Reward prediction error modulates sustained attention

Juliana E. Trach
Department of Psychology
Yale University
New Haven, CT, 06520
juliana.trach@yale.edu

Jed Burde
Department of Psychology
Yale University
New Haven, CT, 06520
jed.burde@yale.edu

Megan T. deBettencourt
Department of Psychology
University of Chicago
Chicago, IL,60637
debetten@uchicago.edu

Angela Radulescu
Center for Data Science
New York University
New York, NY, 10011
angela.radulescu@nyu.edu

Samuel D. McDougle
Department of Psychology
Yale University
New Haven, CT, 06520
samuel.mcdougle@yale.edu

Abstract

Attention and reinforcement learning (RL) are intertwined. While previous work has primarily focused on how your attentional state impacts and shapes RL, how the dynamics of learning might impact your attentional state on a moment-to-moment basis is an open question. Here, we leverage reinforcement learning theory to investigate the moment-to-moment influence of rewards and reward prediction errors on sustained attention. Specifically, we ask how trial-by-trial reward prediction errors might affect ongoing attentional vigilance. Using a task that simultaneously queried people's sustained attention and RL performance, we demonstrate that attentional state is influenced by the magnitude and valence (positive or negative) of recent reward prediction errors. This finding highlights the influence of RL computations on one's attentional state, and provides preliminary evidence for a potential role of the dopaminergic system in meditating the relationship between learning and attentional control.

Keywords: vigilance, sustained attention, reward prediction error

Acknowledgements

We would like to thank the members of the ACT lab at Yale university for productive discussions about this project. J.E.T was funded by the NSF's Graduate Research Fellowship Program.

1 Introduction

The interplay between attention and learning is a crucial aspect of human behavior. Consider learning to play a new song on piano: If you want to play the new piece well, you must be able to sustain your attention as you read and practice the notes. However, inevitably, your attentional state will naturally fluctuate over time [1], waxing and waning between higher and lower attentional states. One possibility is that sustaining a vigilant attention state better will improve learning. But is this interaction a one-way street? That is, can the moment-to-moment dynamics of learning also affect the dynamics of sustained attention?

Previous work indicates that reward and sustained attention do interact such that offering rewards or incentives to participants can substantially boost performance on a sustained attention task [2–5]. For example, Esterman and colleagues [6] found that participants who were incentivized with money or a shortened task duration were more accurate and consistent in their performance on a continuous performance task that taxed sustained attention. While work in this vein broadly suggests that reward can influence sustained attention, it compares aggregate performance between rewarded versus unrewarded blocks of trials, pointing to general motivation as the key factor. This leaves open critical computational questions concerning the interaction of RL and sustained attention in real time: How might trial-by-trial learning outcomes (rewards, prediction errors) modulate sustained attention on a moment-to-moment basis? Do the dynamics of RL computations have a lawful relationship to ongoing attentional vigilance?

In the following study, we integrate a continuous performance task with a probabilistic RL task. The design of our experiment was chosen to assess the impact of core constructs of RL – reward and reward prediction error (RPE) – on moment-to-moment sustained attention. We used RL models to explore how computations in the brain's RL system may influence sustained attention. We tested two potential lawful relations between RPE and sustained attention. One possibility is that especially large unsigned RPEs boost sustained attention performance. In other words, surprise – unexpected rewards (positive RPEs) or unexpected omissions of rewards (negative RPEs) – could boost sustained attention, consistent with the role of surprise in boosting memory encoding [7]. Alternatively, the magnitude and valence (positive or negative) of RPEs could jointly affect sustained attention, such that positive RPEs increase sustained attention relative to negative RPEs. Some straightforward null hypotheses are that RPEs and sustained attention do not dynamically interact, or that RPEs act to distract subjects, perhaps leading to an inverted relationship between (signed or unsigned) RPEs and sustained attention. Our computational approach allows us to test all of the above hypotheses. Behavioral evidence of interactions between RPEs and sustained attention can inspire future investigations into the neural mechanisms that mediate the relationship between RL processes and attention.

2 Hybrid sustained attention/RL task

Thirty participants (N = 19 female; mean age = 20, range = 18-21) were recruited through the subject pool at Yale University and took part in the study for course credit (1 credit/hour).

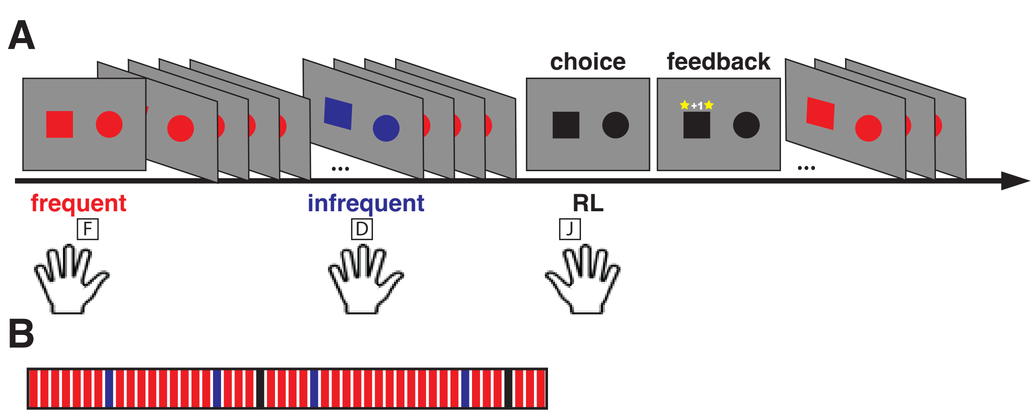

Figure 1: A) Task schematic, depicting frequent and infrequent sustained attention trials, as well as RL trials (choice and feedback). The left hand was used for attention trials and the right hand for RL trials. B) Schematic of trial sequence. Frequent sustained attention trials in red, infrequent sustained attention trials in blue, and RL trials in black.

The task was a combination of a sustained attention task [8] and a probabilistic RL task (to elicit prediction errors; Figure 1A). During a sustained attention task trial, participants saw two adjacent shapes on the screen that were either both orange or both blue. Participants used their left hand to indicate the color of the shapes and were instructed which key corresponded to orange and which to blue at the onset of the task (this assignment was counterbalanced across participants). Each trial was 800ms in duration, regardless of whether the participant responded sooner than that. For this task, on 90% of the trials the shapes were one of the two colors ("frequent" trials) and on the remaining 10% of trials they were the other color ("infrequent" trials). Thus, participants were responding with one button press on the vast majority of trials but occasionally had to inhibit this frequent response to respond correctly on infrequent trials. Participants were not explicitly told about the imbalance in the color frequencies. Following previous work [8, 9], we used accuracy on the infrequent trials to operationalize sustained attention.

Sustained attention trials occurred in pseudorandomized blocks of 17-27 consecutive trials, and the length of each block was not predictable. At the end of a given block, participants would be presented with an RL trial (a probabilistic two-arm bandit task; Figure 1). Here, the shapes would turn black and participants used their right hand to choose one of the two shapes (Figure 1A). There was no cue dissociating the trial types other than the change of stimulus color to black. On RL trials, participants had 1.5 s to make a choice. After they made their response, they received feedback on whether or not they received a reward on that trial (+1 or +0). One shape was associated with an 80% chance of reward and the other was associated with a 20% chance of reward. There were 100 RL trials in total (and thus 100 blocks of attention trials), and the reward probabilities associated with each shape were reversed three times, after 25, 50, and 75 trials. We inserted these reversals so that participants had to continually update the value of the two shapes throughout the task. The task started with a short practice block for each individual task, and then two short blocks of the two tasks combined. Participants then began the main task, which lasted approximately 50 minutes.

3 Computational modeling

We were interested in examining the moment-to-moment relationship between sustained attention and RL computations. To that end, we fit a simple RL model to participants' choices in the probabilistic learning task [10]. We used a standard RL model:

$$Q_{t+1}(s) = Q_t(s) + \alpha\delta_t \tag{1}$$

$$p(s) = \frac{e^{Q(s)\beta}}{\sum_i e^{Q(s_i)\beta}} \tag{2}$$

Where Q represents the value of stimulus s on trial t, δ represents the reward prediction error ($reward_t$–$Q_t(s)$), α represents the learning rate, and β the softmax temperature. During fitting, α was constrained on [0,1] and beta on [0,50], and a Gamma (2,3) prior distribution was used to discourage extreme values of beta (following [11]). We fit two variants of this model, one where a single alpha was used for all trials, and another that allowed asymmetric learning rates for non-rewarded versus rewarded trials. We used the MATLAB function *fmincon* to find parameter values that maximized the log posterior probability of participants' choice data given the model. Fitting runs were conducted 100 times for each model to avoid local minima during optimization, using different randomized starting parameter values over each iteration. The resulting best fit model was used in all further analyses. Model fit quality was evaluated using the Akaike information criteria [12].

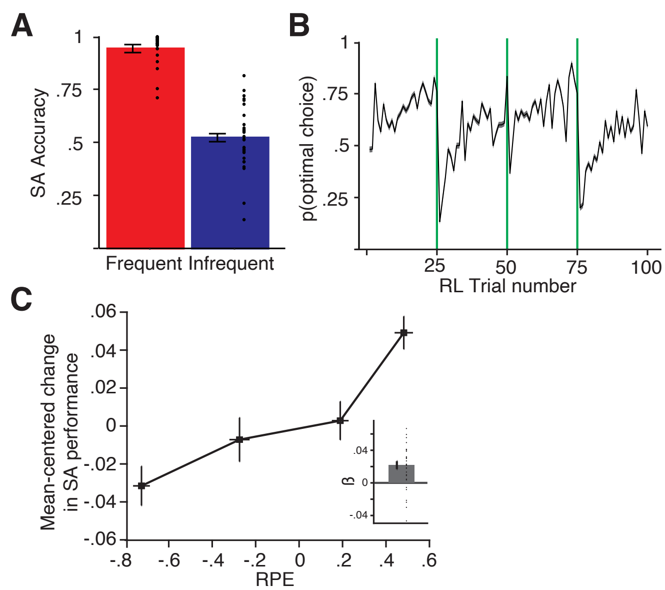

Figure 2: A) Accuracy on frequent and infrequent sustained attention trials. B) learning curve for RL task. Green lines indicate value reversal points. C) Binned RPE quantiles and mean-centered sustained attention performance (i.e., changes in average accuracy on infrequent sustained attention trials following RPEs of various sizes). Inset: regression betas reflecting the impact of RPEs on subsequent sustained attention accuracy. Error bars = 1 SEM.

4 Results

Participants exhibited typical performance on the sustained attention task, being significantly more accurate on frequent trials (M = 94%, SD = 7.6%) as compared to infrequent trials (M = 52%, SD = 15%; $t(29) = 15.71$, $p < .001$; Figure 2A; [8, 9]). In addition, on RL trials, participants learned to select the stimulus most likely to reward them throughout the task (Figure 2B).

Our primary interest was assessing whether RPE modulated sustained attention performance. To do this, we first fit an RL model to each participant's choice data and obtained participant-specific learning rates and temperature parameters (see Methods). We could then extract trial-by-trial RPEs, yielding 100 such values for each subject. Both model variants we tested fit the data well, though we observed a slight advantage for the variant with asymmetric learning rates (summed AIC for single learning rate model = 3407, summed AIC for two learning rate model = 3360; t-test on difference in AIC values; $t(29)$ = -2.12, p= .043). We note that the key results described below were comparable when using the worse-fitting single-rate model.

We performed a linear regression analysis that included all modeled trial-by-trial RPEs as a predictor variable and the mean performance on infrequent sustained attention trials in the subsequent blocks as the dependent variable. Across subjects, we observed a significant effect of RPE on sustained attention (Figure 2C inset; B = 0.026, one sample t-test: $t(29)$ = 4.80, $p < .001$), providing preliminary evidence that signed RPE modulates sustained attention.

To further examine the relationship between RPE and sustained attention, and dissociate it from more basic effects of reward, we binned RPEs into four quantiles (large, negative RPEs, small negative RPEs, small positive RPEs, large positive RPEs) for each subject and compared average sustained attention changes (i.e., relative to the mean sustained attention performance) on the attention task blocks that followed those RPEs. We performed paired t-tests on sustained attention performance between bins of RPEs (Bonferroni correction: α = .05/6 = .0083). The results of this analysis are depicted in Figure 2C. Overall, we found that participants were significantly more accurate on the sustained attention task after large, positive RPEs ($ts(29) > 3.01$, $ps < .0055$), as compared to small positive or negative RPEs. While we also found that participants were more accurate on the attention task after small positive RPEs versus large negative RPEs ($t(28)$ = 2.86, p = .0079). Contrasts between the two middle bins ($t(28)$ = 0.47, p = .64) or the two negative bins ($t(29)$ = 1.3, p = .20) were not significant. We note here that infrequent sustained attention trials never occurred immediately following RL outcomes – at least three frequent attention trials intervened after RL trials before any given infrequent trial within a block. This precludes a localized post-error slowing interpretation of our results. Taken together, these results provide evidence for the hypothesis that signed RPEs influence sustained attention performance, particularly for positive RPEs.

To investigate a more global relationship between RL and sustained attention, we also correlated modeled RL learning rates with average sustained attention performance. We found that sustained attention performance was significantly correlated with learning rates for negative RPE trials (Spearman correlation: ρ = .47, p = .010), and marginally correlated with learning rates for positive RPE trials (Spearman correlation: ρ = .35, p = .055).

Finally, exploratory hypotheses initially included additional predictions of bidirectional effects of sustained attention on RL. That is, we also hypothesized that sustained attention might influence ongoing RL, whereby elevated sustained attention predicts better choice and learning behavior on RL trials. We tested this hypothesis with a logistic regression analysis with accuracy on infrequent attention trials as the predictor variable and accuracy on the subsequent RL trials (operationalized as choosing the shape associated with 80% chance of reward) as the dependent variable. We did not observe a significant effect in this direction (B = .12, $t(29)$ = 0.98, p = .333). Thus, at least in the context of our task, the influence of RL on sustained attention was robust while effects in the opposite direction were not detected.

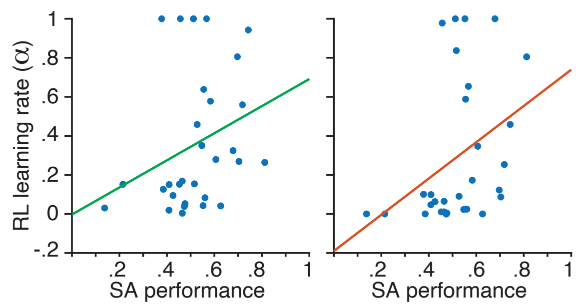

Figure 3: Correlation of learning rate and sustained attention performance. Left: learning rate for rewarded trials, Right: learning rate for unrewarded trials.

5 Discussion and Conclusions

In this study, we sought to investigate the interplay of RL and sustained attention. Overall, we found evidence that signed RPE modulates sustained attention performance in a near-linear manner, such that increasing signed RPEs led to concomitant increases in the agent's attentional state.

This result builds on previous work suggesting that performance-based incentives (rewards) improve sustained attention performance (e.g., [6]). This previous work primarily contributes to debates about theoretical causes of sustained attention lapses. Specifically, the fact that reward can reduce lapses in sustained attention appears to provide evidence against pure "resource" theories of sustained attention (i.e., that lapses in sustained attention are caused by the depletion of limited attentional resources), and is more aligned with "underload" models of sustained attention (i.e., that boredom or lack of motivation causes lapses in sustained attention [6]). We build on this work, showing how reward might boost sustained attention on a moment-to-moment timescale. Critically, our result makes a useful advance by linking sustained attention to reward prediction error, not just reward, providing a more tractable computational explanation.

The fact that signed RPE, rather than unsigned RPE, was related to sustained attention provides initial clues concerning the neural processes that might mediate this relationship. In the context of RPE and memory, effects that correlate with unsigned RPE are thought to be modulated by norepinephrine [13], whereas effects related to signed RPE are thought to be modulated by dopamine [14]. Our results therefore indirectly suggest that the increased phasic dopamine underlying RPEs could boost sustained attention. Of course, the current results do not preclude the involvement of other neurotransmitters in linking attention and RL, such as norepinephrine and acetylcholine. Future work using pharmacological tools is necessary to clarify how each of these systems might interact.

While this work contributes to a mechanistic understanding of the impact of reward on sustained attention, there are a number of limitations to be addressed in the future. First, our results suggest that positive RPE can cause a global boost in sustained attention, considering that the sustained attention and RL tasks used here were unrelated. That is, the RPEs that appeared to impact sustained attention were not directly relevant to the actual sustained attention task. Future work should address whether boosts in sustained attention after RPEs influence any task a subject is performing, or if these effects are more constrained. Second, we curiously did not detect an effect of sustained attention performance on RL. This result is somewhat surprising given the context of previous work on RL and selective attention demonstrating that selective attention significantly impacts both choice and updating processes in a similar RL task [11, 15, 16]. It is possible that the structure of our task, or the context changes between sustained attention and RL trials, attenuated effects of sustained attention on RL. Alternatively, the RL system may be robust to fluctuations in sustained attention.

Overall, our findings demonstrate a tight link between reinforcement learning and moment-to-moment attentional vigilance. This link might be mediated by dopamine [14, 17], the currency of the RL system. How exactly changes in phasic dopamine could induce more tonic effects on sustained attention, and the neural and computational processes supporting this interaction, make for exciting future directions.

References

1. Esterman, M. & Rothlein, D. Models of sustained attention. *Current Opinion in Psychology. Attention & Perception* **29,** 174–180 (2019).
2. Esterman, M. *et al.* Anticipation of Monetary Reward Can Attenuate the Vigilance Decrement. *PLOS ONE* **11.** Publisher: Public Library of Science, e0159741 (2016).
3. Robison, M. K., Unsworth, N. & Brewer, G. A. Examining the effects of goal-setting, feedback, and incentives on sustained attention. *Journal of Experimental Psychology: Human Perception and Performance* **47,** 869–891 (2021).
4. Esterman, M., Poole, V., Liu, G. & DeGutis, J. Modulating Reward Induces Differential Neurocognitive Approaches to Sustained Attention. *Cerebral Cortex* **27,** 4022–4032 (2017).
5. Esterman, M., Reagan, A., Liu, G., Turner, C. & DeGutis, J. Reward Reveals Dissociable Aspects of Sustained Attention. *Journal of experimental psychology. General* **143** (2014).
6. Esterman, M., Rosenberg, M. D. & Noonan, S. K. Intrinsic Fluctuations in Sustained Attention and Distractor Processing. *Journal of Neuroscience* **34,** 1724–1730 (2014).
7. Rouhani, N., Norman, K. A. & Niv, Y. Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **44,** 1430–1443 (2018).
8. deBettencourt, M. T., Keene, P. A., Awh, E. & Vogel, E. K. Real-time triggering reveals concurrent lapses of attention and working memory. *Nature Human Behaviour* **3.** Publisher: Springer US (2019).
9. deBettencourt, M. T., Norman, K. A. & Turk-Browne, N. B. Forgetting from lapses of sustained attention. *Psychonomic Bulletin & Review* **25,** 605–611 (2018).
10. Rescorla, R. & Wagner, A. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory,* 64–99 (1972).
11. Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V. & Niv, Y. Dynamic Interaction between Reinforcement Learning and Attention in Multidimensional Environments. *Neuron* **93.** Publisher: Cell Press, 451–463 (2017).
12. Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19.** Conference Name: IEEE Transactions on Automatic Control, 716–723 (1974).
13. Sara, S. J. The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience* **10.** Number: 3 Publisher: Nature Publishing Group, 211–223 (2009).
14. Schultz, W., Dayan, P. & Montague, P. R. A Neural Substrate of Prediction and Reward. *Science* **275,** 8 (1997).
15. Daniel, R., Radulescu, A. & Niv, X. Y. Intact Reinforcement Learning But Impaired Attentional Control During Multidimensional Probabilistic Learning in Older Adults (2020).
16. Niv, Y. *et al.* Reinforcement learning in multidimensional environments relies on attention mechanisms. *Journal of Neuroscience* **35,** 8145–8157 (2015).
17. Nieoullon, A. Dopamine and the regulation of cognition and attention. *Progress in Neurobiology,* 31 (2002).