Early Sensory Responses Track High-Level Visual Surprise

David Richter (david.richter.work@gmail.com)
Mind, Brain and Behavior Research Center (CIMCYC), University of Granada, Spain
Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, the Netherlands

Paula Pena (paulapena@ugr.es)
Mind, Brain and Behavior Research Center (CIMCYC), University of Granada, Spain

María Ruz (mruz@ugr.es)
Mind, Brain and Behavior Research Center (CIMCYC), University of Granada, Spain

Abstract

Predictive processing theories propose that the brain continuously generates expectations about incoming sensory information and compares these predictions to actual inputs, resulting in sensory prediction errors. However, it remains unclear which stimulus features are predicted by the brain and hence which errors drive neural responses. Here, we addressed this question by recording EEG while participants viewed object images that were expected or unexpected based on probabilistic cues. We used a deep neural network to quantify low- and high-level visual representational distances between expected and unexpected stimuli. Neural activity was then regressed onto these surprise metrics. Results showed a modulation of evoked activity over occipital electrodes approximately 200ms after stimulus onset by high-level, but not low-level, visual surprise. These findings suggest that high-level visual predictions are rapidly integrated into sensory processing.

Keywords: sensory processing; vision; predictive coding; prediction errors; EEG

Introduction

Perception is shaped by our prior knowledge (de Lange et al., 2018; Walsh et al., 2020). Accordingly, predictive processing theories postulate that the brain continuously generates expectations about sensory inputs and computes prediction errors, reflecting the discrepancies between predicted and actual inputs (Friston, 2005; Rao & Ballard, 1999). A critical question in understanding predictive processing is: What features are predicted by the brain, and consequently, what kind of mismatches do sensory prediction errors represent? One hypothesis is that prediction errors arise from features locally represented in each visual cortical area, such that low-level surprise (e.g., orientation) is computed in early visual areas, whereas high-level surprise (e.g., object parts) is computed in higher visual areas. An alternative is that sensory prediction errors largely reflect high-level surprise, computed in higher areas and fed back to earlier areas. In support of the latter account, neural firing rates in macaque V1 (Uran et al., 2022) and fMRI BOLD in human visual cortices, including V1 (Richter et al., 2024), scale with high-level but not low-level surprise. Combined, these findings suggest that predictions may predominantly operate at higher levels of abstraction.

However, little is known about the timing of these predictive processes. Yet, understanding when high-level surprise modulates sensory processing is crucial, as it provides insights into the neurocomputational principles and role of predictions in shaping perception. Early modulations imply that high-level expectations rapidly influence sensory processing, integrating high-level priors with key perceptual mechanisms. In contrast, late modulations indicate that predictions may only modulate later inference stages, possibly related to updating priors, or at post-perceptual stages, such as decision-making.

To address the temporal dynamics underlying perceptual prediction, we exposed participants to object images that were expected or unexpected based on preceding cues, while recording EEG. Using a visual deep neural network (DNN) we quantified low- and high-level visual surprise per trial. Finally, we analyzed whether neural responses scaled with these surprise metrics.

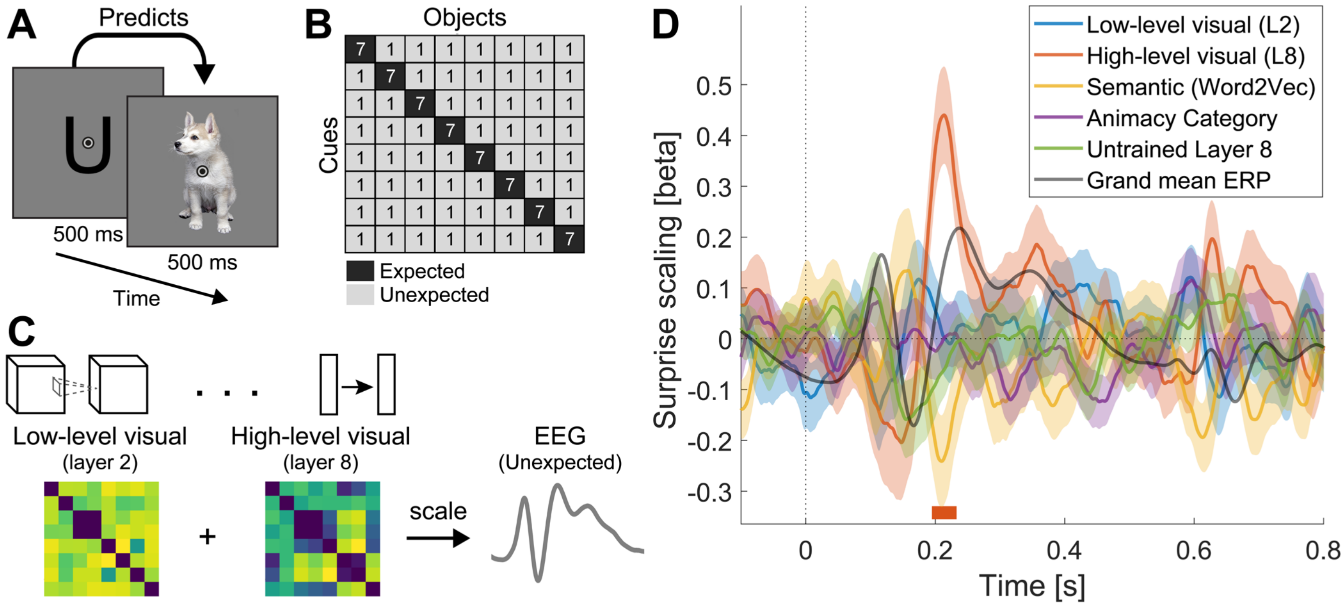

Figure 1: A) Single trial with letter cue (500ms) and predicted object (500ms). **B)** Transitional probability matrix determining the associations between cues and stimuli. Numbers indicate trial numbers within each of the eight blocks. **C)** Analysis rationale illustrating the regression of EEG data onto surprise using the DNN derived representational dissimilarity matrices. **D)** Visually evoked EEG responses scale with high-level visual surprise. Regression of surprise onto ERP amplitudes over parieto-occipital electrodes. Vertical dashed line indicates stimulus onset. Solid lines above the abscissa denote statistically significant clusters.

Methods

Stimuli and Experimental Paradigm. We recorded EEG from 38 participants while they viewed pairs of letter cues and full-color images from various categories. On each trial (Figure 1A), the letter cue probabilistically predicted the identity of the image. Each expected image was seven times more likely to follow its associated cue compared to any other image (Figure 1B). All images appeared as both expected and unexpected stimuli. Participants were not informed about the statistical regularities, but instead were tasked to categorize the entity in the image as animate or inanimate.

Data Analysis and DNN representations. EEG data were recorded using a 64-channel actiCap system (BrainVision). Preprocessing consisted of filtering (0.1 to 128hz), epoching, baseline correction (200ms before cue onset), independent component analysis for artifact removal, interpolation of bad channels, and re-referencing. We focused on visual event related potentials over parieto-occipital electrodes, averaging across Oz, O1, O2, POz, PO7, PO8, PO3, PO4. For each time point (Figure 1C) we regressed EEG amplitudes against visual surprise. Surprise was quantified as the DNN-derived dissimilarity (1 – cosine similarity) of the unexpected seen image from the expected image on that trial. Following previous work (Richter et al., 2024), dissimilarity measures included low-level (layer 2) and high-level (layer 8) visual surprise extracted from AlexNet (Krizhevsky et al., 2017) pretrained on ecoset (Mehrer et al., 2021). Additionally, we included a non-visual semantic surprise model based on word2vec (Mikolov et al., 2013), a task regressor reflecting animacy category surprise, as well as a control model using layer 8 distances from an untrained (randomly initialized) DNN instance.

Results

We assessed whether the amplitudes of visual ERPs elicited by unexpected stimuli were modulated by different levels of surprise. Figure 1D shows that surprise modulated neural responses at parieto-occipital electrodes. A clear peak is evident approximately 200ms post-stimulus onset, showing an upregulation of the evoked response by high-level visual (layer 8) surprise. Indeed, we found only one statistically significant cluster reflecting an increase of the ERP amplitude by high-level visual surprise (t_{max} = 5.39, $p_{corrected}$ = 0.001). No other surprise metric, including low-level visual surprise, attained statistical significance, suggesting that visual processing in our task is predominantly modulated by high-level visual surprise ~200ms post-stimulus onset.

Discussion and Conclusion

Our results revealed that high-level visual surprise modulates neural activity at relatively early stages of sensory processing, approximately 200ms post-stimulus onset, as evidenced by effects on the visually evoked P200 ERP. These findings support and extend previous results (Richter et al., 2024; Uran et al., 2022) and suggest that the brain rapidly incorporates high-level expectations during perceptual inference. Thus, high-level predictions appear to fundamentally influence how sensory processing unfolds instead of (only) post-perceptual processes. The absence of low-level surprise effects further emphasizes the dominant role of high-level information in shaping perception, though this effect could be shaped by task, suggesting that the brain abstracts beyond low-level regularities in favor of high-level information. Relying on high-level predictions may offer evolutionary advantages, because high-level errors can signal important, behaviourally relevant deviations that require rapid belief updating and fast adaptive responses.

Acknowledgments

We thank Tijana Baclic for assistance with data acquisition.

This work was supported by the Marie Skłodowska-Curie Grant “PreVision” (Project 101147241) awarded to DR and grant PID2022.138940NB.I00, funded by MCIN/EI/10.13039/501100011033/ and FEDER, awarded to MR.

References

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends in Cognitive Sciences, 22*(9), 764–779. https://doi.org/10.1016/j.tics.2018.06.002

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84–90. https://doi.org/10.1145/3065386

Mehrer, J., Spoerer, C. J., Jones, E. C., Kriegeskorte, N., & Kietzmann, T. C. (2021). An ecologically motivated image dataset for deep learning yields better models of human vision. *Proceedings of the National Academy of Sciences, 118*(8), e2011417118. https://doi.org/10.1073/pnas.2011417118

Mikolov, T., Chen, K., Corrado, G., Dean, J. (2013) Efficient Estimation of Word Representations in Vector Space. *arXiv*. https://doi.org/10.48550/arXiv.1301.3781

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), 79–87. https://doi.org/10.1038/4580

Richter, D., Kietzmann, T.C., and de Lange F.P. (2024). High-level visual prediction errors in early visual cortex. *PLoS Biology 22*(11): e3002829. https://doi.org/10.1371/journal.pbio.3002829

Uran, C., Peter, A., Lazar, A., Barnes, W., Klon-Lipok, J., Shapcott, K.A., et al. (2022) Predictive coding of natural images by V1 firing rates and rhythmic synchronization. *Neuron 110*(7):1240–1257.e8. https://doi.org/10.1016/j.neuron.2022.01.002

Walsh, K. S., McGovern, D. P., Clark, A., & O’Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences, 1464*(1), 242–268. https://doi.org/10.1111/nyas.14321