AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General


ISSN: 0096-3445

2025, Vol. 154, No. 4, 1063–1079
https://doi.org/10.1037/xge0001727

# Rewards Transiently and Automatically Enhance Sustained Attention

Juliana E. Trach[1], Megan T. deBettencourt[2], Angela Radulescu[3], and Samuel D. McDougle[1, 4]
[1] Department of Psychology, Yale University
[2] Department of Psychology, University of Chicago
[3] Center for Computational Psychiatry, Icahn School of Medicine at Mount Sinai
[4] Wu Tsai Institute, Yale University

Our ability to maintain a consistent attentional state is essential to many aspects of daily life. Still, despite our best efforts, attention naturally fluctuates between more and less vigilant states. Previous work has shown that offering performance-based rewards or incentives can help to buffer against attentional lapses. However, such work is generally focused on long timescales and, critically, does not dissociate between task-based motivation (i.e., where reward is contingent on attention performance) versus more generic motivation or arousal accounts of reward effects. Here, we investigated the influence of reward feedback on attentional vigilance during a simultaneous sustained attention and reinforcement learning (RL) task. Crucially, rewards were tied only to the RL task rather than to attentional performance. We assessed the impact of two core components of RL—reward and surprise—on short-term fluctuations in attentional vigilance. In two experiments ($N = 161$), we demonstrated that intermittent, attention-independent rewards transiently boosted vigilance on a timescale of seconds. We did not find consistent evidence that surprises modulated vigilance. In a third experiment ($N = 135$), we observed that even passively received rewards elicit transient boosts in sustained attention. Together, these findings suggest that rewards transiently buffer against attentional lapses to improve vigilance, likely through generic increases in arousal or motivation. These results point to a fundamental relationship between reward and sustained attention.

***Public Significance Statement***

The waxing and waning of attention is a common experience in daily life. Fluctuations in attention have consequences for cognition and behavior—maintaining a focused attentional state can facilitate learning or support efficient action, whereas lapses in attention can lead to forgetting or errors. Understanding what factors impact the stability of sustained attention has far-reaching implications, from classrooms to operating rooms. In three behavioral studies, we investigate how reward feedback interacts with fluctuations in sustained attention over short timescales. Broadly, we find that sustained attention was transiently enhanced after receiving positive, as compared to negative, feedback, even when the feedback was not contingent on attentional performance. This work highlights one factor that can be leveraged to support sustained attention.

*Keywords:* sustained attention, reinforcement learning, reward, vigilance, cognitive control

*Supplemental materials:* https://doi.org/10.1037/xge0001727.supp



This article was published Online First January 20, 2025.

Joo-Hyun Song served as action editor.

Juliana E. Trach https://orcid.org/0000-0002-3429-3839

Data and code for analysis are available on GitHub (https://github.com/jetrach/SARL_2024). Preprint is available at https://psyarxiv.com/c8zq7. Megan T. deBettencourt is an employee of and has equity in Ruby NeuroTech. The design and results of Experiment 1 were presented in part at the Reinforcement Learning and Decision Making (RLDM) Conference at Brown University in June of 2022 and at the Society for Neuroscience annual conference in San Diego, California, in November 2022 as poster presentations. RLDM additionally included a two-page abstract that is publicly available. All three experiments were presented as a poster at the annual Neural Control of Movement Conference in Dubrovnik, Croatia.

The authors have no known conflicts of interest to report. Juliana E. Trach is supported by the National Science Foundation (Grant GRFP). The authors thank Robb Rutledge and the members of the ACT Lab for helpful discussions about this project. Additionally, they are grateful to Jed Burde and Katherine Chou for their help with data collection for this project.

Juliana E. Trach played a lead role in conceptualization, data curation, formal analysis, investigation, visualization, writing–original draft, and writing–review and editing. Megan T. deBettencourt played a supporting role in formal analysis, writing–original draft, and writing–review and editing and an equal role in conceptualization and methodology. Angela Radulescu played a supporting role in supervision, writing–original draft, and writing–review and editing and an equal role in conceptualization. Samuel D. McDougle played a lead role in conceptualization and funding acquisition, a supporting role in data curation, and an equal role in formal analysis, supervision, visualization, writing–original draft, and writing–review and editing.

Correspondence concerning this article should be addressed to Juliana E. Trach, Department of Psychology, Yale University, 100 College Street, New Haven, CT 06510, United States. Email: Juliana.trach@yale.edu

Imagine practicing a difficult new piece of music on the piano. As you monitor your performance, your attentional vigilance will naturally fluctuate between higher (more vigilant) and lower (less vigilant) states (e.g., Esterman & Rothlein, 2019). Attentional state has important consequences for other cognitive processes, including learning, memory encoding, and memory retrieval (deBettencourt et al., 2018; A. Decker, Dubois, et al., 2023; A. L. Decker, Duncan, & Finn, 2023; Madore et al., 2020). But how attentional vigilance is itself shaped by events in the world is not fully understood. We focus here on one such class of events—reward feedback—asking if the experience of reward (e.g., the sound of playing a musical phrase perfectly vs. poorly) might transiently influence one's attentional state, even when reward is not directly tied to attentional vigilance.

Multiple previous experiments have shown that attention-contingent rewards can affect vigilance: Offering incentives to participants based on their performance in vigilance tasks reliably boosts vigilance performance (Bergum & Lehr, 1964; Esterman et al., 2014, 2016, 2017; Massar et al., 2016; Robison et al., 2021). Generally, such work uses psychometric vigilance tasks (Ariga & Lleras, 2011; Basner & Dinges, 2011; Lim & Dinges, 2008) or variants of continuous performance tasks (deBettencourt et al., 2018; Esterman et al., 2013, 2014, 2016; Esterman & Rothlein, 2019; Manly et al., 1999) to behaviorally track sustained attention. In one example, Esterman et al. (2014) found that participants who were incentivized with money (or a shortened task duration) were more accurate and showed more consistent response times on a continuous performance task designed to measure sustained attention. This type of effect has since been replicated (Esterman et al., 2016; Robison et al., 2021), providing substantive evidence that offering performance-based rewards can boost sustained attention performance.

Such links between reward and sustained attention have important theoretical implications. The positive role of reward in sustained attention is consistent with "cost–benefit" models of sustained attention (Kool & Botvinick, 2014; Kool et al., 2017; Kurzban et al., 2013; Thomson et al., 2015). In these models, attention wanes when the cost of maintaining a consistent attentional state in a task—namely, the opportunity cost of *not* attending to other things (e.g., mind-wandering, other potential tasks) or the effort cost of exerting control to stay attentive—outweighs the benefits (Kool et al., 2017; Kurzban et al., 2013). These models are often contrasted with "resource" theories of sustained attention, which posit that lapses arise from the depletion of a limited cognitive resource that is needed to maintain vigilance (Helton & Russell, 2011, 2013; Helton & Warm, 2008; Muraven & Baumeister, 2000; Warm et al., 2008). Resource theories of sustained attention do not clearly predict reward-related enhancements of sustained attention, as external rewards should not be able to alter the reserves of a finite cognitive resource. However, cost–benefit models of sustained attention assume that performance-based incentives reduce lapses because such incentives directly increase the value of sustaining attention, thus motivating participants to exert more control. Therefore, the current view of the link between reward and sustained attention is that rewards can alter people's estimate of the value of sustaining attention and thereby improve performance.

However, the finding that directly rewarding attention performance boosts performance is arguably somewhat inevitable; indeed, rewarding most behaviors sharpens or improves them. An additional explanation for positive effects of reward on sustained attention is that reward induces a more generic motivation or arousal effect, one that carries over into attentional processes. That is, when we receive a reward from the environment, it may induce a general, transient boost to attention. This idea fits with several findings from the neuroscience literature, including studies showing that neural signals correlated with reward can induce rapid downstream effects on attentional state and drive activity in key attention networks in the brain (Anderson et al., 2016; Sara, 2009; Zhang et al., 2023). In particular, noradrenergic activity in the locus coeruleus, which is associated with attention and arousal (Unsworth & Robison, 2017) and effort mobilization (Unsworth et al., 2022), is also involved in reward processing (Sara, 2009). These findings point to potential mechanisms by which reward automatically influences sustained attention in a manner that goes beyond the more value-based accounts prevalent in the literature. While this generic effect is not mutually exclusive of a "top-down" cost–benefit explanation, it remains to be thoroughly tested.

An additional limitation of previous work is that the effects of reward on sustained attention have been largely reported at the task "block" level, which spans dozens of minutes or more (Esterman et al., 2014, 2016, 2017; Robison et al., 2021). Often, this long-timescale approach is used to model gradual decrements in vigilance over time (e.g., cognitive fatigue). Sustained attention, however, also fluctuates at short timescales concurrently with overall vigilance decrements (Esterman & Rothlein, 2019). Studying a more transient relationship between reward and sustained attention necessitates capturing sustained attention performance on shorter timescales.

Here, we designed a task where rewards were not contingent on attention performance and were received intermittently during a sustained attention task. We developed a novel hybrid task that integrated a continuous performance attention task with an instrumental reinforcement learning (RL) task. During this hybrid task, participants performed blocks of a sustained attention task interleaved with individual trials of an RL task where they could receive reward feedback, and where reward feedback was related only to the RL task rather than performance on the attention task. Our main analyses investigated the impact of reward on people's performance during intervening blocks of attention trials. The task was also designed to assess the impact of both reward prediction error and surprise—two core constructs in RL—on moment-by-moment sustained attention.

We hypothesized two potential interactions between reward feedback during reinforcement learning and sustained attention. One possibility is that reward feedback in our hybrid task would have no impact on subsequent sustained attention performance. This pattern of results would suggest that reward-related boosts in sustained attention documented in the literature are driven by the increased value of sustained attention when it is directly yoked to reward. In other words, this finding would provide evidence that rewards must directly and saliently increase the "payoff" of sustaining attention to elicit a boost in performance. Alternatively, reward could influence sustained attention in a global manner, where rewards boost motivation and arousal and thus obligatorily impact sustained attention. Using our modeling framework, additional hypotheses were tested with respect to surprise (i.e., model-derived unsigned reward prediction errors) and attention. In this way, we could investigate the relationship between reward feedback and attentional vigilance and shed light on the cognitive processes that link reward processing and attention at the scale of seconds.

In Experiments 1 and 2, we asked whether reward feedback during RL influences one's attentional state, and we found robust evidence that rewards unrelated to the attention task transiently enhanced vigilance over a matter of seconds. In Experiment 3, we generalized this finding to passive (i.e., noninstrumental) rewards and again observed a transient facilitation of attentional state. Taken together, these results suggest that the experience of reward (relative to reward omission) facilitates sustained attention even when those rewards do not directly set the value of attentional vigilance. We suggest that these effects reflect a fundamental relationship between reward and attention, where the mere experience of reward rapidly and globally heightens vigilance.

## Experiment 1

### Method

#### *Participants*

Thirty participants were recruited through the subject pool at Yale University and took part in the study for course credit. Demographic information was collected on a tablet during the consenting process. Participants were asked to report their age, sex, and handedness after reviewing the consent form. They were given the options "male," "female," "intersex," or "prefer not to say" to report sex. We did not collect information about race.

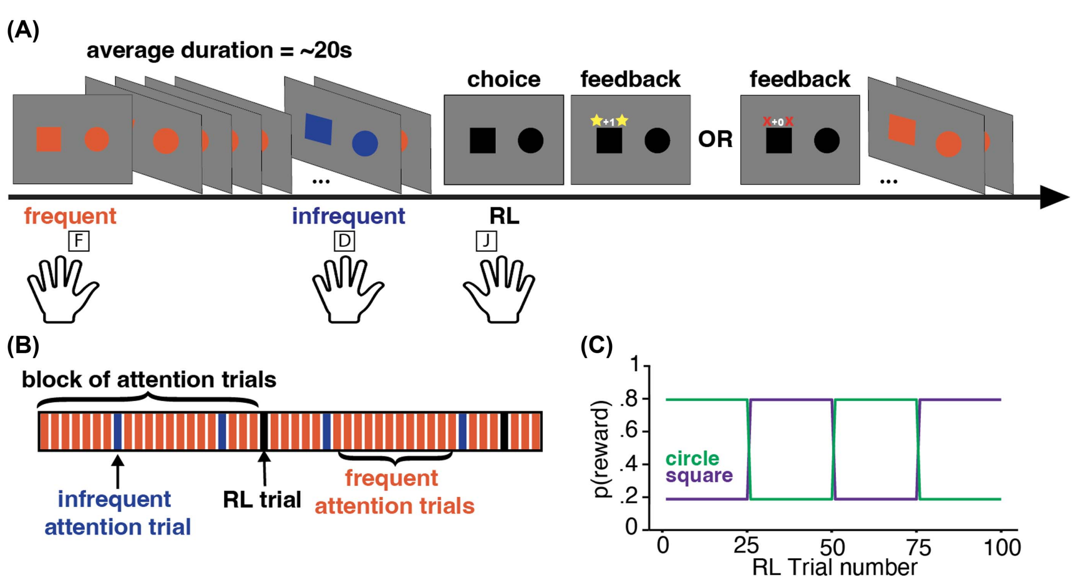
Our sample size was selected a priori based on previous work with hybrid sustained attention–memory tasks (deBettencourt et al., 2018, 2019) and adjusted to account for possible performance-based exclusions. We excluded one participant who responded to fewer than 75% of the RL trials, and an additional participant whose distribution of reward prediction errors did not allow us to do a key analysis (see below), leaving us with a final sample of 28 participants ($N = 18$ reported female, $N = 10$ reported male; $M_{\text{age}} = 19.71$, $SD_{\text{age}} = 2.51$). We planned to exclude participants who had less than 75% accuracy on frequent trials or who made the same motor response (i.e., chose the same shape) on >90% of the RL trials, but no participants met these exclusion criteria.

#### *Procedure*

We interleaved blocks of a sustained attention task (deBettencourt et al., 2019) with individual trials of a probabilistic RL task to administer intermittent rewards during the learning of stimulus values (Figure 1A). The task was programmed using jsPsych (de Leeuw, 2015) and was performed on a Lenovo Ideapad 5 (Ubuntu 22.04).

On sustained attention task trials, participants saw two adjacent shapes on the screen that were either both orange or both blue with a black fixation cross between them. Each shape was presented on the same side throughout the task and counterbalanced across participants. Participants used their left hand to indicate the color of the shapes and were instructed about which action (key press) corresponded to orange and which to blue at the onset of the task (this assignment was counterbalanced across participants). Once the participant made their response, the fixation cross turned white to indicate that their response was registered; however, the stimuli remained on the screen for 800 ms for each trial, regardless of

**Figure 1**
*Experiment 1: Task Design*




*Note.* (A) Task schematic depicting frequent and infrequent sustained attention trials, as well as RL trials (choice and feedback). Each attention trial was displayed for 800 ms, yielding blocks of approximately 20 s on average. The left hand was used for attention trials and the right hand for RL trials. (B) Schematic of a trial sequence. Frequent sustained attention trials in orange (in print, light gray), infrequent sustained attention trials in blue (in print, dark gray), and RL trials in black. Infrequent trials never occurred immediately after RL trials. (C) Example reward schedule for shaped on RL trials. Shapes were either associated with a .2 or a .8 probability of reward. Reward probabilities were reversed three times during the task. RL = reinforcement learning. See the online article for the color version of this figure.

whether the participant responded sooner than that cutoff. Thus, each attention trial was always 800 ms in duration though participants generally responded before the deadline. Trials in which they did not make a response (4.8% of trials, ±7.2) were excluded from the analysis.

Following the typical design of this variant of a sustained attention task (deBettencourt et al., 2019), we varied the frequency of the two colors in order to create "frequent" and "infrequent" (or oddball) trials: one of the two stimulus colors occurred on 90% of the attention trials (frequent trials), and the other appeared on the remaining 10% of trials (infrequent trials). Thus, participants were tasked with making one response on most trials and occasionally had to switch away from this prepotent response to respond correctly on the infrequent trials. Participants were not explicitly informed of the imbalance in the color frequencies.

Sustained attention trials occurred in pseudorandomized blocks of 17–52 consecutive trials, and the length of each block was not predictable (Figure 1B; average block length = 25 trials [22.5 s], ±0.07). At the end of a given block, participants would be presented with a RL trial. We used a probabilistic "two-arm bandit task" to operationalize RL. To signal an RL trial, the shapes would turn black, cueing the participant to now use their right hand to select one of the two shapes (Figure 1A). Importantly, participants used different hands to respond to the attention (left hand) versus RL (right hand) trials. This design ensured that RL choices were not simply driven by participants repeating the prepotent response from the attention trials. 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 (shown for 750 ms) on whether they received a reward on that trial. If they responded too late, they were given feedback that read "Too Slow," and these trials were excluded from analysis. On rewarded trials, two yellow stars and "+1" appeared above the shape that the participant chose (Figure 1A). On unrewarded trials, two red Xs and "+0" appeared above the shape that the participant chose. One shape was associated with an 80% probability of reward, and the other was associated with a 20% probability of reward on individual trials. We implemented this reward structure by generating a random number greater than 0 and less than or equal to 1 once the participant had chosen a shape. On trials where the participant chose the shape with a 20% chance of reward, the random number had to be less than or equal to .2 for them to get positive reward feedback. On trials where the participant chose the shape with an 80% chance of reward, the random number had to be less than or equal to .8 for them to receive positive reward feedback.

There were 100 RL trials in total (and thus 100 intervening blocks of attention trials), and the reward probabilities associated with each shape were reversed three times, after 25, 50, and 75 trials (Figure 1C). We inserted these reversals so that participants had to continually update the value of the two shapes throughout the task, rather than simply perseverating on one of the two shapes that was initially associated with a higher reward probability. Importantly, there was no additional incentive (e.g., bonus money) associated with the feedback or overall performance on RL trials—participants were simply instructed to choose the shape that they thought was most likely to yield positive reward feedback.

Following previous work, we used both accuracy on the infrequent trials and the coefficient of variation (CV) of reaction time (RT) on frequent trials (coefficient of variation = *SD*(RT)/*M*(RT)) to operationalize sustained attention during the task (deBettencourt et al., 2018, 2019; Esterman et al., 2016). Previous work has largely used average RT or RT variability as an index of sustained attention. Here, we opted to focus on RT variability as there is some evidence that this metric is more closely linked to sustained attention (Esterman et al., 2014; Esterman & Rothlein, 2019). Higher infrequent trial accuracy and lower frequent trial RT variability (i.e., lower CV) are both indicative of a vigilant attentional state relative to low accuracy and high RT variability. We calculated accuracy and CV within each block of attention trials to quantify attentional performance for our analyses. For accuracy, we selected all infrequent trials during a given block and calculated the mean accuracy; for CV, we selected all the frequent trials during a block of attention trials and calculated CV over those RTs. In addition to these block-level analyses, we conducted post hoc finer timescale analyses on CV to further characterize interactions between reward and sustained attention. To do this, we divided each block of attention trials into equal-length early, middle, and late phases. We then calculated CV in each of these phases, yielding three CV measurements for each block of attention trials. We note here that given the rarity of infrequent trials, this finer grained analysis was not well-suited to the accuracy metric.

The task started with a short practice block for each individual task, and then two short blocks with the two types of trials interleaved. Participants then began the main task, which lasted approximately 50 min. We pregenerated 10 trial sequences with frequent, infrequent, and RL trials. Participants were randomly assigned to one of the 10 trial sequences. These sequences were constructed in a block-wise manner to ensure that each block of trials had multiple infrequent attention trials and that block lengths were sufficiently varied. To vary block lengths such that they were not predictable to participants, we repeated a vector of all possible block lengths until it was longer than the number of task blocks. Then, we shuffled this vector and selected the appropriate number of blocks. This process allowed participants to have similar overall task length and distribution of attention trials, while still leaving the placement of RL trials sufficiently unpredictable. RL trials always occurred after frequent sustained attention trials, and blocks always began with frequent attention trials. The mean number of frequent attention trials, infrequent attention trials, and RL trials after exclusions per participant was 2,112, 241, and 96, respectively.

## Modeling

In addition to the effect of feedback valence on sustained attention, we were interested in examining the relationship between sustained attention and RL reward prediction error (RPE) computations, which allowed us to dissociate reward valence and surprise. To that end, we fit RL models to participants' choices in the probabilistic learning task. We used standard *Q*-learning models, which updated action values according to a simple δ rule (Rescorla & Wagner, 1972):

$$Q(s)_{t+1} = Q(s)_t + \alpha\delta, \tag{1}$$

$$\delta = r_t - Q(s)_t. \tag{2}$$

Here, $Q(s)_t$ reflects the expected value of stimulus $s$ on trial $t$ (i.e., the reward probability associated with stimulus $s$ at time $t$), α reflects the learning rate, and δ reflects the RPE (Equation 1). RPE is defined as the difference between the observed reward ($r$) and the expected value given the choice of stimulus $s$ (Equation 2). Thus, on a given trial, the expected value for the chosen stimulus ($Q(s)_t$) is updated based on the discrepancy between the received reward and the current expected value of the stimulus, scaled by the learning rate. If the reward feedback was greater than $Q(s)_t$, the RPE would be positive and lead to an *increase* in the expected value of that stimulus after the update. In contrast, if reward feedback was less than $Q(s)_t$, the RPE is negative and $Q(s)_{t+1}$ will *decrease* relative to $Q(s)_t$. While reward feedback is binary, RPEs can vary continuously between −1 and 1 since they are derived from the difference between the binary feedback and the incrementally updated $Q$ values that participants continuously learn.

Action selection between the two presented stimuli was modeled using the softmax function (Equation 3; Daw, 2011):

$$p(s) = \exp(\beta Q(s)) / \sum_i \exp(\beta Q(s_i)), \quad (3)$$

where β reflects the softmax inverse temperature. During fitting, α was constrained on [0, 1] and β on [0, 50], and a Gamma (2,3) prior distribution was used to discourage extreme values of β (following Leong et al., 2017).

Each participant's choice data were fit separately to obtain unique learning rates, inverse temperature parameters, and quality of fit metrics (Akaike information criterion [AIC]). We fit two variants of this model, one where a single learning rate was used for all trials and another that allowed asymmetric learning rates for unrewarded versus rewarded trials (Collins & Frank, 2014; Daw et al., 2002; Frank et al., 2007). We used the MATLAB (2022) function *fmincon* to find parameter values that maximized the log posterior probability of participants' choice data given the model. Fitting runs were conducted 200 times for each model for each participant 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 AIC (Akaike, 1974).

After model fitting, we used subject-specific fitted parameters (rather than average parameters across participants) to simulate the RPEs that participants putatively experienced on each RL trial to use in further analyses. To do this, we initialized the $Q$ value for each stimulus at 0 and then incrementally updated the $Q$ values for each stimulus based on the order of choices the participants made and the feedback that they received on each RL trial, using the model. The RPE on a given trial was calculated as the difference between the expected value of the chosen stimulus on that trial ($Q(s)_t$) and the received reward feedback (0 or 1). Trials where participants did not respond in time were excluded and no updates to $Q$ values were performed on these trials. This process yielded trial-by-trial RPEs for each participant that varied in valence (positive vs. negative) and magnitude. We used RPE magnitude to estimate trial-by-trial surprise.

We also used subject-specific parameters to simulate behavior and depict model fits in Figure 2B. To do this, we used subject-specific parameters to simulate participant choice time courses 100 times. We then calculated the mean probability of selecting the optimal choice (the stimulus that currently held a higher probability of reward) from these simulations. This process created subject-specific learning curves reflecting the probability of making the optimal choice across the length of the RL task. We took the average across these subject-specific curves to obtain the model fit plot shown in red in Figure 2B. Subject-wise model simulations (i.e., before averaging across participants) are depicted in Supplemental Figure S1.

### Analysis

Additional analyses were conducted in R (R Core Team, 2022). Our variables of interest (accuracy on infrequent trials, coefficient of variation in RT on frequent trials, and overall performance on the RL task) were submitted to two-tailed paired $t$ tests (where appropriate), and linear mixed-effects models. We used Pearson correlations to examine the relationship between variables when both variables were continuous and normally distributed and used Spearman correlations for model-derived values or where assumptions of normality were violated. Correlation coefficients were Fisher transformed before performing $t$ tests where necessary. We used Welch $t$ tests were used when equal variance was violated. We used Cohen's $d$ to quantify effect sizes for $t$ tests and report 95% confidence intervals (CI). Linear and logistic mixed-effects models were conducted and tested using the *lme4* and *lmerTest* packages in R.

We conducted primary analyses on both raw and detrended sustained attention metrics to fully examine the short-timescale interactions between RL and sustained attention. Crucially, detrending removes the well-known effect of overall worsening sustained attention performance over the course of an experimental session ("vigilance decrement"), allowing us to isolate the types of low-level fluctuations that we are interested in (e.g., A. Decker, Dubois, et al., 2023). To do this, we used the *detrend* function from the *pracma* library in R to linearly detrend infrequent trial accuracy and frequent trial CV for each subject. We used this same function to detrend finer timescale CV metrics as well.

We include tables of descriptive statistics of our main measurements of interest in the Supplemental Material to this article. In this summary analysis, we calculated split-half reliability for the CV and accuracy measures. We opted to divide the data in half by comparing these measures in odd versus even blocks, instead of comparing first half versus second half, as sustained attention performance is known to gradually deteriorate over the course of a task. To do this analysis, we calculated the average CV and accuracy (both raw and detrended) for odd and even blocks of trials separately. Then, we did a between-subjects Spearman correlation to evaluate the reliability of these metrics over the course of the task. The ρ value from that comparison is listed in the descriptive statistics tables in the Supplemental Material.

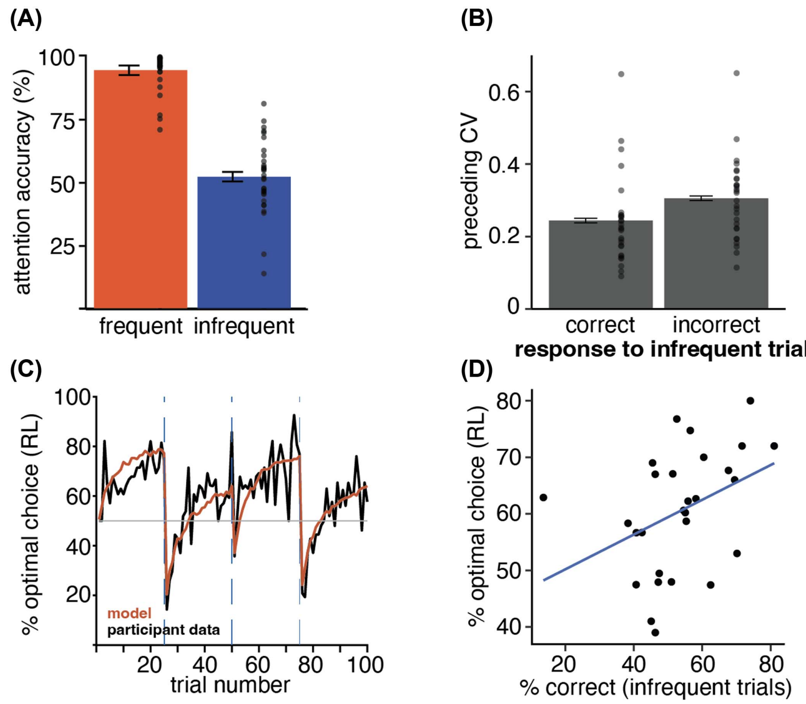
### Transparency and Openness

This study was not preregistered. We include details about sample size and participant exclusions, as well as methodological details to facilitate future replication. Data and analysis code are available at https://github.com/jetrach/SARL_2024.

## Results

Participants' performance on the sustained attention task replicated previous studies that implemented a similar task (deBettencourt et al., 2019): Participants were significantly more accurate on frequent

**Figure 2**
*Experiment 1: Sustained Attention and RL Performance*




*Note*. (A) Accuracy on frequent versus infrequent attention trials. (B) Validation analysis of CV metric of attention. CV on three frequent trials preceding correct versus incorrect responses to infrequent attention trials. (C) Learning curve for RL trials in black. Model fit is depicted in red (in print, light gray). Dashed lines mark value reversals. (D) Correlation between performance on attention trials (accuracy on infrequent trials) and performance on the RL trials (percent of RL trials where the participant chose the shape with a higher reward probability). CV = coefficient of variation; RL = reinforcement learning. See the online article for the color version of this figure.



trials ($M = 94.5\%$, $SD = 7.8\%$) than on infrequent trials, $M = 53.8\%$, $SD = 13.6\%$; contrast: $t(27) = 15.3$, $p < .001$, $d = 2.89$, 95% CI [35.4, 46.2]; Figure 2A. We used two distinct metrics to quantify sustained attention in each block of attention trials: the coefficient of variation of RTs on frequent attention trials (coefficient of variation = standard deviation of RT/mean of RT; Esterman et al., 2016) and accuracy on infrequent attention trials (deBettencourt et al., 2018, 2019). These accuracy- and RT-based sustained attention metrics were correlated, mean within-subject correlation between attention metrics = −.31 ± .15; one-sample $t$ test on Fisher transformed ρ values: $t(27) = -10.45$, $p < .001$, $d = -1.98$, 95% CI [−0.4, −0.27], suggesting that they were likely capturing similar fluctuations in sustained attention. To further validate these measures, we examined whether lapses in the RT metric of sustained attention (i.e., high CV on frequent attention trials) were associated with lapses in the accuracy metric (i.e., incorrect responses on infrequent attention trials). We compared CV over the three frequent attention trials immediately preceding correct versus incorrect responses to infrequent attention probes. If both metrics are capturing attentional state, then we expect higher CVs preceding incorrect responses to infrequent probes and lower CVs preceding correct responses (deBettencourt et al., 2019). We found that CV was indeed significantly higher preceding incorrect responses to infrequent attention trials relative to CV before correct responses, Figure 2D; paired-sample $t$ test: $t(27) = -6.87$, $p < .001$, $d = 1.3$, 95% CI [−0.08, −0.04]. This further indicates that both metrics were capturing similar fluctuations in sustained attention during the task.

Regarding RL performance, participants learned to select the stimulus most likely to reward them throughout the task on RL trials, $M = 60.6\%$ optimal choice, $SD = 10.8\%$; $t(27) = 5.14$, $p < .001$, $d = 0.97$, 95% CI [12.6, 29.4], and did not show a bias to choose one shape or action over the other across the task, paired-sample $t$ test to choose left versus right shape: $t(27) = -0.51$, $p = .614$. Figure 2B depicts participants' learning over the course of the task measured as the probability of selecting the shape associated with a higher probability of reward (for individual model fits, see Supplemental Figure S1). We note that while choosing the more rewarding stimulus 60% of the trials might appear low, the probabilistic nature of the feedback and the multiple value reversals make this a reasonable performance level.

To investigate a global relationship between RL and sustained attention performance, we first correlated participants' mean sustained attention performance (operationalized as their average accuracy on the infrequent attention trials) and mean performance on the RL task (operationalized as the total proportion of trials in which they chose the shape associated with a higher reward probability). We found that sustained attention performance was significantly correlated with RL performance as indexed by accuracy on infrequent attention trials, Figure 2C; Pearson's correlation: $t(26) = 2.14$, $R = .39$, $p = .042$, 95% CI [0.016, 0.66]. The correlation between average CV and RL performance, however, was not significant, Pearson's correlation: $t(26) = -0.91$, $R = -.17$, $p = .373$, 95% CI [−0.51, 0.21]. This coarse analysis provides initial evidence of a relationship between RL and sustained attention in our integrated task. Further, this finding suggests that participants were engaged in both tasks, rather than there being a performance trade-off in this dual-task paradigm.

Predictably, both metrics showed significant decrements over the course of the task: Accuracy on infrequent trials was negatively correlated with trial number, average within-subject Spearman's correlation: $\rho = -.14 \pm .16$, one-sample $t$ test on Fisher transformed $\rho$ values: $\mu = 0$, $t(27) = -4.45$, $p < .001$, $d = 0.84$, 95% CI [−0.21, −0.08], and CV was positively correlated with trial number, average within-subject Spearman correlation: $\rho = .16 \pm .19$, one-sample $t$ test on Fisher transformed $\rho$ values: $\mu = 0$, $t(27) = 4.52$, $p < .001$, $d = 0.85$, 95% CI [0.09, 0.25]. These results are characteristic of canonical vigilance decrements in sustained attention tasks (e.g., Esterman et al., 2013; Kurzban et al., 2013). To isolate the impact of reward on fluctuations in attention for our primary short-timescale analyses, we focus on detrended accuracy and CV metrics in all subsequent analyses (see the Method section). We note that these detrended metrics were similarly correlated, mean within-subject correlation between detrended attention metrics = $-.29 \pm .15$; one-sample $t$ test on Fisher transformed $\rho$ values: $\mu = 0$, $t(27) = -9.79$, $p < .001$, $d = 1.85$, 95% CI [−0.38, −0.26].

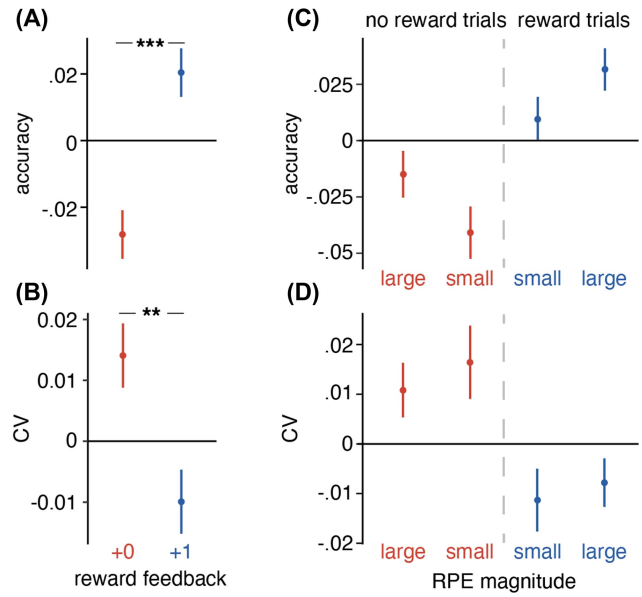
We assessed the short-timescale effects of rewards on attentional vigilance by comparing detrended attention performance after rewarded versus unrewarded RL trials. Participants were more accurate on infrequent attention trials that followed a rewarded RL trial versus an unrewarded RL trial, Figure 3A; detrended accuracy metric: $t(27) = 4.72$, $p < .001$, $d = 0.89$, mean difference = 0.049, 95% CI [0.027, 0.07]. Corroborating this effect, RT coefficients of variation on frequent trials were significantly lower (i.e., less variable) after rewarded versus unrewarded RL trials, Figure 3B; detrended CV metric: $t(27) = 3.22$, $p = .0033$, $d = 0.61$, mean difference = −0.024, 95% CI [−0.039, −0.008]. Both effects were strong when using raw attention metrics as well, raw accuracy metric: $t(27) = 4.83$, $p < .001$, $d = 0.91$, mean difference = 0.052, 95% CI [0.030, 0.075]; $t(27) = 3.1$, $p = .004$, $d = 0.59$, mean difference = −0.023, 95% CI [−0.039, −0.008], despite the presence of an overall performance decrement over the course of the task. These results demonstrate that participants were significantly more vigilant after receiving positive reward feedback, relative to neutral feedback. Moreover, both effects provide robust evidence that reward has a facilitatory effect on attentional vigilance at the single trial level, on a subminute timescale, and, crucially, even when reward is not contingent on attention performance. This finding builds on previous results showing such an effect at much longer task- or block-level timescales and with attention-contingent rewards (Esterman et al., 2014, 2016, 2017; Robison et al., 2021). Therefore, this result suggests that rewards can support sustained attention even when the rewards do not specifically and directly increase the value of attention task performance.

To examine the relationship between learning and attention and quantify surprise during the task, we fit RL models to each participant's RL task data to obtain participant-specific learning rates and inverse temperature parameters (see the Method section for modeling details). We fit two variants of an RL model (Equations 1 and 2), one in which a single learning rate was used to update value representations for all outcome types and one where separate learning rates were fit to rewarded and unrewarded outcomes (Collins & Frank, 2014; Frank et al., 2007). Both models fit the data well, though we observed a consistent advantage for the variant with asymmetric learning rates (summed AIC for single learning rate model = 3,215, summed AIC for two-learning rate model = 3,176). We thus performed our main computational analyses using the asymmetric learning rate model to approximate trial-by-trial prediction errors (we note that the key results described below were comparable when using the worse-fitting single-rate model).

The average fit of the RL model to participants' behavior in the RL task is shown in Figure 2B (with individual subject model fits in Supplemental Figure S1). Consistent with the significant relationship between behavioral measures of performance across the two tasks, RL learning rates were significantly correlated with attention task performance (positive learning rate: Spearman correlation: $\rho = .47$, $p = .012$; negative learning rate: Spearman's correlation: $\rho = .44$, $p = .021$). The inverse temperature parameter was also

**Figure 3**

*Experiment 1: Effect of Reward on Sustained Attention*




*Note.* (A) Accuracy (detrended) on infrequent attention trials following rewarded versus unrewarded reinforcement learning trials. All error bars reflect 1 standard error of the mean. (B) RT variability (measured by detrended coefficient of variation) on frequent attention trials following rewarded versus unrewarded reinforcement learning trials. (C) Accuracy on infrequent trials after large and small magnitude and positive/negative reward prediction errors (RPEs), which were computed via our modeling analysis. Unrewarded trials on left (red) and rewarded trials on right (blue). (D) RT variability (measured by the coefficient of variation) on frequent attention trials after large and small magnitude RPEs. NUnrewarded trials on left and rewarded trials on right. CV = coefficient of variation; RT = reaction time. See the online article for the color version of this figure.
** $p < .01$. *** $p < .001$.

significantly correlated with attention task performance (Spearman's correlation: $\rho = -.46$, $p = .015$). Taken together, these correlations provide model-driven evidence relating people's performance on the two interleaved tasks and indicate that participants who were better learners also tended to maintain a better attentional state than poor learners.

Next, we asked whether attention performance was affected by the magnitude of prediction errors. To extract trial-by-trial RPEs for each participant and quantify surprise, we simulated the best fitting model using each participant's observed sequence of choices and optimized model parameters. We examined the impact of RPE magnitude on sustained attention separately for rewarded and unrewarded trials because of the large effect of reward on sustained attention discussed previously. We first grouped attention blocks by whether a rewarded or unrewarded trial occurred at the start of that block. We then performed a median split on the RPEs associated with those trials. In other words, we labeled RL trials as reflecting four different outcomes: a "large" negative RPE, which reflects an unpredicted *lack* of reward; a small negative RPE, which reflects a less surprising lack of reward; a small positive RPE, which reflects a predicted reward; and a large positive RPE, which reflects a surprising reward. Although there were numerical differences, we did not find evidence that large versus small RPEs differently affected detrended accuracy on infrequent attention trials, Figure 3C; rewarded trials: $t(27) = -1.68$, $p = .104$, mean difference = 0.02, 95% CI [−0.05, 0.005]; unrewarded trials: $t(27) = 1.62$, $p = .116$, mean difference = 0.03, 95% CI [−0.007, 0.06], nor detrended CV, Figure 3D; rewarded trials: $t(27) = -0.54$, $p = .595$, mean difference = −0.002, 95% CI [−0.05, 0.005]; unrewarded trials: $t(27) = -0.65$, $p = .52$, mean difference = −0.006, 95% CI [−0.02, 0.01]. Thus, these results indicate that RPE magnitude (i.e., surprise) may not significantly impact attentional state above and beyond the robust effects of reward itself.

Finally, to further investigate the relationship between RPE and sustained attention, we used four linear mixed-effects models fit to the detrended sustained attention metrics (accuracy on infrequent trials or CV) with either trial-by-trial RPEs or with feedback valence (+1 or +0) as predictors. All models included random intercepts and slopes for each subject. We found that RPE and RPE valence were both significant predictors of sustained attention ($t$s > 3.20, $p$s ≤ .001). However, we did not find evidence that including continuous RPEs as the predictor variables improved model fit versus feedback valance alone (accuracy ~ reward: AIC = 1,378; accuracy ~ RPE: AIC = 1,384; CV ~ reward: AIC = −3,030; CV ~ RPE: AIC = −3,023), suggesting that observed boosts in sustained attention were primarily driven by the simple presence of reward. Taken together, these results suggest that reward feedback temporarily increases attentional vigilance, even when rewards are purely symbolic in nature and unconnected to attentional performance.

What is the time course of this effect within the attention task blocks? In post hoc analyses, we sought to characterize the effect of reward feedback on attentional state at a finer timescale. To do this, we divided each block of attention trials into three equal phases (early, middle, late) and calculated CV on frequent trials within each phase of each attention block. We then detrended this metric over the full length of the experiment in the same way as the block-level CVs (see the Method section). We compared the average CV in the early, middle, and late phases of the attention blocks (each representing a 5.4–8.1 s window of time within a block) following rewarded versus unrewarded RL trials, subtracting the average CV after unrewarded trials from the average CV after rewards. This analysis revealed a significant boost in sustained attention during the earliest attention trials, early phase: one-sample $t$ test: $M = 0.036$; $t(27) = 3.57$, $p < .001$, $d = 0.7$, 95% CI [0.015, 0.056], that diminished over the course of the block, middle phase: $M = 0.018$; one-sample $t$ test: $t(27) = 2.35$, $p = .026$, $d = 0.44$, 95% CI [0.002, 0.03], until attention performance was not significantly affected by rewards the end of the attention block, late phase: $M = 0.003$; one-sample $t$ test: $t(27) = 0.26$, $p = .799$, 95% CI [−0.019, 0.024]. This result suggests that reward induces a transient boost in sustained attention—on the order of several to a dozen seconds—that quickly diminishes with time.

In addition to our planned analyses on the impact of reward and surprise on attention, we conducted exploratory analyses concerning the possible effects of sustained attention on RL performance. That is, we asked whether sustained attention might affect ongoing RL behavior in our task such that elevated sustained attention in a given block predicts better choices on a subsequent RL trial. We tested this hypothesis with logistic regression mixed-effects models, with each of the detrended attention metrics as the predictor variable and accuracy on the single subsequent RL trial (operationalized as choosing the shape associated with the currently higher chance of yielding reward) as the dependent variable. We did not observe a significant effect in this direction (RL trial accuracy ~ detrended accuracy: fixed effect of accuracy on infrequent trials on RL performance: $B = 0.12$, $z = 0.95$, $p = .34$; RL trial accuracy ~ detrended CV: fixed effect of CV on RL performance: $B = -0.16$, $z = -0.61$, $p = .545$). Thus, at least in the context of our integrated task, the influence of rewards on sustained attention was robust, while effects in the opposite direction were not detected.

Overall, the results of Experiment 1 provide evidence of a short-timescale relationship between rewards earned via instrumental learning and sustained attention. In contrast, we did not find evidence that the surprise associated with those rewards (i.e., prediction error) influenced attention performance. These results extend beyond motivational state-based accounts for the influence of reward on sustained attention, implying a more rapid, dynamic mechanism for this interaction that does not rely on rewards directly modulating the value of attention performance.

Importantly, reward feedback is often not simply binary. In most everyday cases of instrumental learning, reward is a scalar variable that can take on a range of values (e.g., amounts of money, enjoyment of different foods, subtle or effusive social encouragement). Thus, to both replicate and extend the results of Experiment 1, in Experiment 2, we implemented a modified task with continuous rewards and recruited a large online sample, again investigating short-timescale interactions between reward and sustained attention.

## Experiment 2

### Method

#### *Participants*

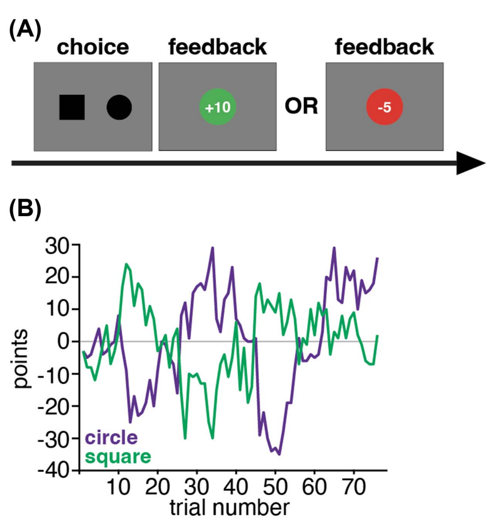
We recruited 146 participants from the online platform Prolific to complete the study. Demographic information was collected via participants' records through Prolific. To report sex, participants were asked "What is your sex, as recorded on legal/official documents?" with the options "male" and "female" as responses. Participants were also asked about their ethnicity; however, we do not report that information here for parity with demographic information collected in Experiment 1.

Exclusion criteria were identical to those in Experiment 1. We excluded participants that had less than 75% accuracy on frequent trials ($N = 8$) or made responses on less than 75% of RL trials (i.e., <57 RL trials; $N = 9$), indicating that they were not engaged in the task. In addition, we excluded two participants who made the same response on RL trials for over 90% of trials, suggesting that they were not trying to learn which shape would yield greater reward. Several participants met more than one exclusion criterion and our final sample included 133 participants ($N = 44$ reported female, $N = 89$ reported male; $M_{\text{age}} = 32.7$, $SD_{\text{age}} = 4.96$).

### Procedure

The online task closely resembled the task in Experiment 1 with two main differences. First, we shortened the task to include only 76 RL trials for each participant. This change shortened the task duration to approximately 30 min, making the task more amenable to remote participation while still preserving sufficient data for our analyses. Second, rather than having probabilistic binary reward feedback on RL trials (+1 or +0), we varied rewards continuously across positive and negative "point" values. During feedback, participants saw the number of points they were awarded in the center of the screen on a red circle if the point value was negative or a green circle if the point value was positive (Figure 4A). In this task, participants were instructed to choose the shape that they thought would yield more points. The same two reward schedules were used for all participants; however, the location and shape associated with each reward schedule were counterbalanced across participants. These reward schedules were generated by using a random walk between −50 and 50 points to create a set of possible schedules.

**Figure 4**
*Experiment 2: Task Design*




*Note.* (A) Modified feedback for Experiment 2. Attention trials were identical to those in Experiment 1. (B) Reward schedules. Assignment of reward schedule to shape was counterbalanced across participants. See the online article for the color version of this figure.

The chosen schedules were selected to vary across positive and negative point values and have multiple reversals in which shape yielded more reward. Average point values for each schedule were relatively equal as well (Schedule A: $M = 0.013$ points ± 11 $SD$; Schedule B: $M = 0.013$ points ± 16 $SD$) to limit biases toward choosing the stimulus that was overall more rewarding. Final reward schedules varied between −35 and 29 points and are depicted in Figure 4B.

Blocks of attention trials varied between 19 and 27 trials in length (average block length = 23 trials [20.7 s], ±0.042 $SD$). We pregenerated 10 sequences of frequent, infrequent, and RL trials in the same way as in Experiment 1 in order to facilitate online data collection. Each participant was randomly assigned to one of these 10 trial sequences. Overall, this yielded 1,517 frequent trials, 168 infrequent trials, and 75 RL trials on average after excluding trials where participants did not respond (2.7% of trials excluded, ±3.6 $SD$).

### Modeling

The modeling procedure was the same as in the previous experiment. We note that the change in reward feedback from binary feedback to points that span a wider range did not alter the modeling approach; RPEs are calculated as the difference between received and expected reward (see Equation 2). Participants' choices were fit individually to capture subject-specific learning parameters. As in Experiment 1, we used these subject-specific parameters to simulate RPEs on each trial. Additionally, we used these subject-specific parameters to simulate behavior and plot model fit in Figure 5D using the same procedure as in Experiment 1.

### Analysis

We followed the same analysis plan for Experiment 2 as detailed in the Methods section of Experiment 1.

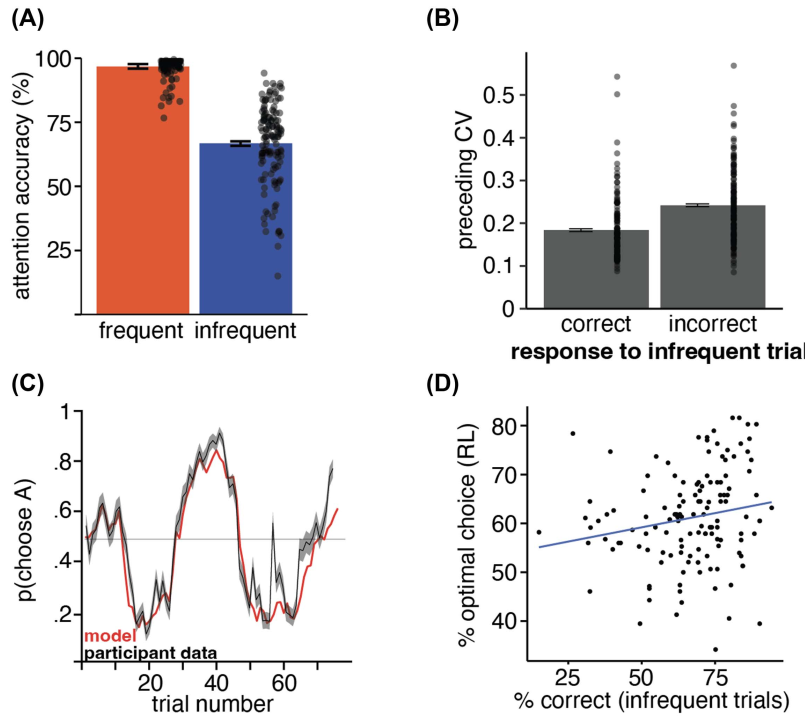
### Transparency and Openness

This study was not preregistered. We include details about sample size and participant exclusions, as well as methodological details to facilitate future replication. Data and analysis code are available at https://github.com/jetrach/SARL_2024.

## Results

Attention task performance was comparable to Experiment 1: Participants were more accurate on frequent trials ($M = 96.7\%$, $SD = 4.08\%$) than they were on infrequent trials, $M = 66.8\%$, $SD = 15.7\%$; two-sample paired $t$ test: $t(132) = 24.07$, $p < .001$, $d = 2.09$, 95% CI [0.27, 0.32]; Figure 5A, replicating previous studies with a similar attention task (deBettencourt et al., 2019). Accuracy on infrequent attention trials and CV on frequent attention trials were significantly correlated, mean within-subject correlation between attention metrics: $\rho = -.23 \pm .15$; one-sample $t$ test: $\mu = 0$, $t(132) = -16.4$, $p < .001$, $d = 1.42$, 95% CI [−0.29, −0.22]; mean within-subject correlation between detrended attention metrics: $\rho = -.21 \pm .15$; one-sample $t$ test: $\mu = 0$, $t(132) = -15.89$, $p < .001$, $d = 1.38$, 95% CI [−0.24, −0.19], and CV was significantly higher preceding incorrect responses to infrequent attention trials

**Figure 5**
*Experiment 2: Sustained Attention and RL Performance*




*Note.* (A) Accuracy on frequent and infrequent attention trials. All error bars represent ±1 standard error of the mean. (B) Validation analysis of CV metric of attention. CV on three frequent trials preceding correct versus incorrect responses to infrequent attention trials. (C) Probability of selecting the shape associated with purple reward schedule (circle) in Figure 4B. Participant behavior in black. Simulated probability of choosing the circle from average fitted parameters. (D) Correlation between performance on attention trials (accuracy on infrequent trials) and performance on the RL trials (percent of RL trials where the participant chose the shape with a higher point value). CV = coefficient of variation; RL = reinforcement learning. See the online article for the color version of this figure.

(relative to correct responses), Figure 5D; paired-sample $t$ test: $t(132) = -12.26$, $p < .001$, $d = -1.06$, 95% CI [−0.07, −0.05], in this sample as well, indicating that both metrics were similarly capturing performance on the attention task. We again found evidence of a significant vigilance decrement over time, with individuals becoming less accurate, Average Within-Subject Correlation Accuracy × Block Number: $\rho = -.13 \pm .18$ *SD*; one-sample $t$ test on Fisher transformed $\rho$ values: $\mu = 0$, $t(132) = -8.27$, $p < .001$, $d = -0.72$, 95% CI [−0.17, −0.10], and more variable in their RTs, Average Within-Subject Correlation CV × Block Number: $\rho = .083 \pm .23$; one-sample $t$ test on Fisher transformed $\rho$ values: $\mu = 0$, $t(132) = 4.19$, $p < .001$, $d = 0.36$, 95% CI [0.05, 0.13], over the course of the task. We performed primary analyses with detrended attention metrics.

On RL trials, participants showed evidence of learning, selecting the shape that was associated with more points on 60.9% of trials (*SD* = 9.7%), a value significantly greater than chance guessing, one-sample $t$ test: $\mu = .5$, $t(132) = 13.09$, $p < .001$, $d = 1.13$, 95% CI [59.5, 62.8]. Further, participants did not show a bias to choose one shape over the other across the task, paired-sample $t$ test to choose left versus right shape: $t(132) = 1.61$, $p = .111$. Again, the two-learning rate model was a better fit for participant behavior (AIC = 10,869) than the one-learning rate model (AIC = 11,764), so we used the better fitting model for our analysis. Figure 5B depicts the probability of the participant choosing the shape associated with Schedule A in black and the model fit with the simulated probability of choosing Bandit A in red.

We examined overall correlations between performance on the concurrent tasks and assessed the influence of reward feedback and surprise on attentional vigilance. As in Experiment 1, CV was negatively correlated with overall RL performance but the correlation was not statistically reliable (Spearman's correlation: $\rho = -.13$, $p = .15$). In contrast, sustained attention performance indexed by accuracy on infrequent trials was correlated with overall RL performance (Figure 5C; Spearman's correlation: $\rho = .23$, $p = .0084$) and learning rate (positive learning rate: Spearman's correlation: $\rho = .18$, $p = .0396$; negative learning rate: Spearman's correlation: $\rho = .25$, $p = .0041$), replicating the results of Experiment 1 in a larger sample with a modified task. The correlation between accuracy on infrequent trials and subject-wise inverse temperature parameters was also negative, as in Experiment 1; however, it was not statistically reliable (Spearman's correlation: $\rho = -.15$, $p = .081$).

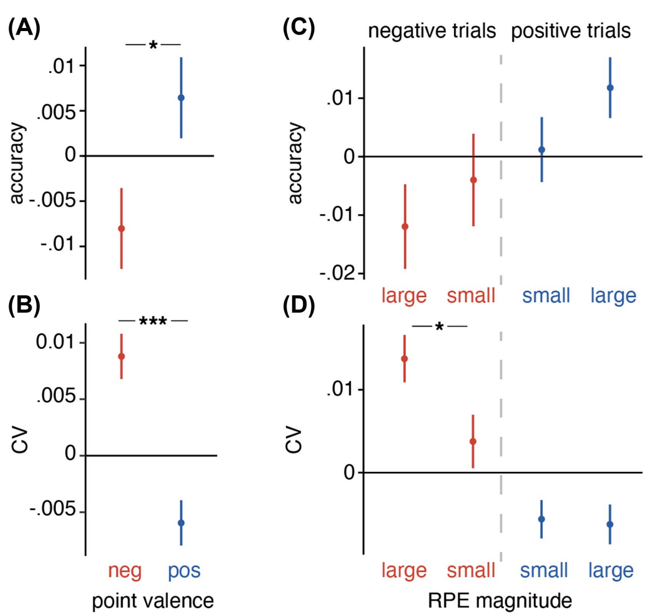
This replication demonstrates the robustness of the relationship between performance on the two tasks.

We compared attention performance after RL trials with positive point values to attention performance after RL trials with negative point values to examine the impact of rewards on attention. Participants exhibited significantly boosted sustained attention after positive point trials relative to negative point trials in both accuracy, Figure 6A, $t(132) = 2.29$, $p = .024$, $d = -0.20$, 95% CI [0.002, 0.027], and CV, Figure 6B, $t(132) = -5.22$, $p < .001$, $d = 0.45$, 95% CI [−0.02, −0.009]. This finding thus replicates the reward analyses presented in Experiment 1.

In addition, we examined the relationship between RPE magnitude and attention performance after positive and negative point value trials. We simulated trial-by-trial RPEs individually for each participant following the same procedure detailed in Experiment 1. We grouped trials based on whether the point value on the previous RL trial had been positive or negative. Then, we did a median split on RPEs for each group of trials and compared attention performance after large or small magnitude RPEs for both negative and positive feedback. We found no significant differences in attentional state after large versus small RPEs on positive point trials, Figure 6C, detrended accuracy: $t(132) = -1.55$, $p = .122$, 95% CI [−0.02, 0.003]; Figure 6D, detrended CV: $t(132) = 0.21$, $p = .832$, 95% CI [−0.005, 0.006], nor in accuracy on negative point trials, detrended accuracy: $t(132) = -1.56$, $p = .120$, 95% CI [−0.04, 0.004]. After negative points trials, however, participants did exhibit higher CV (i.e., worse sustained attention) after especially large negative RPEs relative to small negative RPEs, $t(132) = 2.15$, $p = .033$, $d = 0.19$, 95% CI [0.0008, 0.02].

**Figure 6**

*Experiment 2: Effect of Reward on Sustained Attention*




*Note.* (A) Accuracy (detrended) on infrequent attention trials following rewarded versus unrewarded reinforcement learning trials. All error bars reflect ±1 standard error of the mean. (B) RT variability (measured by detrended coefficient of variation) on frequent attention trials following rewarded versus unrewarded reinforcement learning trials. (C) Accuracy on infrequent trials after large and small magnitude RPEs. Unrewarded trials on left and rewarded trials on right. (D) RT variability (measured by coefficient of variation) on frequent attention trials after large and small magnitude RPEs. Unrewarded trials on left and rewarded trials on right. RPE = reward prediction error; CV = coefficient of variation; RT = reaction time. See the online article for the color version of this figure.
* $p < .05$. *** $p < .001$.

For completeness, we compared four linear mixed-effects models to assess the impact of reward and RPE on the two metrics of sustained attention. Each model was fit to either infrequent trial accuracy or CV using (a) point value from previous RL trial, (b) simulated RPE from previous RL trial, (c) point value *valence* (0 for negative, 1 for positive), or (d) RPE *valence* (0 for negative, 1 for positive) as predictors. Additionally, models included random intercepts and slopes for each subject. All models were significantly predictive of attentional state ($t$s > 2.25, $p$s < .0245), replicating the findings of Experiment 1. Model fit was similar across variants of the model (accuracy ~ points: AIC = 4,286; accuracy ~ points valence: AIC = 4,282; accuracy ~ RPE: AIC = 42,584; accuracy ~ RPE valence: AIC = 4,280; CV ~ points: AIC = −14,492; CV ~ points valence: AIC = −14,481; CV ~ RPE: AIC = −14,464; CV ~ RPE valence: AIC = −14,467), likely due to the high correlation between points and RPEs across trials (average correlation across participants = .84, $SD$ = .16).

We again followed these analyses with a post hoc analysis to characterize the dynamics of this attentional boost within blocks of attention trials, as in Experiment 1. Replicating the findings of Experiment 1, we found that the boost in sustained attention performance was strongest during the earliest phase of the block, early trials: $M = 0.031$; one-sample $t$ test: $t(132) = 9.04$, $p < .001$, $d = 0.78$, 95% CI [0.024, 0.038], and quickly diminished, middle phase: $M = 0.003$; one-sample $t$ test: $t(132) = 0.77$, $p = .442$, 95% CI [−0.004, 0.01]; late phase: $M = 0.006$; one-sample $t$ test: $t(132) = 1.60$, $p = .112$, 95% CI [−0.001, 0.013]. This analysis provides further evidence of a quite transient boost in sustained attention following rewarding feedback.

As in Experiment 1, we assessed whether attentional state was predictive of learning in our task. We performed a logistic regression that used the detrended attention metrics to predict whether participants selected the more valuable shape on a given trial. We did not find that attentional state was predictive of choice performance in the task (RL trial accuracy ~ detrended accuracy: fixed effect of accuracy on infrequent trials on RL performance: $B = -0.02$, $z = -0.29$, $p = .772$; RL trial accuracy ~ detrended CV: fixed effect of CV on RL performance: $B = 0.064$, $z = 0.363$, $p = .716$), perhaps suggesting that overall relationships between performance on the two tasks was primarily due to the influence of ongoing RL on attention, rather than a robust bidirectional relationship.

## Experiment 3

The results of Experiments 1 and 2 demonstrate a positive effect of reward feedback on attentional vigilance. This effect is transient, only lasting several seconds, and, critically, does not require rewards to be directly tied to attention task performance. In the following study, we modified our approach to assess whether simply processing reward feedback itself was sufficient to elicit boosts in sustained attention, or whether the observed effect was contingent on people actively engaging in instrumental learning (i.e., using the reward feedback to guide future behavior).

To compare these different interpretations, in Experiment 3, we replaced the instrumental RL task with a "passive" reward task where participants were cued to simply press a button to reveal reward feedback. We closely matched the rewards that participants received with the statistics of rewards received in Experiment 2 (see the Method section) to enable comparison across the two experiments. We examined attentional performance after positive and negative reward trials as in Experiment 2. In this way, we can test whether reward feedback in the absence of instrumental learning also yields boosts in attentional vigilance.

## Method

### Participants

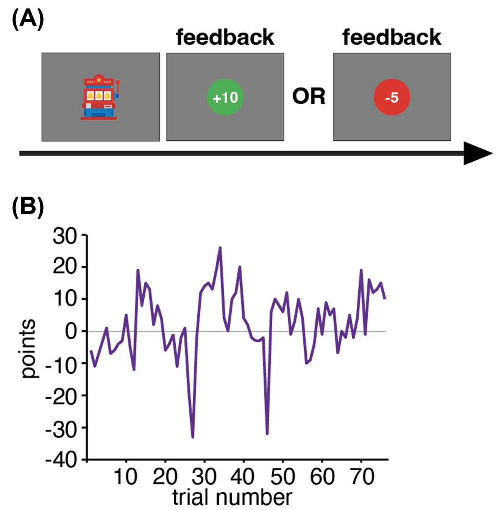
We recruited 150 participants from Prolific to participate in this study. Demographic information was collected through Prolific, as in Experiment 2. For parity with Experiment 2, we excluded participants who were less than 75% accurate on frequent attention trials ($N = 7$) and participants who did not respond promptly (<1.5 s) to at least 75% of reward trials ($N = 9$, four of which also met the first exclusion criterion). We included the second exclusion criterion to match the second attention check in Experiment 2, although participants needed to respond to each reward trial in Experiment 3 before the task could proceed. The deadline was based on the RL trial durations in Experiments 1 and 2. In addition to these two performance-based exclusions, three data files were never received due to participant connectivity issues, leaving us with a final sample of 135 participants ($N = 56$ reported female, $N = 79$ reported male; $M_{\text{age}} = 32.3$, $SD_{\text{age}} = 5.2$).

### Procedure

The task in Experiment 3 resembled that in Experiment 2: Participants experienced blocks of 19–27 attention trials followed by individual reward task trials. However, on the reward task trials, participants saw a single slot machine cartoon at the center of the screen and were instructed to press the "J" key with their right hand to reveal reward feedback. The trial did not end until the participant made a response. Once they pressed the J key, the points that they received on that trial were displayed in a green circle if it was a positive point value or a red circle if it was a negative point value. Points were displayed for 750 ms before the next block of attention trials began (Figure 7A). There were 76 of these passive reward trials. Again, participants used their left hand, positioned on the F and D keys, to respond to attention trials. We generated 10 trial sequences of frequent, infrequent, and RL trials prior to launching the experiment online as in Experiment 2. One of these pregenerated trial sequences was then randomly selected for each participant. The task took approximately 30 min to complete.

All participants received the same reward sequence in Experiment 3 (Figure 7B). We used the rewards that participants received in Experiment 2 to construct the reward sequences that participants experienced in Experiment 3. To do this, we first selected the point value that most participants in Experiment 2 received on each trial to create a sequence of 76 rewards. Then, we took the average of each participants' reward sequence and subtracted that value from each

**Figure 7**
*Experiment 3: Task Design*




*Note*. (A) Modified reinforcement learning trials and feedback for Experiment 3. Slot machine image from https://www.Flaticon.com. (B) Reward schedule. See the online article for the color version of this figure.

point value in the originally constructed reward sequence. This step is necessary since participants in Experiment 2 tended to choose the higher value shape as they learned, and we wanted to ensure that Experiment 3 participants were experiencing both positive and negative point trials. Finally, we rounded this sequence to the nearest integer for ease of presentation during the task. This process yielded a reward schedule with a range of −33 to 26 points, a mean of 2.46 points, and a cumulative value of 187 points which is comparable to the points received by participants in Experiment 2 (range = −35 to 29, $M = 2.75$, cumulative value = 205 points).

### Analysis

Our primary analyses mirrored the approach in Experiment 2. Specifically, we compared detrended attention performance (accuracy on infrequent trials and CV on frequent trials) after negative point trials and positive point trials. We used a mixed-factor analysis of variance to test for differences between performance on Experiments 2 and 3. Here, we report effect size using partial eta squared ($\eta^2$).

We excluded trials where participants did not respond, yielding 1,528 frequent and 169 infrequent trials per participant on average. All reward trials were included as participants needed to respond before the task progressed on these trials.

### Transparency and Openness

This study was not preregistered. We include details about sample size and participant exclusions, as well as methodological details to facilitate future replication. Data and analysis code are available at https://github.com/jetrach/SARL_2024.

## Results

We again replicated attention task behavior in Experiment 3: Participants were significantly more accurate on frequent versus infrequent attention trials, Figure 8A; paired-sample $t$ test: $t(134) = 22.46$, $p < .001$, $d = 1.93$, 95% CI [0.30, 0.36], the two metrics of sustained attention were significantly correlated, mean within-subject correlation between attention metrics: $\rho = -.26 \pm .15$, one-sample $t$ test: $\mu = 0$, $t(134) = -18.59$, $p < .001$, $d = -1.60$, 95% CI [−0.30, −0.25]; mean within-subject correlation between detrended attention metrics: $\rho = -.23 \pm .14$, one-sample $t$ test: $\mu = 0$, $t(134) = -18.09$, $p < .001$, $d = -1.56$, 95% CI [−0.27, −0.22], and CV was significantly higher (i.e., more variable) preceding incorrect responses, versus correct, on infrequent attention trials, paired-sample $t$ test: $t(134) = -13.38$, $p < .001$, $d = -1.15$, 95% CI [−0.07, −0.05].

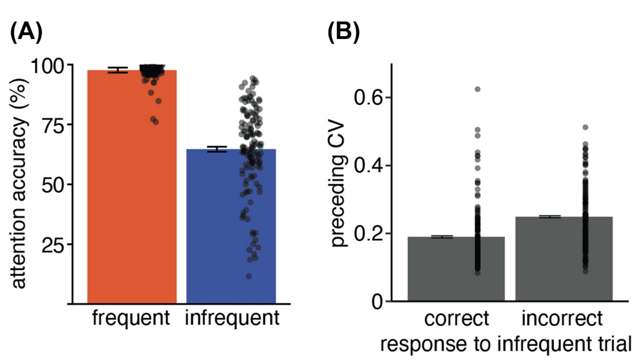
Performance on the attention task was similar across Experiments 2 and 3: There were no overall differences in accuracy between the two tasks, Experiment 2: $M = 93.1\%$, Experiment 3: $M = 94.4\%$; Welch two-sample $t$ test: $t(261.8) = 1.23$, $p = .220$, 95% CI [−0.02, 0.004], nor specifically accuracy on infrequent attention trials, Experiment 2: $M = 66.8\%$, Experiment 3: $M = 64.7\%$; Welch two-sample $t$ test: $t(261.2) = 1.03$, $p = .305$, 95% CI [−0.02, 0.06]. However, participants did tend to be slightly faster to respond on average in Experiment 3 as compared to Experiment 2, Experiment 2: $M = 345$ ms, Experiment 3: $M = 321$ ms; Welch two-sample $t$ test: $t(265.8) = 3.69$, $p < .001$, $d = 0.45$, 95% CI [11.29, 37.18]. This could have arisen since there was no extra load of a learning task in Experiment 3. Importantly, CV was not different across the two experiments, Experiment 2: $M = 0.31$, Experiment 3: $M = 0.31$; Welch two-sample $t$ test: $t(255.8) = 0.29$, $p = .772$, 95% CI [−0.03, 0.02].

Comparison of attention performance after positive versus negative point trials suggests a significant, but less robust, boost in attentional vigilance after reward trials relative to Experiment 2. There were no differences in accuracy after positive versus negative point trials, detrended accuracy: $t(134) = 0.70$, $p = .485$, 95% CI [−0.007, 0.02]; Figure 8A, although individuals were numerically more accurate after positive point trials as compared to negative point trials (positive point trials: $M = 0.0024$; negative point trials: $M = -0.0015$). Participants were, however, significantly less variable in their response times after positive point versus negative point trials, detrended CV: $t(134) = 3.64$, $p < .001$, $d = 0.31$, 95% CI [−0.01, −0.004]; Figure 8B, providing some evidence of a reward-related boost in attention performance, at least in our RT-related measures of vigilance. Finally, we conducted the same post hoc analysis as in Experiments 1 and 2 to examine the within-block dynamics of this facilitatory effect, dividing each block of attention trials into three phases. As in Experiments 1 and 2, this analysis revealed a transient boost in sustained attention after rewarded trials, early phase: $M = 0.011$; one-sample $t$ test: $t(134) = 3.39$, $p < .001$, $d = 0.29$, 95% CI [0.005, 0.02], which quickly diminished, middle phase: $M = 0.0006$; one-sample $t$ test: $t(134) = 0.18$, $p = .858$, 95% CI [−0.006, 0.007]; late trials: $M = -0.004$; one-sample $t$ test: $t(134) = -1.01$, $p = .315$, 95% CI [−0.011, 0.004].

We conducted a 2 × 2 analysis of variance with one between-subject factor (experiment) and one within-subject factor (reward valence) to assess if the reward facilitation effect (on CV) in Experiment 3 was similar to that seen in than in Experiment 2. CV did not differ across the two experiments, experiment: $F(1, 266) = 2.48$, $p = .116$, and CV was significantly modulated by reward valence, reward valence: $F(1, 266) = 40.04$, $p < .001$, $\eta_p^2 = 0.13$. Crucially, we found that there was no interaction between experiment and reward valence, Experiment × Reward Valence: $F(1, 266) = 2.34$, $p = .127$. Thus, we can infer that the influence of reward valence was comparable between the two studies, at least in its effect on CV.

**Figure 8**

*Experiment 3: Sustained Attention Performance*




*Note.* (A) Accuracy on frequent versus infrequent attention trials. (B) Validation analysis of CV metric of attention. CV on three frequent trials preceding correct versus incorrect responses to infrequent attention trials. CV = coefficient of variation. See the online article for the color version of this figure.

## Discussion

In this study, we sought to investigate short-timescale interactions between rewards and attentional vigilance when the rewards were not directly contingent on vigilance performance. Overall, we found evidence that reward transiently facilitated attentional vigilance, extending previous results (Esterman et al., 2014, 2016, 2017; Massar et al., 2016; Robison et al., 2021) to a timescale of seconds. We did not find consistent evidence that surprise reliably modulated attention performance. Crucially, the observed facilitatory effects of reward were robust even though reward was never contingent on attentional performance and did not directly increase the value of attentional performance for the participant. This was most dramatically seen in Experiment 3, where reward facilitation effects persisted even when rewards were received passively, requiring no learning or active choice by the participant. Taken together, our results thus point to an automatic link between reward and sustained attention.

The finding that noncontingent rewards boost sustained attention both builds on and adds specificity to previous work suggesting that performance-based incentives improve sustained attention (Esterman et al., 2014, 2016, 2017; Massar et al., 2016; Robison et al., 2021). Generally, the fact that rewards can reduce lapses in sustained attention is thought 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; Helton & Russell, 2011, 2013; Muraven & Baumeister, 2000; Thomson et al., 2015; Warm et al., 2008). Instead, it is more aligned both with "underload" models (i.e., that a lack of motivation leads to lapses in sustained attention; Manly et al., 1999; for a review of models of sustained attention, see Esterman & Rothlein, 2019) and opportunity cost models of sustained attention (i.e., where lapses occur when the cost of maintaining a consistent attentional state overshadows the expected value of the task; Kool et al., 2017; Kurzban et al., 2013; Thomson et al., 2015).

Our results are unique relative to the extant literature for two reasons: First, the RL task reward feedback was not contingent on performance in the attention task, indicating that this facilitatory effect can emerge even when rewards do not directly increase the value of attentional performance. Second, our reward stimuli were only symbolic (i.e., we did not give any additional monetary reward based on performance in either task), suggesting that even subtle reward signals or abstract goals can affect attention. Thus, our results suggest that reward feedback need not increase the value of sustained attention to induce performance improvements.

We note that our results stand in contrast to a previous study that suggested that rewards *must* be performance-contingent to trigger improvements in sustained attention (Massar et al., 2016). Massar et al. (2016) specifically compared performance during a sustained attention task, where participants received no rewards (baseline), performance-based rewards, or random (i.e., not connected to task performance) rewards during different blocks of trials. They found that while sustained attention was improved in the performance-based reward block relative to baseline, there was a decrease in performance during the random reward block relative to baseline. These results, however, are difficult to directly compare to the present study, as Massar and colleagues were looking at average sustained attention measures over a block of performance-contingent reward trials versus a block of random reward trials, rather than comparing sustained attention performance after rewarded versus unrewarded trials within each condition. It is possible that individual rewards were boosting performance in both cases, however such transient boosts in sustained attention following random rewards would be obscured by this block-level analysis.

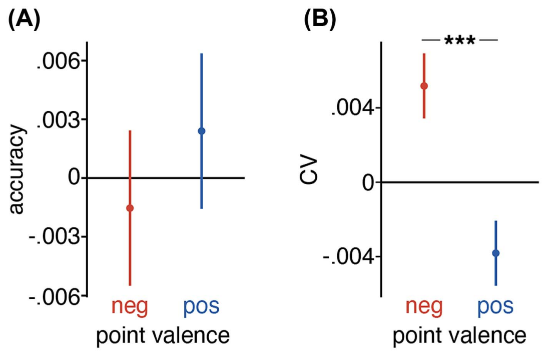
The fact that we still observed reward-related facilitation in Experiment 3 where rewards were only received passively (albeit only for the RT-related attention metrics) argues against the idea that the effects seen in Experiments 1 and 2 could be explained by an indirect contingency, where earning reward in the RL tasks boosts attention so that participants can make better choices and thus earn more reward in the long run. Further evidence against this interpretation comes from our analyses looking at the effects of attentional state on RL performance—we found no reliable evidence that people made better choices (operationalized as choosing the bandit with a higher reward probability) when they were in high versus low vigilance states. Finally, our finding across all three experiments that reward facilitation was most reliable in the earliest phases of the subsequent attention blocks, which suggests a rather transient effect on the order of several seconds, also argues against the idea that people optimized their attentional state to prepare for later RL trials. If that were the case, we may have expected to see the opposite timecourse effect (i.e., an increase, or "ramping," of the reward facilitation effect over time).

An alternative interpretation of our results is that the interspersed RL trials in our task may act as consistent feedback that promotes general task engagement. Previous work has shown that increasing task engagement (i.e., reducing the monotony of the sustained attention task) can reduce time-on-task vigilance effects (Pop et al., 2012; Robison et al., 2021). In regard to feedback and task engagement, recent work suggests that providing intermittent performance feedback to participants can also support sustained attention (Robison et al., 2021). Robison et al. (2021) examined the effect of goal-setting and performance-based feedback on sustained attention. Participants were assigned to one of four conditions where they were given a performance goal (or not) and/or where they received feedback (or not). Interestingly, they found that participants who both had a goal and received feedback had better sustained attention performance and reported higher levels of motivation and less mind-wandering during task execution, relative to participants who received only one of the manipulations or neither. While this result does provide evidence that feedback can boost task engagement and sustained attention, it differs from our results in important ways. First, feedback in previous experiments was directly related to attentional performance, in contrast to our design. Second, our effects are tied to a specific feature of feedback—rewarded versus unrewarded—rather than a generalized effect of any feedback.

We speculate that our results reflect a low-level mechanism where reward responses in the brain automatically and transiently boost attention. Rewards might boost sustained attention via projections from subcortical dopaminergic structures to prefrontal and parietal circuits implicated in attentional vigilance (Chudasama & Robbins, 2004; Esterman et al., 2013; Granon et al., 2000; Nieoullon, 2002; Totah et al., 2013; Westbrook & Braver, 2016). Behaviorally, this hypothesis would be strengthened by evidence that RPEs parametrically modulated attentional state in our task, a relationship we found to be unreliable. It is possible that our design was not sensitive enough to uncover this relationship in behavior (indeed, we saw some subtle signs of this effect). Further, we used computational models to simulate the latent RPEs that participants might be experiencing on a trial-by-trial basis, but other techniques (e.g., neuroimaging) could provide a more direct measure of RPE strength, which is itself a direct correlate of dopamine release. Noradrenergic activity in the locus coeruleus is also a potential mediator of a relationship between reward and attention (Aston-Jones & Cohen, 2005; Sara, 2009; Zhang et al., 2023). The locus coeruleus has long been implicated in regulating vigilance and arousal (Aston-Jones & Bloom, 1981; Aston-Jones et al., 1991), and it is also interconnected with the dopaminergic circuitry involved in reward processing during reinforcement learning (Sara, 2009). Thus, this network is well positioned to coordinate reward inputs and attentional processes. Future work is necessary to clarify how different types of reward and feedback differentially impact attentional state and the associated brain networks (e.g., deBettencourt et al., 2015).

We included Experiment 3 to examine whether benefits to attentional state observed in Experiments 1 and 2 were contingent on instrumental RL, or whether they represented a more generalized effect of reward on sustained attention. We found that participants still showed improved attention after positive rewards, although they were not experiencing the reward feedback in the context of an instrumental learning task. While this boost in attention from reward was evident and comparable in magnitude to the effect in Experiment 2, we did not find significant modulation of our other metric of sustained attention, accuracy on infrequent attention trials (Figure 9). One possible explanation for this difference could be that participants were not paying as close attention to the reward trials in Experiment 3 since they did not need to use the reward feedback to guide any decisions. We believe this to be an unlikely explanation, as reward feedback was displayed for the same amount of time as in Experiment 2, the subsequent block of attention trials began immediately after the presentation of reward feedback, and attention performance was comparable between the two experiments. Alternatively, the fact that there was no choice behavior on

**Figure 9**
*Experiment 3: Effect of Reward on Sustained Attention*




*Note.* (A) Accuracy (detrended) on infrequent attention trials following rewarded versus unrewarded reinforcement learning trials. All error bars reflect ±1 standard error of the mean. (B) RT variability (measured by the detrended coefficient of variation) on frequent attention trials after rewarded versus unrewarded reinforcement learning trials. CV = coefficient of variation; neg = negative; pos = positive; RT = reaction time. See the online article for the color version of this figure.
*** $p < .001$.

RL trials (i.e., Experiment 3 was not a dual task) might have allowed participants to deploy more cognitive resources to protect against errors during the attention blocks, despite being unable to control subtler variations in CV.

Our task design allowed us to examine the relationship between reward, surprise, and fluctuations in sustained attention within-subject and on a finer timescale than previous behavioral work. We show that individual trials can influence subsequent sustained attention on the scale of tens of seconds within the flow of the task, rather than aggregating across much longer periods of time (i.e., the length of an experimental session). In post hoc analyses, we attempted to get even finer temporal resolution by dividing attention blocks (~13–20 s in duration) into distinct phases to understand the dynamics of attentional facilitation after rewards. In Experiment 1, this post hoc analysis revealed significant reward effects in the middle of the block of attention trials but not in the last third. In Experiments 2 and 3, the reward effect was most reliable in the phase immediately following the reward feedback but not in the second and third phases. Thus, the facilitating effect of reward on attention appeared to be rather transient, which in our view further points to an automatic link between reward and attention in our task, rather than a deliberative accounting of the value of sustaining attention.

One consideration in interpreting our findings is that we are describing boosts in sustained attention performance after rewards, relative to performance after unrewarded trials, as we did not collect a baseline attention period to compare to. It is thus ambiguous whether we are seeing an inhibitory effect of punishment on attention rather than a facilitatory effect of reward on attention. This issue is somewhat semantic in our view—all feedback, or even a lack of feedback, can be said to have some kind of valence when compared to explicitly valanced feedback received at neighboring time points. For example, while the feedback in Experiment 1 is not explicitly negative (+0, rather than negative point values in Experiments 2 and 3), such feedback will still elicit a negative prediction error in the participant's brain if they are expecting to receive rewarding feedback. Vigilance decrements over the course of the task present a problem for collecting a baseline measurement of participants' attentional performance: collecting a baseline attention measurement at the beginning of the session when participants are not fatigued would overestimate attentional performance, whereas a baseline period at the end of the session when participants are fatigued would underestimate attentional performance. Thus, we cannot say definitively if rewards boost sustained attention relative to some baseline or if nonrewards and negative feedback impair attention performance, or if both processes occur simultaneously. Future work could engage with this question; however, the results presented here should be interpreted as a relative difference in performance, rather than a unidirectional effect.

We curiously did not detect robust effects of attentional state on RL performance. 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 (Leong et al., 2017; Niv et al., 2015). In contrast to research on selective attention and RL, recent work has indicated that attentional lapses are beneficial for learning in certain circumstances (A. Decker, Dubois, et al., 2023). It is possible that the more rapid attentional dynamics of our integrated task, or the context changes that occurred between sustained attention and RL trials, attenuated the potential effects of sustained attention on RL. Alternatively, the RL system may be robust to minor fluctuations in sustained attention—particularly in a relatively simple choice task such as our own. Sustained attention may, for example, be required when adaptive choice requires maintenance of a more complex model of the task (Otto et al., 2015), when reward feedback is more abstract or requires working memory (McDougle et al., 2022), or when the space of possible features to attend to is larger (Wise et al., 2024).

Finally, on a methodological note, we think our results also speak to interpretability issues with RL model parameters. Recent work has suggested that some RL model parameters can be hard to interpret and are affected by task context (Eckstein et al., 2022; Vrizzi et al., 2023). Our results may speak to difficulties in interpreting specifically the learning rate parameter of RL models—we found robust between-subject correlations between sustained attention and RL learning rates in our experiments. Although we did not see a reliable trial-by-trial influence of sustained attention on RL choices, our findings still suggest that learning rates cannot easily be assumed to only represent precise neurocognitive variables (e.g., corticostriatal plasticity) because they may be shaped by other psychological factors like attentional state or mood (Jangraw et al., 2023). These state variables will be important to measure or model to improve the interpretability of model parameters, especially in applications such as computational psychiatry.

## Constraints on Generality

Our results should, in theory, be generalizable to the broader population. While the initial sample from Experiment 1 was recruited from the Yale University undergraduate population, which could limit generalizability, we followed up with two large online samples that were more diverse. Still, all our samples are limited to primarily young adults, so it is unclear how these interactions between sustained attention and reward might evolve with typical aging. In addition, we opted to employ variations on the same basic task in order to facilitate comparison across experiments. Future work could implement a wider variety of tasks to fully characterize interactions between reward and vigilance.

## Conclusion

Our findings demonstrate a tight link between reward feedback and attentional dynamics, even when rewards are not contingent on behavior. Future work could explore the nature of this interaction further; for example, can our effect be explained by a kind of automatic cognitive accounting, where reward inputs globally signal that general environmental conditions are improving and that the cost of greater attentional vigilance might be worth possible ensuing rewards? Might our effects also be interpreted at a lower level, where reward feedback is correlated with phasic dopamine, which is then automatically broadcast to frontoparietal regions where it may act as a "gain" signal for attentional vigilance? How changes in automatic effort accounting and/or phasic dopamine induce short-timescale modulations of sustained attention, and the neural and computational systems supporting this process, remain open questions that are relevant across domains of human performance.

## References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Anderson, B. A., Kuwabara, H., Wong, D. F., Gean, E. G., Rahmim, A., Brašić, J. R., George, N., Frolov, B., Courtney, S. M., & Yantis, S. (2016). The role of dopamine in value-based attentional orienting. *Current Biology*, *26*(4), 550–555. https://doi.org/10.1016/j.cub.2015.12.062

Ariga, A., & Lleras, A. (2011). Brief and rare mental "breaks" keep you focused: Deactivation and reactivation of task goals preempt vigilance decrements. *Cognition*, *118*(3), 439–443. https://doi.org/10.1016/j.cognition.2010.12.007

Aston-Jones, G., & Bloom, F. E. (1981). Activity of norepinephrine-containing locus coeruleus neurons in behaving rats anticipates fluctuations in the sleep-waking cycle. *The Journal of Neuroscience*, *1*(8), 876–886. https://doi.org/10.1523/JNEUROSCI.01-08-00876.1981

Aston-Jones, G., Chiang, C., & Alexinsky, T. (1991). Discharge of noradrenergic locus coeruleus neurons in behaving rats and monkeys suggests a role in vigilance. *Progress in Brain Research*, *88*(Pt. C), 501–520. https://doi.org/10.1016/S0079-6123(08)63830-3

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of the locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Basner, M., & Dinges, D. F. (2011). Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. *Sleep*, *34*(5), 581–591. https://doi.org/10.1093/sleep/34.5.581

Bergum, B. O., & Lehr, D. J. (1964). Monetary incentives and vigilance. *Journal of Experimental Psychology*, *67*(2), 197–198. https://doi.org/10.1037/h0040966

Chudasama, Y., & Robbins, T. W. (2004). Dopaminergic modulation of visual attention and working memory in the rodent prefrontal cortex. *Neuropsychopharmacology*, *29*(9), 1628–1636. https://doi.org/10.1038/sj.npp.1300490

Collins, A. G. E., & Frank, M. J. (2014). Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychological Review*, *121*(3), 337–366. https://doi.org/10.1037/a0037015

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In M. R. Delgado, E. A. Phelps, & T. W. Robbins (Eds.), *Decision making, affect, and learning: Attention and performance XXIII* (pp. 3–38). Oxford Academic. https://doi.org/10.1093/acprof:oso/9780199600434.003.0001

Daw, N. D., Kakade, S., & Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks*, *15*(4–6), 603–616. https://doi.org/10.1016/S0893-6080(02)00052-7

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

deBettencourt, M. T., Cohen, J. D., Lee, R. F., Norman, K. A., & Turk-Browne, N. B. (2015). Closed-loop training of attention with real-time brain imaging. *Nature Neuroscience*, *18*(3), 470–475. https://doi.org/10.1038/nn.3940

deBettencourt, M. T., Keene, P. A., Awh, E., & Vogel, E. K. (2019). Real-time triggering reveals concurrent lapses of attention and working memory. *Nature Human Behaviour*, *3*(8), 808–816. https://doi.org/10.1038/s41562-019-0606-6

deBettencourt, M. T., Norman, K. A., & Turk-Browne, N. B. (2018). Forgetting from lapses of sustained attention. *Psychonomic Bulletin & Review*, *25*(2), 605–611. https://doi.org/10.3758/s13423-017-1309-5

Decker, A., Dubois, M., Duncan, K., & Finn, A. S. (2023). Pay attention and you might miss it: Greater learning during attentional lapses. *Psychonomic Bulletin & Review*, *30*(3), 1041–1052. https://doi.org/10.3758/s13423-022-02226-6

Decker, A. L., Duncan, K., & Finn, A. S. (2023). Fluctuations in Sustained Attention Explain Moment-to-Moment Shifts in Children's Memory Formation. *Psychological Science*, *34*(12), 1377–1389. https://doi.org/10.1177/09567976231206767

Eckstein, M. K., Master, S. L., Xia, L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. E. (2022). The interpretation of computational model parameters depends on the context. *eLife*, *11*, Article e75474. https://doi.org/10.7554/eLife.75474

Esterman, M., Grosso, M., Liu, G., Mitko, A., Morris, R., & DeGutis, J. (2016). Anticipation of monetary reward can attenuate the vigilance decrement. *PLOS ONE*, *11*(7), Article e0159741. https://doi.org/10.1371/journal.pone.0159741

Esterman, M., Noonan, S. K., Rosenberg, M., & Degutis, J. (2013). In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cerebral Cortex*, *23*(11), 2712–2723. https://doi.org/10.1093/cercor/bhs261

Esterman, M., Poole, V., Liu, G., & DeGutis, J. (2017). Modulating reward induces differential neurocognitive approaches to sustained attention. *Cerebral Cortex*, *27*(8), 4022–4032. https://doi.org/10.1093/cercor/bhw214

Esterman, M., Reagan, A., Liu, G., Turner, C., & DeGutis, J. (2014). Reward reveals dissociable aspects of sustained attention. *Journal of Experimental Psychology: General*, *143*(6), 2287–2295. https://doi.org/10.1037/xge0000019

Esterman, M., & Rothlein, D. (2019). Models of sustained attention. *Current Opinion in Psychology*, *29*, 174–180. https://doi.org/10.1016/j.copsyc.2019.03.005

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(41), 16311–16316. https://doi.org/10.1073/pnas.0706111104

Granon, S., Passetti, F., Thomas, K. L., Dalley, J. W., Everitt, B. J., & Robbins, T. W. (2000). Enhanced and impaired attentional performance after infusion of D1 dopaminergic receptor agents into rat prefrontal cortex. *The Journal of Neuroscience*, *20*(3), 1208–1215. https://doi.org/10.1523/JNEUROSCI.20-03-01208.2000

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experimental Brain Research*, *212*(3), 429–437. https://doi.org/10.1007/s00221-011-2749-1

Helton, W. S., & Russell, P. N. (2013). Visuospatial and verbal working memory load: Effects on visuospatial vigilance. *Experimental Brain Research*, *224*(3), 429–436. https://doi.org/10.1007/s00221-012-3322-2

Helton, W. S., & Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica*, *129*(1), 18–25. https://doi.org/10.1016/j.actpsy.2008.04.002

Jangraw, D. C., Keren, H., Sun, H., Bedder, R. L., Rutledge, R. B., Pereira, F., Thomas, A. G., Pine, D. S., Zheng, C., Nielson, D. M., & Stringaris, A. (2023). A highly replicable decline in mood during rest and simple tasks. *Nature Human Behaviour*, *7*(4), 596–610. https://doi.org/10.1038/s41562-023-01519-7

Kool, W., & Botvinick, M. (2014). A labor/leisure tradeoff in cognitive control. *Journal of Experimental Psychology: General*, *143*(1), 131–141. https://doi.org/10.1037/a0031048

Kool, W., Shenhav, A., & Botvinick, M. M. (2017). Cognitive control as cost–benefit decision making. In T. Egner (Ed.), *The Wiley handbook of cognitive control* (pp. 167–189). Wiley. https://doi.org/10.1002/9781118920497.ch10

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(6), 661–679. https://doi.org/10.1017/S0140525X12003196

Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., & Niv, Y. (2017). Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron*, *93*(2), 451–463. https://doi.org/10.1016/j.neuron.2016.12.040

Lim, J., & Dinges, D. F. (2008). Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences*, *1129*(1), 305–322. https://doi.org/10.1196/annals.1417.002

Madore, K. P., Khazenzon, A. M., Backes, C. W., Jiang, J., Uncapher, M. R., Norcia, A. M., & Wagner, A. D. (2020). Memory failure predicted by attention lapsing and media multitasking. *Nature*, *587*(7832), 87–91. https://doi.org/10.1038/s41586-020-2870-z

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia*, *37*(6), 661–670. https://doi.org/10.1016/S0028-3932(98)00127-4

Massar, S. A. A., Lim, J., Sasmita, K., & Chee, M. W. L. (2016). Rewards boost sustained attention through higher effort: A value-based decision making approach. *Biological Psychology*, *120*, 21–27. https://doi.org/10.1016/j.biopsycho.2016.07.019

MATLAB. (2022). *9.12.0.2009381 (R2022a)*. MathWorks. https://www.mathworks.com/

McDougle, S. D., Ballard, I. C., Baribault, B., Bishop, S. J., & Collins, A. G. E. (2022). Executive function assigns value to novel goal-congruent outcomes. *Cerebral Cortex*, *32*(1), 231–247. https://doi.org/10.1093/cercor/bhab205

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin*, *126*(2), 247–259. https://doi.org/10.1037/0033-2909.126.2.247

Nieoullon, A. (2002). Dopamine and the regulation of cognition and attention. *Progress in Neurobiology*, *67*(1), 53–83. https://doi.org/10.1016/S0301-0082(02)00011-4

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. *The Journal of Neuroscience*, *35*(21), 8145–8157. https://doi.org/10.1523/JNEUROSCI.2978-14.2015

Otto, A. R., Skatova, A., Madlon-Kay, S., & Daw, N. D. (2015). Cognitive control predicts use of model-based reinforcement learning. *Journal of Cognitive Neuroscience*, *27*(2), 319–333. https://doi.org/10.1162/jocn_a_00709

Pop, V. L., Stearman, E. J., Kazi, S., & Durso, F. T. (2012). Using engagement to negate vigilance decrements in the NextGen environment. *International Journal of Human–Computer Interaction*, *28*(2), 99–106. https://doi.org/10.1080/10447318.2012.634759

R Core Team. (2022). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Lassical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.

Robison, M. K., Unsworth, N., & Brewer, G. A. (2021). Examining the effects of goal-setting, feedback, and incentives on sustained attention. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(6), 869–891. https://doi.org/10.1037/xhp0000926

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, *10*(3), 211–223. https://doi.org/10.1038/nrn2573

Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. *Perspectives on Psychological Science*, *10*(1), 82–96. https://doi.org/10.1177/1745691614556681

Totah, N. K. B., Kim, Y., & Moghaddam, B. (2013). Distinct prestimulus and poststimulus activation of VTA neurons correlates with stimulus detection. *Journal of Neurophysiology*, *110*(1), 75–85. https://doi.org/10.1152/jn.00784.2012

Unsworth, N., Miller, A. L., & Aghel, S. (2022). Effort mobilization and lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience*, *22*(1), 42–56. https://doi.org/10.3758/s13415-021-00941-6

Unsworth, N., & Robison, M. K. (2017). A locus coeruleus-norepinephrine account of individual differences in working memory capacity and attention control. *Psychonomic Bulletin & Review*, *24*(4), 1282–1311. https://doi.org/10.3758/s13423-016-1220-5

Vrizzi, S., Najar, A., Lemogne, C., Palminteri, S., & Lebreton, M. (2023, October 6). *Comparing the test–retest reliability of behavioral, computational and self-reported individual measures of reward and punishment sensitivity in relation to mental health symptoms*. Open Science Framework. https://doi.org/10.31234/osf.io/3u4gp

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, *50*(3), 433–441. https://doi.org/10.1518/001872008X312152

Westbrook, A., & Braver, T. S. (2016). Dopamine does double duty in motivating cognitive effort. *Neuron*, *89*(4), 695–710. https://doi.org/10.1016/j.neuron.2015.12.029

Wise, T., Emery, K., & Radulescu, A. (2024). Naturalistic reinforcement learning. *Trends in Cognitive Sciences*, *28*(2), 144–158. https://doi.org/10.1016/j.tics.2023.08.016

Zhang, Y., Chen, Y., Xin, Y., Peng, B., & Liu, S. (2023). Norepinephrine system at the interface of attention and reward. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *125*, Article 110751. https://doi.org/10.1016/j.pnpbp.2023.110751

Received November 25, 2023
Revision received December 2, 2024
Accepted December 14, 2024 ■