Foveating sensing with **KNN-convolutional neural networks** based on **isotropic cortical magnification**

Nicholas M. Blauch, George A. Alvarez, Talia Konkle

Harvard University, Department of Psychology and Kempner Institute for Natural and Artificial Intelligence

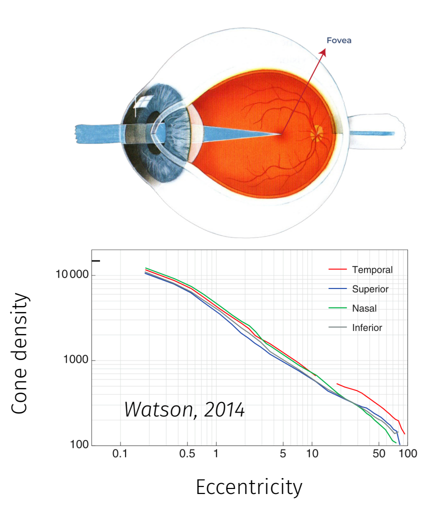
correspondence: nblauch@fas.harvard.edu, talia_konkle@harvard.edu

Kempner INSTITUTE For the Study of Natural & Artificial Intelligence at Harvard University

Vision Sciences Lab

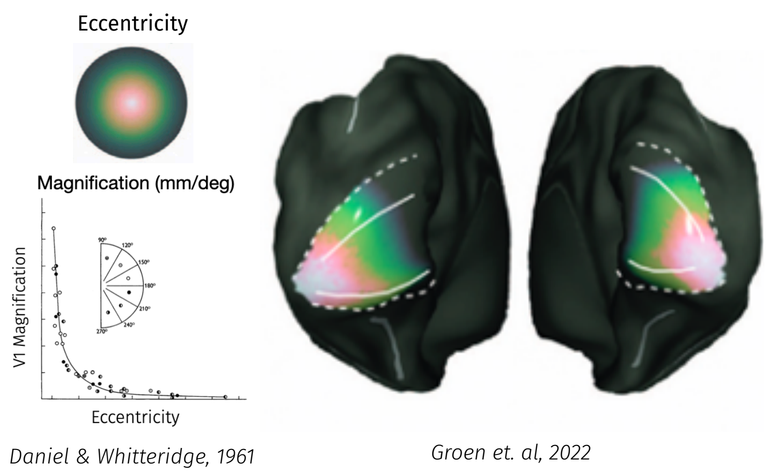
Introduction

Human vision is foveated



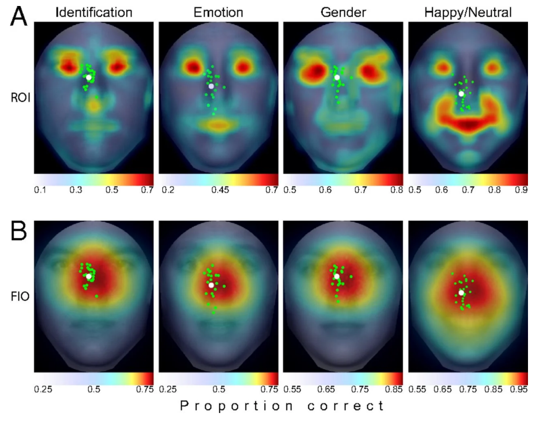
Foveation profoundly shapes cortical organization and visual cognition

Retinotopic organization & cortical magnification



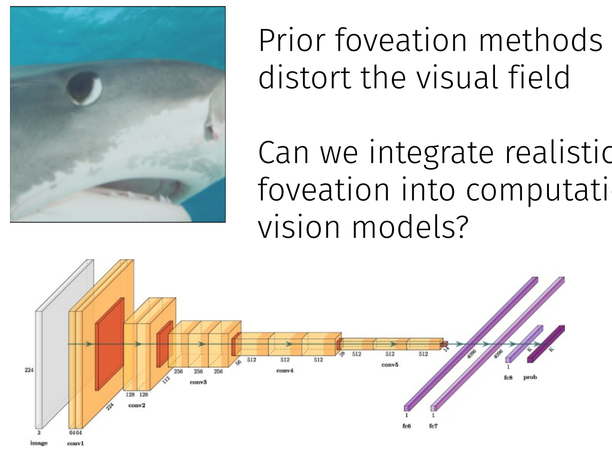
Daniel & Whitteridge, 1961 *Groen et. al, 2022*

Optimized active eye movements



Peterson & Eckstein, 2012

Computer vision models lack mechanisms for realistic foveation

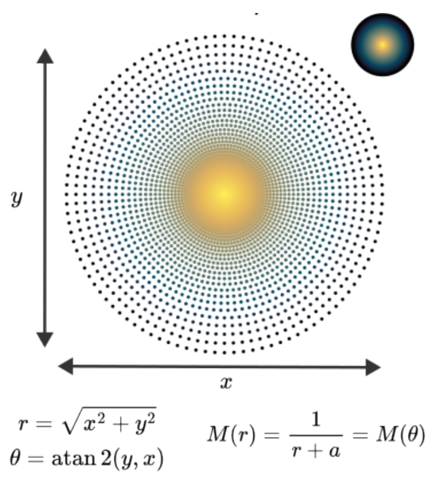


Prior foveation methods distort the visual field

Can we integrate realistic foveation into computational vision models?

Foveated sensing based on **isotropic cortical magnification**

Sampling in visual space

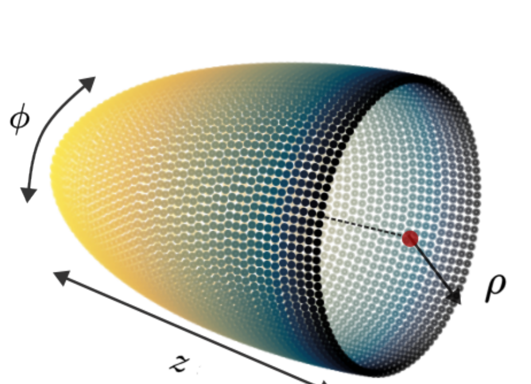


$r = \sqrt{x^2 + y^2}$ $\quad M(r) = \frac{1}{r+a} = M(\theta)$

$\theta = \text{atan}\,2(y, x)$

Sensor manifold

Rovamo & Virsu, 1984



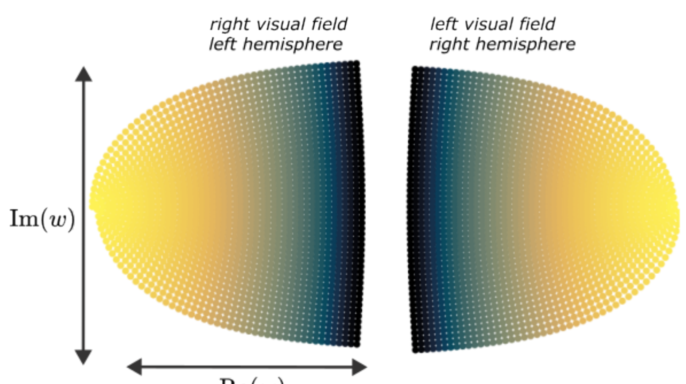
$\rho = M(r)\sin(r)$

$z = \int_0^w \left[M(r)^2 - (d\rho/dr)^2\right]^{0.5} dr$

$\phi = \theta$

Flat sensor manifold

(visualization only)

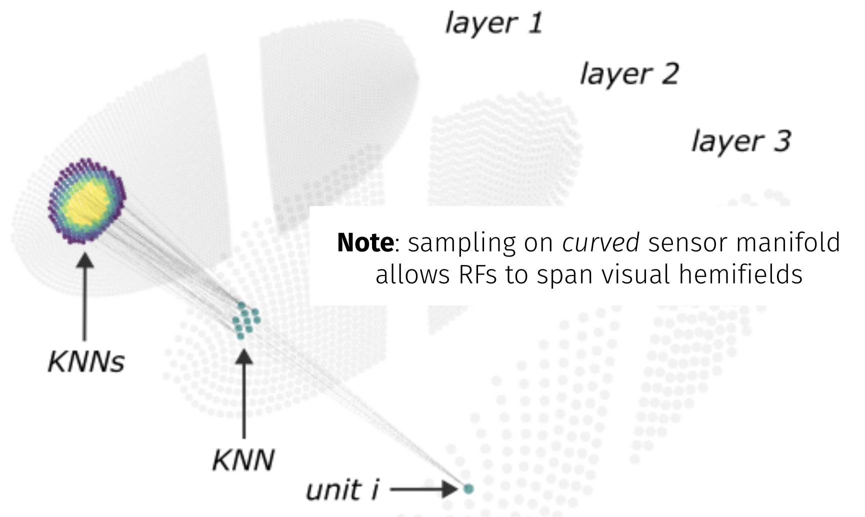


Note: *this is the famous Schwartz (1980) model*

$z = x + iy$

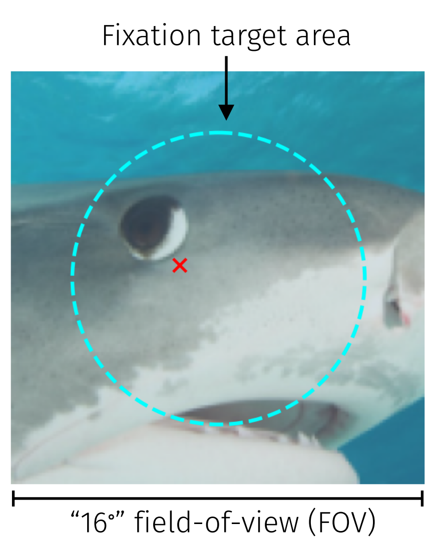
$w = \log(z + a)$

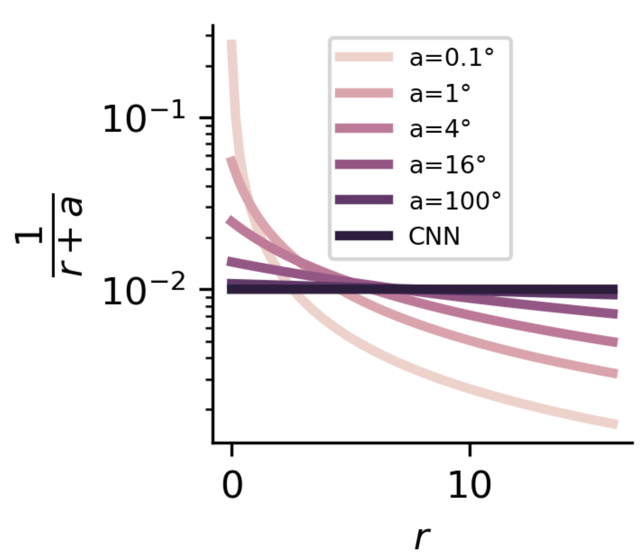
K-Nearest-Neighbor ConvNets (KNN-CNNs)

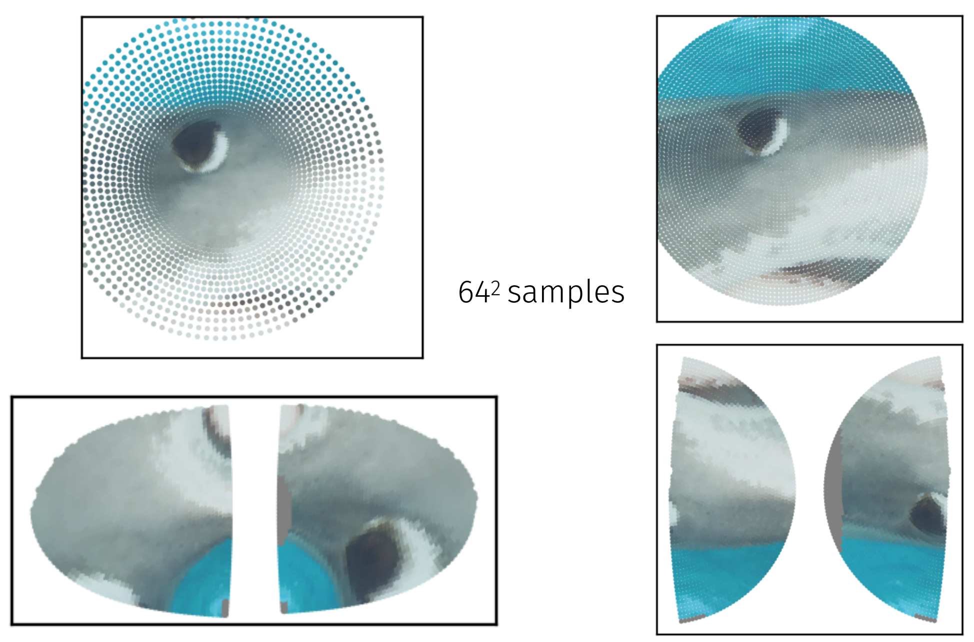


Note: sampling on *curved* sensor manifold allows RFs to span visual hemifields

Result 2: strong performance in foveated networks

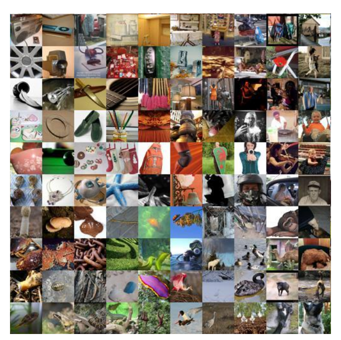


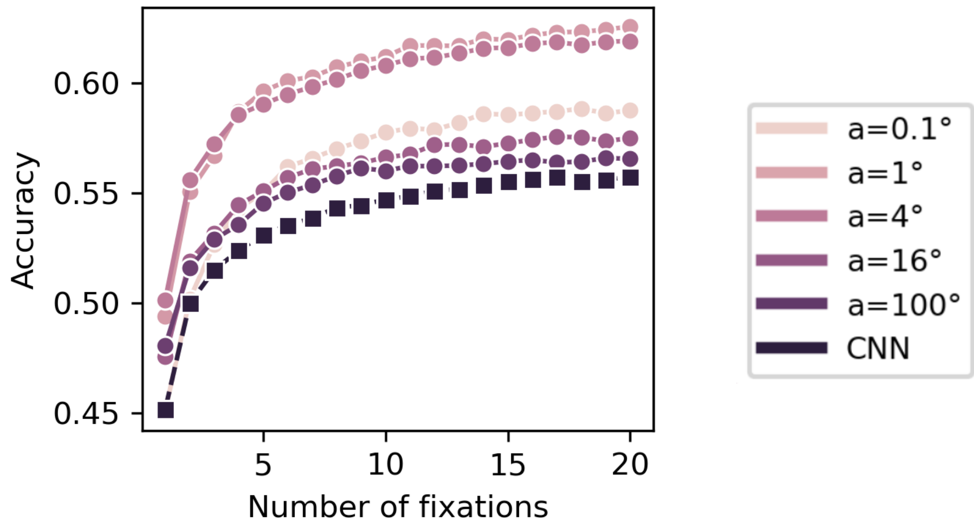




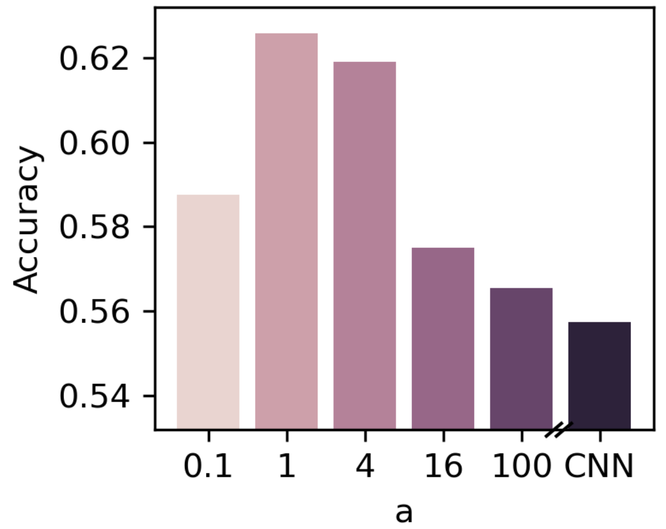
Performance increases with multiple fixations

ImageNet-100



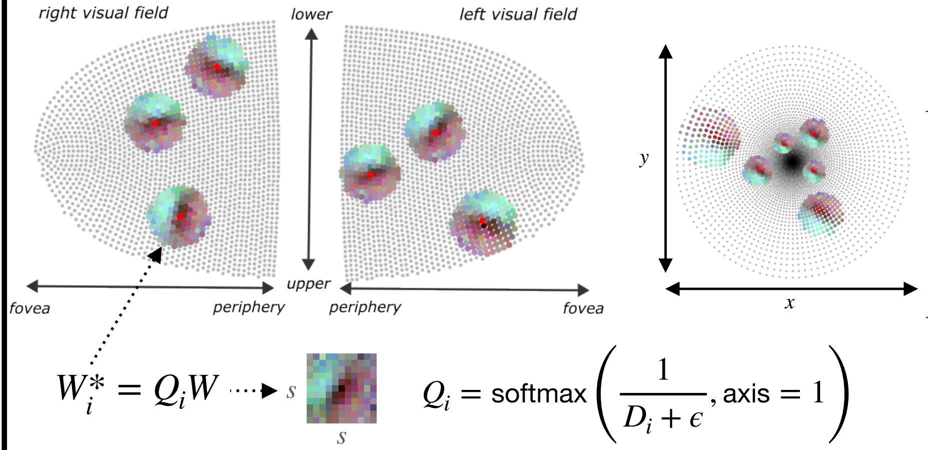


Optimal performance for strong-intermediate foveation



Result 1: foveated model accounts for biological receptive field properties

Receptive fields (RFs) are uniform circular samples on sensor manifold



$W^*_i = Q_i W$ $\quad Q_i = \text{softmax}\left(\frac{1}{D_i + \epsilon}, \text{axis} = 1\right)$

W^*_i: weight in ith neighborhood

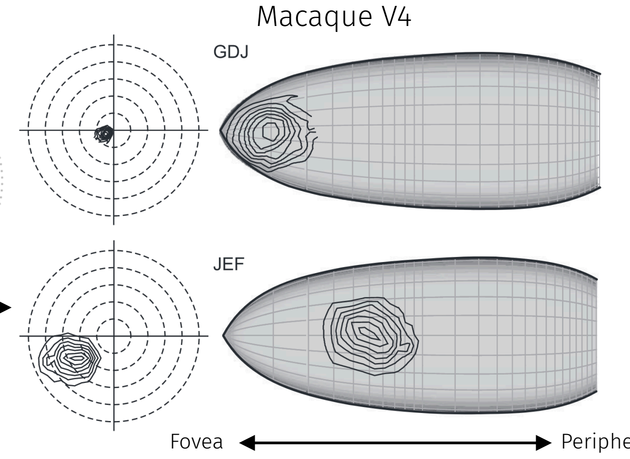
W: reference kernel

Q_i: mapping matrix

D_i: distance from normalized neighborhood coords to reference coords

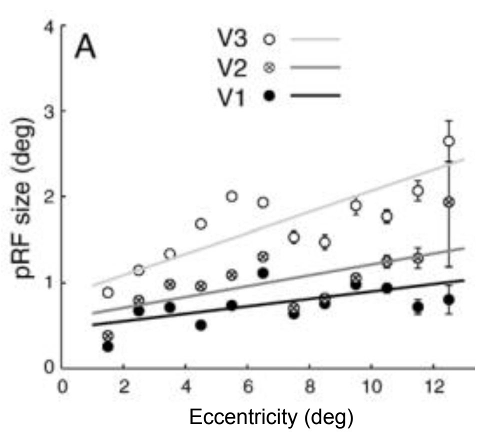
Visual RF size increases with eccentricity and hierarchical level

Macaque V4



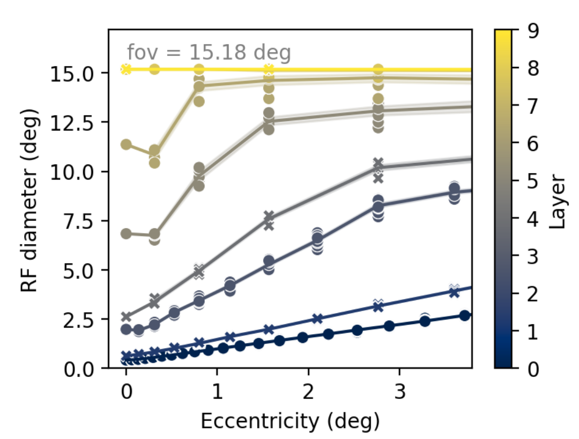
Motter, 2009

Human V1-V3



Dumoulin & Wandell, 2008

Model



Existing grid-like foveated sensors warp receptive fields

$M(r) = \frac{1}{r+a}$ $\quad n_r = n_\theta$

$w = \log(r + a)$

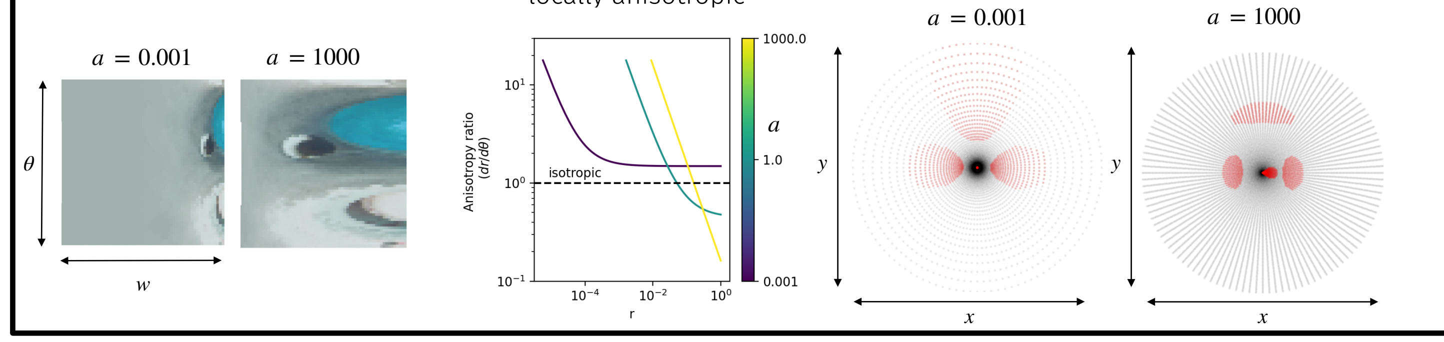
(grid-like simplification of Schwartz (1980) model)

Log-polar image foveation

e.g. Jérémie et. al, 2024

Sampling is locally anisotropic

Eccentricity-dependent RF aspect ratios



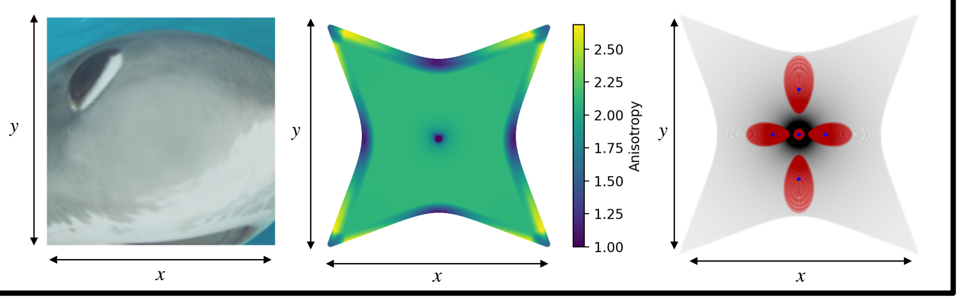
Warped cartesian foveation

e.g. Wang et. al, 2021; see also Da Costa et. al, 2024

Foveated cartesian image

Sampling is locally anisotropic

Eccentricity-dependent RF shapes



Conclusions

Human vision is **foveated** — non-uniform sampling is mapped to a uniform layout in V1 via **isotropic cortical magnification**

We integrate foveation into computational models with a **magnification function,** appropriate **sensor manifold,** and **KNN-CNN architecture**

This leads to **biologically-plausible receptive field** properties, and **strong performance**

References:

Da Costa, D., Kornemann, L., Goebel, R., & Senden, M. (2024). Scientific Reports, 14(1), 8980.
Daniel, P. M., & Whitteridge, D. (1961). The Journal of Physiology, 159(2), 203–221.
Dumoulin, S. O., & Wandell, B. A. (2008). NeuroImage, 39(2), 647–660.
Jérémie, J.-N., Daucé, E., & Perrinet, L. U. (2024). arXiv.
Motter, B. C. (2009). The Journal of Neuroscience, 29(18), 5749–5757.
Rovamo, J., & Virsu, V. (1984). Vision Research, 24(3), 283–286.
Schwartz, E. L. (1980). Vision Research, 20(8), 645–669.
Wang, B., Mayo, D., Deza, A., Barbu, A., & Conwell, C. (2021). arXiv
Watson, A. B. (2014). Journal of Vision, 14(7), 15.
Groen, I. I. A., Dekker, T. M., Knapen, T., & Silson, E. H. (2022). *Trends in Cognitive Sciences*, *26*(1), 81–96.
Peterson, M. F., & Eckstein, M. P. (2012). *PNAS*, *109*(48).