The Influence of Expectations on Recurrent Processing During Challenging Image Recognition

Paulo A. Ortiz (paulo.ortizfonseca@donders.ru.nl)
Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherland
Kapittelweg 29, 6525 EN Nijmegen, The Netherlands

Lea-Maria Schmitt (lea-maria.schmitt@donders.ru.nl)
Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherland
Kapittelweg 29, 6525 EN Nijmegen, The Netherlands

Floris P. de Lange (floris.delange@donders.ru.nl)
Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherland
Kapittelweg 29, 6525 EN Nijmegen, The Netherlands

Abstract

Recent work on challenging visual conditions has highlighted the role of recurrent processing in supporting robust perception. Predictive processing frameworks propose that prior knowledge plays a role in disambiguating sensory input, potentially reducing the need for recurrence under strong expectations. In an object recognition paradigm, we examined whether expectations influence neural processing of challenging and non-challenging visual input. Using recent adversarial attack techniques, we generated a set of perceptually challenging stimuli. We then created perceptual expectations for some stimuli by presenting them in a predictable order. Preliminary behavioral data suggested that stronger expectations enhance both the accuracy and speed of categorical judgement. We will present MEG data that will elucidate how expectations modulate temporal signatures linked to recurrent processing. This work aims to shed light on how predictive mechanisms shape the neural dynamics of robust perception.

Keywords: predictive processing; recurrence; object recognition; deep neural networks; MEG

Background and Motivation

Recurrent processing has been proposed as a central mechanism for robust perception in biological systems (Groen et al., 2018; Kar et al., 2019; van Bergen & Kriegeskorte, 2020). Core object recognition can often be solved by relying solely on an initial feedforward sweep of processing from lower to higher brain regions (DiCarlo et al., 2012; van Bergen & Kriegeskorte, 2020). However, recent work has shown that challenging visual conditions--such as occlusion, and degradation--engage additional top-down and lateral recurrent computations over time (Kar et al., 2019; Rajaei et al., 2019; Seijdel et al., 2021, Tang et al., 2018).

A theoretical framework that may account for the role of recurrent activity is predictive processing (de Lange et al., 2018; Friston, 2005; Keller & Mrsic-Flogel, 2018; Rao & Ballard, 1999). In brief, it posits that the brain uses top-down information to generate predictions about external causes in the world, which are then compared against incoming sensory input enabling fast and robust perception. When sensory information is challenging and not easily mapped onto a specific cause, expectations might restrict the range of possible interpretations. By iteratively testing a limited set of expectations to sensory data, biological systems could converge on a suitable interpretation faster. Recurrent dynamics are thought to play a key role in facilitating such an interplay between expectations and sensory information by propagating them from higher to lower brain regions to disambiguate perception (Gilbert & Li, 2013; van Bergen & Kriegeskorte, 2020). This framework suggests that recurrence is dynamically engaged based on both the challengingness of visual input and the specificity of these predictions.

To examine whether recurrent processing fulfils the role proposed by these predictive theories, we investigate whether expectations enhance sensory processing particularly during the perception of challenging visual input. To this end, we complemented a set of non-challenging images with a set of challenging images by leveraging recent advances in adversarial attack techniques developed for artificial neural networks (ANNs; Gaziv et al., 2023). This technique embeds key high-level features from another object class into original images to disrupt their object recognition performance. Additionally, we manipulated expectations by varying the transitional probabilities between consecutive images. We are currently assessing the impact of stimulus challengingness and expectations on human performance using behavioral metrics. Next, using magnetoencephalography (MEG), we aim to examine the temporal dynamics of image information across the cortex, using decoding analysis of object identity.

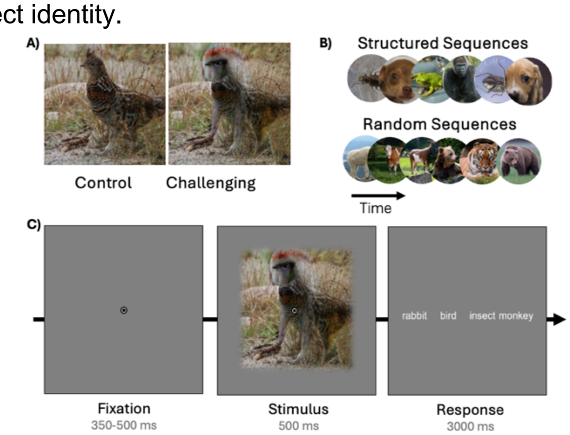

Figure 1: Experimental Design. A: Challenging Image Manipulation. B: Expectation Manipulation. C: Experimental Set-Up

Materials and Methods

Challenging images.

To generate challenging images, we adopted Gaziv et al.'s (2023) approach using adversarial image perturbations (**Fig. 1A;** *i.e.,* small, targeted pixel-level changes designed to disrupt recognition). In their work they showed that ANNs trained with adversarial examples during their optimization generated perturbations that not only disrupted network performance but also human perception.

Our stimulus set consisted of a subset of ImageNet (Russakovsky et al., 2015) classes mapped into a custom ten basic animal categories (e.g., dog, bear, etc.; Engstrom et al., 2019). Challenging mages were generated using *targeted* modulations, modifying an image with features of a specific distractor class. We used a pre-trained ResNet50 (He et al., 2015) adversarially trained with an $\ell_2$norm pixel budget of *3.0* (i.e., the ℓ_2 norm distance between the original and perturbed image was constrained to a maximum of *3.0*; Engstrom et al., 2019; Gaziv et al., 2023). In total, we generated 360 images with varying levels of ℓ_2-norm budget: 7.5, and 10. Unperturbed versions of the same image served as non-challenging controls.

Expectation manipulation.

We manipulated expectations by controlling the predictability between consecutive images in a sequence (**Fig 1B**). In the structured condition, five animal classes were presented in a fixed repeating sequence, creating strong expectations about upcoming stimuli. In the random condition, the remaining five classes appeared in random order, such that no order expectations could be formed. Previous work has employed similar manipulations, showing that humans are able to extract such regularities of the environment and use these expectations to enhance both the speed and accuracy of recognition (de Lange et al., 2018; Schapiro & Turk-Browne, 2015).

Experimental procedure.

Participants (*n=7*) viewed 1800 animal images sequentially. Each image was shown for 500 ms, followed by a 350–500 ms ISI. On a subset of trials (11% training, 22% testing), participants completed a 4AFC task identifying the most recent animal as accurately and quickly as possible. Images were presented in sequence blocks, belonging either to the structured or random condition. In each sequence block, participants viewed five animal classes in either a fixed (structured) or randomized (random) order. The specific image exemplars varied throughout the experiment (**Fig. 1C**).

The experiment consisted of a training and a testing phase. During training, all images were initially presented unperturbed, allowing to learn the transition probabilities. As training progressed, half of the images were gradually made challenging with increasing levels of perturbation. In the testing phase, the task and timing remained identical. However, all challenging images were presented at a perturbation level aiming to achieve a disruption on classification performance, targeting 75% accuracy at a group level. This ensured that challenging stimuli introduced sufficient perceptual ambiguity while remaining identifiable. The images used during training differed from those used in testing. Class-to-condition assignments, challenging image selection, and block order were randomized per participant.

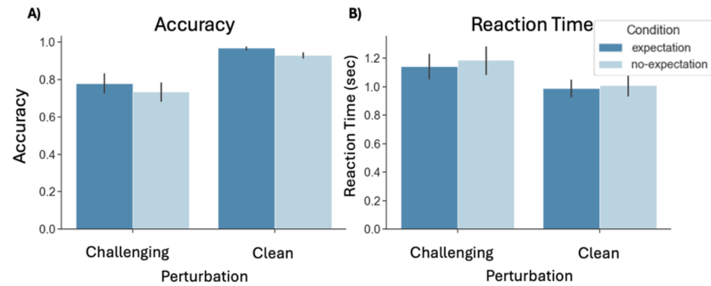

Figure 2. Main Results: Behavioral performance across expectation conditions and image perturbation levels. A: Accuracy. B: RT

Preliminary Results and Future Directions

The project is currently in the data collection phase. Thus far, only behavioral data has been acquired from a small subsample of participants. In line with our expectations, preliminary observations suggest a main effect of image challengingness, with higher accuracy and faster reaction times for clean compared to challenging images. Additionally, stronger expectations appear to show similar trends, with greater accuracy and shorter reaction times for structured versus random sequences. Lastly, a potential interaction may be present, suggesting that stronger expectation may enhance both the accuracy and speed of category judgements for both clean and challenging conditions (**Fig 2**). Given the project's early stage and the limited sample, no statistical analyses have been conducted.

After completing behavioral data collection and examining whether the effects are still present, we will extend the experiment to MEG. Previous operationalizations of recurrence focus on the timing and strength of decodable information in the brain, with delayed peaks indicating greater reliance on recurrent computations (Kar et al., 2019; Kietzmann et al., 2019). Our MEG analysis will enable us to investigate whether expectations enhanced the neural processing of challenging information. Specifically, we will leverage these signatures of recurrence to quantify how prior knowledge modulates these dynamics under challenging conditions.

Altogether, this work aims to advance our understanding of how prior expectations interact with challenging visual input and how recurrent processing may be adaptively engaged in the brain.

Acknowledgements

The work is supported by the Dutch Research Council (NWO) via the VICI Grant (231.043) allocated to F.P.d.L.

References

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? Trends in Cognitive Sciences, 22(9), 764–779.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How Does the Brain Solve Visual Object Recognition? Neuron, 73(3), 415–434.

Engstrom, L., Ilyas, A., Salman, H., Santurkar, S., & Tsipras, D. (2019). Robustness (Python library) [Computer software]. https://github.com/MadryLab/robustness

Friston, K. (2005). A theory of cortical responses. Philosophical Transactions of the Royal Society B: Biological Sciences, 360(1456), 815–836.

Gaziv, G., Lee, M. J., & DiCarlo, J. J. (2023). Robustified ANNs Reveal Wormholes Between Human Category Percepts (No. arXiv:2308.06887). arXiv.

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. Nature Reviews Neuroscience, 14(5), 350–363.

Groen, I. I. A., Jahfari, S., Seijdel, N., Ghebreab, S., Lamme, V. A. F., & Scholte, H. S. (2018). Scene complexity modulates degree of feedback activity during object detection in natural scenes. PLOS Computational Biology, 14(12), e1006690.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition (No. arXiv:1512.03385). arXiv..

Kar, K., Kubilius, J., Schmidt, K., Issa, E. B., & DiCarlo, J. J. (2019). Evidence that recurrent circuits are critical to the ventral stream's execution of core object recognition behavior. Nature Neuroscience, 22(6), 974–983.

Keller, G. B., & Mrsic-Flogel, T. D. (2018). Predictive Processing: A Canonical Cortical Computation. Neuron, 100(2), 424–435.

Kietzmann, T. C., Spoerer, C. J., Sörensen, L. K. A., Cichy, R. M., Hauk, O., & Kriegeskorte, N. (2019). Recurrence is required to capture the representational dynamics of the human visual system. Proceedings of the National Academy of Sciences, 116(43), 21854–21863.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. Frontiers in Systems Neuroscience, 2.

Rajaei, K., Mohsenzadeh, Y., Ebrahimpour, R., & Khaligh-Razavi, S.-M. (2019). Beyond core object recognition: Recurrent processes account for object recognition under occlusion. PLOS Computational Biology, 15(5), e1007001.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. Nature Neuroscience, 2(1), Article 1.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 115(3), 211–252.

Schapiro, A., & Turk-Browne, N. (2015). Statistical Learning. In Brain Mapping (pp. 501–506). Elsevier.

Seijdel, N., Loke, J., Klundert, R. van de, Meer, M. van der, Quispel, E., Gaal, S. van, Haan, E. H. F. de, & Scholte, H. S. (2021). On the Necessity of Recurrent Processing during Object Recognition: It Depends on the Need for Scene Segmentation. Journal of Neuroscience, 41(29), 6281–6289.

Tang, H., Schrimpf, M., Lotter, W., Moerman, C., Paredes, A., Ortega Caro, J., Hardesty, W., Cox, D., & Kreiman, G. (2018). Recurrent computations for visual pattern completion. Proceedings of the National Academy of Sciences, 115(35), 8835–8840

van Bergen, R. S., & Kriegeskorte, N. (2020). Going in circles is the way forward: The role of recurrence in visual inference. Current Opinion in Neurobiology, 65, 176–193.