Time-Resolved EEG Decoding Reveals a Flip from Enhanced Expected to Unexpected Action Outcomes

Kirsten Rittershofer (k.rittershofer@ucl.ac.uk)
Department of Experimental Psychology, University College London
London, United Kingdom

Daniel Yon (d.yon@bbk.ac.uk)
School of Psychological Sciences, Birkbeck, University of London
London, United Kingdom

Clare Press (c.press@ucl.ac.uk)
Department of Experimental Psychology, University College London
London, United Kingdom

Abstract

It is widely proposed that our brains use expectations about what is likely in order to perceive, but it remains unclear how exactly they shape perception. Bayesian accounts propose that perception is biased towards expected events, ensuring quick and veridical experiences, whereas cancellation accounts argue that unexpected inputs are perceptually prioritized because they are informative. Here, we tested a recent proposal reconciling these views, which suggests that expectations initially bias perception towards what is expected, followed by reactive enhancement of only particularly surprising inputs that are informative for model updating. Using time-resolved decoding of EEG (electroencephalography) data, we provide evidence for this account – enhanced neural representations of expected action outcomes early in time, which later reversed to favour unexpected outcomes. These results suggest that expectations make perception both veridical and informative by exerting distinct influences across time.

Keywords: visual perception; expectations; action; EEG; decoding

Introduction

Our sensory systems are constantly bombarded with noisy input and it has been widely suggested that our brains must use expectations about what is likely in order to perceive (de Lange, Heilbron, & Kok, 2018). However, it is unclear how exactly prior expectations shape perception and neural processing. Bayesian accounts propose that perception is biased towards what we expect (Kersten, Mamassian, & Yuille, 2004). This helps resolve ambiguities in sensory inputs, quickly generating veridical experiences as expected events are, by definition, more likely to occur. Supporting this idea are findings showing enhanced perception and sharper neural representations for expected events (e.g., Stein & Peelen, 2015; Yon, Gilbert, de Lange, & Press, 2018).

In contrast, cancellation accounts argue that perception prioritizes unexpected events, as these are informative (i.e., tell us something we did not already know) and hence important for learning and model updating (Blakemore, Wolpert, & Frith, 1998). This is supported by findings showing enhanced perception and neural representations for unexpected stimuli (e.g., Rideaux et al., 2024; Blank & Davis, 2016). The opposing process theory proposes an alternative to these conflicting accounts, suggesting that expectations can render perception both veridical and informative without sacrificing one adaptive advantage for the other (Press, Kok, & Yon, 2020): Initially, perception is biased towards what we expect with subsequent reactive enhancement of only particularly surprising inputs that are informative for model updating. In the present study, we test this account using EEG to look at how the neural representations of expected and unexpected action outcomes evolve across time.

Methods

Procedure

In each main task trial, participants ($N = 36$) performed an action – either an index or little finger abduction – and observed an outcome on the screen: An avatar hand executing a movement that was congruent (expected) with their own movement on 50% of trials and incongruent (unexpected) on the remaining 50%. They were then asked to identify the observed finger movement (e.g., *Did the index finger move?*). Additionally, in separate localizer blocks, participants passively viewed index and little finger abductions on the screen while performing a detection task at fixation. EEG was recorded throughout both tasks.

Decoding Analysis

We trained a classifier (linear discriminant analysis; LDA) on the localizer data from all 27 electrodes to discriminate

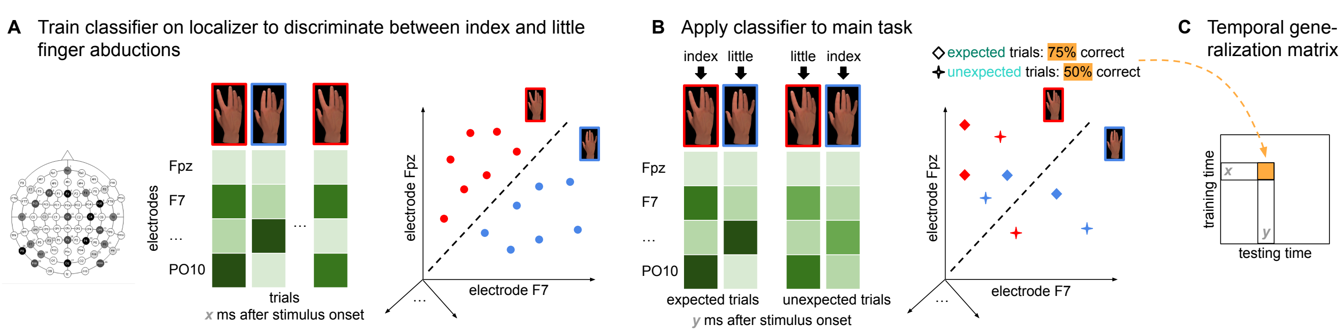

Figure 1: **Analysis approach. (A)** A linear classifier was trained on localizer data from all 27 electrodes at time point x after stimulus onset to distinguish between index and little finger abductions. **(B)** The trained classifier was then applied to the main task data of expected and unexpected trials at time point y. **(C)** The obtained decoding accuracy forms the entry at training time x and testing time y in the resulting temporal generalization matrix. This procedure was repeated for all combinations of training and testing time from -100 to 300 ms relative to stimulus onset, yielding one temporal generalization matrix for each condition.

between index and little finger abductions (Fig. 1A). This classifier was then applied to the main task data, separately for all expected and unexpected trials (Fig. 1B). Decoding was performed in a time-resolved manner: The classifier was trained and tested on each time point from 100 ms before to 300 ms after the onset of the abducted hand in steps of 5 ms, resulting in a temporal generalization matrix (Fig. 1C). Using the training time window when within-localizer decoding peaked (135-175 ms) (following Kok, Mostert, and de Lange, 2017), we compared decoding accuracies between the expected and unexpected condition using cluster-based permutation tests.

Results

Behavioural

Participants were more accurate in judging the identity of the presented finger abduction on expected compared to unexpected trials ($p = .022$; Fig. 2A). However, they responded more quickly on unexpected trials ($p = .010$; Fig. 2B). These differences between conditions suggest that our expectation manipulation was effective.

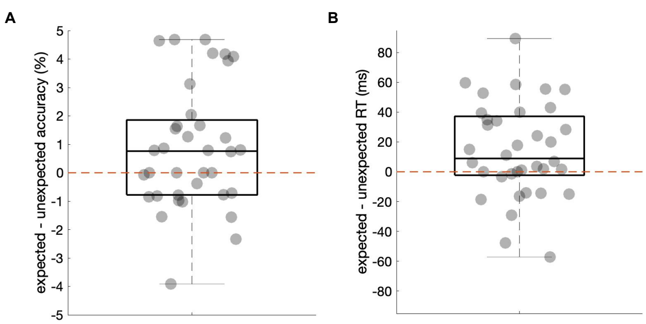

Figure 2: **Behavioural results. (A)** Participants' accuracy difference in identifying the presented finger abduction on expected and unexpected trials. **(B)** Participants' reaction time (RT) difference on correct trials.

EEG Decoding

The diagonal of the temporal generalization matrices shows that we can decode the presented finger abduction from around 100 ms in both the expected and unexpected condition (Fig. 3A and B). Looking at the differences in decoding between the two conditions reveals that early in time (and already before stimulus onset – in line with predicted pre-activation of expectations) decoding is higher in the expected condition. This subsequently reverses such that around 200 ms, post-stimulus decoding is significantly higher in the unexpected condition (Fig. 3C and D).

Conclusion

Our decoding results show enhanced neural representations for expected events early in time, with the condition differences before stimulus onset in line with predicted pre-activation of the expected action outcome (see Kok et al., 2017). Later in time, the effect reverses such that unexpected

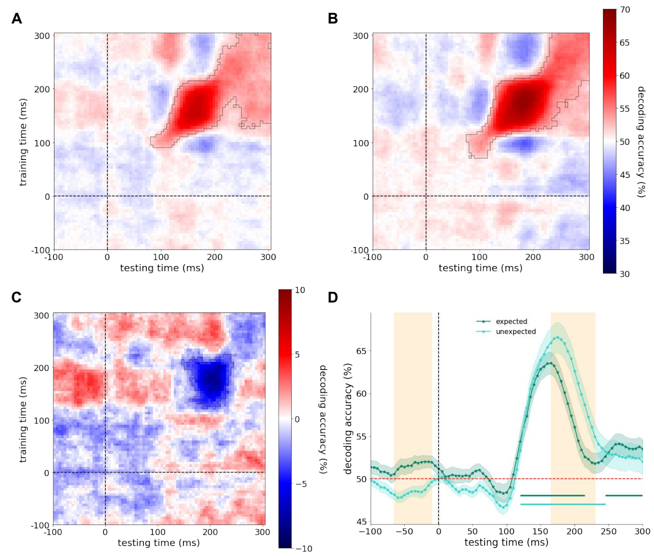

Figure 3: **Decoding results. (A)** Temporal generalization matrix for expected trials with 0 ms being the onset of the abducted hand. **(B)** Temporal generalization matrix for unexpected trials. **(C)** Expected-unexpected differences. **(D)** Decoding traces for expected and unexpected trials averaged over the training time 135-175 ms. Horizontal lines show when decoding is significantly different from chance (50%). Shaded areas mark significant differences between conditions.

events are better represented. These results are in line with the opposing process theory and a previous behavioural study showing enhanced perception of expected action outcomes at 50 ms, followed by enhanced perception of unexpected outcomes at 200 ms (Yon & Press, 2017), as well as with findings from another recent EEG study (McDermott, De Martino, Schwiedrzik, & Auksztulewicz, 2024). Such a temporal dynamic may explain how expectations can make perception veridical, while still allowing for the accurate perception of particularly unexpected events that are important for learning, bridging gaps between earlier conflicting studies and models.

Acknowledgments

This work was supported by a European Research Council (ERC) consolidator grant (101001592) under the European Union's Horizon 2020 research and innovation programme, awarded to CP.

References

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*(7), 635–640.

Blank, H., & Davis, M. H. (2016). Prediction errors but not sharpened signals simulate multivoxel fMRI patterns during speech perception. *PLoS Biology*, *14*(11), 1–32.

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception? *Trends in Cognitive Sciences*, *22*(9), 764–779.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annu. Rev. Psychol.*, *55*(1), 271–304.

Kok, P., Mostert, P., & de Lange, F. P. (2017). Prior expectations induce prestimulus sensory templates. *Proceedings of the National Academy of Sciences*, *114*(39), 10473–10478.

McDermott, H. H., De Martino, F., Schwiedrzik, C. M., & Auksztulewicz, R. (2024). Dissociable dynamic effects of expectation during statistical learning. *bioRxiv*, 1–25.

Press, C., Kok, P., & Yon, D. (2020). The perceptual prediction paradox. *Trends in Cognitive Sciences*, *24*(1), 13–24.

Rideaux, R., Dang, P., Jackel-David, L., Buhmann, Z., Rangelov, D., & Mattingley, J. B. (2024). Violated predictions enhance the representational fidelity of visual features in perception. *bioRxiv*.

Stein, T., & Peelen, M. V. (2015). Content-specific expectations enhance stimulus detectability by increasing perceptual sensitivity. *Journal of Experimental Psychology: General*, *144*(6), 1089.

Yon, D., Gilbert, S. J., de Lange, F. P., & Press, C. (2018). Action sharpens sensory representations of expected outcomes. *Nature Communications*, *9*(1), 1–8.

Yon, D., & Press, C. (2017). Predicted action consequences are perceptually facilitated before cancellation. *Journal of Experimental Psychology: Human Perception and Performance*, *46(3)*, 1073.