Efficient Allocation of Working Memory Resource for Utility Maximization in Humans and Recurrent Neural Networks

Qingqing Yang and Hsin-Hung Li (2025)

The Ohio State University

Introduction

- Working memory (WM) is the neural and cognitive process that temporarily maintain task-relevant information online (Baddeley, 2003).
- WM has limited capacity and is inherently noisy (Luck and Vogel, 1997; Panichello et al., 2019). Thereby, the ability to flexibly and efficiently allocate WM resource is essential for adaptive behaviors.
- However, recent studies have reported mixed results regarding whether WM resources can be modulated based on reward alone (Brissenden et al., 2023, Van den Berg et al., 2023, Weiss et al., 2025, Manga et al., 2020).

Preview

We investigate a fundamental and unresolved question:

Whether WM resources are allocated efficiently to maximize reward (utility)?

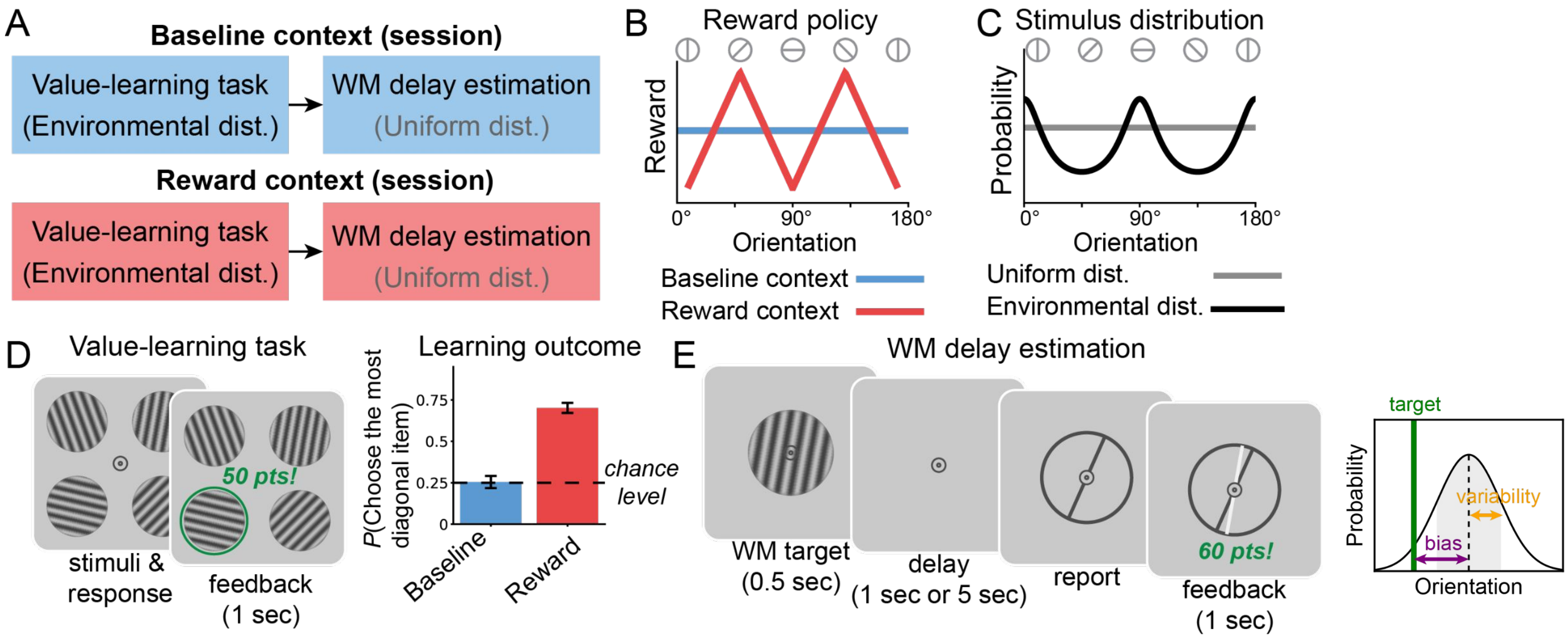
Our approach

1. We combined experiments with humans and recurrent neural networks (RNNs) trained with varying prior and reward information.

2. We extended efficient coding theory to provide a normative account of how WM resources can be allocated over time to optimize expected utility.

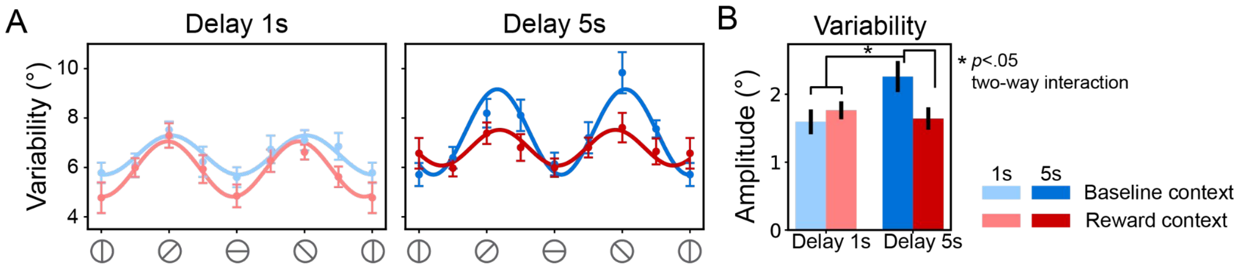
Behavioral tasks

Orientation estimation exhibits distinct patterns of bias and variability (Taylor and Bays, 2018, Wei and Stocker, 2017, Girshick et al., 2011).



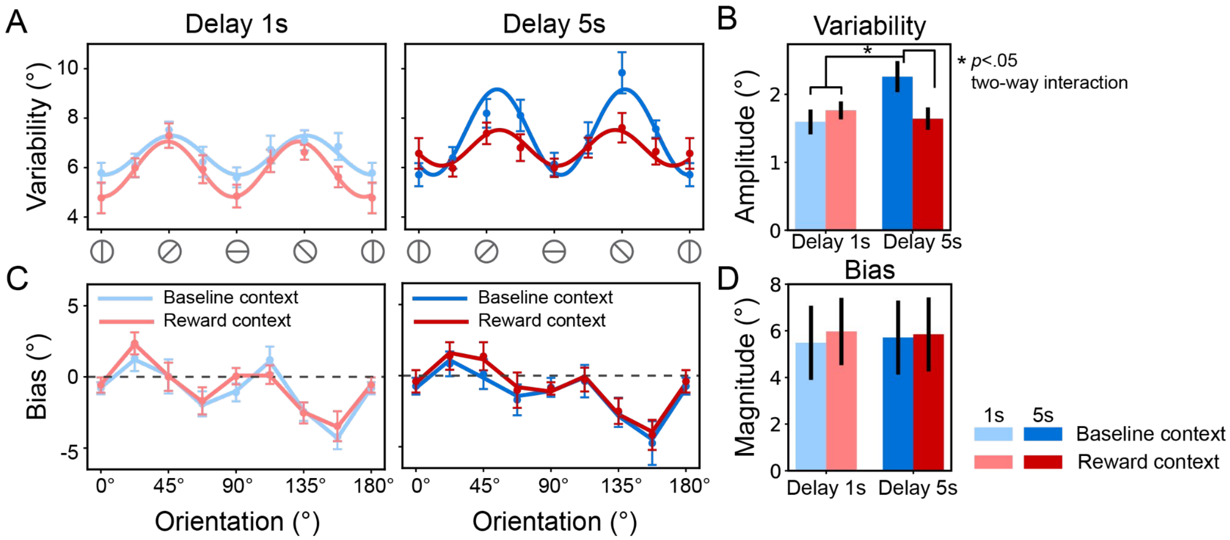
Human WM resource

Memory representations were more stable for stimuli associated with higher reward, as well as for those more probable in the natural environment.



Human WM resource

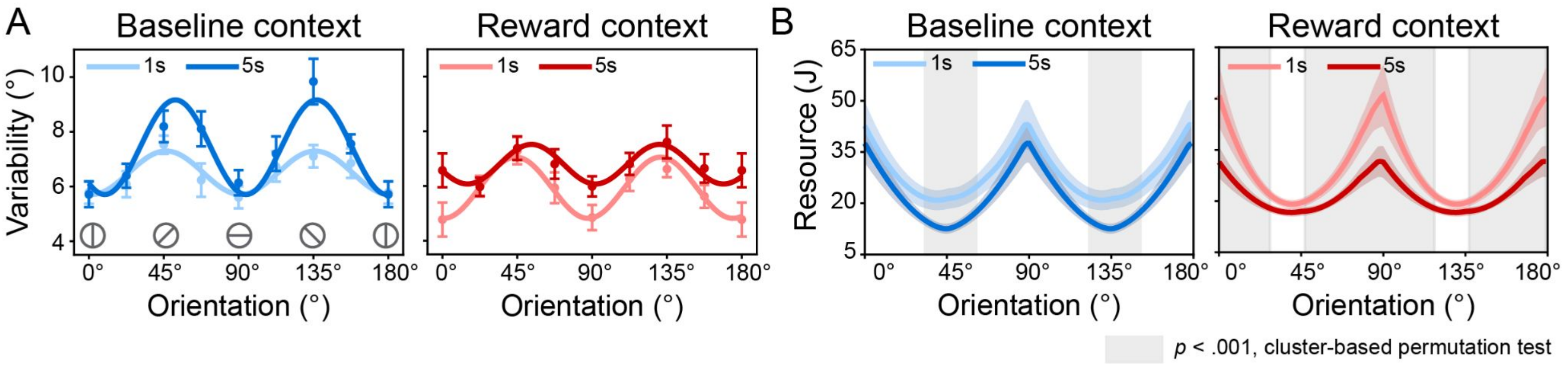
Bias was not influenced by reward policy or delay.



Human WM resource

We quantify the theoretical resource quantity using Fisher information *J* by the Variable Precision (VP) model (Van den Berg et al., 2012).

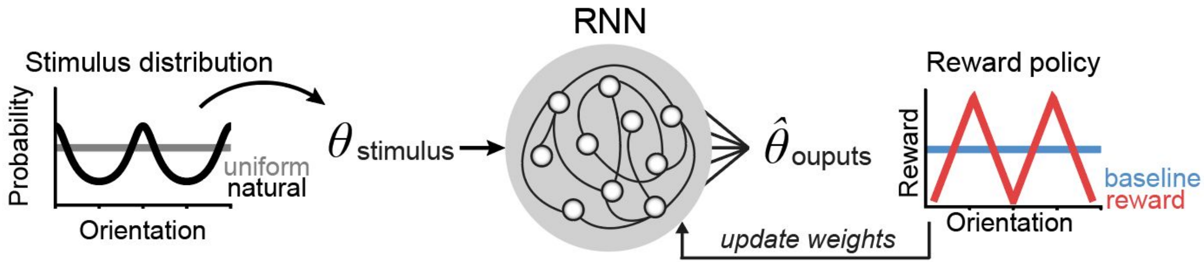
The estimated resource roughly followed the environmental distribution of orientations, but <u>rewards affected how WM resources evolved across delay</u>.



RNN architecture

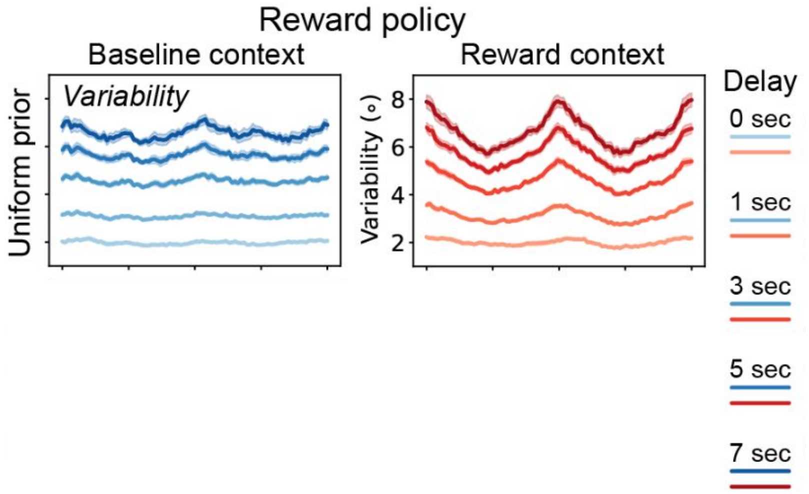
RNNs are trained in a 2 x 2 design:

- Training stimulus distribution (uniform and natural distribution);
- Loss functions (maximize reward or minimize error).



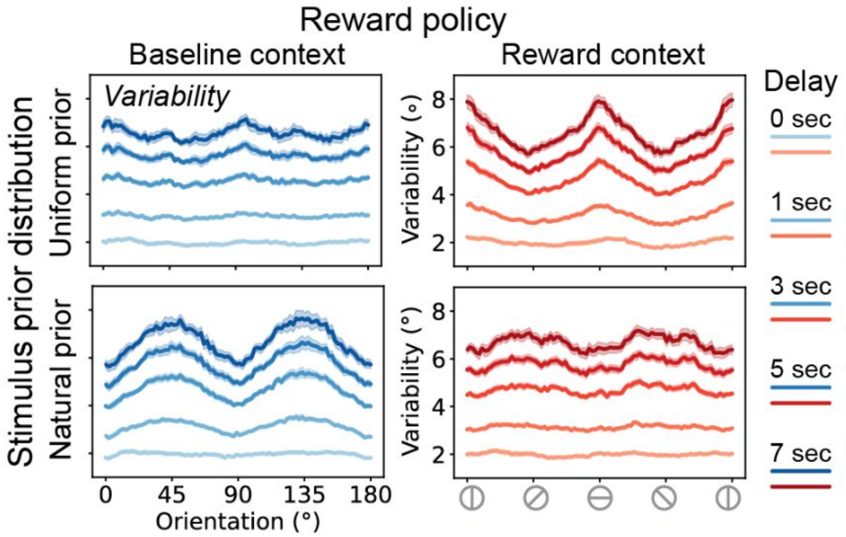
RNN WM resource

Joint effect of prior and reward in RNN WM resource allocation.



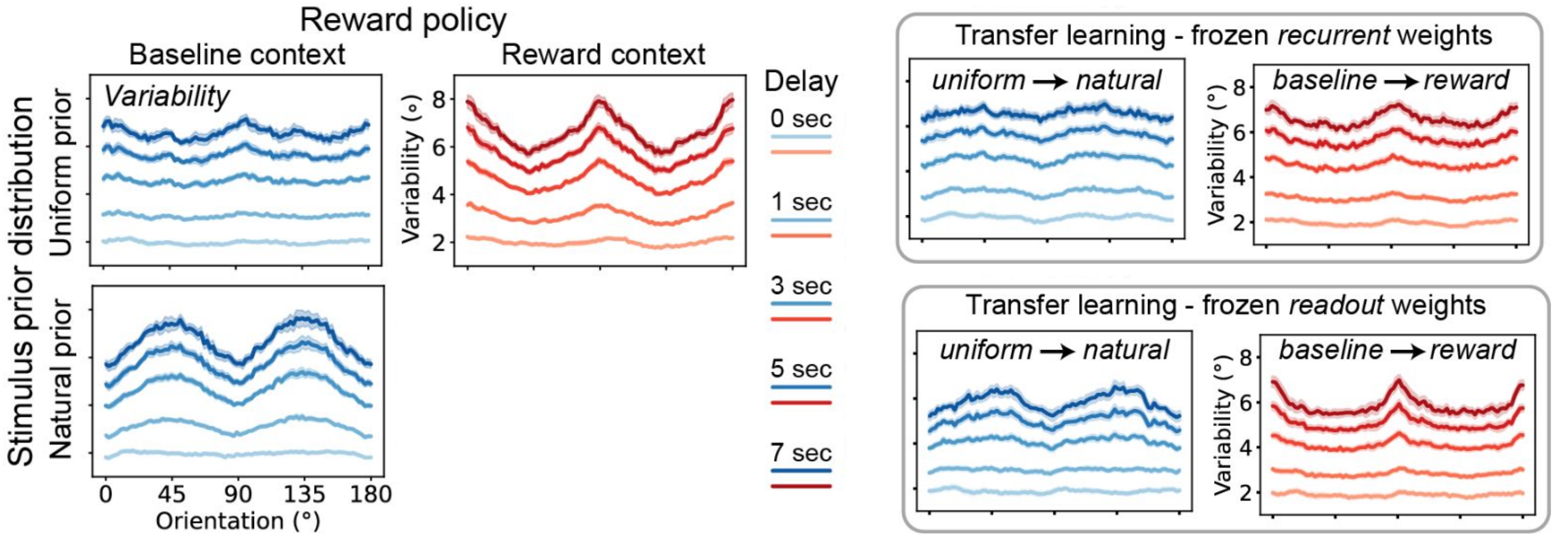
RNN WM resource

Joint effect of prior and reward in RNN WM resource allocation.



Dissecting the contributions of recurrent dynamics

The recurrent dynamics plays a more important role in allowing the RNNs to adapt to different stimulus distributions or reward policies



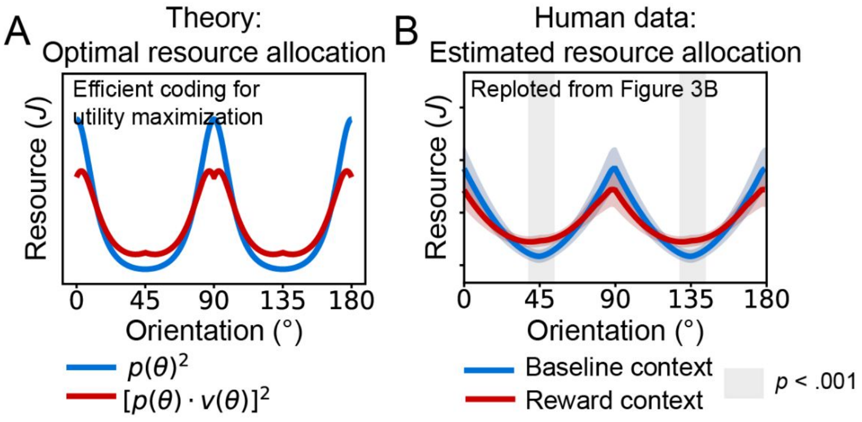
Normative theory

- The efficient coding theory – perceptual system should allocate resource based on the statistical structure of the environment, to ***minimize the expected loss*** (Wei and Stocker, 2015; Hahn and Wei 2024; Morais and Pillow 2018), with the optimal resource: $J_{\text{opt}} \propto p(\theta)^q$
- We extended to the tasks where stimuli θ carry context-dependent rewards, and shift the objective to ***maximize expected utility***, with optimal resource solution: $J_{\text{opt}} \propto [p(\theta)\, v(\theta)]^q$

Normative theory

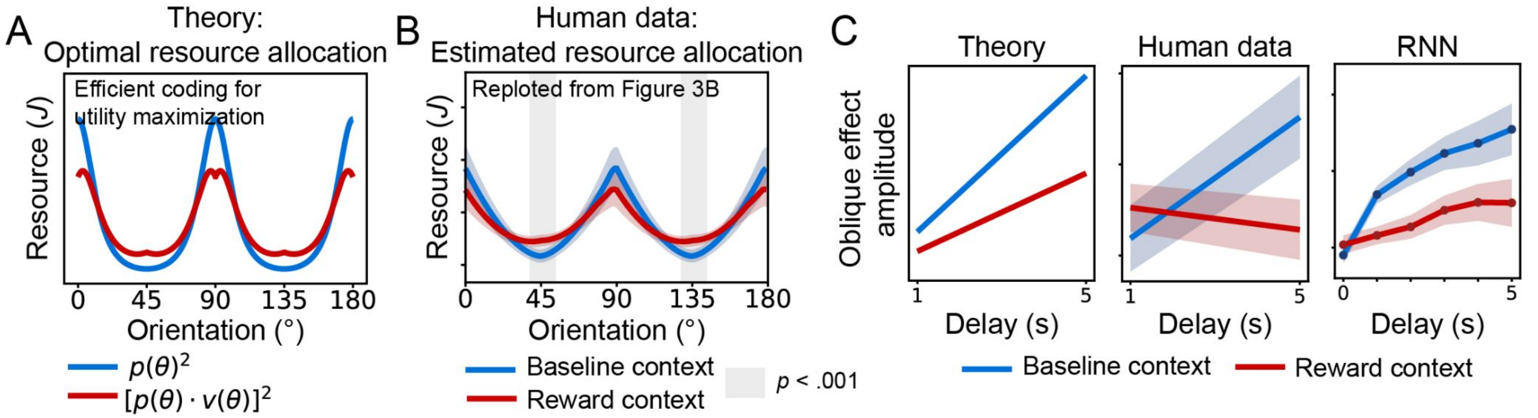
- The efficient coding theory – perceptual system should allocate resource based on the statistical structure of the environment, to ***minimize the expected loss*** (Wei and Stocker, 2015; Hahn and Wei 2024; Morais and Pillow 2018), with the optimal resource: $J_{\text{opt}} \propto p(\theta)^q$

- We extended to the tasks where stimuli θ carry context-dependent rewards, and shift the objective to ***maximize expected utility***, with optimal resource solution: $J_{\text{opt}} \propto [p(\theta)\, v(\theta)]^q$



Normative theory

- We extended the theory by incorporating the time domain in the loss function, and the solution of the optimal resource allocation depends only on the prior distribution and reward.



Summary

1. With behavioral experiments, we found that after value learning, humans allocate WM resources efficiently according to stimulus-reward association in order to maximize utility.
2. RNNs trained to maximize reward exhibit human-like behaviors, integrating both prior and reward information in a manner consistent with the human data.
3. We extended efficient coding theory to the time domain and reformulated the objective from minimizing error to utility maximization, providing a normative account of how WM resources can be allocated over time to optimize expected utility.

Thank you for listening!

Details in manuscript.

Contact yang.6118@osu.edu for questions!

Welcome chat at poster session:

Dec 4th, 4:30-7:30 pm PST, Exhibit Hall C-E.