Arousal optimizes behavior by promoting latent state transitions

Harrison Marble (harrison_marble@brown.edu)[*]
Carney Institute for Brain Science, Brown University, 164 Angell Street
Providence, RI, 02912 USA

Tiantian Li (tiantian.li@maxplanckschools.de)[*]
Max Planck School of Cognition, Stephanstr. 1A
Leipzig, D-04103 Germany

Matthew R. Nassar (matthew_nassar@brown.edu)
Carney Institute for Brain Science, Brown University, 164 Angell Street
Providence, RI, 02912 USA

[*] co-first author

Abstract:

People are often faced with surprising events that defy expectations. These events are thought to elicit transient activity in the locus coeruleus/norepinephrine system and elevation of peripheral arousal markers such as the P300 in EEG and pupil dilation, but the function served by these arousal signals remains unclear. We propose that they facilitate latent state transitions that dynamically control the mental context governing learning and perception. To test this theory, we collected EEG and pupillometry data in a novel color prediction and reproduction task in two complementary contexts that prescribe opposite relationships between latent state transitions and learning. We found that stimuli with high state transition probability elicited pupil dilation and amplification of several event-related potentials including the P3a. Trial-to-trial variability in these signals was related to perceptual biases and learning, with heightened physiological signatures of arousal corresponding to decreased perceptual bias and context-dependent learning effects. Our findings support the theory that arousal-based LC/NE system signals optimize perception and learning by promoting global latent state transitions.

Keywords: arousal; pupil dilation; P3; learning; bias

Introduction

Navigating a dynamic world inevitably involves the violation of expectations, a phenomenon accompanied by surprise signals that trigger transient arousal responses. Arousal is in part mediated by the locus coeruleus/norepinephrine (LC/NE) system and can be indexed through peripheral physiological markers including pupil dilation and EEG signals such as the P3 event-related potential (ERP) (Nieuwenhuis, Aston-Jones, & Cohen 2005; Joshi et al. 2016; Vazey et al. 2018; Joshi & Gold 2020). Despite the behavioral and neural consequences of these phasic activations in the LC/NE system, their exact computational function remains unclear. Existing theories have proposed normative roles for these fluctuations in arousal. One theory suggests that the arousal system directly influences the lowered perceptual, memory, and choice biases (Urai et al. 2017; de Gee et al. 2017; Krishnamurthy et al. 2017). An alternative theory proposes that the LC/NE-related arousal signals play a role in modulation of learning (Devauges & Sara 1991; Yu and Dayan 2005; Nassar et al. 2012; Ghosh et al. 2021). However, recent empirical findings suggest that at least some relationships between transient arousal markers and learning depend on environmental context (Nassar, Bruckner, & Frank 2019). These findings have motivated the development of a new unifying theory for the arousal system which builds on theoretical and empirical research suggesting that learning and perceptual inference are improved by maintaining explicit representations of the latent causal process giving rise to observable data, which we refer to as latent states (Gershman & Niv 2013; Collins & Frank 2016). This new framework posits that transient activation of the LC/NE system partitions representations of latent states in time which optimizes both learning and perceptual inference (Razmi & Nassar 2022). The current work tested this unifying theory by leveraging statistical environments involving different latent state transition structures in a novel color reproduction and prediction task with concurrent EEG and pupil measurements.

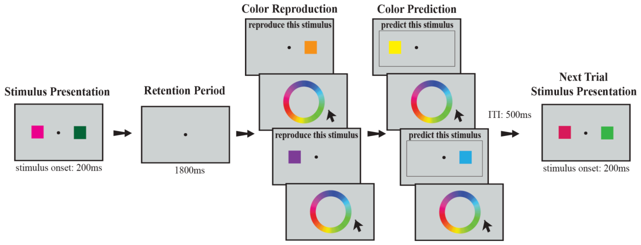

Figure 1: Task schematic. On each trial, participants go through three phases: stimuli presentation, color reproduction, and color prediction.

Methods

We collected behavioral, EEG (Brain Vision actiCHamp Plus 64, 1000Hz), and pupillometry (SR Research Eyelink 1000 Plus, 1000Hz) data from 57 subjects performing a novel color perception and prediction task (Li et al. 2023). On each trial, participants are presented with two colored stimuli for 200ms, and after a delay, prompted to reproduce the colors sequentially. After reproduction, participants predict the upcoming colors, which is possible because colors at each location are generated through an independent sequential process following either the changepoint (CP) or oddball (OB) state transition structure, depending on the block condition. Color reproductions were used to estimate biases toward predicted colors on each trial, which can be thought of as Bayesian integration of prior and likelihood information, and might emerge during perception or working memory storage (Krishnamurthy 2017). Color prediction updates, along with color prediction errors, were used to compute single trial measures of learning rate (Nassar 2012). To infer normative measures of state transition probability (STP) and uncertainty (belief entropy) we simulated behavior of a Bayesian ideal observer.

To investigate relationships between task parameters and pupil/EEG measures, we fit a linear model that contained terms for model calculated entropy, STP, block condition (1 CP, -1 OB), and a STP*condition interaction. Clusters were then formed by spatially or temporally connected data points exceeding a cluster forming threshold of P<0.005 in each regressor's t-statistic map, and corrected for multiple comparisons using permutation testing.

Results

Consistent with arousal systems encoding a surprise signal, higher STP elicited increased pupil dilation and several ERPs, including the P3a signal, when compared to standard stimuli. STP-mediated pupil dilation was largest around three seconds after stimulus onset (Figure 2B). We identified five ERP clusters that related to surprising stimuli, including a signal that resembled a frontal P3a (Figure 2C). These ERPs occurred ~300-1100 ms after stimulus onset (Figure 2D).

The trial-by-trial STP-related signals were correlated with decreased bias and adjusted learning to fit context. Individual trial pupil and EEG effects were quantified as the dot product of the measured signal on a trial and the cluster's t-statistic map. Both the pupil and EEG trial effects were normalized and summed to give an aggregate physiological signal (Figure 3A). When this single trial signal was separated into quantiles, we found that increased STP signal was associated with decreased bias in both conditions (p=0.001 in CP, p<0.001 in OB; Figure 3C), and affected learning bidirectionally, reflecting increased learning in the CP block (p<0.001) and decreased learning in the OB block (p=0.028; Figure 3B). Consistent with this observation, the aggregate signal could be added to a regression model to predict trial by trial adjustments in participant bias (Figure 3D) and learning (Figure 3E), with the latter best explained when the aggregate signal was allowed to affect learning in a condition-dependent manner. Individual ERPs had more nuanced relationships to behavior, with some, such as the P3a showing more prominent negative relationships to bias (Figure 3F), and others, such as a late parietal negative and a late frontal positive cluster relating more closely to contextual learning rate adjustments (Figure 3G).

Discussion

Taken together, our results demonstrate that peripheral markers of transient LC-NE arousal responses represent latent state transitions and predict behavioral dynamics related to learning and perceptual bias that are consistent with LC/NE playing a functional role in optimizing behavior by facilitating latent state updates. However, the current work relies on proxy measures of the arousal system which challenges the specificity of our LC/NE related signal interpretations. We hope our results motivate further investigations exploiting more direct measures of LC and pharmacological manipulations to test the specific biological and causal predictions of our arousal-latent state theory.

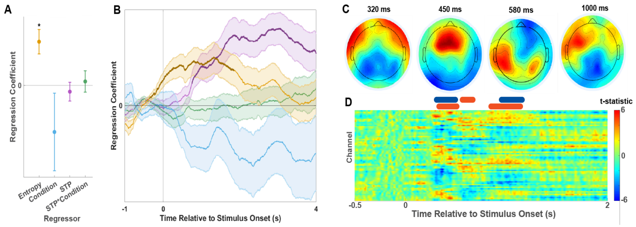

Figure 2: Pupil and EEG linear model results (A) Baseline pupil regression (B) Running pupil regression coefficients for STP (purple), entropy (yellow), condition (blue) and STP*condition (green) (C) topographical maps of the EEG STP coefficients (D) t-statistic map for STP coefficient of EEG regression, horizontal lines indicate cluster timepoints.

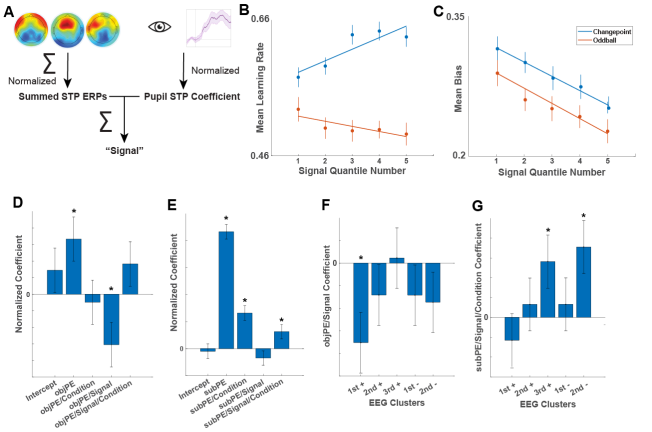

Figure 3: EEG and pupil signal data relates to behavior (A) EEG and pupil are combined to create total signal value (B-C) Learning rate and bias by individual trial signal quantile (D-E) Bias and learning circular model results for combined signal (F-G) Individual cluster regression coefficients for (F) objPE term of bias regression and (G) subPE term of learning regression.

Acknowledgments

We would like to thank Timothy Chen, Adit Sabnis and Sienna Bruinsma for help with EEG, pupillometry and behavioral data collection. This work is funded by NIH grant R01MH126971. The funder has no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

References

Collins, A. G. E. & Frank, M. J. (2016). Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning. Cognition 152, 160–169.

Devauges, V. & Sara, S. J. (1991). Memory retrieval enhancement by locus coeruleus stimulation: evidence for mediation by β-receptors. Behav. Brain Res. 43, 93–97.

de Gee, J. W. et al. (2017). Dynamic modulation of decision biases by brainstem arousal systems. Elife 6, 1–36.

Dimigen, O. (2020). Optimizing the ICA-based removal of ocular EEG artifacts from free viewing experiments. Neuroimage 207, 116117.

Gershman, S. J. & Niv, Y. (2013). Perceptual estimation obeys Occam's razor. Front. Psychol. 4, 1–11.

Ghosh, A., Massaeli, F., Power, K. D., Omoluabi, T., Torraville, S. E., Pritchett, J. B., Sepahvand, T., Strong, V. D., Reinhardt, C., Chen, X., Martin, G. M., Harley, C. W., & Yuan, Q. (2021). Locus Coeruleus Activation Patterns Differentially Modulate Odor Discrimination Learning and Odor Valence in Rats. Cerebral Cortex Communications, 2(2).

Joshi, S., Li, Y., Kalwani, R. M. & Gold, J. I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. Neuron 89, 221–234.

Joshi, S. & Gold, J. I. (2020). Pupil Size as a Window on Neural Substrates of Cognition. Trends Cogn. Sci. 24, 466–480.

Krishnamurthy, K., Nassar, M. R., Sarode, S. & Gold, J. I. (2017). Arousal-related adjustments of perceptual biases optimize perception in dynamic environments. Nat. Hum. Behav. 1.

Li, T., Marble, H., Contreras-Carerra, M., & Nassar, M. R. (2023). Does arousal optimize behaviour by promoting latent state transitions? Retrieved from osf.io/2wcxt

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. Nature Neuroscience, 15(7), 1040–1046.

Nassar, M. R., Bruckner, R. & Frank, M. J. (2019). Statistical context dictates the relationship between feedback-related EEG signals and learning. Elife 8, 1–26.

Nieuwenhuis, S., Aston-Jones, G. & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. Psychol. Bull. 131, 510–532.

Razmi, N., & Nassar, M. R. (2022). Adaptive Learning through Temporal Dynamics of State Representation. The Journal of Neuroscience, 42(12), 2524–2538.

Urai, A. E., Braun, A. & Donner, T. H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. Nat. Commun. 8, 1–11.

Vazey, E. M., Moorman, D. E. & Aston-Jones, G. (2018). Phasic locus coeruleus activity regulates cortical encoding of salience information. Proc. Natl. Acad. Sci. U. S. A. 115, E9439–E9448.

Yu, A. J. & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. Neuron 46, 681–692.