A Neural Network Model of Visual Attention Integrating Biased Competition and Reinforcement Learning

Jonathan Morgan* Badr Albanna* James P. Herman*

November 9, 2024

Abstract

We present a neural network model of visual attention (NNMVA) that integrates biased competition and reinforcement learning to capture key aspects of attentional behavior. The model combines self-attention mechanisms from Vision Transformers (ViTs), Long Short-Term Memory (LSTM) networks for working memory, and an actor-critic reinforcement learning framework to map visual inputs to behavioral outputs. The self-attention mechanism simulates biased competition among artificial neural representations, determining their influence on working memory, which in turn provides top-down feedback to guide attention. Trained end-to-end with reinforcement learning using reward feedback—paralleling learning processes in non-human primates—the NNMVA replicates key behavioral signatures of attention, such as improved performance and faster reaction times for cued stimuli. Manipulating the model's attention mechanisms affects performance in ways analogous to experimental manipulations in primate frontal eye fields and superior colliculus. Additionally, artificially inducing attentional biases alters value estimates and temporal difference (TD) errors, offering predictions about how attention may interact with dopamine-related learning signals in the brain. Our findings suggest that reward-driven behavior alone can account for several key correlates of attention, providing a computational framework to explore the interplay between attention and reinforcement learning in both biological and artificial systems.

*University of Pittsburgh, Department of Neuroscience, Pittsburgh, PA, USA

Contents

1 Introduction **3**

2 Results **5**
2.1 Orientation Change Detection Task Environment 6
2.2 Model Overview 7
2.3 Model Behavior 7
2.4 Evaluating Attention Bias in the Model 9

3 Effects of Artificially Modulating Attention Bias on Performance **11**
3.1 Manipulating Bias Affects Response Rate and Reaction Times 11
3.2 Manipulating Bias Influences Criterion and Sensitivity 13

4 Influence of Induced Bias on Value Estimates and Temporal Difference Errors **16**
4.1 Effects of Bias Manipulation on Temporal Difference Errors . . 16
4.2 Effects of Bias Manipulation on Value Estimates 17

5 Discussion **18**

6 Methods **20**
6.1 Model Overview 20
6.2 Biased Competition Module 21
6.3 Working Memory Module 23
6.4 Reinforcement Learning 24
6.5 Task Difficulty 25

1 Introduction

Visual attention is a fundamental cognitive function that enables organisms to selectively process some sensory information and ignore the rest. At the behavioral level, attended stimuli are associated with improved perceptual judgements and faster reaction times, compared to unattended [1–5]. At the neuronal level, attended stimuli are associated with increased spiking activity and decreased spike-count correlations compared to unattended [6–10]. Understanding the mechanisms underlying visual attention is crucial for bridging our knowledge of behavior and neural computations.

In the lab, Visual attention has been studied extensively in both humans and non-human primates (NHPs) using tasks incorporating cues which indicate behaviorally relevant stimulus locations [11]. Importantly, spatial cues are shown transiently before the presentation of visual stimuli to be detected or discriminated. Temporal separation of cues and stimuli is vital because cued and uncued stimuli do not differ in appearance. This allows any difference in behavioral responses to cued versus uncued stimuli to be ascribed exclusively to the subject's internal attentional state. However, for spatial cues to confer a behavioral benefit, it is thought that their location must be stored in visual working memory (VWM) during the interval between cue and stimulus presentation (REF).

The need to temporarily store the location of a spatial cue is just one example of the intimate interconnections between attention and VWM [12–14]. Attention plays a crucial role in selecting which information is encoded into working memory, and conversely, the contents of working memory can guide attentional focus [15–21]. This bidirectional relationship allows for the dynamic updating of information based on task demands and environmental changes [22–24].Visuo-spatial attention, in particular, may be involved in selecting spatial locations for encoding into working memory [12,25,26]. Moreover, some studies have shown that directing attention back to the spatial locations of memorized items enhances memory performance, suggesting that spatial attention helps refresh or strengthen working memory representations [27–30]. This spatial coupling allows for more efficient processing and recall of visual information by prioritizing locations of interest and integrating new sensory inputs with existing memory representations [21].

A large number of theories and models have been advanced to explain visual attention (for a comprehensive reivew, see REF). However, a major question that remains unaddressed is whether the behavioral hallmarks of

attention arise as a result of reward-seeking behavior. Some theories propose that attention and working memory may be guided by reward maximization [31–34], while others suggest factors such as minimizing surprise [35], reducing uncertainty [36], maximizing information [37,38], or minimizing free energy [39]. Experiments in monkeys have shown that reward manipulations affect neuronal responses and behavior similarly to attentional cues [40–42]. Work in humans has also revealed that reward value can guide attentional selection (for a review see Anderson 2015).

Traditionally, theories of attention and reward have emphasized how value estimates can influence attentional allocation, with attention being directed towards stimuli of higher expected value [32]. However, recent evidence shows that manipulating attention affects stimulus processing in basal ganglia circuits known to participate in value-based decision making [43]. This perspective opens up new avenues for understanding the neuronal mechanisms underlying attention and reward learning.

The findings highlighted above suggest that reward could account for neuronal and behavioral correlates of attention in humans and NHPs. However, no existing model (of which we are aware) integrates reward-based learning with attentional control. A model that learns to allocate attention based on reward feedback would open several new lines of research. First, it would closely parallel the training of NHPs, allowing for direct comparisons between model and monkey attention task behavior. Second, it would provide testable predictions about how experimentally manipulating attention in NHPs might affect reward value estimates, thereby exploring how attentional states can modulate reward learning—a concept that challenges traditional views and could reshape our understanding of attention-reward interactions. Finally, integrating reward-based learning with attentional control could have practical applications in artificial intelligence and machine learning, enhancing algorithms that require efficient allocation of computational resources.

Here we present a model incorporating an attention mechanism, memory, and reinforcement learning that is trained end-to-end on reward feedback in an attention task commonly used in NHPs. The model combines elements from Vision Transformers (ViTs) [44], Long Short-Term Memory (LSTM) networks [45, 46], and an actor-critic reinforcement learning framework [47]. The model replicates key behaviors observed in NHP experiments. When trained on a cued orientation-change detection task [11, 48], the model exhibits enhanced performance and faster reaction times for cued stimuli compared to uncued ones [1–5]. Manipulating model attention affects perfor-

mance in qualitatively similar ways to experimental manipulations in primate frontal eye fields (FEF) [49] and superior colliculus (SC) [50, 51]. Finally, we show that artificially inducing biases in the model modulates value estimates in a manner that depends on the attentional "state" of the model. Our findings suggest that reward-driven behavior alone may explain several key correlates of attention in NHPs, offering new insights into the relationship between attention, learning, and reward.

2 Results

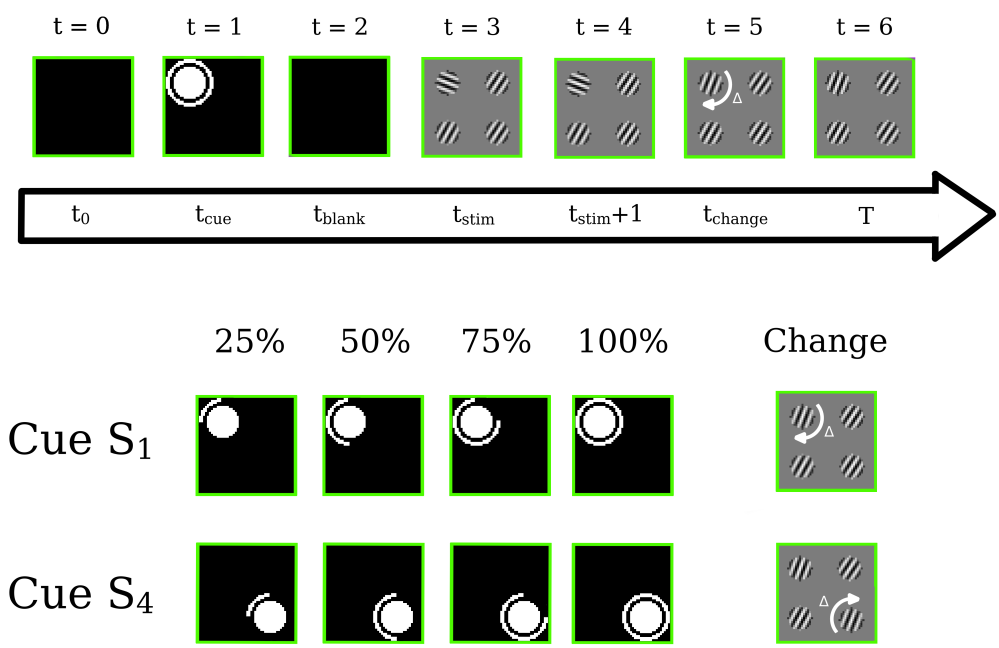

Figure 1: Each trial in the task comprises seven time steps. In each time step, a 50×50 grayscale image is input to the model. Black images are shown at $t = 0$ and $t = 2$. The cue is shown at $t = 1$ and can either be at S_1 (top left) or S_4 (bottom right). The cue can take four configurations, where the ring around the center disk changes, indicating the probability (25%, 50%, 75%, or 100% chance) that the change will appear at the cued location if the trial is a change trial.

2.1 Orientation Change Detection Task Environment

We trained our model on an orientation change detection task in which the agent must report if a change has occurred during each trial Figure 1. Each trial comprised 7 time steps. At each time step, a 50x50 grayscale image was shown to the agent. At $t = 0$, the trial began with a black image. At $t = 1$, a spatial cue was displayed. The cue appeared at stimulus position 1 (Cue S_1) on one half of the trials and at position 4 (Cue S_4) on the other half. At $t = 2$, a black screen was again displayed. At $t = 3$, stimulus onset occurred, displaying four "Gabor" stimuli in randomly chosen orientations at fixed positions: top left (S_1), bottom left (S_2), top right (S_3), and bottom right (S_4). At $t = 4$, the stimuli remained unchanged with the exception of orientation "noise" added in each time step of stimulus presentation to control task difficulty (see Methods). If the trial was a no-change trial, stimuli remain unchanged from $t = 4$ to $t = 6$. At $t = 5$, in a change trial, the orientation of one of the four stimuli changed by Δ degrees (where Δ varied from trial to trial). The orientations then remained unchanged from $t = 5$ to $t = 6$. Half of all trials were "change trials" and and the other half were "no-change trials" (balanced across cue presentation positions).

The spatial cue provided information about the probability of an orientation-change event occurring at its position. Four distinct cues indicated 25%, 50%, 75%, or 100% probability of an orientation change occurring at the cued stimulus position (conditioned on the trial being a "change trial"). The cue's orientation-change probability was indicated by a white arc that subtended 25%, 50%, 75% or 100% of a central disc's circumference (Figure 1). We adopt the convention established by previous work in referring to the event probability indicated by the cue as "cue validity" [11, 52–54].

At each time step, the agent could either choose to wait ($a^{(t)}$ = 'wait') or declare a change ($a^{(t)}$ = 'declare change'). If $t < 6$, waiting resulted in no reward ($r = 0$), and the trial advanced to the next time step. If $t = 6$, waiting rewarded the agent $r = 1$ in a no-change trial and $r = 0$ in a change trial. Declaring a change always ended the trial. If the agent correctly declared a change at $t \geq 5$, it received a reward $r = 1$. In all other cases, declaring a change resulted in $r = 0$.

Before describing the model's behavior, we first briefly overview model components to facilitate interpretation of the model's attention dynamics.

2.2 Model Overview

Our NNMVA consists of three primary modules: the *working memory module*, the *competition module*, and the *reinforcement learning module* (Figure 2). The input image, X, is explicitly split into 4 visual patches: $X = \{x_i\}_{i=1}^{4}$, and this 4-patch structure is retained in the competition and working memory modules. The primary benefit of our choice to structure the model's processing around stimulus positions is that it renders the competition module highly interpretable. Specifically, it allows the model's "attention map" to be viewed as a 4x4 array in which each element defines how each input patch, x_i, is weighted in its contribution to each of the 4 outputs of the competition module $\alpha_i \xi_i$. For example, if all the model's "attention" is on S_1, the output of the competition module will only contain information about the stimulus at S_1. Importantly, because of the "non-selective pathway" that bypasses the competition module, even when attention is fully allocated to a single position, information from other stimulus positions is retained. Full model details can be found in the Methods section.

2.3 Model Behavior

Our model exhibited orderly "psychometric" and "chronometric" functions with characteristic sigmoidal shapes commonly observed in human and NHP experiments Figure 3. Larger orientation-change Δ values were associated with increased hit rates and decreased reaction times, qualitatively comparable to those seen in countless human and NHP psychophysics experiments. Performance was similar for positive (clockwise) orientation changes and negative (counterclockwise) orientation changes. This "sanity check" indicates that the model was not learning to exploit idiosyncrasies in the stimulus representation but rather was detecting orientation changes as intended.

We found a clear "cueing effect" on model hit rates and reaction times, again recapitulating results from human and NHP literature [1, 5, 55–60].

As shown in Figure 3, cues assigning higher probabilities to a location increase the response rate when the change occurs at that location (Figure 3A) and slightly decrease reaction times (Figure 3B). The most significant effects are observed when the cue location differs from the change location: high-validity cues at non-target locations significantly decrease response rates (Figure 3C) and increase reaction times (Figure 3D). This suggests that the cue presented at time $t_{\text{cue}} = 1$ biases the competition among ANRs at the

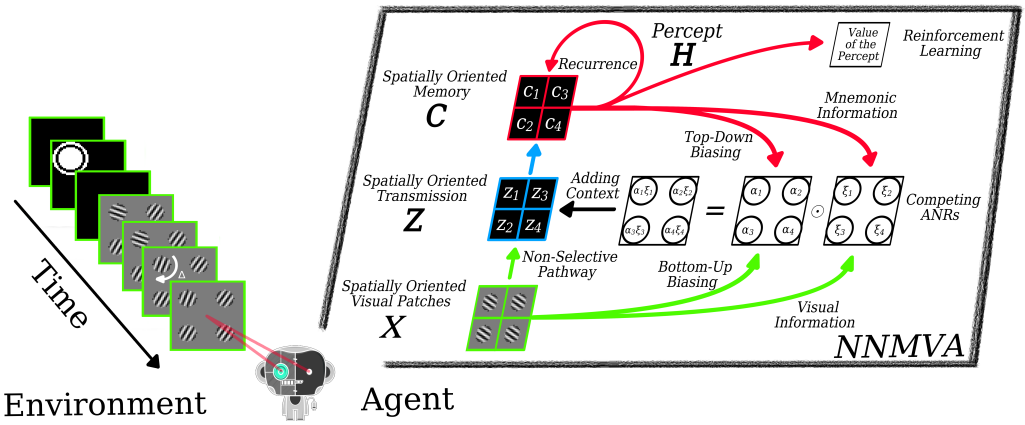

Figure 2: High level outline of the agent structure and NNMVA. The environment consists of the imagery described in Figure 1. The agent takes a single image as input at each timestep. The image (X) is initially split into four visual patches, each patch is associated with a stimulus (or cue) location at stimulus onset (or cue onset). These patches are pre-processed by a pre-trained convolutional neural network, resulting in the visual representation X, which is of lower dimension, but still spatially aligned with the original image. The visual representation X is then processed in parallel to construct bottom-up biases and supply immediate visual information to the competing ANRs, ξ_i. The previous percept H is used to supply top-down biasing as well as Mnemonic information to the competing representations. Importantly, the ANRs ξ_i competing for expression are generated at every timestep from memory and visual inputs. The global bias α_i determines the expression of the competitor ξ_i in the context projection, where α_i is the proportion of the context projection occupied by the competitor ξ_i. The visual input is integrated with this context projection via a non-selective pathway, i.e., no filtering of immediate visual inputs. Together, the context and visual inputs make up the transmission Z to working memory. The transmission and the previous percept are used to update the VWM state C, then a new percept is selected from the VWM state and this percept is projected to the reinforcement learning module, where one neural network determines the value of the percept, and another (not shown) determines the action to take which maximizes the possibility of reward.

time of change ($t_{\text{change}} = 5$).

2.4 Evaluating Attention Bias in the Model

In Figure 3G, we present averaged self-attention maps (in the form of heatmaps) over trials where we displayed cue at location S_1 and forced an orientation change of $\Delta = 15°$ at the S_1 location at the time of change ($t = 5$). The model learns to allocate a high bias ($\alpha_1^{(t_{\text{cue}})} \approx 1$) to the ANR corresponding to the cued location ($\xi_1^{(t_{\text{cue}})}$), transmitting cue information to all working memory components. This bias persists even when a blank image is shown at $t = 2$, indicating that the percept component, or mnemonic representation $h_1^{(t)}$, is driving the bias toward the ANR $\xi_1^{(t)}$, even in the absence of visual inputs (visual patch $x_1^{(t)} = \mathbf{0}$). This feature of our model resembles experimental results showing that spatially oriented working memory and spatially oriented attention rely on common neural processes [61–63]. At the time of the stimulus change, we observe the integration of stimulus-driven bias (due to the orientation change) and bias derived from the percept (resulting from higher-validity cues), consistent with theories of attentional control [64–69]. More specifically, we observe that even with a constant orientation change (Δ), cues with higher validity (down the rows in Figure 3G) at the time of cue result in higher bias for the cued location at the time of change (column of heatmaps at $t = 5$).

In Figure 3E and F we quantitatively demonstrate how the bias with respect to the S_1 location, $\alpha_1^{(t_{\text{change}})}$, relates to the magnitude of the orientation change Δ. When the change occurs at S_1, $\alpha_1^{(t_{\text{change}})}$ increases with larger Δ values (black curves), reaching near-maximal levels with high validity cues on S_1 in Figure 3F. Conversely, when the change occurs at S_4, higher Δ values lead to a decrease in $\alpha_1^{(t_{\text{change}})}$ (blue curves in Figure 3E and F). Low validity cues at the S_1 location and changes at the S_4 location result in decreased bias with respect to the S_1 location. For large Δ, this can cause the bias $\alpha_1^{(t)}$ to approach 0 (blue curve in Figure 3E). However, when a high validity cue occurs at the S_1 location and a change occurs at the S_4 location, the bias associated with the S_1 location is significantly higher, even as the orientation change is increased at the S_4 location (blue curve in Figure 3F). Together, these results illustrate that spatial processing is biased by the magnitude of the orientation change at the time of change and by the cue shown at the time of cue.

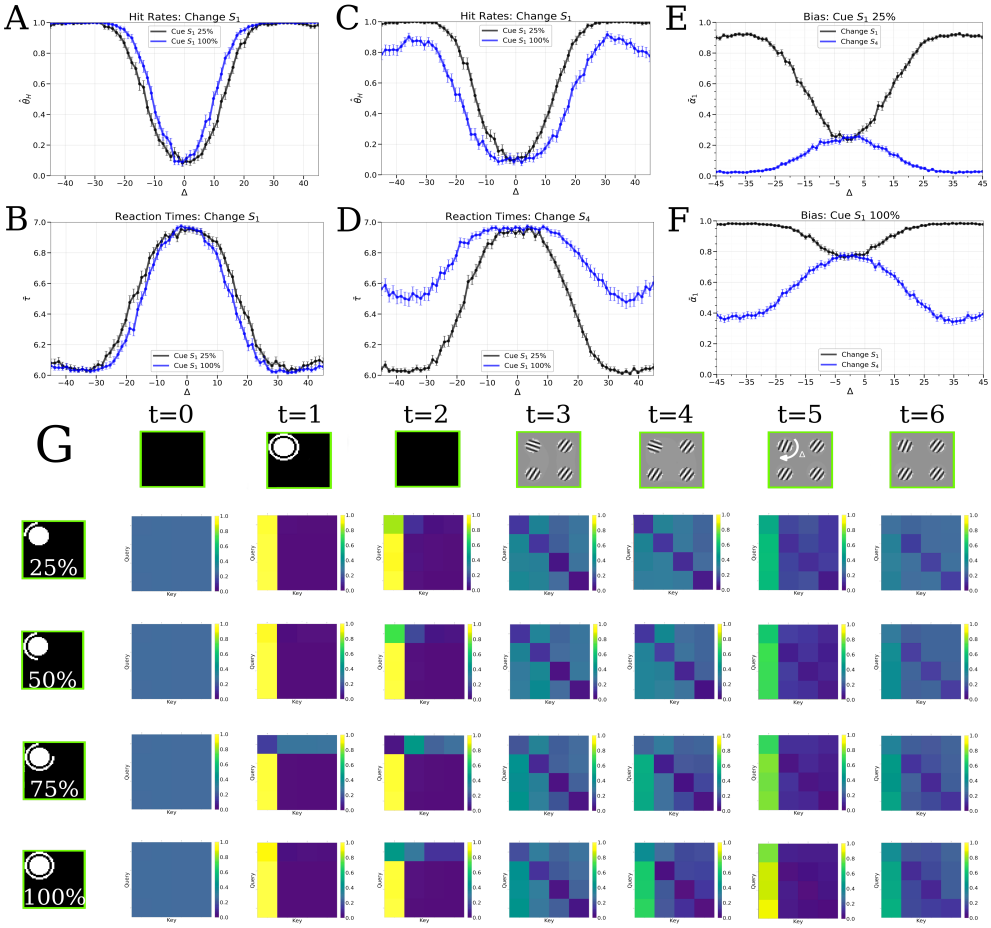

Figure 3: **A**–**F** shows the psychometric (**A** and **C**) and chronometric (**B** and **D**) evaluation of our agent for either a 25% or 100% cue on the S_1 location and either a change on S_1 or a change on S_4. Each data point was 500 trials where Δ specifies the magnitude and direction of the orientation change. The hit-rate was computed as $\hat{\theta_H} = n_H/(n_H + n_M)$, where n_H is the total number of hits and n_M is the total number of misses. The reaction times were computed as $\hat{\tau} = 1/500 \sum_i \tau_i$, where τ_i is the time the trial ended, either by the agent declaring a change or waiting through the final timestep. **A** and **B** show the effects of a 25% (blue curves) or 100% (black curves) cue on hit-rates (**A**) and reaction times (**B**) for changes at the S_1 location. **C** and **D** show the same for changes at the S_4 location. **E** and **F** plot the global bias $\alpha_1^{(t_{change})}$ on the competing ANR $\xi_1^{(t_{change})}$ for a 25% cue on S_1 (**E**) and a 100% cue on S_1 (**F**). Changes occurred at location S_1 (black curves) or location S_4 (blue curves). **G** shows the self-attention maps (heatmaps), describing the bias assigned to different competing ANRs, averaged over 500 trials. Each row of heatmaps corresponds to a specific cue and each column to a specific trial time. Only cues and changes at S_1 are shown. An orientation change of $\Delta = 15°$ occurred in every trial.

3 Effects of Artificially Modulating Attention Bias on Performance

3.1 Manipulating Bias Affects Response Rate and Reaction Times

Experimental studies have shown that microstimulation of brain regions involved in attentional control, such as the frontal eye fields (FEF), lateral intraparietal area (LIP), and superior colliculus (SC), can bias visual processing toward specific retinotopic locations in the visual scene [49–51, 70]. For example, microstimulation of the FEF enhances the neural representation of stimuli at corresponding spatial locations, leading to improved detection performance and faster reaction times for stimuli presented at those locations [49]. Similarly, microstimulation of the SC has been shown to modulate both perceptual sensitivity and decision criterion, affecting how stimuli are processed and how responses are generated [51].

In our model, artificially increasing the bias $\alpha_i^{(t)}$ assigned to specific competing artificial neural representations (ANRs) $\xi_i^{(t)}$ within the competition module can generate similar effects. By increasing the bias toward a particular compeititor $\xi_i^{(t)}$, we enhance the expression of the corresponding ANR in the transmission $Z^{(t)}$ to the working memory module. This results in the prioritization of information from the visual patch $x_i^{(t)}$ and the percept component $h_i^{(t)}$. As demonstrated in Figure 4A–C, when we artificially increase the bias in favor of the change location (e.g., setting $\alpha_1^{(t_{\text{change}})} = 1$ at the time of the orientation change), we observe a slight increase in hit rates for detecting the change. This suggests that enhancing the attentional bias toward the relevant location facilitates the processing of the stimulus change, leading to improved detection performance.

Conversely, when we increase the bias toward a location different from where the change occurs, we see a significant decrease in hit rates (Figure 4A–C). Directing bias away from the relevant stimulus location where a change occurred impairs detection performance, due to the ANR spatially oriented with the change location being absent in the transmission to the working memory module. The inhibitory effect of misallocated attention is further reflected in reaction times (Figure 4D–F), where inducing a high bias at non-change locations results in significant increases in response latency. These findings are consistent with behavioral studies showing that invalid

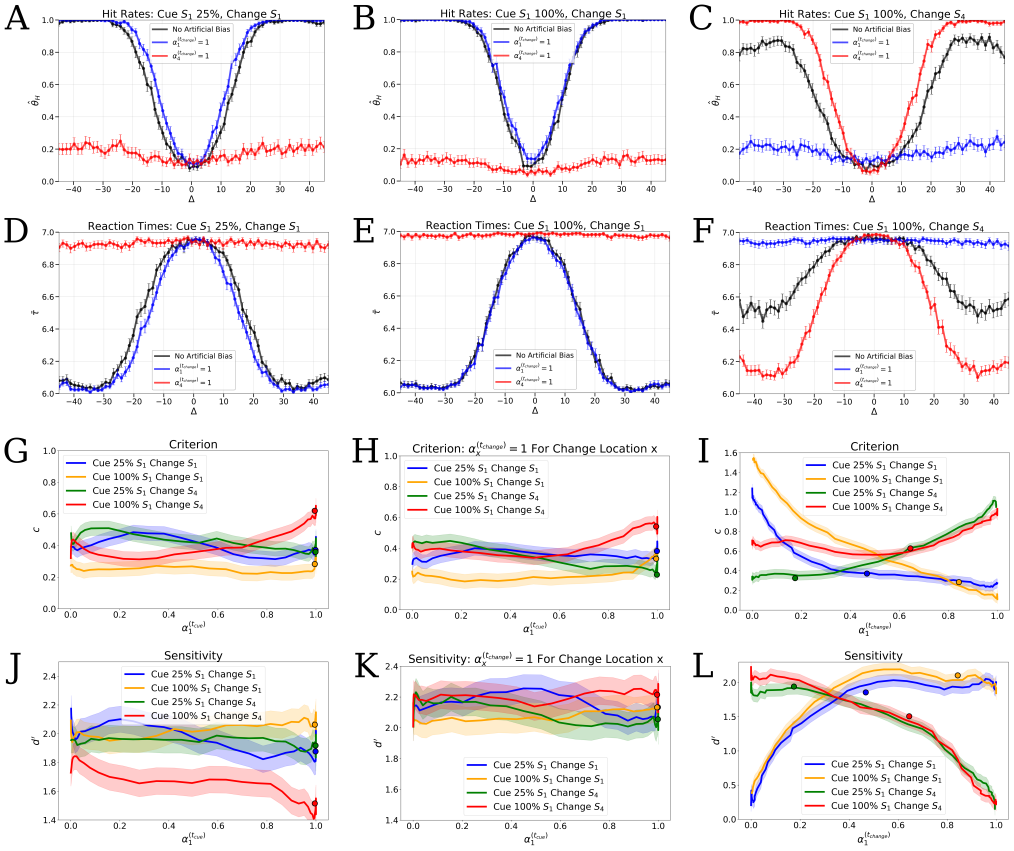

Figure 4: Plots showing the effect of artificially modulating the bias. All data points are the result of an average over 500 trials. Artificial modulation involves inducing a high bias in one of the columns of the self-attention map (Figure 3G). In all cases, the bias is induced on the competitors with respect to the S_1 ($\alpha_1^{(t)}$ and $\xi_1^{(t)}$) location or the S_4 ($\alpha_4^{(t)}$ and $\xi_4^{(t)}$) location. In **A**–**C**, we plot the hit rates versus the orientation change Δ. If $\alpha_i^{(t_{\text{change}})} = 1$, this indicates that the transmission $Z^{(t_{\text{change}})}$ has been completely biased toward $\xi_i^{(t_{\text{change}})}$, i.e., there are no traces of other competitors in the transmission. We manipulate the bias to this extreme for either the $\xi_1^{(t_{\text{change}})}$ or $\xi_4^{(t_{\text{change}})}$ competitors in **A**–**C** and then again for the reaction time plots in **D**–**F**. In **G**–**I**, we show the criterion measured (Equation 2) after manipulating the bias with respect to the S_1 location at the time of cue ($\alpha_1^{(t_{\text{cue}})}$ in **G** and **H**) or at the time of orientation change ($\alpha_1^{(t_{\text{change}})}$ in **I**). The colored dots indicate the average criterion measured without bias manipulations. The criterion was evaluated over 500 trials where there was a 50% chance of the change occurring at the indicated locations; otherwise, no change occurred. This procedure is repeated in **J**–**L** with the exception that we are measuring the sensitivity (Equation 1). In **H** and **K**, we modulate the bias with respect to the S_1 location at the time of cue ($\alpha_1^{(t_{\text{cue}})}$), but then immediately following the cue, we maximize the bias for the location where the change will occur ($\alpha_x^{(t_{\text{change}})}$ for a change on S_x, $x \in \{1, 4\}$). This is to prevent the secondary biasing effects that the cue induces at change time ($t = 5$ column in Figure 3G) from affecting our comparisons.

spatial cues or misdirected attention can slow down responses and reduce accuracy [11, 52–54].

3.2 Manipulating Bias Influences Criterion and Sensitivity

To quantify the effects of bias modulation further, we analyzed the decision criterion and sensitivity measures under different bias manipulations. In signal detection theory, an observer's ability to distinguish between signal and noise is measured using two key metrics: sensitivity (d') and criterion (c). Sensitivity measures how well the observer can distinguish between signal and noise, calculated using the hit rate ($\hat{\theta}_H$) and the false alarm rate ($\hat{\theta}_{FA}$). The formula for d' is:

$$d' = Z(\hat{\theta}_H) - Z(\hat{\theta}_{FA}), \tag{1}$$

where Z represents the inverse cumulative distribution function (CDF) of the standard normal distribution, also known as the z-score (not to be confused with the transmission $Z^{(t)}$ in our model). Criterion measures the observer's decision bias, indicating the tendency to favor one type of response over another:

$$c = -\frac{1}{2}\left[Z(\hat{\theta}_H) + Z(\hat{\theta}_{FA})\right]. \tag{2}$$

We assessed these metrics for individual stimulus locations by modifying the task environment so that only the stimulus location in question could experience an orientation change, with change trials occurring with probability 0.5. For example, if we are interested in the criterion of the agent's response with respect to stimulus location S_4 given a 100% validity cue at S_1, then we run multiple trials with the cue at S_1 and only allow orientation changes to occur at S_4. We then compute Equations (1) and (2) over the obtained behavioral data.

In the literature, there is ongoing debate regarding how attention modulates criterion and sensitivity in visual tasks. Some studies have found that attention can modulate both the decision criterion and sensitivity, indicating that attention enhances perceptual processing and decision-making [41]. Specifically, Luo and Maunsell [41] showed that attentional cues not only improve the ability to discriminate between stimuli (increasing sensitivity) but also affect the observer's response bias (criterion). Contrastingly, recent experimental work suggests that presaccadic attention may predominantly

influence the decision criterion rather than sensitivity. For instance, Gupta et al. [71] found that shifts of attention just before eye movements adjust the criterion without significantly affecting sensitivity, implying that attention reallocations linked to saccades might alter decisional strategies more than perceptual encoding. Other studies propose that strategically modulating the decision criterion can indirectly enhance sensitivity. Wang et al. [72] argue that criterion shifts, when optimized, can lead to improvements in sensitivity by better aligning the decision boundary with the signal distribution, particularly in tasks with asymmetric signal distributions or unequal prior probabilities. This suggests a complex interplay between criterion and sensitivity, where attentional mechanisms can influence perceptual performance both directly and indirectly.

Our findings contribute to this debate by showing that artificially manipulating the bias in our model can differentially affect criterion and sensitivity. In Figure 4G and J, we manipulated the bias at the time of cue t_{cue} with respect to the cue location (e.g., adjusting $\alpha_1^{(t_{\text{cue}})}$ for a cue at S_1). We observed that maximizing the bias at the cue location, for a high validity assigning cue, increases the criterion and decreases sensitivity when the change occurs at a different location (red curves in Figure 4G and J). This suggests that the model's attentional focus at the cue time can affect its ability to detect changes elsewhere. Notably, under natural conditions without artificial bias manipulations, attention is already maximized at the cue location (dots in Figures 4G, H, J, K).

We know from Figure 3G that spatial bias at the time of change is affected by the attributes of the cue. For example, a 100% valid cue at the S_1 location (bottom row in Figure 3G) at the time of cue ($t = 1$ column of heatmaps) leads to higher bias on the cued location at the time of change ($t = 5$ column in Figure 3G). To isolate the effect of bias at the time of cue from secondary effects at the time of change, we conducted additional experiments where we maximized the bias for the change location at t_{change} (Figures 4H and K). This negated the secondary bias effect resulting from the cue and offers more equality in comparisons. We found that the cue's effect on the criterion remained largely unchanged (Figure 4H), but the impact on sensitivity was abolished (Figure 4K), indicating that the decrease in sensitivity observed in Figure 4J is a result of spatial bias being directed away from the change location at the time of change.

In Figures 4I and L, we demonstrate that manipulating the bias at the

time of change ($\alpha_1^{(t_{\text{change}})}$) significantly influences critetion and sensitivity. Increasing bias toward the change location decreases the criterion and increases sensitivity (blue and gold curves in Figure 4I and L), while increasing bias toward non-change locations has the opposite effect (green and red curves in Figure 4I and L). This occurs because increasing $\alpha_i^{(t_{\text{change}})}$ amplifies the expression level of the spatially corresponding ANR $\xi_i^{(t)}$. If the change occurred at stimulus location S_i at time $t = t_{\text{change}}$, then the change information would initially be localized in $\xi_i^{(t_{\text{change}})}$, since $xi_i^{(t)}$ is the only competing ANR that consists of information from the immediate visual patch $x_i^{(t)}$. Increasing $\alpha_i^{(t_{\text{change}})}$ increases the signal strength, or expression, of this change information in the transmission $Z^{(t)}$ to working memory. Since biasing the ANR associated with the change location is directly amplifying the most task relevant information in the transmission to VWM, it seems more fitting to describe attention at the time of change as mostly modulating sensitivity, and that criterion decreases is a result of conversions from misses to hits. This effect mirrors the enhanced perceptual sensitivity observed in microstimulation experiments [49, 51].

It is important to note that these results are not capable of resolving the debate about the attentional effects on criterion and sensitivity. Rather, they show that the effects of attention on criterion and sensitivity can be quite complex, and possibly task dependent. For example, the criterion modulation shown in Figure 4 G and H is a result of cue information being transmitted to working memory, where this information has changed the structure of the ANRs in VWM. The cue information has changed the decodability of percepts transmitted to the reinforcement learning module, affecting behavior in a way that results in a criterion shift. Bias at the time of change mostly results in sensitivity changes. However, bias assignment at the time of change involves multiple interactions. When percepts are drawn from VWM Figure 2 at the time of change, they interact with immediate visual inputs to bias attention on the spatial ANRs associated with the cue location. Moreover, to observe an orientation change there must be stored information of the original stimulus orientations in working memory (since these are drawn randomly). This information also interacts with visual inputs to assign bias to potential change locations Figure 3E and F. These results and the dynamics displayed by our model demonstrate that criterion and sensitivity effects may involve multiple pathways and interactions between working memory and attention.

4 Influence of Induced Bias on Value Estimates and Temporal Difference Errors

Reinforcement learning (RL) provides a computational framework for modeling how agents learn to make decisions through interactions with their environment [47]. In neuroscience, RL has been utilized to explain how animals and humans adapt their behavior based on reward feedback [73, 74]. A key component of RL is the *value function* $V(H^{(t)})$ (Equation 9), which estimates the expected cumulative future reward from a given percept $H^{(t)}$ at time t. Here, $H^{(t)}$ is the percept constructed from working memory (Figure 2). The value function is updated using the *temporal difference (TD) error* δ_t (Equation 10), which quantifies the discrepancy between expected and received rewards. In the brain, dopaminergic neurons have been associated with encoding TD errors [75, 76]. Fluctuations in dopamine release have been found to correspond to reward prediction errors, influencing synaptic plasticity and learning processes [77]. Attention mechanisms, which prioritize certain sensory inputs over others, could be motivated by value [31, 32, 34, 78, 79] while also potentially influencing value estimates [80–82]. Biasing attention toward specific stimuli may thus modulate the computation of TD errors.

4.1 Effects of Bias Manipulation on Temporal Difference Errors

In Figure 5, we present the TD errors and value estimates from our model under various conditions. Figure 5A compares the TD errors when a cue is presented at location S_1, and a change occurs either at S_1 or S_4. When a high-validity cue is shown at a location different from where the change occurs (dashed red line in Figure 5A), we observe significant suppression of the TD errors. Figures 5B and 5C demonstrate the influence of bias manipulations on the TD errors. For low-validity cues, maximizing the bias on the change location (setting $\alpha_1^{(t_{change})} = 1$: red curves in Figure 5B and C) results in significant increases in the TD error (red line in Figure 5B). However, for high-validity cues, even a maximized bias has negligible effects (overlapping curves in Figure 5C). When we force a high bias at locations other than the change location (Figures 5E and 5F), we observe a nearly complete suppression of the TD errors. In summary, the TD errors significantly increase with respect to the bias when the biased location and the change location coincide

(red, blue, and green dashed lined in Figure 5D). However, when the change location and the assigned bias do not coincide, or a high validity cue was shown at a location different than the change location, we see a significant suppression of TD errors (dashed blue, dashed red, and dashed purple lines in Figure 5D).

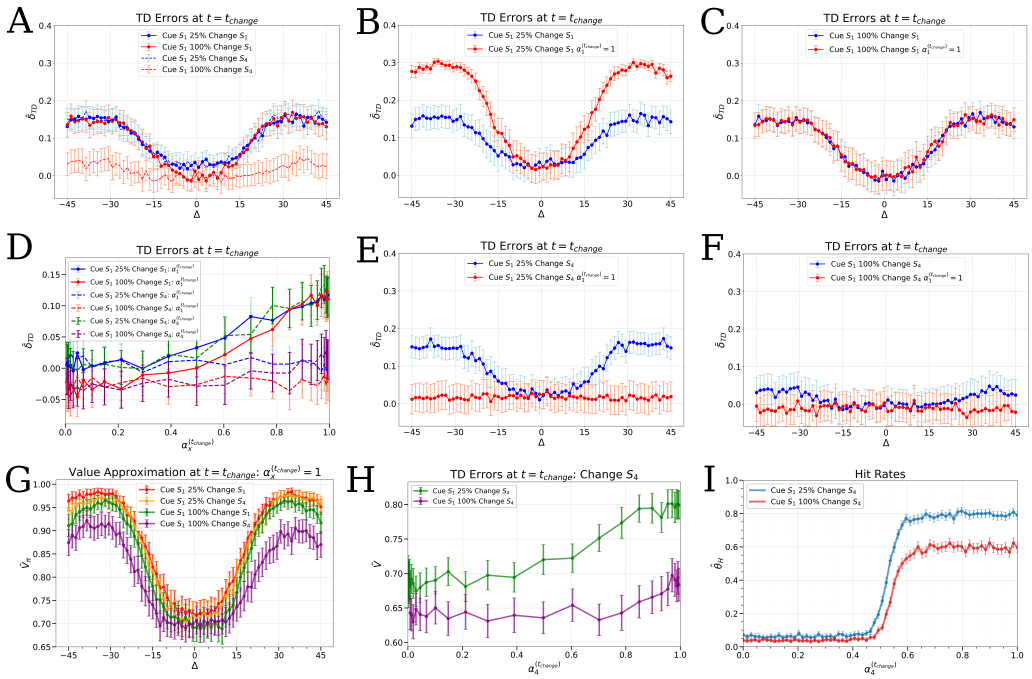

Figure 5: Plots showing the TD errors (δ_{TD} from Equation 10 in **A**–**F**) and value estimates (V_π from Equation 9 in **G** and **H**). Data points are averaged over 500 trials. Cues always occur at the S_1 location and changes occur at either the S_1 or S_4 location. In **D**, bias is modulated with respect to the bias indicated in the legend, i.e., $\alpha_x^{(t_{change})} = \alpha_4^{(t_{change})}$ for the dashed purple line. In **G**, bias is maximized with respect to the change location for all Δ shown, i.e., for the purple curve $\alpha_4^{(t_{change})} = 1$. In **I** the hit-rate is shown for modulation of bias $\alpha_4^{(t_{change}}$ and changes on S_4

4.2 Effects of Bias Manipulation on Value Estimates

The cue location, probability assignment by the cue, and the induced bias all affect the TD errors at the time of change. Similar to the method used to gen-

erate Figures 4H and 4K, we sought to isolate the effects of the cue from the secondary bias manipulation the cue induces at the time of change. Therefore, in Figure 5G, we varied the cue and change conditions but maximized the bias on the change location at the time of change (setting $\alpha_1^{(t_{\text{change}})} = 1$ for a change occurring at location S_1). A high-validity cue on a location different from the change location causes a decrease in value estimates, even for large orientation changes Δ (purple lines in Figure 5G). We also show that the value estimate increases approximately linearly with the bias, $\alpha_4^{(t_{change})}$, when there is a low-validity cue at an opposing location and the change location (S_4) differs from the cued location (green line in Figure 5H). The value estimation is significantly suppressed when there is a high-validity cue on a location different from the change (purple line in Figure 5H).

Considering the results shown in Figures 5G and 5H together with those in Figure 4H, we might be tempted to attribute the increased decision criterion resulting from high validity cues to a decrease in perceived value, since our model's action selection is based solely on value estimates. However, in Figure 5I, we demonstrate that there is a sharp increase in response rates before any such increase in value estimates (red curve in Figure 5I compared to the purple curve in Figure 5H). This suggests that there is some level of generalization beyond that of value estimates in our model's action selection policy.

5 Discussion

In this study, we developed a neural network model of visual attention (NNMVA) that embodies the functional components of attention as described by Knudsen [64], integrating both bottom-up and top-down processes within a biased competition framework [83]. Our model leverages self-attention mechanisms inspired by Vision Transformers [44, 84] to simulate competition among artificial neural representations (ANRs), and incorporates an LSTM-like working memory module [45, 46] that provides feedback and top-down biases.

By training the model using reinforcement learning with reward feedback [47], we emulated how attentional biases can be shaped by reward contingencies, mirroring learning processes in the brain. The NNMVA exhibits key behavioral signatures of visual attention observed in experimental studies, such as improved performance and faster reaction times for cued stimuli

compared to uncued ones [1–5]. Furthermore, we found that the self-attention map demonstrated priority map-like configuration from which we could easily measure and manipulate the bias towards some representations over others.

Artificially manipulating the bias within the competition module affected the model's performance in ways analogous to microstimulation experiments in the frontal eye fields (FEF) and superior colliculus (SC) of primates [49–51]. These manipulations could increase or decrease response rates and reaction times, depending on the location of an orientation change and the ANRs being biased. Furthermore, manipulating the bias in our model also affected criterion and sensitivity measurements. While the exact modulations are likely task dependent, we found that our model accomplished these modulation through multiple interacting pathways. Assining a high bias on the cue location at the time a high validity cue was shown resulted in a criterion increase and sensitivity decrease at non-cued locations. However, the sensitivity decrease was caused by a high bias being assigned to non-cued locations during the time of change, an effect from cue information being stored in working memory. We also demonstrated that manipulating the bias at the time of change significantly affects sensitivity, and that this is a result of either amplifying or suppressing the relevant change detection signals that the reinforcement learning module requires to make optimal decisions.

Notably, inducing bias in favor of certain spatially oriented ANRs affected the model's value estimates and temporal difference (TD) errors, depending on which representations were biased. This finding aligns with the idea that attention can influence reinforcement learning signals [32, 80, 82], suggesting that biased competition mechanisms may interact with value-based learning in the brain. Our model may offer predictions about microstimulation in a biased competition framework, where artificially biasing neural activity could affect dopamine-related TD error signals, with the specific impact depending on the neural representations being biased.

By integrating the biased competition model into an end-to-end computational framework, our work contributes to a deeper understanding of the computational principles underlying visual attention. The model's incorporation of spatial context and recurrent feedback mirrors biological attention mechanisms, providing a platform for generating testable predictions about neural and behavioral responses to attentional manipulations. However, our model has limitations that warrant further investigation. The simplicity of the task environment and the discretization of visual inputs may not capture the full complexity of natural visual scenes. Additionally, while our model

is biologically inspired, it does not account for all the intricacies of neural processing in the brain. Future work could focus on extending the model to more complex and naturalistic tasks, incorporating more detailed neural architectures, and exploring the role of attention in multisensory integration.

The implications of our findings extend to both artificial intelligence and cognitive neuroscience. For AI, incorporating biologically plausible attention mechanisms, such as top-down biases to the self-attention mechanism, could enhance the performance of transformer neural networks in reinforcement learning tasks involving partial observability. In neuroscience, our model may offer a computational tool for exploring how attention and reward-based learning interact, potentially informing experimental designs and the interpretation of neurophysiological data. Our neural network model provides a computationally and biologically grounded framework for understanding visual attention through the lens of biased competition. By demonstrating how attentional biases can influence both perception and value-based learning, we offer insights into the neural mechanisms of attention and open avenues for collaborative advances in computational modeling and empirical research.

6 Methods

6.1 Model Overview

Our NNMVA consists of three primary modules: the *working memory module*, the *competition module*, and the *reinforcement learning module*. These modules interact dynamically to simulate the processes of visual attention, where multiple stimuli compete for representation and attentional biases influence the outcome of this competition. In this section, we define the key components of our model: the visual inputs X (green boxes in Figure 2), the competitors ξ_i, the transmission Z (blue boxes in Figure 2), the visual working memory (VWM) state C (red boxes in Figure 2), and the percept H (red arrows in Figure 2). Except for the competitors, each of these is composed of subcomponents: $X = \{x_i\}_{i=1}^{n_{\text{patch}}}$, $Z = \{z_i\}_{i=1}^{n_{\text{patch}}}$, $C = \{c_i\}_{i=1}^{n_{\text{patch}}}$, and $H = \{h_i\}_{i=1}^{n_{\text{patch}}}$.

In line with the theory and experimental observations discussed above, working memory is spatially oriented. In implementation, this means we maintain multiple recurrent hidden states (c_i in Figure 2), each one associated with a spatial region in the visual scene (x_i in the imagery generated by the

environment). In our orientation change task, we split the image into 4 visual patches. Thus, we utilize 4 VWM components (each box is a component in X, Z, and C). Utilizing multiplicative interactions, we construct the competitors (ξ_i) utilizing the percept and visual components. These are also spatially oriented (ξ_i is derived from x_i and h_i). Two parallel pathways make up the construction of the transmission, Z. The first is a non-selective pathway where the visual patches x_i enter into the transmission components z_i in a spatially oriented manner. The second pathway is one that adds "context" to these visual inputs via the biased selection of the competitors (self-attention [84]). Notably, this is the only point in the NNMVA where spatial information is mixed. Without the self-attention mechanism, the VWM components would have no information outside of their associated visual patches. The transmission, along with the previous percept is used to update the VWM state. Finally, the percept H serves as the "state" for the reinforcement learning module, which approximates the value and chooses an action based on said state. From here on, we will also add the superscript t to each of the components and states to indicate the trial time. This is an important distinction because the visual state, transmission, VWM state, and competitors evolve with respect to time.

6.2 Biased Competition Module

The competitors $\xi_i^{(t)}$ are generated at every timestep. These artificial neural representations (ANRs) are constructed by a function $\xi_i^{(t)} = \Xi(x_i^{(t)}, h_i^{(t-1)})$ that takes as input the current visual patch $x_i^{(t)}$ and the previous percept component $h_i^{(t-1)}$, where Ξ is parameterized by a neural network. Thus, the competitor $\xi_i^{(t)}$ is spatially oriented toward the patch $x_i^{(t)}$ in the visual scene, with the content of $\xi_i^{(t)}$ influenced by both immediate visual inputs and memory. The competitors compete for expression in an outgoing transmission $Z^{(t)}$ from the competition module to the VWM module. While this is standard in transformer architectures [44], it also shares relevance to the biased competition framework [83] and Knudsen's functional components of attention [64]. Information from the visual scene, $S^{(t)}$, is expressed in the transmission $Z^{(t)}$ via a non-selective pathway. The non-selective pathway is similar to that described in the Guided Search 6.0 model [85, 86].

The competition for expression is mediated by self-attention weights $a_{ij}^{(t)}$, which determine how much each competitor $\xi_j^{(t)}$ influences the transmission

component $z_i^{(t)}$ with respect to i^th visuo-spatial location:

$$z_i^{(t)} = x_i^{(t)} + \sum_{j=1}^{n_{\text{patch}}} a_{ij}^{(t)} \, \xi_j^{(t)}, \tag{3}$$

The attention weights $a_{ij}^{(t)}$ are computed based on the compatibility between the visual inputs and their spatially corresponding percepts, similar to the self-attention mechanism in vision transformers [44, 84]:

$$a_{ij}^{(t)} = A\left(x_i^{(t)}, x_j^{(t)}, h_i^{(t-1)}, h_j^{(t-1)}\right), \tag{4}$$

where $A(\cdot)$ is a function parameterized by a neural network that computes the attention weights using the input features and percept components. This mechanism aligns with the biased competition framework of attention [83], where multiple stimuli compete for neural representation, and attentional biases influence the outcome of this competition. The self-attention mechanism with feedback from the percept follows the five forms of guidance outlined by the Guided Search 6.0 model [86]. Bottom-up guidance is implemented via the bias derived from the immediate visual patches $x_i^{(t)}$ and $x_j^{(t)}$, when $i = j$. Top-down guidance is implemented via the bias derived via the percept components $h_i^{(t-1)}$ and $h_j^{(t-1)}$. Prior biases affect the current biases through the percept, i.e., bias at an earlier point in time can affect the percept (and the bias derived from a percept) at later points in time. Finally, the self-attention mechanism also assigns bias relative to the structure in the visual scene (bias derived from the relationship between x_i and x_j, when $i \neq j$ in Equation 4).

The attention weights are normalized for each i:

$$\sum_{j=1}^{n_{\text{patch}}} a_{ij}^{(t)} = 1. \tag{5}$$

This normalization [87] is the implementation of competition for expression in each of the outgoing transmission components. To quantify the overall expression of each competitor $\xi_j^{(t)}$ in the transmission $Z^{(t)}$, we define the bias as:

$$\alpha_j^{(t)} = \frac{1}{n_{\text{patch}}} \sum_{i=1}^{n_{\text{patch}}} a_{ij}^{(t)}, \tag{6}$$

where the bias $\alpha_j^{(t)}$ represents the total influence of $\xi_j^{(t)}$ across all transmission components $z_i^{(t)}$ at time t.

6.3 Working Memory Module

The working memory module is responsible for maintaining and updating the visual working memory (VWM) state $C^{(t)}$ over time and constructing the outgoing percept $H^{(t)}$. Each VWM component $c_i^{(t)}$ is updated based on the previous percept component $h_i^{(t-1)}$ and the input from the transmission component $z_i^{(t)}$, using the update rule described by Hochreiter's LSTM architecture [45, 46]:

$$c_i^{(t)} = f_i^{(t)} \odot h_i^{(t-1)} + u_i^{(t)} \odot \psi_i^{(t)}, \tag{7}$$

where $\odot$ denotes element-wise multiplication. Here:

- $f_i^{(t)} = F\left(z_i^{(t)}, h_i^{(t-1)}\right)$ is the forget gate determining how much of the previous memory to retain, with $0 \leq f_i^{(t)} \leq 1$. The forget gate acts as a decay rate for each entry of the VWM component $c_i^{(t)}$, consistent with findings on adaptive decay in working memory [88, 89].
- $u_i^{(t)} = U\left(z_i^{(t)}, h_i^{(t-1)}\right)$ is an interaction term deciding how new information should interact with existing information, with $-1 \leq u_i^{(t)} \leq 1$.
- $\psi_i^{(t)} = \Psi\left(z_i^{(t)}, h_i^{(t-1)}\right)$ is the candidate memory content derived from the transmission component $z_i^{(t)}$ and the previous percept component $h_i^{(t-1)}$.

Each of these functions is parametrized by a neural network. By extracting information from the visual representations $\psi_i^{(t)}$ that have spatial identifiers and using this to update the corresponding VWM components, the network can store both the "what" and the "where" in the visual scene [90] along with the spatial and temporal context provided by the transmission component $z_i^{(t)}$.

We derive the percept $H^{(t)}$ [91] from the VWM state $C^{(t)}$. Each of the percept components $h_i^{(t)}$ is constructed by the following rule:

$$h_i^{(t)} = \phi_i^{(t)} \odot \left(\frac{c_i^{(t)}}{n_i^{(t)}}\right), \tag{8}$$

with:

- $\phi_i^{(t)} = \Phi\left(z_i^{(t)}, h_i^{(t-1)}\right)$ acting as an output gate controlling the flow of information from the memory component $c_i^{(t)}$.
- $n_i^{(t)}$ being a normalization factor ensuring that the percept components are appropriately scaled [46].

Here, the function $\Phi(\cdot)$ is another neural network. The percept $H^{(t)} = \{h_i^{(t)}\}$ provides top-down feedback to the competition module while also serving as input to the reinforcement learning module. This aligns with the ideas presented by Knudsen [64], in which a working memory representation is the basis for decision making and planning.

6.4 Reinforcement Learning

To emulate learning processes observed in non-human primates (NHPs) and humans, our model is trained using a reinforcement learning (RL) framework [47]. In this framework, the agent interacts with its environment by observing visual stimuli and taking actions to maximize the expected cumulative rewards over time. The goal of the RL framework is to enable the model to learn optimal policies that maximize future rewards based on the agent's perceptual inputs and previous experiences. The input to the RL module is the percept $H^{(t)}$, which encapsulates the relevant features of the visual scene as represented in the working memory module. This percept serves as the input to both the action-selection policy $\pi(H^{(t)})$ and the value function $V_\pi(H^{(t)})$, both of which are parameterized by neural networks in our model.

The action-selection policy $\pi(H^{(t)})$ maps the current percept to an action that the agent will take at time t. In our environment, the agent has two possible actions:

$$\begin{aligned} \pi(H^{(t)}) &= 0 \quad (\text{“wait” action}), \\ \pi(H^{(t)}) &= 1 \quad (\text{“declare change” action}). \end{aligned}$$

The “wait” action implies that the agent decides not to make a response and continues to process further information from the environment. The “declare change” action represents the agent's decision to identify a change in the visual stimulus. The policy network is trained to maximize the expected future rewards by choosing the action that is predicted to have the highest

value. The value function $V_\pi(H^{(t)})$ estimates the expected cumulative future reward from the current percept $H^{(t)}$:

$$V(H^{(t)}) = \mathbb{E}\left[\sum_{\tau=t}^{T} \gamma^{\tau-t} r_\tau \,\middle|\, H^{(t)}\right], \tag{9}$$

where $\gamma \in [0, 1]$ is the discount factor, r_τ is the reward received at time step τ, and T is the terminal time step for the task. The value function predicts how much reward the agent expects to receive by following its learned policy π from the current percept.

Learning in the RL module is driven by the temporal difference (TD) error, which measures the difference between the predicted value and the actual reward received at each time step:

$$\delta_t = r_t + \gamma V(H^{(t+1)}) - V(H^{(t)}), \tag{10}$$

where r_t is the reward received at time t and δ_t is the TD error. This error signal is used to update the value function $V_\pi(H^{(t)})$ and the action-selection policy $\pi(H^{(t)})$ to better predict future rewards and make more optimal decisions. In our model, both the value function and the policy are parameterized by neural networks. The agent's performance improves over time as it receives rewards and updates its predictions based on experience, thereby learning to allocate bias and make decisions that maximize cumulative rewards.

6.5 Task Difficulty

To control task difficulty, Gabor stimuli were corrupted with rotational "noise". Defnining θ_i^* as the "true" Gabor orientation for S_i, the orientation in the input image shown to the agent is:

$$\theta_i = \theta_i^* + \delta_{it}$$

where $\delta_{it} \sim \mathcal{N}(0, \sigma)$ is the rotational noise at time step t. If the stimulus is selected for change, then at $t = 5$ and $t = 6$:

$$\theta_i = \theta_i^* + \Delta + \delta_{it}$$

The orientation noise parameter σ is set to 5. The orientation change parameter Δ is a random variable drawn at the beginning of a change trial,

with $\Delta \sim \mathrm{U}(-k, k)$, where k is adjusted based on the agent's performance, starting at $k = 65$ and decreasing as performance improves to increase task difficulty.

References

[1] Marisa Carrasco. Visual attention: The past 25 years. *Vision research*, 51(13):1484–1525, 2011.

[2] Kelsey Clark, Ryan Fox Squire, Yaser Merrikhi, and Behrad Noudoost. Visual attention: Linking prefrontal sources to neuronal and behavioral correlates. *Progress in neurobiology*, 132:59–80, 2015.

[3] James E Hoffman. Visual attention and eye movements. *Attention*, pages 119–153, 2016.

[4] Roopali Bhatnagar and Jacob L Orquin. A meta-analysis on the effect of visual attention on choice. *Journal of Experimental Psychology: General*, 151(10):2265, 2022.

[5] Nicole C Rust and Marlene R Cohen. Priority coding in the visual system. *Nature Reviews Neuroscience*, 23(6):376–388, 2022.

[6] Carrie J McAdams and John HR Maunsell. Effects of attention on the reliability of individual neurons in monkey visual cortex. *Neuron*, 23(4):765–773, 1999.

[7] Carrie J McAdams and John HR Maunsell. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area v4. *Journal of Neuroscience*, 19(1):431–441, 1999.

[8] Alexander Thiele and Mark A Bellgrove. Neuromodulation of attention. *Neuron*, 97(4):769–785, 2018.

[9] Marlene R Cohen and John HR Maunsell. Attention improves performance primarily by reducing interneuronal correlations. *Nature neuroscience*, 12(12):1594–1600, 2009.

[10] Douglas A Ruff and Marlene R Cohen. Attention increases spike count correlations between visual cortical areas. *Journal of Neuroscience*, 36(28):7523–7534, 2016.

[11] Michael I Posner, Charles R Snyder, and Brian J Davidson. Attention and the detection of signals. *Journal of experimental psychology: General*, 109(2):160, 1980.

[12] Edward Awh, Edward K Vogel, and S-H Oh. Interactions between attention and working memory. *Neuroscience*, 139(1):201–208, 2006.

[13] Adam Gazzaley and Anna C Nobre. Top-down modulation: bridging selective attention and working memory. *Trends in cognitive sciences*, 16(2):129–135, 2012.

[14] Anastasia Kiyonaga and Tobias Egner. Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychonomic bulletin & review*, 20:228–242, 2013.

[15] Klaus Oberauer. Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3):411, 2002.

[16] Fiona McNab and Torkel Klingberg. Prefrontal cortex and basal ganglia control access to working memory. *Nature neuroscience*, 11(1):103–107, 2008.

[17] Nancy B Carlisle, Jason T Arita, Deborah Pardo, and Geoffrey F Woodman. Attentional templates in visual working memory. *Journal of neuroscience*, 31(25):9315–9322, 2011.

[18] Dirk van Moorselaar, Jan Theeuwes, and Christian NL Olivers. In competition for the attentional template: Can multiple items within visual working memory guide attention? *Journal of Experimental Psychology: Human Perception and Performance*, 40(4):1450, 2014.

[19] Nick Berggren and Martin Eimer. Visual working memory load disrupts template-guided attentional selection during visual search. *Journal of Cognitive Neuroscience*, 30(12):1902–1915, 2018.

[20] Nancy B Carlisle and Árni Kristjánsson. How visual working memory contents influence priming of visual attention. *Psychological Research*, 82:833–839, 2018.

[21] Freek van Ede, Sammi R Chekroud, and Anna C Nobre. Human gaze tracks the focusing of attention within the internal space of visual working memory. *Journal of Vision*, 19(10):133b–133b, 2019.

[22] A Caglar Tas, Steven J Luck, and Andrew Hollingworth. The relationship between visual attention and visual working memory encoding: A dissociation between covert and overt orienting. *Journal of experimental psychology: human perception and performance*, 42(8):1121, 2016.

[23] Brett Bahle, Valerie M Beck, and Andrew Hollingworth. The architecture of interaction between visual working memory and visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 44(7):992, 2018.

[24] Daniela Gresch, Sage EP Boettcher, Freek van Ede, and Anna C Nobre. Shifting attention between perception and working memory. *Cognition*, 245:105731, 2024.

[25] Mary E Wheeler and Anne M Treisman. Binding in short-term visual memory. *Journal of experimental psychology: General*, 131(1):48, 2002.

[26] Fabiano Botta and Juan Lupiáñez. Spatial distribution of attentional bias in visuo-spatial working memory following multiple cues. *Acta psychologica*, 150:1–13, 2014.

[27] Edward Awh, John Jonides, and Patricia A Reuter-Lorenz. Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3):780, 1998.

[28] Edward Awh, Lourdes Anllo-Vento, and Steven A Hillyard. The role of spatial selective attention in working memory for locations: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, 12(5):840–847, 2000.

[29] Melonie Williams, Pierre Pouget, Leanne Boucher, and Geoffrey F Woodman. Visual–spatial attention aids the maintenance of object representations in visual working memory. *Memory & cognition*, 41:698–715, 2013.

[30] Sami R Yousif, Monica D Rosenberg, and Frank C Keil. Using space to remember: Short-term spatial structure spontaneously improves working memory. *Cognition*, 214:104748, 2021.

[31] Okihide Hikosaka, Kae Nakamura, and Hiroyuki Nakahara. Basal ganglia orient eyes to reward. *Journal of neurophysiology*, 95(2):567–584, 2006.

[32] John HR Maunsell. Neuronal representations of cognitive state: reward or attention? *Trends in cognitive sciences*, 8(6):261–265, 2004.

[33] Clayton Hickey, Leonardo Chelazzi, and Jan Theeuwes. Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience*, 30(33):11096–11103, 2010.

[34] Michel Failing and Jan Theeuwes. Selection history: How reward modulates selectivity of visual attention. *Psychonomic bulletin & review*, 25(2):514–538, 2018.

[35] Simone Vossel, Christoph Mathys, Klaas E Stephan, and Karl J Friston. Cortical coupling reflects bayesian belief updating in the deployment of spatial attention. *Journal of Neuroscience*, 35(33):11532–11542, 2015.

[36] Ilya E Monosov. How outcome uncertainty mediates attention, learning, and decision-making. *Trends in neurosciences*, 43(10):795–809, 2020.

[37] Wayne W Schmaedeke. Information-based sensor management. In *Signal processing, sensor fusion, and target recognition II*, volume 1955, pages 156–164. SPIE, 1993.

[38] Jacqueline Gottlieb, Mary Hayhoe, Okihide Hikosaka, and Antonio Rangel. Attention, reward, and information seeking. *Journal of Neuroscience*, 34(46):15497–15504, 2014.

[39] Thomas Parr and Karl J Friston. Working memory, attention, and salience in active inference. *Scientific reports*, 7(1):14678, 2017.

[40] Jalal K Baruni, Brian Lau, and C Daniel Salzman. Reward expectation differentially modulates attentional behavior and activity in visual area v4. *Nature neuroscience*, 18(11):1656–1663, 2015.

[41] Thomas Zhihao Luo and John HR Maunsell. Attentional changes in either criterion or sensitivity are associated with robust modulations in lateral prefrontal cortex. *Neuron*, 97(6):1382–1393, 2018.

[42] Thomas Zhihao Luo and John HR Maunsell. Attention can be subdivided into neurobiological components corresponding to distinct behavioral effects. *Proceedings of the National Academy of Sciences*, 116(52):26187–26194, 2019.

[43] James P Herman, Fabrice Arcizet, and Richard J Krauzlis. Attention-related modulation of caudate neurons depends on superior colliculus activity. *eLife*, 9:e53998, sep 2020.

[44] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

[45] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[46] Maximilian Beck, Korbinian Pöppel, Markus Spanring, Andreas Auer, Oleksandra Prudnikova, Michael Kopp, Günter Klambauer, Johannes Brandstetter, and Sepp Hochreiter. xlstm: Extended long short-term memory. *arXiv preprint arXiv:2405.04517*, 2024.

[47] Richard S Sutton. Reinforcement learning: An introduction. *A Bradford Book*, 2018.

[48] Ramanujan Srinath, Douglas A Ruff, and Marlene R Cohen. Attention improves information flow between neuronal populations without changing the communication subspace. *Current Biology*, 31(23):5299–5313, 2021.

[49] Tirin Moore and Katherine M Armstrong. Selective gating of visual signals by microstimulation of frontal cortex. *Nature*, 421(6921):370–373, 2003.

[50] James Cavanaugh and Robert H Wurtz. Subcortical modulation of attention counters change blindness. *Journal of Neuroscience*, 24(50):11236–11243, 2004.

[51] James Cavanaugh, Bryan D Alvarez, and Robert H Wurtz. Enhanced performance with brain stimulation: attentional shift or visual cue? *Journal of Neuroscience*, 26(44):11347–11358, 2006.

[52] Robert Egly, Jon Driver, and Robert D Rafal. Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, 123(2):161, 1994.

[53] Tormod Thomsen, Karsten Specht, Lars Ersland, and Kenneth Hugdahl. Processing of conflicting cues in an attention-shift paradigm studied with fmri. *Neuroscience letters*, 380(1-2):138–142, 2005.

[54] Benoit Brisson and Pierre Jolicoeur. Express attentional re-engagement but delayed entry into consciousness following invalid spatial cues in visual search. *PLoS One*, 3(12):e3967, 2008.

[55] Hermann J Müller and John M Findlay. Sensitivity and criterion effects in the spatial cuing of visual attention. *Perception & Psychophysics*, 42(4):383–399, 1987.

[56] Harold L Hawkins, Steven A Hillyard, Steven J Luck, Mustapha Mouloua, Cathryn J Downing, and Donald P Woodward. Visual attention modulates signal detectability. *Journal of Experimental Psychology: Human Perception and Performance*, 16(4):802, 1990.

[57] IJ Saltzman and WR Garner. Reaction time as a measure of span of attention. *The Journal of psychology*, 25(2):227–241, 1948.

[58] Jerry S Carlson, C Mark Jensen, and Keith F Widaman. Reaction time, intelligence, and attention. *Intelligence*, 7(4):329–344, 1983.

[59] William Prinzmetal, Christin McCool, and Samuel Park. Attention: reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology: General*, 134(1):73, 2005.

[60] Deborah A Jehu, Alyssa Desponts, Nicole Paquet, and Yves Lajoie. Prioritizing attention on a reaction time task improves postural control and reaction time. *International Journal of Neuroscience*, 125(2):100–106, 2015.

[61] Timothy J Silk, Mark A Bellgrove, Pia Wrafter, Jason B Mattingley, and Ross Cunnington. Spatial working memory and spatial attention rely on common neural processes in the intraparietal sulcus. *Neuroimage*, 53(2):718–724, 2010.

[62] Akiko Ikkai and Clayton E Curtis. Common neural mechanisms supporting spatial working memory, attention and motor intention. *Neuropsychologia*, 49(6):1428–1434, 2011.

[63] Muhammet Ikbal Sahan, Tom Verguts, Carsten Nicolas Boehler, Gilles Pourtois, and Wim Fias. Paying attention to working memory: Similarities in the spatial distribution of attention in mental and physical space. *Psychonomic bulletin & review*, 23:1190–1197, 2016.

[64] Eric I Knudsen. Fundamental components of attention. *Annu. Rev. Neurosci.*, 30(1):57–78, 2007.

[65] Stefan Van der Stigchel, Artem V Belopolsky, Judith C Peters, Jasper G Wijnen, Martijn Meeter, and Jan Theeuwes. The limits of top-down control of visual attention. *Acta psychologica*, 132(3):201–212, 2009.

[66] Farhan Baluch and Laurent Itti. Mechanisms of top-down attention. *Trends in neurosciences*, 34(4):210–224, 2011.

[67] Carola Dolci, C Nico Boehler, Elisa Santandrea, Anneleen Dewulf, Suliann Ben-Hamed, Emiliano Macaluso, Leonardo Chelazzi, and Einat Rashal. Integrated effects of top-down attention and statistical learning during visual search: An eeg study. *Attention, Perception, & Psychophysics*, 85(6):1819–1833, 2023.

[68] Craig Hedge, Klaus Oberauer, and Ute Leonards. Selection in spatial working memory is independent of perceptual selective attention, but they interact in a shared spatial priority map. *Attention, Perception, & Psychophysics*, 77:2653–2668, 2015.

[69] Agnes Rosner, Michael Schaffner, and Bettina von Helversen. When the eyes have it and when not: How multiple sources of activation combine to guide eye movements during multiattribute decision making. *Journal of Experimental Psychology: General*, 151(6):1394, 2022.

[70] Koorosh Mirpour, Wei Song Ong, and James W Bisley. Microstimulation of posterior parietal cortex biases the selection of eye movement goals during search. *Journal of neurophysiology*, 104(6):3021–3028, 2010.

[71] Priyanka Gupta and Devarajan Sridharan. Presaccadic attention does not facilitate the detection of changes in the visual field. *PLoS Biology*, 22(1):e3002485, 2024.

[72] Lupeng Wang, James P Herman, and Richard J Krauzlis. Neuronal modulation in the mouse superior colliculus during covert visual selective attention. *Scientific Reports*, 12(1):2482, 2022.

[73] Peter Dayan and Nathaniel D Daw. Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4):429–453, 2008.

[74] Yael Niv. Reinforcement learning in the brain. *Journal of Mathematical Psychology*, 53(3):139–154, 2009.

[75] Wolfram Schultz and Anthony Dickinson. Neuronal coding of prediction errors. *Annual review of neuroscience*, 23(1):473–500, 2000.

[76] Clay B Holroyd and Michael GH Coles. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological review*, 109(4):679, 2002.

[77] Paul W Glimcher. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, 108(supplement_3):15647–15654, 2011.

[78] Brian A Anderson, Hiroto Kuwabara, Dean F Wong, Emily G Gean, Arman Rahmim, James R Brašić, Noble George, Boris Frolov, Susan M Courtney, and Steven Yantis. The role of dopamine in value-based attentional orienting. *Current Biology*, 26(4):550–555, 2016.

[79] Brian A Anderson. The attention habit: How reward learning shapes attentional selection. *Annals of the new York Academy of Sciences*, 1369(1):24–39, 2016.

[80] Seung-Lark Lim, John P O'Doherty, and Antonio Rangel. The decision value computations in the vmpfc and striatum use a relative value code that is guided by visual attention. *Journal of Neuroscience*, 31(37):13214–13223, 2011.

[81] John HR Maunsell. Neuronal mechanisms of visual attention. *Annual review of vision science*, 1(1):373–391, 2015.

[82] Yuan Chang Leong, Angela Radulescu, Reka Daniel, Vivian DeWoskin, and Yael Niv. Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron*, 93(2):451–463, 2017.

[83] Robert Desimone, John Duncan, et al. Neural mechanisms of selective visual attention. *Annual review of neuroscience*, 18(1):193–222, 1995.

[84] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[85] Jeremy M Wolfe, Melissa L-H Võ, Karla K Evans, and Michelle R Greene. Visual search in scenes involves selective and nonselective pathways. *Trends in cognitive sciences*, 15(2):77–84, 2011.

[86] Jeremy M Wolfe. Guided search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, 28(4):1060–1092, 2021.

[87] John H Reynolds and David J Heeger. The normalization model of attention. *Neuron*, 61(2):168–185, 2009.

[88] Pierre Barrouillet, Annick De Paepe, and Naomi Langerock. Time causes forgetting from working memory. *Psychonomic bulletin & review*, 19:87–92, 2012.

[89] Joost de Jong, Sophia Wilhelm, and Elkan G Akyürek. Adaptive forgetting speed in working memory. *Psychonomic Bulletin & Review*, pages 1–10, 2024.

[90] Ingrid R Olson and Christy Marshuetz. Remembering "what" brings along "where" in visual working memory. *Perception & Psychophysics*, 67:185–194, 2005.

[91] Ziyi Duan and Clayton E Curtis. Visual working memories are abstractions of percepts. *Elife*, 13:RP94191, 2024.