint doi: https://doi.org/10.1101/072678; this version posted November 3, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We approximate the distribution over the diffusion variance q with a discrete set of points, $\{q^1, \ldots, q^K\}$. Specifically, the diffusion variance is drawn according to $P(q = q^k) \propto \exp\{ak^{-b}\}$ where a and b are constants and q^k ranges over K equally spaced values between 1 and 30 (note that the k superscript is an index, not a power). This distribution embodies the assumption that a slow rate of change is *a priori* more probable. For all our simulations, we used $K = 50, a = 8$ and $b = 5$, but the results are not sensitive to small variations in these values.

When q is known, the posterior over w_t can be approximated as a Gaussian using the extended Kalman filter [21], leading to the following updates for the mean and variance:

$$\textbf{Posterior mean}\text{: } \hat{w}_t = f(\hat{w}_{t-1}) + \alpha_t x_t (y_t - \hat{w}_{t-1} x_t) \tag{4}$$

$$\textbf{Posterior variance}\text{: } v_t = (1 - \alpha_t x_t^2)\sigma_t^2 \tag{5}$$

where $\sigma_t^2 = [f'(\hat{w}_{t-1})]^2 v_{t-1} + q$ is the predictive variance after application of the transition function $f(\hat{w}_{t-1})$ but before observing (x_t, y_t), and $f'(w)$ is the derivative of the transition function.

The learning rate α_t is given by $\alpha_t = \sigma_t^2/\lambda_t$, where $\lambda_t = x_t^2\sigma_t^2 + r$ is the variance of the posterior marginal likelihood:

$$P(y_t|\mathbf{x}_{1:t-1}, \mathbf{y}_{1:t-1}, q) = \mathcal{N}(y_t; \hat{w}_{t-1}x_t, \lambda_t). \tag{6}$$

Intuitively, λ_t expresses the overall unpredictability (or noisiness) of the data; this unpredictability grows monotonically with the noise variance. Learning will be faster to the extent that the prediction uncertainty is high (large σ_t^2) relative to the overall unpredictability (low λ_t).

With q unknown, the posterior marginal over w_t is a mixture of Gaussians:

$$P(w_t|\mathbf{x}_{1:t}, \mathbf{y}_{1:t}) = \sum_{k=1}^{K} P(q = q^k|\mathbf{x}_{1:t}, \mathbf{y}_{1:t})\mathcal{N}(w_t; \hat{w}_t^k, v_t^k). \tag{7}$$

where we have indexed $\hat{w}$ and v by k to indicate their implicit dependence on $q = q^k$. To obtain $P(q = q^k|\mathbf{x}_{1:t}, \mathbf{y}_{1:t})$, we apply the update equations 4 and 5 for each q^k in parallel, and use Bayes' rule to compute the posterior over q:

$$P(q = q^k|\mathbf{x}_{1:t}, \mathbf{y}_{1:t}) \propto P(q = q^k) \prod_{\tau=1}^{t} \mathcal{N}(y_\tau; \hat{w}_{\tau-1}^k x_\tau, \lambda_\tau^k). \tag{8}$$

In [7], a nonlinear transition function was specified that captures two properties of synaptic plasticity: (1) asymptotic decay to 0, and (2) slower decay for stronger associations. The following functional form was used:

$$f(w) = \frac{w}{1 + \exp\{-\beta w^2\}}, \tag{9}$$

bioRxiv preprint doi: https://doi.org/10.1101/072678; this version posted November 3, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

where $\beta \geq 0$ is a parameter that governs the nonlinear relationship between w and the decay rate; in particular, smaller absolute values of w decay more rapidly than larger values. Following [7], we used $\beta = 10$. Note that in the *true* generative process (unknown to participants), $f(w)$ was the identity function. We employ the nonlinear transition function for consistency with prior work, but our results do not depend on this assumption.

In the simulations, all parameters were the same as in the original report [7], except that we set r equal to its true experimental value (25 or 100 depending on the experimental condition). Participants did not know this true value, but we chose it for expedience; our results are robust to large variations in this parameter value.

Acknowledgments

This work was supported in part by the Center for Brains, Minds and Machines (CBMM), funded by NSF STC award CCF-1231216.

References

[1] Turrigiano GG. The self-tuning neuron: synaptic scaling of excitatory synapses. Cell. 2008;135:422–435.

[2] Schuman EM, Madison DV. Locally distributed synaptic potentiation in the hippocampus. Science. 1994;263:532–536.

[3] Frey U, Morris RG. Synaptic tagging and long-term potentiation. Nature. 1997;385:533–536.

[4] Frey U, Morris R. Weak before strong: dissociating synaptic tagging and plasticity-factor accounts of late-LTP. Neuropharmacology. 1998;37:545–552.

[5] Redondo RL, Morris RG. Making memories last: the synaptic tagging and capture hypothesis. Nature Reviews Neuroscience. 2011;12:17–30.

[6] Sajikumar S, Frey JU. Late-associativity, synaptic tagging, and the role of dopamine during LTP and LTD. Neurobiology of Learning and Memory. 2004;82:12–25.

[7] Gershman SJ. The penumbra of learning: A statistical theory of synaptic tagging and capture. Network: Computation in Neural Systems. 2014;25:97–115.

[8] Moncada D, Viola H. Induction of long-term memory by exposure to novelty requires protein synthesis: evidence for a behavioral tagging. The Journal of Neuroscience. 2007;27:7476–7481.

bioRxiv preprint doi: https://doi.org/10.1101/072678; this version posted November 3, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[9] Merhav M, Rosenblum K. Facilitation of taste memory acquisition by experiencing previous novel taste is protein-synthesis dependent. Learning & Memory. 2008;15:501–507.

[10] Wang SH, Redondo RL, Morris RG. Relevance of synaptic tagging and capture to the persistence of long-term potentiation and everyday spatial memory. Proceedings of the National Academy of Sciences. 2010;107:19537–19542.

[11] Duncan K, Sadanand A, Davachi L. Memory's penumbra: Episodic memory decisions induce lingering mnemonic biases. Science. 2012;337:485–487.

[12] Gershman SJ. On the blessing of abstraction. The Quarterly Journal of Experimental Psychology. 2016;p. 1–5.

[13] Behrens TE, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. Nature Neuroscience. 2007;10:1214–1221.

[14] Wilson RC, Nassar MR, Gold JI. Bayesian online learning of the hazard rate in change-point problems. Neural Computation. 2010;22:2452–2476.

[15] Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI. Rational regulation of learning dynamics by pupil-linked arousal systems. Nature Neuroscience. 2012;15:1040–1046.

[16] Roesch MR, Esber GR, Li J, Daw ND, Schoenbaum G. Surprise! Neural correlates of Pearce–Hall and Rescorla–Wagner coexist within the brain. European Journal of Neuroscience. 2012;35:1190–1200.

[17] Pearce JM, Hall G. A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychological Review. 1980;87:532–552.

[18] Moncada D, Ballarini F, Martinez MC, Frey JU, Viola H. Identification of transmitter systems and learning tag molecules involved in behavioral tagging during memory formation. Proceedings of the National Academy of Sciences. 2011;108:12931–12936.

[19] Angela JY, Dayan P. Uncertainty, neuromodulation, and attention. Neuron. 2005;(4):681–692.

[20] Aston-Jones G, Cohen JD. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annual Review of Neuroscience. 2005;28:403–450.

[21] Anderson BD, Moore JB. Optimal Filtering. Prentice-Hall; 1979.