An evidence accumulation model of motivational and developmental influences over sustained attention

Harrison Ritz[1], **Joseph DeGutis**[2,3], **Michael J. Frank**[1], **Michael Esterman**[2,4], **& Amitai Shenhav**[1]
1. Department of Cognitive, Linguistic, and Psychological Sciences; Carney Institute for Brain Science, Brown University.
2. Geriatric Research, Education, & Clinical Center, VA Boston Healthcare System.
3. Department of Psychiatry, Harvard Medical School.
4. Department of Psychiatry, Boston University School of Medicine.

* Corresponding Author: hritz@brown.edu

Abstract

Sustaining focus is difficult, but it is under our control. Previous research has found that people's ability to sustain attention depends on external incentives and changes over the lifespan. However, previous research has made limited progress in characterizing the specific cognitive mechanisms involved in sustained attention. These mechanisms are investigated in the current experiment, which uses drift diffusion modeling to re-analyze a series experiments on sustained attention. In Experiment 1, we found that incentives influence information processing (noise) but not decision strategy (threshold). In Experiment 2, we found that noise and threshold have distinct development trajectories, and that while older adults have noisier accumulation, they are better at sustaining attention. These results help provide mechanistic insight into recent findings in sustained attention.

Keywords: attention; drift diffusion; motivation; development

Introduction

Achieving our goals often requires sustained attention, such as when ignoring distractions while writing. This focus is known to wane with time, a phenomenon called vigilance decrement (Fortenbaugh, DeGutis, & Esterman, 2017). This decrement is not a fixed limitation, as it can be reduced or even eliminated by incentives (Esterman, Reagan, Liu, Turner, & DeGutis, 2014; Esterman et al., 2016). Sustained attention also changes over the lifespan, coinciding with changes in cognitive ability and motivation (Fortenbaugh et al., 2015). However, the cognitive mechanisms driving differences in sustained attention remain unclear, including how and when people control information processing or response strategies.

Recently developed theories have suggested that sustained attention depends on actively inhibiting our default mind-wandering states (Fortenbaugh et al., 2017). These theories propose that intrusive thoughts introduce noise into task performance, and are actively suppressed by domain-general cognitive control. While this framework has provided a strong theoretical foundation for understanding the interaction between different cognitive systems, the attentional and decision processes driving sustained attention have largely evaded formal analysis (though see Hawkins, Mittner, Forstmann, & Heathcote, 2019).

A common approach for measuring sustained attention is to see how performance in speeded reaction time tasks change over time (Esterman, Noonan, Rosenberg, & Degutis, 2013; Figure 1a). These tasks can be modeled as an evidence accumulation process, notably with the drift diffusion model (DDM; Ratcliff (1978); Figure 1b). Critically, the DDM can distinguish information processing (e.g., drift rate or accumulation noise) from response strategies (e.g., decision threshold or starting point) based on the joint distribution of reaction time and accuracy.

Prior theoretical work has suggested that participants control their decision threshold (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006) or accumulation noise (Manohar et al., 2015) in order to maximize rewards and minimize effort (Shenhav, Botvinick, & Cohen, 2013). Given the general role of motivated control in sustained attention, and the specific role that this control may play in protecting information processing from the task-irrelevant noise that is central to the drift diffusion process, we sought to use the DDM to develop a process-level model of sustained attention. We re-analyzed a collection of recent experiments examining how sustained attention changes under incentives and over the lifespan, measuring how the different components processes involved in sustained attention are modified by internal and external control demands.

Experiment 1

The goals of Experiment 1 were to (1) validate our computational model of sustained attention and (2) measure how incentives influence different components of our model. These data were previously reported in two previous publications on the influence of reward on sustained attention (Esterman et al., 2014, 2016).

Methods

Participants 106 individuals participated in Experiment 1. Participants' accuracy was very high and negatively-skewed (median participant: 97%), so we excluded 7 participants who had less than 80% accuracy (bottom 5th percentile), leaving 99 participants for the final analysis. All participants provided informed consent in accordance with their local university's institutional review boards.

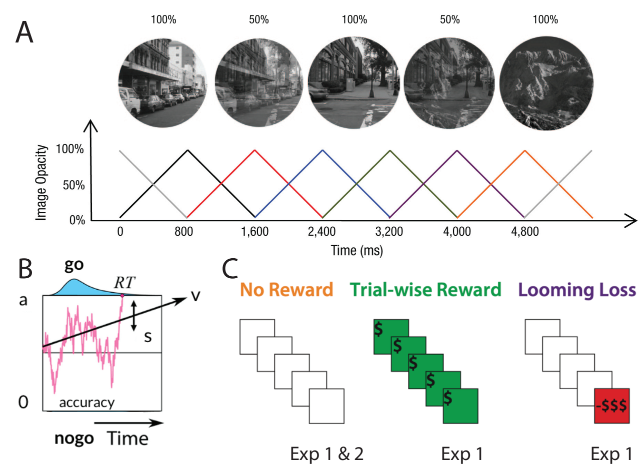

Figure 1: Task and model. (A) Participants saw images morph into one other, pressing a key for city images (go trial; 90%) and withholding a response to mountain images (nogo trial; 10%). Adapted from Fortenbaugh et al. (2017). (B) In the drift diffusion model, evidence is noisily accumulated over time with a drift rate *v* (here fixed to 1) and accumulation nose *s*. When the accumulated evidence hits the threshold level *a*, participants either make a response (upper bound) or withhold a response (lower bound). Adapted from Manohar et al. (2015). (C) In Experiment 1, participants were randomly assigned to one of three incentive conditions, either self-motivated ('no reward'), earning money on each trial for good performance ('trial-wise reward'), or faced a large loss for a critical commission error ('looming loss').

gradCPT task All of the experiments used the gradual continuous performance task (gradCPT; Esterman et al., 2013), a variant of the go-nogo task. Participants saw images of cities and mountains within a central aperture. Each image linearly morphed into the next one, with full opacity at 800ms (Figure 1a). Participants were required to make a key press for city images ('go'; 90% of trials), and withhold a keypresss for mountain images ('nogo'; 10% of trials). Participants had one minute of practice, followed by a single 10 minute session encompassing 750 trials.

Reward Manipulation Participants performed the gradCPT task under one of three different reward manipulations. Across all three conditions, the trial structure was identical: participants did not receive any reward feedback during the experiment. Participants in the 'no-reward' group (n=33) were not given any external incentives (i.e., were self-motivated).

Participants in the 'trial-wise reward' group (n=35) earned rewards on each trial for good performance. In one sub-group (n=18), participants received $\pm$ \$0.01 for accurate/inaccurate performance on go trials, and $\pm$ \$0.10 for accurate/inaccurate performance on nogo trials. In the second sub-group (n=17), participants received an endowment of \$18, and were told that they would lose \$0.25 for every commission error (response during nogo trials), and nogo-preceding omission error (no response during go trials). We pooled across these reward conditions, as they had similar first-order performance (see Esterman et al. (2016), figure 1), and similarly depended on a per-trial reward contingency.

Finally, participants in the 'looming loss' condition (n=31) received an endowment of \$18, and were instructed that they would lose their endowment if they made an error on a critical nogo trial, which would be indicated by a colored border. This critical trial always occurred after the first 10 minutes of the task, with only these first 10 minutes analyzed.

Evidence Accumulation Model We modelled participant's go-nogo performance using the drift-diffusion model (Ratcliff, 1978), a process model for decision-making that has previously been applied to the go-nogo task (Gomez, Ratcliff, & Perea, 2007). In our DDM, participants noisily accumulate evidence about the stimulus category (city or mountain), and make the corresponding go or nogo response when this evidence reaches a fixed threshold. Both the rate of accumulation ('drift rate') and threshold are normalized by accumulation noise. On trials in which participants made a response, we estimated the approximate joint probability of their response and reaction time (RT) under our model (Navarro & Fuss, 2009), and on trials where participants did not make a response, we estimated the probability of their response under our model.

Previous experiments have proposed that there is motivated control over accumulation noise during decision-making (Manohar et al., 2015), analogously to how motivated control reduces motor variability (Manohar, Finzi, Drew, & Husain, 2017; Pekny, Izawa, & Shadmehr, 2015), and consistent with neuroscientific evidence for top-down control over sensory noise (Nakajima, Schmitt, & Halassa, 2019). These findings, alongside observations from the original experiments that RT variability (but not mean RT) was sensitive to incentives and fatigue, motivated our focus on accumulation noise.

Our model contained six free parameters: threshold, go trial accumulation noise ('go-noise'), nogo trial accumulation noise ('nogo-noise'), lapse rate, and the (linear) rate of change in threshold and go-noise across the session (Figure 1b). We fixed the drift rate to 1 in all models, as it cannot be uniquely estimated alongside threshold and noise. The lapse rate determines the probability that response RTs were uniformly distributed and that no-response choices were random, improving the robustness of the model fit and an additional index of (in)attention. The threshold change and noise change parameters capture the difference in parameters between the final and first trial. We interpolated the overall threshold and noise parameters to the midpoint of the session for ease of interpretation.

We fit separate noise parameters for go and nogo trials to capture generic differences between these conditions, such as differences in familiarity or reactive stopping processes, similarly to previous models for the go-nogo task

(Ratcliff, Huang-Pollock, & McKoon, 2018). While the strong imbalance of trial types (90% go trials) would make starting point an alternative candidate, the small number of commission errors made this parameter difficult to reliably estimate.

We fit our model's parameters to participants' behavior using hierarchical Expectation-Maximization (EM; Huys et al. (2011)). EM alternates between maximum likelihood estimation of the parameters under a group-level multivariate normal prior (M-step), and updating this group prior based on the point estimates of participants' parameters (E-step), repeating these two steps until convergence. We reinitialized the parameters three times within each maximization step to improve robustness. We removed the first 16 trials of each session, as these were much slower and more variable than the rest of the session, as well as rare trials where RTs were faster than 400ms or slower than 1600ms.

Results

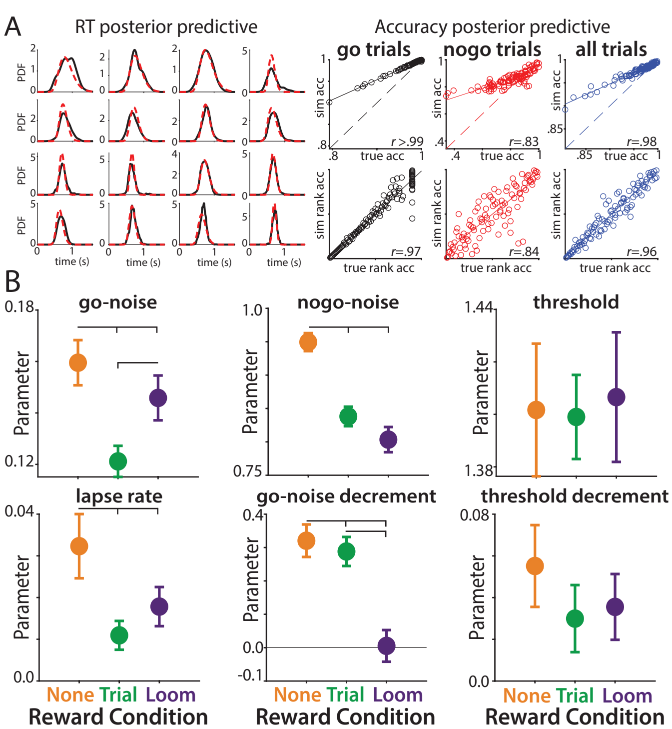

Figure 2: Experiment 1. (A) Left, RT distributions for a set of participants evenly spaced across the range of model likelihood (black), overlaid with the model-predicted RT distribution (red). Right, the correlation between individual differences in observed and model-predicted accuracy, both for real performance (top) and ranked performance (bottom). (B) DDM parameters across the three reward conditions. Contrasts indicate significant simple effects ($p < .05$); error bars are SEM.

Model Validation To validate our model of sustained attention, we first confirmed that we were able to successful recover our parameters. We were able to accurately estimate the parameters that generated simulated behavior within the range observed in participants (correlation between simulated and recovered parameters: all rs $> .85$).

Next, we confirmed that our model parsimoniously captured participants behavior. We fit a flexible model in which both drift rate and threshold parameters varied between go and nogo trials. We found that drift rate and threshold were highly correlated within go trials ($r = .74$), and differed across conditions by a similar ratio ($r = .69$), consistent with a noise parameter that normalizes both of these terms and differed across conditions.

Finally, we performed posterior predictive checks to confirm that our model captures the underlying behavior (Gelman, Meng, and Stern (1996); Frank et al. (2015). Our model appeared to capture participants' RT distribution across the range of model fit (Figure 2a). We also found that our model's predictions of individual differences in accuracy were highly correlated with ground-truth performance (go trials: $r > .99$; nogo trials: $r = .83$). However, our model systematically overestimated participants' accuracy, especially during nogo trials (Figure 2a). These discrepancies likely arise due to nogo trials' relatively small contribution to the overall fit (only making up 10% of trials), and trade-offs that fit participants' peaked RT distributions.

Reward Manipulation We found that go-noise ($F_{2,96} = 6.32$, $p = .0026$) and nogo-noise ($F_{2,96} = 23.4$, $p = 5.3\text{e-}9$; Figure 2b) significantly differed by reward condition. Participants in the no-reward conditions were significantly noisier than those in the reward conditions during both go and nogo trials, consistent with motivated control over accumulation noise (Manohar et al., 2015). In contrast, participants had highly similar thresholds across reward conditions ($F < 1$; $BF_{10} = .095$; Reward Group $\times$ Parameter (noise, threshold): $F_{2.7,131} = 8.59, p = 5.5\text{e-}5$).

We found inconclusive evidence for whether lapse rate varied across groups ($F_{2,96} = 3.96, p = .022$, below Holm- or Bonferroni-corrected thresholds). Lapse rate was correlated with both go-noise ($r = .61$) and nogo-noise ($r = .49$), but condition differences for noise parameters remained significant when controlling for lapse rate. It is plausible that noise terms and lapse rate similarly reflect task-irrelevant distraction.

Participants' go-noise decrement depended on their reward condition ($F_{2,96} = 13.0, p = 7.21\text{e-}6$). Participants in the no-reward and trial-wise reward conditions became noisier over the course of the experiment, whereas participants in the looming loss condition maintained the same level of performance through-out the experiment (Figure 2b). In contrast, there was little difference across reward conditions in how participants' threshold changed over time ($F < 1$, $BF_{10} = 0.15$; Figure 2b), with all participants becoming more conservative over the course of the experiment (MAP t-test:

$t_{98} = 3.79, p = 2.7\text{e-}4$; Reward Group $\times$ Parameter (noise, threshold): $F_{2,96} = 4.35, p = .016$).

Discussion

This experiment validated our model of sustained attention. We found that our model was identifiable, and did a good job capturing participants' reaction times and individual differences in accuracy, albeit with optimistic expectations for their accuracy. These results suggest that our relatively simple model was sufficient to largely capture effects of interest, and particularly how these effects depend on between-participant differences in reward condition.

Our model revealed that incentives influenced different components of the decision process. Whereas participants had a similar decision thresholds across reward conditions, we found that participants decreased their overall noise in reward conditions, and eliminated the change in noise in the looming loss condition. These effects demonstrate that our model was able to capture participants' selective control over different facets of the decision process in response to heightened motivation.

Given that our model appeared to reflect reliable and dissociable mechanisms involved in sustained attention, we next sought to replicate this model to a large online sample, using the diagnostic power of this model to investigate how sustained attention changes in response to the motivational and cognitive changes that occur over the lifespan.

Experiment 2

The goal of Experiment 2 was to (1) replicate our model findings in a large, well-powered sample (2) to measure how sustained attention changes over the lifespan. Changes in sustained attention over the lifespan may reflect several facets of attentional control, such as differences in motivation, judgements of ability, and subjective effort costs (Swirsky & Spaniol, 2019). Previous research has consistently found that older participants have higher decision thresholds (Starns & Ratcliff, 2010; Ratcliff, Thapar, & McKoon, 2001), but little is known about about how aging influences the decrement of mechanistic decision parameters over the course of an experimental session. By measuring how different components of our process model change over the lifespan, this experiment can help identify targets for control. Model-agnostic analyses of a subset of these data were previously reported in (Fortenbaugh et al., 2015).

Methods

Participants 21,409 participants took part in Experiment 2. Participants performed the gradCPT task online at TestMyBrain.org through voluntary sign-up. Participants did not complete any other tasks. We excluded 138 participants for accuracy less than 80%, a sub-group that roughly matched the age distribution of the sample as a whole. Participants were representative of a broad age range (Figure 3a). This experiment was approved by a local institutional review board.

Procedure and Analysis Participants performed the same task as in Experiment 1, hosted online at TestMyBrain.org. The web version of this task was 4 minutes long, encompassing 300 trials. We used an identical likelihood function as in Experiment 1, except that we did not use hierarchical priors for this analysis due to the large sample size, and to better measure individual differences. While these data are fundamentally cross-sectional, our large sample size provides good power to examine population-level changes across the lifespan.

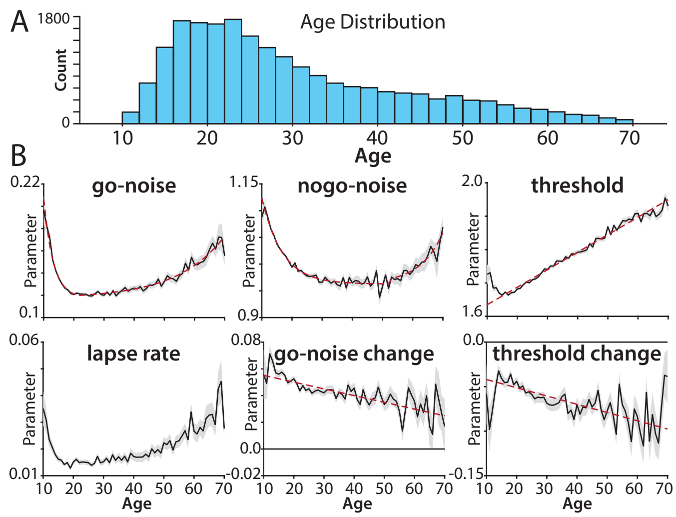

Figure 3: Experiment 2. (A) Age distribution across our sample. (B) Participants' estimated parameters plotted as a function of their age. Red dashed lines are predicted parameters from (non)linear regression models. Grey shaded area indicates SEM.

Results

Model Validation We found that the quality of model fit in Experiment 2 largely replicated our observations from Experiment 1. Model-generated performance was highly correlated with individual differences in accuracy (go trials: $r = .99$; nogo trials: $r = .88$), but was similarly optimistic.

Threshold and noise demonstrate dissociable relationships with age We observed qualitatively different developmental trajectories for threshold and noise (see Figure 3b). While threshold rapidly declined for the youngest participants, the trend was dominated by a linear increase in threshold over the lifespan. In contrast, noise developed non-linearly, decreasing early in development and increasing in development. Comparing the relationship between these parameters and age, we found that threshold had a stronger linear relationship with age than either go-noise (coefficient test: $F > 3000$) or nogo-noise ($F > 4000$; Figure 3b). These observations are consistent with observations that older adults have larger thresholds (value accuracy over speed; Starns and Ratcliff (2010)), and have noisier reaction time distributions (Ratcliff et al., 2001).

Different forms of noise have differential trajectories over the lifespan Closer inspection of the change in go- and nogo-noise over the lifespan reveal that they exhibited distinct developmental trajectories (Figure 3b). We modelled these trajectories using a six parameter piecewise exponential function, fitting separate growth curves before and after estimated transition ages using non-linear regression:

$$\text{noise} = \begin{cases} a \times \exp(b \times age) + c, & \text{if age} < t \\ a' \times \exp(b' \times (age - t)) + c, & \text{if age} \geq t \end{cases}$$

Go-noise rapidly decreased during adolescence (predicted age to reach 99% of asymptote: 24.6 years old), whereas nogo-noise developed more slowly (99% of asymptote: 41.2 years old; between-parameter t-test on early growth rate: $t = 43.2, p << .001$). In contrast, go-noise increased more slowly over the lifespan, whereas nogo-noise quickly increased later in life (between-parameter t-test on late growth rate: $t = 3.03, p = .0012$). These effects were robust to including lapse rate as a covariate.

Developmental trajectory of vigilance While participants in both experiments became noiser over the session, the rate at which participants got noisier decreased over the lifespan ($t = -13.7, p = 7.9\text{e-}43$). Interestingly, we found that within-session changes in threshold and noise traded-off against each other, such that participants were showed greater threshold change exhibited weaker go-noise change ($r = -.27$; Figure 3b), a feature that was absent in simulated behavior.

Whereas participants' Experiment 1 became more conservative within a session, participants in Experiment 2 became more liberal within a session ($t = -58.2, p << .001$). We found a linear trend across the lifespan indicating that older participants became liberal more quickly ($t = -11.8, p = 8.0\text{e-}32$; Figure 3b).

Discussion

We found that the same model which fit behavior in our laboratory experiment generalized to a large online sample, here additionally teasing apart developmental trajectories for different facets of sustained attention.

We first found that the decision-related (threshold) and attention-related (noise) components of task performance exhibited dissociable developmental trajectories. Whereas threshold exhibited a largely continuous increase over the lifespan, accumulation noise exhibited a U-shaped developmental course. These results support observations in the original paper that first-order behaviors cluster into 'strategy' (mean RT, criterion) and 'ability' (RT variability, sensitivity) factors (Fortenbaugh et al., 2015). Critically, the current analyses provide process-level insight into this factor structure.

While the trajectory of both noise parameters differed from threshold, they also differed from each other in important ways. Whereas go-noise developed quickly and decayed slowly, nogo-noise developed slowly and decayed quickly. These results are consistent with distinct control strategies for facilitation and suppression, and motivate future experiments examining how these different control strategies change over the lifespan.

Finally, we found gradual developmental changes in how attention was sustained within a session. Surprisingly, older adults were better than younger adults at sustaining their level of go-noise, contrary to classic research that found increased vigilance decrements over the lifespan (Parasuraman, Nestor, & Greenwood, 1989). While older adults were overall noisier than younger adults, they sustained attention better than similarly noisy adolescents. A potential explanation for this finding is that the older adults may have been more motivated to monitor their performance. Consistent with this interpretation, the original experiment found that older adults had stronger post-error slowing in this task (Fortenbaugh et al., 2015). This explanation is speculative, and future experiments should specifically investigate process-level contributions to vigilance over the lifespan.

Older participants' also showed a different threshold timecourse, shifting liberal more quickly than younger participants. Older participants' shift coincided with their higher overall thresholds, potentially reflecting a gradual optimization of threshold (though see: (Starns & Ratcliff, 2010)). Future work should compare participant against the reward-optimizing configuration (Bogacz et al., 2006; Manohar et al., 2015).

General Discussion

Across these two experiments, we found that our adapted drift diffusion model usefully captured participants' performance in an established sustained attention task, revealing dissociable changes in the underlying decision process elicited by different motivational states and developmental stages. These experiments demonstrate the utility of formalizing task performance with computational models that can carve up the underlying cognitive processes.

In Experiment 1, we found that participants reduced their accumulation noise when they were incentivized, and they could eliminate the vigilance decrement when facing a large looming loss. Our model was able to identify the targets of motivated control, dissociating the changes in decision threshold and accumulation noise that researchers have proposed are optimized for reward (Bogacz et al., 2006; Manohar et al., 2015). What our model does not address is how control is specified, such as why the looming loss condition eliminates the vigilance decrements. However, a potential explanation may come from the optimal control literature that motivates DDM optimization models (Manohar et al., 2015; Shenhav et al., 2013). An optimal policy depends on both the current reward and the expected future returns (Bellman, 1957), expected future returns that decrease over the course of the session for trial-wise rewards, but which plausibly remain constant or increase for the looming loss condition. This hypothesis should be formalized and tested in future experiments.

In Experiment 2, we found that our model parameters were further dissociated by their distinct developmental trajectories. Again, our approach usefully isolates the trajectory of different processes, but doesn't address why they change. Developmental changes in sustained attention likely reflect a myriad of causes, such as changes in strategy, motivation to engage in the experiment, and cognitive ability. Further progress towards addressing these fundamental questions require large-scale longitudinal experiments that combine psychometric and cognitive modelling of how control changes over the lifespan (Ferdinand & Czernochowski, 2018).

While the current set of experiments demonstrate the utility of this model, there are notable limitations to our approach. Our model strongly over-estimates participants' accuracy. One extension of our model could be to include a collapsing decision threshold, allowing us to capture participants' peaked reaction-time distributions at higher levels of accumulation noise. Another expansion could be to better capture the local adjustment to performance, such as repetition biases, rhythmic responding, post-error slowing, that may contribute to participants' accuracy (Urai, de Gee, Tsetsos, & Donner, 2019). Future work should provide a more complete picture of within- and across-trial dynamics, and confirm our characterization of how parameters depend on incentives and age.

These experimental results are consistent with models of attentional control that connect the motivated reduction of accumulation noise during decision-making to motivated reductions in motor noise during actions (Manohar et al., 2015, 2017). The current set of results validate and extend these theories, showing that the across-trial maintenance of noise is also under motivated control. These findings raise interesting theoretical questions about the algorithmic similarity of optimization processes across motor and cognitive control (Ritz, Frömer, & Shenhav, 2019).

References

Bellman, R. (1957). A markovian decision process. *Journal of Mathematics and Mechanics*, *6*(5), 679–684.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.*, *113*(4), 700–765.

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), e0159741.

Esterman, M., Noonan, S. K., Rosenberg, M., & Degutis, J. (2013). In the zone or zoning out? tracking behavioral and neural fluctuations during sustained attention. *Cereb. Cortex*, *23*(11), 2712–2723.

Esterman, M., Reagan, A., Liu, G., Turner, C., & DeGutis, J. (2014). Reward reveals dissociable aspects of sustained attention. *J. Exp. Psychol. Gen.*, *143*(6), 2287–2295.

Ferdinand, N. K., & Czernochowski, D. (2018). Motivational influences on performance monitoring and cognitive control across the adult lifespan. *Front. Psychol.*, *9*, 1018.

Fortenbaugh, F. C., DeGutis, J., & Esterman, M. (2017). Recent theoretical, neural, and clinical advances in sustained attention research. *Ann. N. Y. Acad. Sci.*, *1396*(1), 70–91.

Fortenbaugh, F. C., DeGutis, J., Germine, L., Wilmer, J. B., Grosso, M., Russo, K., & Esterman, M. (2015). Sustained attention across the life span in a sample of 10,000: Dissociating ability and strategy. *Psychol. Sci.*, *26*(9), 1497–1510.

Frank, M. J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J. F., & Badre, D. (2015). fMRI and EEG predictors of dynamic decision parameters during human reinforcement learning. *J. Neurosci.*, *35*(2), 485–494.

Gelman, A., Meng, X.-L., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Stat. Sin.*, 733–760.

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *J. Exp. Psychol. Gen.*, *136*(3), 389–413.

Hawkins, G. E., Mittner, M., Forstmann, B. U., & Heathcote, A. (2019). Modeling distracted performance. *Cogn. Psychol.*, *112*, 48–80.

Huys, Q. J. M., Cools, R., Gölzer, M., Friedel, E., Heinz, A., Dolan, R. J., & Dayan, P. (2011). Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. *PLoS Comput. Biol.*, *7*(4), e1002028.

Manohar, S. G., Chong, T. T.-J., Apps, M. A. J., Batla, A., Stamelou, M., Jarman, P. R., … Husain, M. (2015). Reward pays the cost of noise reduction in motor and cognitive control. *Curr. Biol.*, *25*(13), 1707–1716.

Manohar, S. G., Finzi, R. D., Drew, D., & Husain, M. (2017). Distinct motivational effects of contingent and noncontingent rewards. *Psychol. Sci.*, *28*(7), 1016–1026.

Nakajima, M., Schmitt, L. I., & Halassa, M. M. (2019). Prefrontal cortex regulates sensory filtering through a basal Ganglia-to-Thalamus pathway. *Neuron*.

Navarro, D. J., & Fuss, I. G. (2009). Fast and accurate calculations for first-passage times in wiener diffusion models. *J. Math. Psychol.*.

Parasuraman, R., Nestor, P., & Greenwood, P. (1989). Sustained-attention capacity in young and older adults. *Psychol. Aging*, *4*(3), 339–345.

Pekny, S. E., Izawa, J., & Shadmehr, R. (2015). Reward-dependent modulation of movement variability. *J. Neurosci.*, *35*(9), 4015–4024.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.*, *85*(2), 59–108.

Ratcliff, R., Huang-Pollock, C., & McKoon, G. (2018). Modeling individual differences in the go/no-go task with a diffusion model. *Decisions*, *5*(1), 42.

Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychol. Aging*, *16*(2), 323–341.

Ritz, H., Frömer, R., & Shenhav, A. (2019). Bridging motor and cognitive control: It's about time! *Trends Cogn. Sci.*, *0*(0).

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240.

Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise: Boundary optimality in the diffusion model. *Psychol. Aging*, *25*(2), 377–390.

Swirsky, L. T., & Spaniol, J. (2019). Cognitive and motivational selectivity in healthy aging. *Wiley Interdiscip. Rev. Cogn. Sci.*, e1512.

Urai, A. E., de Gee, J. W., Tsetsos, K., & Donner, T. H. (2019). Choice history biases subsequent evidence accumulation. *Elife*, *8*.