CORTEX 138 (2021) 191–202

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

The time-course of prediction formation and revision in human visual motion processing

Tessel Blom [a,*], Stefan Bode [a,b] and Hinze Hogendoorn [a]

[a] Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, VIC, Australia
[b] Department of Psychology, University of Cologne, Cologne, Germany

ARTICLE INFO

Article history:
Received 29 April 2020
Reviewed 10 July 2020
Revised 27 August 2020
Accepted 5 February 2021
Published online 19 February 2021

Keywords:
Prediction
Visual motion processing
EEG
Multivariate pattern analysis

ABSTRACT

Establishing the real-time position of a moving object poses a challenge to the visual system due to neural processing delays. While sensory information is travelling through the visual hierarchy, the object continues moving and information about its position becomes outdated. By extrapolating the position of a moving object along its trajectory, predictive mechanisms might effectively decrease the processing time associated with these objects. Here, we use time-resolved decoding of electroencephalographic (EEG) data from an apparent motion paradigm to demonstrate the interaction of two separate predictive mechanisms. First, we reveal predictive latency advantages for position representations as soon as the second object in an apparent motion sequence – even before the stimulus contains any physical motion energy. This is consistent with the existence of omni-directional, within-layer waves of sub-threshold activity that bring neurons coding for adjacent positions closer to their firing threshold, thereby reducing the processing time of the second stimulus in one of those positions. Second, we show that an additional direction-specific latency advantage emerges from the third sequence position onward, once the direction of the apparent motion stimulus is uniquely determined. Because the receptive fields of early visual areas are too small to encompass sequential apparent motion positions (as evidenced by the lack of latency modulation for the second stimulus position), this latency advantage most likely arises from descending predictions from higher to lower visual areas through feedback connections. Finally, we reveal that the same predictive activation that facilitates the processing of the object in its expected position needs to be overcome when the object's trajectory unexpectedly reverses, causing an additional latency disadvantage for stimuli that violate predictions. Altogether, our results suggest that two complementary mechanisms interact to form and revise predictions in visual motion processing, modulating the latencies of neural position representations at different levels of visual processing.

* *Corresponding author*. Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, VIC 3010, Australia.
E-mail address: tblom@student.unimelb.edu.au (T. Blom).

https://doi.org/10.1016/j.cortex.2021.02.008
0010-9452/

1. Introduction

One of the fundamental tasks of the visual system is to localize objects in the outside world. However, when an object is moving, establishing its position at a given instant is complicated by neural transmission delays; while incoming sensory information from the eyes is travelling to visual areas in the brain, the object continues moving. If the visual system did not somehow compensate for its own processing time, the perceived position of a moving object would always lag its true position.

The view that brains are essentially prediction machines that are continuously attempting to match the incoming sensory inputs to expectations or predictions has become increasingly influential in the cognitive sciences (Clark, 2013). By implementing predictive mechanisms through motion extrapolation, the visual system could overcome the challenge of neural transmission delays. Indeed, a wide range of evidence from electrophysiological, neuroimaging, and behavioural studies over the past decades supports the existence of motion extrapolation mechanisms in the visual system (see Hogendoorn, 2020, for a recent review). Perhaps most prominently, motion extrapolation has been argued to underlie a visual illusion known as the Flash Lag Effect (FLE; Nijhawan, 1994). In the FLE, a flash is temporally and spatially aligned with a moving bar, but is perceived to lag the position of the moving bar. The exact neural mechanisms have been debated in the years following Nijhawan's presentation of the illusion in terms of extrapolation mechanisms. In this interpretation, the predictability of the trajectory of the moving bar is used by the visual system to spatially extrapolate the location of the object into the present (thereby reducing the impact of neural delays). The stationary flash on the other hand cannot be extrapolated, which leads to a spatial misalignment between the moving object and the stationary flash (Nijhawan, 1994, 2008). A prominent alternative explanation for the illusion explains the effect in terms of differential latencies. This account poses that moving objects are processed more rapidly than stationary objects, such that the position of a moving object is more up-to-date than the position of the flash, which again leads to a perceived spatial misalignment (Whitney & Murakami, 1998; Whitney, Murakami, & Cavanagh, 2000). Although often pitted against each other, these different accounts do not need to be conflicting. Where the differential latencies account speaks to a purely temporal, local, low-level neural mechanism, extrapolation implies a spatial mechanism that might operate through lateral or descending predictions and the two explanations could be complementary in a motion context (Maiche, Budelli, & G'omez-Sena, 2007).

A large body of evidence speaks to the spatial extrapolation account involving cortical feedback mechanisms. For instance, motion prediction is observed not only for continuous motion, but also for apparent motion (Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010; Blom, Feuerriegel, Johnson, Bode, & Hogendoorn, 2020; Hogendoorn & Burkitt, 2018; Hogendoorn, Carlson, & Verstraten, 2008), where a stimulus is presented sequentially in discrete positions but appears to 'move' through the intervening empty space. Because receptive fields in the early visual system are smaller than the distance between the positions in the apparent motion array, these stimuli do not trigger an early motion signal (Muckli, Kohler, Kriegeskorte, & Singer, 2005; Smith, Singh, Williams, & Greenlee, 2001). Nevertheless, a latency benefit can be observed for neural representations of stimuli along predictable apparent motion trajectories (Hogendoorn & Burkitt, 2018) as well as pre-activation of neural representations of the anticipated position ahead of the arrival of corresponding sensory information (Alink et al., 2010; Blom et al., 2020). These effects have been argued to result from later feedback connections from higher visual areas, particularly MT. When an illusory percept of motion is induced by subsequently presented stimuli, V1 regions not directly activated by the motion-inducing stimuli but located along the apparent motion path have been shown to display a BOLD response as if there was actual motion (Muckli et al., 2005; Sterzer, Haynes, & Rees, 2006). This neural activity might be caused by feedback connections from hMT+/V5 as shown through effective connectivity (Sterzer et al., 2006) and the fact that this neural activity in V1 disappears when TMS is applied over hMT+/V5 (Vetter, Grosbras, & Muckli, 2015). Additionally, V1 has been found to perform pattern completion, where the presentation of the beginning of a temporal sequence evoked neural patterns that represent the entire sequence (Ekman, Kok, & De Lange, 2017; Xu, Jiang, Poo, & Dan, 2012). More generally, prior cognitive expectations about the identity of a stimulus have been found to evoke neural activity in early sensory cortices that correspond to the expected stimulus in multiple modalities, even when that stimulus is eventually omitted, again supporting the idea of a top-down pathway for prediction (Andersen & Lundqvist, 2019; Demarchi, Sanchez, & Weisz, 2019; Kok, Failing, & de Lange, 2014; Kok, Mostert, & De Lange, 2017).

Moreover, Wibral, Bledowski, Kohler, Singer, and Muckli (2009) showed that the C1 component of the event-related potential (ERPs), with an onset of 60 ms post-stimulus, was not influenced by whether a stimulus was presented within an apparent motion context or not. Although the presence of apparent motion could be detected in V5 from an early EEG time component at around 100 ms, the response to the stimulus position was only modulated in later time windows of 110–150 ms, again showing that feedback connections are necessary for predictive modulations in early visual cortex (Wibral et al., 2009). This parallels the pattern of results from Hogendoorn and Burkitt (2018), who also used apparent motion and likewise showed that the latency of late (140–150 ms), but not early (80–90 ms) neural representations were affected by the predictability of the motion sequence (Hogendoorn & Burkitt, 2018). More generally, no modulations of the C1 component were found in a study looking at which components of the ERP were modulated by prediction. Conversely, later components were modulated by the predictability of the stimulus (Alilovíc, Timmermans, Reteig, Van Gaal, & Slagter, 2019).

Other lines of evidence demonstrate that predictive mechanisms are already at play at the earliest stages of visual processing reducing processing times for moving objects. A moving bar has been shown to elicit a wave of spiking activity in the retinae of salamanders and rabbits, such that the

population activity travels near the leading edge of the bar, effectively extrapolating its anticipated position (Berry, Brivanlou, Jordan, & Meister, 1999). With the retina's ability to spatially extrapolate moving stimuli, all visual layers downstream will be able to represent the location of a (predictably) moving stimulus with a shorter latency, simply because sensory information coding for the stimulus in a certain location becomes available earlier.

In addition, long-range horizontal connections between modules that process similar features of stimuli are well established within the visual cortex (Gilbert & Wiesel, 1989; Toth, Rao, Kimt, Somers, & Sur, 1996). This connectivity leads to a spread of activity that is not able to drive responses, but can have a facilitatory influence (Frégnac, Bringuier, & Chavane, 1996). This sub-threshold propagating activity has been suggested to act as a pre-activation of adjacent cortical regions, hence effectively preparing the cortex for the object potentially starting to move (in any direction). By bringing neurons representing the position ahead of the stimulus closer to their firing threshold, neuronal processing time could be reduced, and a moving object could be represented with a shorter latency (Benvenuti et al., 2020; Chemla et al., 2019; Jancke, Chavane, & Grinvald, 2010; Jancke & Erlhagen, 2010; Series, Lorenceau, & Frégnac, 2003). V1 indeed exhibits a shorter time-to-peak latency for moving than stationary stimuli in both cats (Jancke, Erlhagen, Schöner, & Dinse, 2004) and macaques (Subramaniyan et al., 2018). A build-up of anticipatory spiking responses occurs long before a moving bar enters the classical receptive field in V1 neurons (Benvenuti et al., 2020), and this pre-activation of cortical regions ahead of the arrival of sensory input from the moving stimulus has also been shown to contribute to the FLE (Jancke, Chavane, Naaman, & Grinvald, 2004; Jancke & Erlhagen, 2010; Jancke et al., 2010). Furthermore, Maiche et al. (2007) demonstrated that the pre-activations of two objects that would collide summate and lead to a larger FLE by decreasing the processing time of a moving object even more. This suggests that early, low-level mechanisms also generate a latency advantage for moving stimuli through lateral connections that can help counteract neural transmission delays.

Taken together, it seems likely that predictive mechanisms are implemented at multiple levels of the visual hierarchy, with the retina and the earliest visual areas potentially implementing prediction in feedforward and lateral signals for (smooth) motion, and higher cortical areas implementing prediction through top-down, feedback mechanisms. This interpretation is consistent with the conclusion of a recent behavioural paradigm implicating prediction mechanisms at both monocular and binocular levels of the visual pathway (van Heusden, Harris, Garrido, & Hogendoorn, 2019). A core function of these extrapolation mechanisms is presumably to allow us to perceive (and interact with) moving objects with greater temporal precision; however, it remains unknown how quickly predictive signals are formed when viewing a moving object and how predictive mechanisms at these different levels might interact. There are two closely related aspects to the question of when predictive signals arise during visual processing: when can we find neural representations influenced by prediction during visual processing? And when presented with a new moving object, how rapidly does the visual system form a predictive representation of its position?

This second question, how rapidly motion prediction signals develop over time, has received much less investigation. It has been argued that for an object in smooth motion, the necessary motion signals can be extracted during feedforward processing within 2–3 ms, and so extrapolation could take place nearly instantly (Nijhawan, 2008). The Fröhlich effect, a visual illusion where the initial position of a moving object that suddenly appears is perceived shifted in the direction of its motion (Fröhlich, 1924; Kirschfeld & Kammer, 1999), corroborates the suggestion that a motion extrapolation signal might be formed nearly instantly. However, the same cannot be true for apparent motion paradigms for two reasons. Firstly, as noted before, receptive fields in early visual areas are too small to extract a motion signal from widely-spaced stimuli in an apparent motion display. Secondly, before the presentation of the second stimulus, an apparent motion display contains no motion yet. It therefore minimally requires the time in-between two stimulus presentations before a motion signal can be extracted from such a display. It seems likely that predictive mechanisms at different levels of the hierarchy might integrate motion information over different temporal windows, with early, low-level mechanisms potentially generating extrapolation signals very rapidly following initial stimulation through propagating activity, and later mechanisms integrating more trajectory information. However, to our knowledge there is no available empirical evidence that speaks to this question.

Here, we used a time-resolved EEG decoding paradigm to investigate how rapidly motion predictions are formed for different neural representations. We adopted a circular apparent motion paradigm (as previously used in Hogendoorn and Burkitt (2018) and Blom et al. (2020)), which allowed us to separate the early feedforward and lateral neural activity from neural patterns which could include activity from higher areas and feedback mechanisms. We then asked how much information about the motion trajectory was required for a latency advantage to develop for different neural representations. By investigating both the start of motion sequences (in which predictions are first formed) as well as after unexpected motion reversals (in which previously formed predictions are violated, and new predictions must be formed), we could compare the formation and revision of motion predictions.

2. Methods and materials

The current experiment includes data collected using two slightly different protocols. Protocols did not differ in any parameters relevant to the current analyses, since the stimulus presentations analysed here are present in both protocols. Rather, these differences involved the relative frequency of the different types of trials, and how those trials were divided over blocks and sessions for the comfort of the observer. In control analyses split by study protocol, we further confirmed that there was no difference in the general pattern of results (data not shown). Except where otherwise stated, the following details pertain to both protocols. Study methods and analyses were not pre-registered prior to the

research being conducted. MATLAB code used to run the experiment and analyse the data will be available at https://osf.io/9ms6k/at the time of publication. We report how we determined our sample size, all data exclusions (if any), all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

2.1. Participants

Twelve observers (6 female, average age 25) with normal or corrected-to-normal vision participated under the first protocol, and twenty observers (12 female, average age 23) with normal or corrected-to-normal vision participated under the second protocol. Sample size was based on previous research using the same paradigm (Blom et al., 2020). Both protocols were approved by the human research ethics committee of the University of Melbourne (Ethics ID 1954628), Australia, and conducted in accordance with the Declaration of Helsinki. All observers signed informed consent before participating in the experiment and were reimbursed AUD15 per hour for their time.

2.2. Stimuli

The stimulus was a black, truncated wedge presented on a uniform 50% grey background. The stimulus could be presented in one of eight equally-spaced locations around a white central fixation point, at 22.5°, 67.5°, 112.5°, 157.5°, 202.5°, 247.5°, 292.5°, and 337.5° of polar angle from the vertical. Inner and outer edges of the wedge were 6.3 and 7.7 degree of visual angle (dva) away from fixation, respectively. The wedge covered 11° of polar angle, with 1.3 dva at the inner and 1.5 dva at the outer edge. The stimulus was presented for 66 ms, with an inter-stimulus interval of 33 ms and an inter-trial interval of 400 ms between sequences. Stimuli were presented on an ASUS ROG PG258 monitor with a resolution of 1920 × 1080 running at 120 Hz. The monitor was controlled by a HP EliteDesk 800 G3 TWR PC running Matlab R2017b with PsychToolbox 3.0.14 extensions (Brainard, 1997). Observers viewed the stimuli from a headrest at a distance of 60 cm.

2.3. Behavioral task

Observers were instructed to respond with a button press on the keyboard whenever a wedge was presented in red instead of black. This occurred a total of 32 times per block under Protocol 1 and 50 times per block under Protocol 2. The purpose of the task was to keep observers engaged with the stimulus, and behavioural data was not analysed. Under Protocol 1, trials with targets were discarded, and target trials were rerun at the end of each block. Under protocol 2, trials with targets were simply discarded.

2.4. Experimental design

Twelve observers completed 6 blocks of sequences across 3 testing sessions (Protocol 1). Twenty observers completed 2 blocks across 2 testing sessions (Protocol 2).

Under Protocol 1, each block contained the following types of trials, randomly interleaved:

- Sequences with 1, 2, or 3 consecutive presentations (3 sequence lengths x 8 starting positions x 2 directions x 10 repetitions = 480 trials).
- Sequences with 4, 5, 6, 7, or 8 consecutive presentations (5 sequence lengths x 8 starting positions x 2 directions x 2 repetitions = 160 trials).
- Sequences with 16, 20, 24, 28, 32, 36, 40, or 44 consecutive presentations (8 sequence lengths x 8 starting positions x 2 directions = 128 trials).
- Sequences with 16, 20, 24, 28, 32, 36, 40, or 44 consecutive presentations followed by a reversal and continuation in the opposite direction for 8–16 (randomly determined) additional presentations (8 sequence lengths x 8 starting positions x 2 directions = 128 trials).

A target stimulus was randomly presented in 32 trials per block. Because these trials were appended to the trial list, each block encompassed 928 trials (split up into 16 sets of 58 trials). Each new set was initiated with a button press by the observer. Each observer completed two blocks per session and a block lasted approximately 30 min. Taken together, each observer completed 5,568 trials.

Under Protocol 2, all types of trials were combined in a single block, randomly interleaved:

- Sequences with 4, 5, 6, 7, or 8 consecutive presentations (5 sequence lengths x 8 starting positions x 2 directions x 8 repetitions = 640 trials).
- Sequences with 9, 10, 11, 12, 13, 14, 15, or 16 consecutive presentations (8 sequence lengths x 8 starting positions x 2 directions x 4 repetitions = 512 trials).
- Sequences with 9, 10, 11, 12, 13, 14, 15, or 16 consecutive presentations followed by a reversal and continuation in the opposite direction for 1–8 (randomly determined) additional presentations (8 sequence lengths x 8 starting positions x 2 directions x 4 repetitions = 512 trials).

In each block a target was randomly presented in 50 trials and these trials were discarded. Each block was split up into 13 sets and each new set was initiated with a button press by the observer. In a session, observers completed one block, taking approximately 90 min. Taken together, each observer completed 3,328 trials.

2.5. EEG acquisition and preprocessing

64-channel EEG data, placed according to the 10/20 system, and data from 6 EOG and 2 mastoid channels were acquired using a BioSemi ActiveTwo EEG system sampling at 2048 Hz. EEG data were re-referenced offline to the average of the two mastoid electrodes and resampled to 512 Hz. Eleven observers had one bad channel during one of the sessions. This channel was spherically interpolated using EEGlab (Delorme & Makeig, 2004).

All data was epoched relative to stimulus onset. For classification analyses, we make a distinction between two types of epochs: training and test epochs. Training epochs were

used to train classifiers to distinguish stimulus positions, whereas test epochs were used to test whether the classifiers generalized to unseen data.

Under both protocols, the training epochs were time-locked to the first presentation in a sequence. The first presentation in each sequence was random and has no history and therefore its position could not be anticipated. Half of the first presentation epochs were assigned to the training set and the other half was assigned to the test set. The training epochs were taken from 800 ms before stimulus onset to 500 ms after and were baseline-corrected to the mean of the 200 ms period before stimulus onset.

The test set consisted of all other stimuli (i.e., all consecutive presentations), and epochs were again taken from 800 ms before stimulus onset to 500 ms after and were baseline-corrected to the mean of the 800 ms period before stimulus onset under Protocol 1. This baseline period was chosen such that it was consistent across all epochs and contained a full-cycle of motion for the majority of the epochs in order to avoid introducing stimulus specific differences as much as possible. Under Protocol 2, epochs were baseline-corrected to the mean of the 200 ms before the first stimulus presentation. Sequences under Protocol 2 were much shorter making it possible to baseline to a period without stimulus presentations. The number of epochs was matched across test sets (e.g., the same number of epochs for sequence position 1, 2, 3, 4).

All epochs with large artifacts were removed using an automated rejection procedure. Epochs for which the standard deviation of amplitudes across all electrodes and all time-points exceeded 50 μV were rejected for analysis. Across all observers, 35.2% (SD = 16.0%) of epochs were removed in this way. This conservative rejection procedure was chosen to avoid introducing systematic patterns in the data which the classifiers might exploit.

2.6. *Multivariate pattern analysis*

Epochs from the training sets were used to train time-resolved linear discriminant analysis (LDA) classifiers (Grootswagers, Wardle, & Carlson, 2017). The classifier was first trained to separate neural patterns associated with the presentation of the stimulus in two different positions using the amplitudes from all 64 EEG channels as features. To calculate classification performance, a set of LDA classifiers was trained on each pairwise combination of eight stimulus positions at each individual time-point in the epoch for each individual observer. Each time-point was defined as the average of the surrounding 10 ms time-window. The classifier was trained on half of the data (to estimate a model) and subsequently tested on the other half (to test the quality of the model, i.e., how well it generalized). This procedure was repeated independently after swapping the roles of the two halves and both classification performances were averaged. An above-chance classifier performance at a certain time-point indicates that the distributed pattern of EEG signal contained information that allowed the classifier to distinguish the two positions of the stimulus. The mean classification performance across all pairwise comparisons was statistically tested using permutation testing. A null distribution was created by generating 1,000 datasets with shuffled position labels (i.e., the position within the stimulus array) and each permuted dataset was analysed in the exact same way as the original dataset. Time-points at which the observed classification performance fell in the highest or lowest .5% of the null-distribution were taken as significant ($p < .01$, two-tailed, uncorrected). In addition, a temporal generalization matrix (TGM) was generated for all combination of training and test time points. This allowed investigating when a distinct neural representation was formed and for how long it remained stable (King & Dehaene, 2014).

To investigate how quickly predictions build up and where along the visual hierarchy latency advantages emerge, classifiers were trained on distinguishing stimulus locations in the training set (where the stimulus locations were unpredictable) and were tested on the same stimulus locations in the test set. Because the predictability of the location of the stimulus increased when it was presented within a motion sequence (instead of being randomly presented in one of eight locations in the training set), this analysis allowed us to investigate how quickly a predictive latency advantage arose both when predictions were newly formed (at the start of a sequence) and when predictions needed to be revised (following reversals, see Fig. 1). In the testing procedure, each stimulus location was tested pairwise against the respective other 7 possible stimulus locations, revealing 7 measures of neural evidence in favour of the one location compared to the other locations. Summing classification assignments in favour of the location being presented, and repeating this for all epochs containing all 8 locations, yielded a measure called Classification Anisotropy Index (CAI) (Hogendoorn & Burkitt, 2018). A positive CAI at any given time-point indicated that there was neural evidence in favour of the presented location, a negative CAI indicated neural evidence against the presented location, and a CAI of zero meant there was no information regarding the presented location in the neural representation. For stimuli presented later on in the sequence, the pairwise comparison to the previously presented stimulus is included in the CAI. However, any potential lingering neural activity in response to that previously presented stimulus would not be expected to have a net effect on the CAI. This is because the CAI includes pairwise comparisons against other positions further away (which might introduce a positive bias) as well as the previously presented position (which would provide a similar but negative bias). Hence, overall CAI is unlikely to be affected by any position biases. Classifiers trained on training time-points of 70, 90, 110, 130 and 150 ms from the training set were tested on all time-points from different sequence position epochs. This analysis resulted in a CAI curve for each training time point and each sequence position, showing which neural representation becomes active when, depending on the position of the stimulus in the sequence.

Subsequently, in order to compare the latency of the decoding peaks for the different sequence positions, we calculated the latency at which the CAI curves first exceeded the threshold value of 80% of the proportion of peak CAI performance, for each sequence position, for each training time-point and for each individual participant (the exact value of the threshold did not affect the interpretation of the results, as demonstrated in Figure S1 in the Supplemental

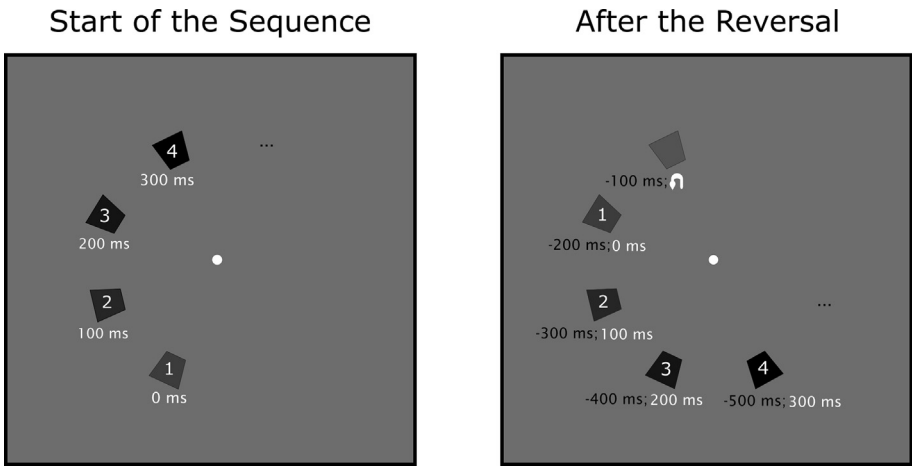

Fig. 1 – Stimuli were presented for 33 ms with an inter stimulus interval of 66 ms. To investigate how quickly prediction builds up, we tested classifiers that were trained on all neural representations on epochs from the first four stimuli at the start of every sequence (left panel). To see how quickly predictive mechanisms recovered after a prediction violation, we tested classifiers on epochs from the first four stimuli moving in the opposite direction to the initial sequence direction (right panel).

Information). For four participants, it proved impossible to establish the peak CAI performance, and they were excluded from this analysis. The latency analyses therefore included a total of 28 participants. Latency plots were developed using the Gramm data visualization toolbox (Morel, 2018).

3. Results

3.1. *Classification performance*

Fig. 2A shows the mean classification performance averaged across all pairwise comparisons and all observers. Classifiers were able to decode the presented stimulus location starting 50 ms after stimulus onset, meaning that from 50 ms onward there was information in the neural signal regarding stimulus position. To explore the temporal structure of neural representations activated by the presentation of the unpredictable stimulus (i.e., the first stimulus in the sequence), Fig. 2B shows the temporal generalization matrix (TGM). This shows that distinct neural representations were activated (red) and deactivated (blue) sequentially. The first representation cluster between approximately 50 and 100 ms did not cross-generalize to 100–150 ms (as indicated by blue clusters off the diagonal), and conversely, the 100–150 ms representation did also not cross-generalize to 50–100 ms. This can be interpreted as these neural representations being distinctly different and probably representing different stages of visual processing.

3.2. *Classification Anisotropy Index*

Classifiers trained on distinguishing stimulus locations in the (unpredictable) training set were tested on the same stimulus locations when they were presented within the apparent motion sequence. Summing the evidence in favour of the presented stimulus location (deviation from chance level) across all pairwise comparisons yields a measure of Classification Anisotropy Index (CAI) (Hogendoorn & Burkitt, 2018). Classifiers trained on training time-points of 70, 90, 110, 130 and 150 ms from the training set (indicated by the vertical lines in Fig. 2A and B) were tested on all time-points from different sequence position epochs at the start of the sequence and after the reversal (an illustration of stimulus positions can be found in Fig. 1).

Fig. 3 shows when different neural representations from the training set were activated depending on the position of the stimulus in the sequence. From top to bottom, the different rows show the performance of classifiers trained on progressively later neural representations. Markers below each plot indicate the time point at which the corresponding curve peaks. For early representations (70–90 ms training time; first two rows), both at the start of the sequence (left panels) as well as after the reversal (right panels), maximum CAI was observed at similar time-points. This implies that early on, neural representations follow a similar time-course regardless of the position of the stimulus in the sequence. However, with increasing training times (>110 ms, third to fifth row), the CAI curve for the first sequence position appeared to peak later than the curves for the second, third and fourth sequence positions, both at the start of the sequence as well as after the reversal. Note that the CAI amplitude of the first stimulus after the reversal is much smaller than at the start of the sequence, which could be attributable to a diminished response due to repetition suppression or a lower signal-to-noise ratio since the stimulus was presented just 200 ms prior.

3.3. *Latency comparisons of CAI peaks*

To quantify the differences in latency between the CAI curves of different sequence positions, we calculated the latency at which the CAI first exceeds 80% of peak performance for each training time-point and each sequence position (in

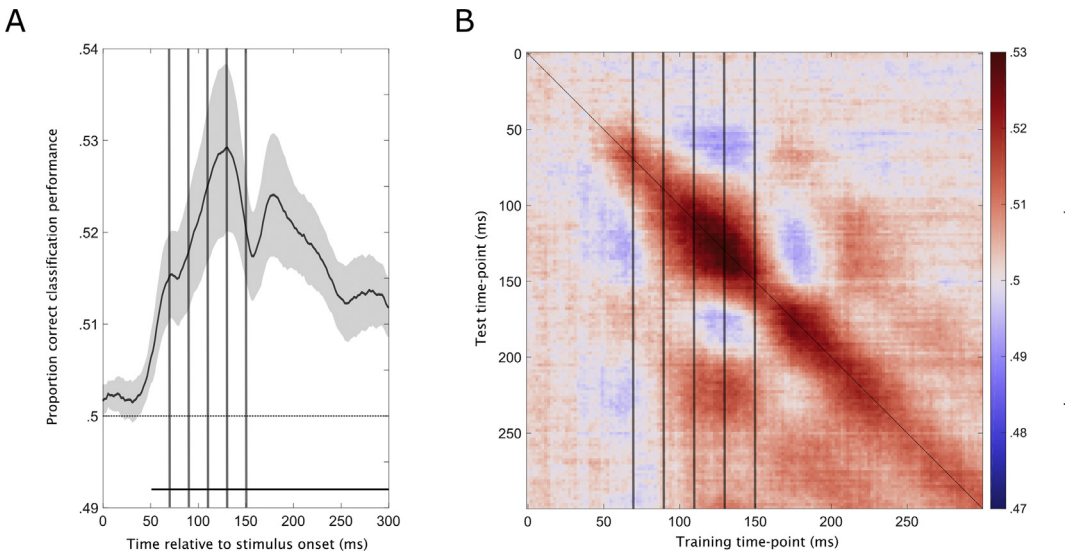

Fig. 2 – Mean classification performance over time for multivariate classifiers trained to distinguish neural patterns evoked by the presentation of a black wedge in each of eight possible positions. Stimuli were presented in sequences, and classifiers were trained and tested only on the first stimulus of each sequence. A) Decoding position from the neural signal. Classification Performance is calculated for each participant separately and subsequently averaged across all participants. The plotted classifier performance reflects the diagonal line from panel B where training and test time-points are the same. Shaded areas indicate the 95% confidence intervals. Time-points at which classification performance significantly exceeded a permuted-labels null-distribution ($p < .01$) are indicated by the line below the curve. Classification performance first rose above chance at 50 ms post-stimulus. Vertical lines indicate the training time-points further investigated in follow-up analyses. B) A temporal generalization matrix (TGM) showing classification performance across all possible pairwise comparisons at each combination of training and testing time-points. Above chance classification performance (red) indicate points in training-testing-time space where neural representations for the presented stimulus location were activated. Vertical lines indicate the training time-points further investigated in follow-up analyses.

accordance with Hogendoorn & Burkitt, 2018; note that changing the threshold value did not change the results in a meaningful way, see Figure S1 in the Supplemental Information for other threshold values). The left panel in Fig. 4 shows the 80% peak latencies for Sequence Position one (SP1), Sequence Position two (SP2), Sequence Position three (SP3) and Sequence Position four (SP4) relative to the start of the sequence. We did not statistically analyse sequence positions of five or greater, because progressively fewer trials were available with increasing sequence position. This caused classification performance to become prohibitively noisy and latency estimates to become unreliable. The pattern of results for SPs 5–7, however, was almost identical to SP3 and SP4 (see Figure S2 in the Supplemental Information). On average, the stimulus location could be decoded from 50 ms onward, but the onset of classification performance varied between participants. This means that for nine participants (five of the otherwise included participants) it was not possible to establish the CAI peak latency until 70 ms. To present only results from a consistent sample, we therefore only display the results starting at 70 ms. To test for latency differences between sequence positions, a paired-samples *t*-test was conducted to test for which training time-points the latency of SP1 was significantly different from later SPs (top rows of dots at the bottom of the panel; $p < .05$, uncorrected) and for which training time-points SP2 was significantly different from later SPs (bottom rows of dots at the bottom of the panel; $p < .05$, uncorrected). For early training time-points (e.g., approximately 70 - 95 ms), there was no difference in latency between any of the later sequence positions compared to SP1. From approximately 95 ms onward, there was a significant difference in latency for SP3 and SP4 compared to SP1. From approximately 125 ms onward, also SP2 showed a significant latency difference when compared to SP1. Additionally, SP3 and SP4 showed a latency difference compared to SP2 for early training time-points (e.g., approximately 110 - 125 ms and 90 - 130 ms respectively). In other words, the early representations coding for SP2 appeared to be processed like SP1 (with no latency advantage), but the later representations appeared to be processed like SP3 (with a latency advantage). For SP3 and later positions, both early and late representations showed a latency advantage over SP1.

On Reversal trials, the motion sequence reversed direction after continuing in the original direction for at least one or two cycles (Protocol 2 and Protocol 1 respectively), after which a stable motion prediction would have been formed (see Fig. 1). The latency at which CAI reached 80% of peak was calculated for each sequence position after reversal in the same way as for the different sequence positions at the start of the sequence (Fig. 4, right panel). We observed that after the

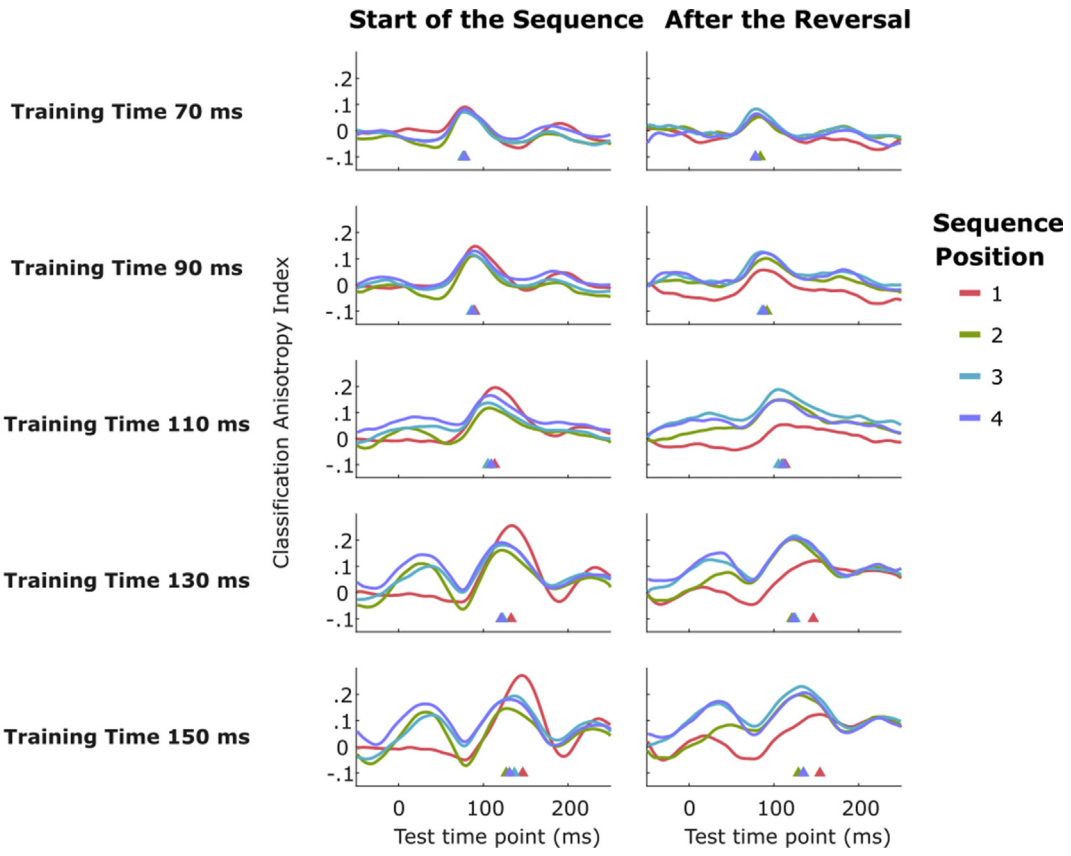

Fig. 3 – Classification Anisotropy Index (CAI; i.e., summed evidence in favour of the presented stimulus location compared to the other stimulus locations) as a function of time for the first, second, third and fourth position in the sequence at the start of the sequence (left column) or after a motion reversal (right column). Each row shows the CAI curves for increasing training times. The different CAI curves in each panel represent different sequence positions. CAI curves are calculated for each participant separately and subsequently averaged across all participants. The markers at the bottom show the time-point at which the corresponding CAI curve peaks.

reversal, all sequence positions after SP1 had a significant latency difference from SP1, for all neural representations starting at approximately 100 ms (as indicated by the dots at the bottom of the graph in Fig. 4, $p < .05$, uncorrected).

Comparing the pattern of latency advantage for different sequence positions between stimuli presented at the start of the sequence and following a reversal, we observed a difference for SP2. Whereas at the sequence start SP2 diverged from SP1 only for later neural representations, for post-reversal trials SP2 diverged from SP1 already for early neural representations. As a result, following a reversal there was no longer a difference between SP2 and SP3 or any of the following sequence positions.

Finally, we directly compared the first sequence positions after a reversal and at the start of the sequence in Fig. 5. We observed latency differences between SP1 at the start and SP1 after a reversal between approximately 110 - 130 ms. Both SP1s were unpredictable, but after the reversal there was also a prediction violation for SP1 that was naturally not present at the start of a sequence. Our results show that this prediction violation led to a latency disadvantage for mid-level representations of the violation stimulus compared to the same representations of an unpredictable stimulus.

4. Discussion

Establishing the position of a moving object poses a challenge to the visual system due to inherent neural transmission delays. Predictive mechanisms modulating the latency of neural representations that could help counteract some of these delays have been established in both animals and humans (Subramaniyan et al., 2018; Jancke, Erlhagen, et al., 2004; Hogendoorn & Burkitt, 2018), and here we extend these findings. Using a multivariate pattern classification approach applied to time-resolved EEG recordings, we investigated which neural representations first exhibit latency advantages, and how rapidly these predictive latency advantages arise.

We find that the very earliest neural representations of incoming visual information do not show latency advantages for predictable, apparent motion stimuli, which is consistent with previous research (Wibral et al., 2009; Alilović et al., 2019; Hogendoorn & Burkitt, 2018). These neural representations (between approximately 50-95 ms post-stimulus) reflect the first detectable neural activity containing information about the position of the stimulus. Since we used an apparent motion paradigm, it is likely that the distance between the stimulus positions is too large for lateral connections to

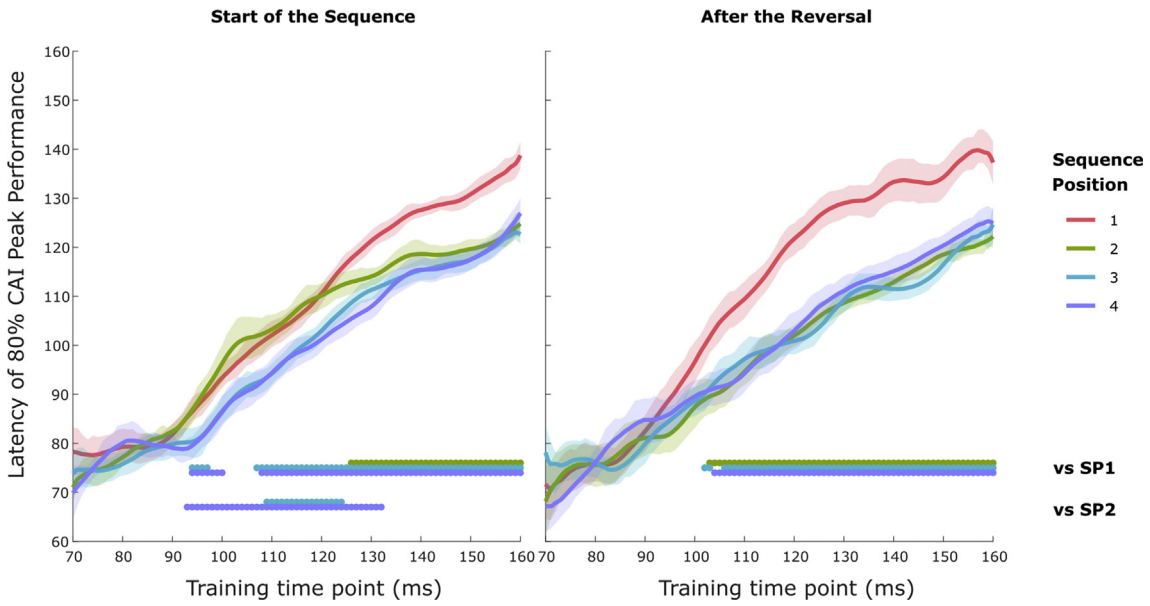

Fig. 4 – Latency of 80% CAI peak performance for each training time-point and sequence position at the start of the sequence and after the reversal. CAI curves are calculated for each participant separately and subsequently averaged across all participants. Shaded areas indicate Standard Error of the Mean. Dots at the bottom indicate individual time-points at which the peak latency for Sequence Positions 2–4 was significantly different from either Sequence Position 1 (SP1) or Sequence Position 2 (SP2) as tested with a paired-sample *t*-test at $p < .05$, uncorrected.

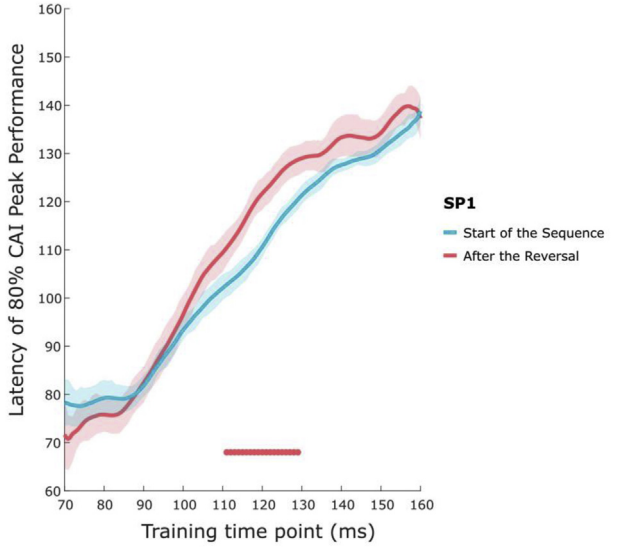

Fig. 5 – Latency of 80% CAI peak performance for each training time-point for Sequence Position 1 (SP1) at the start of the sequence and after the reversal. CAI curves are calculated for each participant separately and subsequently averaged across all participants. Shaded areas indicate Standard Error of the Mean. Dots at the bottom indicate individual time-points at which the peak latency between the latency curves was significantly different as tested with a paired-sample *t*-test at $p < .05$, uncorrected.

modulate the latency of these neural representations. Neural representations between 100 - 150 ms do not cross-generalize to the earlier test time-points, meaning that they are distinct from the initial 50–95 ms representations. We find that this next level of visual processing can benefit from predictions induced by apparent motion, since neural representations of predictable stimuli exhibit latency advantages over neural representations of unpredictable stimuli.

Our second question concerned how much information about an object's past trajectory is required to generate a prediction. Our results suggest that the answer to this question is not completely straight forward. We showed that at the start of a sequence, when a prediction was first formed, Sequence Position 2 already exhibited latency advantages (for later neural representations >125 ms). This is later than the latency advantage observed for later sequence positions (i.e., Sequence Position 3 and further), where a latency advantage already emerges around 95 ms. In our apparent motion paradigm, the location of Sequence Position 1 is 12.5% predictable, the location of Sequence Position 2 is 50% predictable and the locations of Sequence Position 3 and later are 100% predictable. Even though the location of the second stimulus is only 50% predictable, we still observed a latency advantage, albeit for slightly later neural representations than the latency advantages observed for the third stimulus and thereafter. One interpretation of this finding is that descending predictions regarding the location of the second sequence position may not have been formed yet (due to the relatively high level of uncertainty remaining), such that the latency advantage for the later neural representations for this second stimulus are more likely caused by horizontal connections. The first sequence position at the start of the sequence could

induce a travelling wave of activity in all directions of space (Jancke, Erlhagen, et al., 2004; Jancke & Erlhagen, 2010). Receptive field size increases with further processing steps (Dumoulin & Wandell, 2008; Smith et al., 2001), and it is possible that for later neural representations, the sub-threshold wave of neural activity caused by the stimulus in the location of the first sequence position reaches the neurons representing the stimulus in the location of the second sequence position (Chemla et al., 2019). This could cause a latency advantage for Sequence Position 2, even though no definite direction-specific predictions could have been formed yet. From Sequence Position 3 onward, we observed a latency advantage also for earlier neural representations. Since these representations showed no latency advantage for Sequence Position 2, we believe this latency advantage is unlikely to be caused by sub-threshold traveling waves of activity reaching the receptive field of the next presented stimulus, since we otherwise would have observed them for Sequence Position 2 as well. Instead, predictions regarding the motion trajectory of the stimulus may have been formed once the trajectory is definitively established, making the predictability of the next stimulus location 100% and subsequently projecting to lower visual areas, potentially causing the shorter latency for even earlier neural representations observed here.

We subsequently investigated how previously-formed predictions are revised in the case of a prediction violation. We found that after an unpredictable motion reversal, the first Sequence Position no longer had a latency advantage, as expected – instead, we observed a temporary latency disadvantage (Fig. 5). Importantly, we observed that the latency advantages for subsequent sequence positions were restored almost instantaneously. The first Sequence Position after the motion reversal was the stimulus that signaled the reversal, and the first point in time at which the prediction violation could be detected. The neural representations of the second sequence position, which was presented only 100 ms later, already show a latency advantage compared to the first stimulus. Surprisingly, and in contrast to stimuli presented at the start of the sequence, even early neural representations of the second sequence position benefited from this latency advantage.

One difference between Sequence Position 2 in the reversal and start conditions is the available trajectory information. At the start of the sequence, Sequence Position 2 is the first stimulus providing information regarding the direction of the motion trajectory. However, following a reversal, the previous trajectory (i.e., the final stimulus before the reversal) and the first stimulus after the reversal together provide enough information to determine the direction of the new motion trajectory. Accordingly, the location of Sequence Position 2 can be fully predicted using the locations of those two stimuli. This might explain the difference in time-course between Sequence Position 2 at the start of the sequence and after a motion reversal.

Interestingly, neural representations representing an unpredictable stimulus (Sequence Position 1 at the start of the sequence) or a prediction violation (Sequence Position 1 after a motion reversal) exhibited latency differences for mid-level representations, but not for earlier or later neural representations. These neural representations approximately correspond to the neural representations exhibiting latency differences between Sequence Position 2 and later sequence positions. As discussed above, no predictions could be formed for Sequence Position 2 yet, but predictions for Sequence Position 3 and later could descend from higher areas to this layer through feedback connections. We have previously shown that following a motion reversal, these neural representations represent the anticipated but never presented stimulus location, before correctly representing the actually presented stimulus location with the arrival of afferent sensory information (Blom et al., 2020). It could therefore be possible that since the wrong neural representations in this processing layer have been pre-activated through descending predictions, not only was there no latency advantage for this sequence position, there even might have been a latency *disadvantage* since erroneous pre-activations needed to be overcome first.

Sequence Position 1 after the reversal had also been activated 200 ms prior, making repetition suppression another potential explanation. There has been considerable discussion on how this phenomenon relates to prediction, since both repetition and expectation can reduce neuronal responses, but have been found to have separable effects on neural representations in the visual system (Tang, Smout, Arabzadeh, & Mattingley, 2018). Grill-Spector, Henson, and Martin (2006) proposed multiple neural models of repetition suppression, one of which proposes that the diminished neural response to a repeated stimulus is caused by faster processing of the stimulus. However, what we observe here is not faster, but slower processing of the repeated stimulus, making the observed latency disadvantage more likely to be caused by a wrong prediction than by stimulus repetition. Furthermore, Todorovic and de Lange (2012) showed that in the auditory system expectation suppression and repetition suppression could be dissociated on the basis of their time-courses, with repetition suppression influencing the early and late time-windows, but not the intermediate time-windows. The intermediate time-window was only affected by expectation (Todorovic & de Lange, 2012). This is relevant for what we observed for Sequence Position 1 after the reversal in Fig. 5: no latency disadvantage was observed for the early and late time-windows, but there was a latency disadvantage for the intermediate time-window. To dissociate the effects of prediction violation and repetition, a follow-up experiment could be conducted where the violation is not signaled by a motion reversal, but rather a change of direction in an orthogonal direction towards a stimulus that had not been presented previously.

For stimuli presented later on in the sequence, lingering neural activity from the previous stimulus, presented only 100 ms earlier, might be present in the EEG signal. However, this activity should not bias the classifier. The classifiers are trained on a static signal, the onset of the sequence, and then tested on the stimulus location within the apparent motion sequence. The EEG signal in the test set always contains a mixture of neural signal related to the same stimulus when it appears earlier and later in the sequence. However, the classifier is trained on the onset of the stimulus without any prior stimulus presentations contaminating the training data. In other words, when estimating the model, the classifier can therefore not be biased by any lingering activity from before the stimulus onset of interest in the test data.

The training set would contain neural signals for stimuli presented later in the sequence, but since the motion direction was counterbalanced, the trained classifier could not utilize this information for prediction. Future experiments could control for this aspect even more directly by altering the speed of the apparent motion sequence between trials. Such a manipulation would jitter any preceding neural activity between trials, which would fully eliminate the possibility of any systematic impact on lingering neural activity on the classifier.

In summary, we provide evidence for the interaction between two separate mechanisms of prediction formation and revision in human visual motion processing using time-resolved EEG decoding and an apparent motion paradigm. We found that neural representations of the second sequence position already exhibited a latency advantage and even earlier neural representations from the third sequence position onward as well. These additional latency advantages could be caused by descending predictions regarding the object's trajectory from higher visual areas through feedback connections, which may not had been formed yet for the second sequence position. Consistent with this interpretation, these earlier neural representations exhibited a latency disadvantage following a prediction violation compared to when the stimulus was unpredictable. The latency advantages observed for the later neural representations of the second sequence position most likely arose from a propagating wave of sub-threshold neural activity through long-range horizontal connections, that prepared the neurons representing the stimulus in the second location for the arrival of the stimulus. Altogether, these results reveal a pattern of latency advantages for predictable stimuli, as well as additional delays when those predictions are violated, which might be realized through a combination of long-range horizontal and feedback connections. It would be of interest to investigate whether the formation of these latency effects is influenced by the speed of the apparent motion stimulus and whether a temporal violation (the presentation time of the stimulus) instead of a spatial violation (a motion reversal) would yield similar results to the ones observed here.

Open practices

The study in this article earned Open Materials and Open Data badges for transparent practices. Materials and data for the study are available at https://osf.io/9ms6k/.

Credit author statement

Tessel Blom: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project administration. **Stefan Bode:** Resources, Writing – Review & Editing, Supervision. **Hinze Hogendoorn:** Conceptualization, Methodology, Resources, Writing – Review & Editing, Supervision, Funding acquisition.

Acknowledgements

This work was supported by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (DP180102268) awarded to H.H.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2021.02.008.

REFERENCES

Alilovíc, J., Timmermans, B., Reteig, L. C., Van Gaal, S., & Slagter, H. A. (2019). No evidence that predictions and attention modulate the first feedforward sweep of cortical information processing. *Cerebral Cortex, 29*(5), 2261–2278. https://doi.org/10.1093/cercor/bhz038

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *Journal of Neuroscience, 30*(8), 2960–2966. https://doi.org/10.1523/JNEUROSCI.3730-10.2010

Andersen, L. M., & Lundqvist, D. (2019). Somatosensory responses to nothing: An MEG study of expectations during omission of tactile stimulations. *Neuroimage, 184*, 78–89. https://doi.org/10.1016/j.neuroimage.2018.09.014

Benvenuti, G., Chemla, S., Boonman, A., Perrinet, L., Masson, G., & Chavane, F. (2020). Anticipatory responses along motion trajectories in awake monkey area V1. *bioRxiv*. https://doi.org/10.1101/2020.03.26.010017

Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature, 398*(6725), 334–338. https://doi.org/10.1038/18678

Blom, T., Feuerriegel, D., Johnson, P., Bode, S., & Hogendoorn, H. (2020). Predictions drive neural representations of visual events ahead of incoming sensory information. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1917777117, 201917777–201917777.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Chemla, S., Reynaud, A., Volo, M. D., Zerlaut, Y., Perrinet, L., Destexhe, A., et al. (2019). Suppressive traveling waves shape representations of illusory motion in primary visual cortex of awake primate. *Journal of Neuroscience, 39*(22), 4282–4298. https://doi.org/10.1523/JNEUROSCI.2792-18.2019

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Demarchi, G., Sanchez, G., & Weisz, N. (2019). Automatic and feature-specific prediction-related neural activity in the human auditory system. *Nature Communications, 10*(1). https://doi.org/10.1038/s41467-019-11440-1, 3440–3440.

Dumoulin, S. O., & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *Neuroimage, 39*(2), 647–660. https://doi.org/10.1016/j.neuroimage.2007.09.034

Ekman, M., Kok, P., & De Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex.

Nature Communications, 8. https://doi.org/10.1038/ncomms15276

Frégnac, Y., Bringuier, V., & Chavane, F. (1996). Synaptic integration fields and associative plasticity of visual cortical cells in vivo. *Journal of Physiology-Paris, 90*(5–6), 367–372. https://doi.org/10.1016/S0928-4257(97)87921-1

Fröhlich, F. W. (1924). Über die messung der empfindungszeit. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere, 202*(1), 566–572.

Gilbert, C. D., & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *Journal of Neuroscience, 9*(7), 2432–2442.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10*(1), 14–23. https://doi.org/10.1016/j.tics.2005.11.006

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time-series neuroimaging data. *Journal of Cognitive Neuroscience, 29*(4), 677–697. https://doi.org/10.1017/CBO9781107415324.004

Hogendoorn, H. (2020). Motion extrapolation in visual processing: Lessons from 25 Years of flash-lag debate. *The Journal of Neuroscience, 40*(30), 5698–5705. https://doi.org/10.1523/JNEUROSCI.0275-20.2020

Hogendoorn, H., & Burkitt, A. N. (2018). Predictive coding of visual object position ahead of moving objects revealed by time-resolved EEG decoding. *Neuroimage, 171*, 55–61. https://doi.org/10.1016/j.neuroimage.2017.12.063

Hogendoorn, H., Carlson, T. A., & Verstraten, F. A. (2008). Interpolation and extrapolation on the path of apparent motion. *Vision Research, 48*(7), 872–881. https://doi.org/10.1016/j.visres.2007.12.019

Jancke, D., Chavane, F., & Grinvald, A. (2010). Stimulus localization by neuronal populations in early visual cortex: Linking functional architecture to perception. In *Dynamics of visual motion processing: Neuronal, behavioral, and computational approaches* (pp. 95–116). Springer US.

Jancke, D., Chavane, F., Naaman, S., & Grinvald, A. (2004). Imaging cortical correlates of illusion in early visual cortex. *Nature, 428*(6981), 423–426. https://doi.org/10.1038/nature02396

Jancke, D., & Erlhagen, W. (2010). Bridging the gap: A model of common neural mechanisms underlying the Fröhlich effect, the flash-lag effect, and the representational momentum effect. In *Space and time in perception and action* (pp. 422–440). https://doi.org/10.1017/CBO9780511750540.025

Jancke, D., Erlhagen, W., Schöner, G., & Dinse, H. R. (2004). Shorter latencies for motion trajectories than for flashes in population responses of cat primary visual cortex. *Journal of Physiology, Paris, 556*(3), 971–982. https://doi.org/10.1113/jphysiol.2003.058941

King, J. R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences, 18*(4), 203–210. https://doi.org/10.1016/j.tics.2014.01.002

Kirschfeld, K., & Kammer, T. (1999). The frohlich effect: A consequence of the interaction of visual focal attention and metacontrast. *Vision Research, 39*(22), 3702–3709. https://doi.org/10.1016/S0042-6989(99)00089-9

Kok, P., Failing, M. F., & de Lange, F. P. (2014). Prior expectations evoke stimulus templates in the primary visual cortex. *Journal of Cognitive Neuroscience, 26*(7), 1546–1554.

Kok, P., Mostert, P., & De Lange, F. P. (2017). Prior expectations induce pre-stimulus sensory templates. *Proceedings of the National Academy of Sciences of the United States of America, 114*(39), 10473–10478. https://doi.org/10.1073/pnas.1705652114

Maiche, A., Budelli, R., & Gomez-Sena, L. (2007). Spatial facilitation is involved in flash-lag effect. *Vision Research, 47*(12), 1655–1661. https://doi.org/10.1016/j.visres.2007.02.008

Morel, P. (2018). Gramm: Grammar of graphics plotting in Matlab. *The Journal of Open Source Software, 3*(23). https://doi.org/10.21105/joss.00568, 568–568.

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *Plos Biology, 3*(8). https://doi.org/10.1371/jour-nal.pbio.0030265

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature, 370*(6487), 256–257. https://doi.org/10.1038/370256b0

Nijhawan, R. (2008). Visual prediction: Psychophysics and neurophysiology of compensation for time delays. *Behavioral and Brain Sciences, 31*(2), 179–198. https://doi.org/10.1017/S0140525X08003804

Series, P., Lorenceau, J., & Frégnac, Y. (2003). The "silent" surround of V1 receptive fields: Theory and experiments. *Journal of Physiology-Paris, 97*(6), 453–474. https://doi.org/10.1016/j.jphysparis.2004.01.023

Smith, A., Singh, K., Williams, A., & Greenlee, M. (2001). Estimating receptive field size from fMRI data in human striate and extrastriate visual cortex. *Cerebral Cortex, 11*(12), 1182–1190. https://doi.org/10.1093/cercor/11.12.1182

Sterzer, P., Haynes, J. D., & Rees, G. (2006). Primary visual cortex activation on the path of apparent motion is mediated by feedback from hMT+/V5. *Neuroimage, 32*(3), 1308–1316. https://doi.org/10.1016/j.neuroimage.2006.05.029

Subramaniyan, M., Ecker, A. S., Patel, S. S., Cotton, R. J., Bethge, M., Pitkow, X., & Tolias, A. S. (2018). Faster processing of moving compared with flashed bars in awake macaque V1 provides a neural correlate of the flash lag illusion. *Journal of Neurophysiology, 120*(5), 2430–2452. https://doi.org/10.1152/jn.00792.2017

Tang, M. F., Smout, C. A., Arabzadeh, E., & Mattingley, J. B. (2018). Prediction error and repetition suppression have distinct effects on neural representations of visual information. *eLife, 7*. https://doi.org/10.7554/eLife.33123

Todorovic, A., & de Lange, F. P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *Journal of Neuroscience, 32*(39), 13389–13395. https://doi.org/10.1523/JNEUROSCI.2227-12.2012

Toth, L. J., Rao, S. C., Kimt, D.-S., Somers, D., & Sur, M. (1996). Subthreshold facilitation and suppression in primary visual cortex revealed by intrinsic signal imaging. *Neurobiology, 93*, 9869–9874.

van Heusden, E., Harris, A. M., Garrido, M. I., & Hogendoorn, H. (2019). Predictive coding of visual motion in both monocular and binocular human visual processing. *Journal of Vision, 19*(1), 1–12. https://doi.org/10.1167/19.1.3

Vetter, P., Grosbras, M. H., & Muckli, L. (2015). TMS over V5 disrupts motion prediction. *Cerebral Cortex, 25*(4), 1052–1059. https://doi.org/10.1093/cer-cor/bht297

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience, 1*(8), 656–657. https://doi.org/10.1038/3659

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research, 40*(2), 137–149. https://doi.org/10.1016/S0042-6989(99)00166-2

Wibral, M., Bledowski, C., Kohler, A., Singer, W., & Muckli, L. (2009). The timing of feedback to early visual cortex in the perception of long-range apparent motion. *Cerebral Cortex, 19*(7), 1567–1582. https://doi.org/10.1093/cercor/bhn192

Xu, S., Jiang, W., Poo, M. M., & Dan, Y. (2012). Activity recall in a visual cortical ensemble. *Nature Neuroscience, 15*(3), 449–455. https://doi.org/10.1038/nn.3036