

Evidence of Serial Dependence from Decoding of Visual Evoked Potentials

Giacomo Ranieri,¹ Alessandro Benedetto,² Hao Tam Ho,¹ David C. Burr,¹ and  Maria Concetta Morrone²

¹Department of Neuroscience, Psychology, Pharmacology, and Child Health, University of Florence, 50135 Florence, Italy and ²Department of Translational Research on New Technologies in Medicine and Surgery, University of Pisa, 56123 Pisa, Italy

It is well known that recent sensory experience influences perception, recently demonstrated by a phenomenon termed “serial dependence.” However, its underlying neural mechanisms are poorly understood. We measured ERP responses to pairs of stimuli presented randomly to the left or right hemifield. Seventeen male and female adults judged whether the upper or lower half of the grating had higher spatial frequency, independent of the horizontal position of the grating. This design allowed us to trace the memory signal modulating task performance and also the implicit memory signal associated with hemispheric position. Using classification techniques, we decoded the position of the current and previous stimuli and the response from voltage scalp distributions of the current trial. Classification of previous responses reached full significance only 700 ms after presentation of the current stimulus, consistent with retrieval of an activity-silent memory trace. Cross-condition classification accuracy of past responses (trained on current responses) correlated with the strength of serial dependence effects of individual participants. Overall, our data provide evidence for a silent memory signal that can be decoded from the EEG potential, which interacts with the neural processing of the current stimulus. This silent memory signal could be the physiological substrate subserving at least one type of serial dependence.

Key words: decoding; EEG; serial dependence; vision; working memory

Significance Statement

The neurophysiological underpinnings of how past perceptual experience affects current perception are poorly understood. Here, we show that recent experience is reactivated when a new stimulus is presented and that the strength of this reactivation correlates with serial biases in individual participants, suggesting that serial dependence is established on the basis of a silent memory signal.

Introduction

As natural scenes are relatively stable over time, the brain employs computational strategies to exploit spatiotemporal redundancies, reducing the complexity of perceptual processing. One strategy, first suggested by Helmholtz (1867) and championed by Gregory (1968) is to build expectations or priors of the structure of the world from previous perceptual history and test these against current sensory input. This concept is well formalized within the Bayesian framework, where perceptual expectations can be considered priors to be combined with current input (Wolpert et al., 1995; Kersten et al., 2004).

The effects of perceptual history have been long studied, largely by a technique known as priming, (introduced by Lashley, 1951),

occurring when exposure to one stimulus influences the response to subsequent stimuli without conscious guidance or intention (Weingarten et al., 2016). Priming is an important phenomenon observed in a range of cognitive studies, from linguistics to social psychology, and particularly in perceptual research (Maljkovic and Nakayama, 1994; Kristjánsson and Ásgeirsson, 2019). Priming typically results in improvement in accuracy and reaction times. However, it has been shown that previous stimuli cannot only speed and facilitate current perception but also bias responses away from veridicality. This is particularly apparent with a paradigm termed “serial dependence,” where responses are strongly biased toward recent perceptual experience (Cicchini et al., 2014; Fischer and Whitney, 2014; Cicchini et al., 2017). This bias has been observed for a variety of basic stimulus features, such as orientation and position (Bliss et al., 2017; Manassi et al., 2018; Fritsche et al., 2020), as well as for more complex perceptions, such as numerosity, facial identity, and scene gist (Cicchini et al., 2014; Liberman et al., 2014; Manassi et al., 2017) and probably also working memory (Kiyonaga et al., 2017).

The physiological mechanisms underlying serial dependence are still poorly understood. Many studies emphasize the high-

Received Sep. 17, 2021; revised July 19, 2022; accepted July 21, 2022.

Author contributions: A.R., A.B., H.T.H., D.C.B., M.C.M. designed research; A.B., H.-T.H. and G.R. performed research; A.B., H.-T.H. and G.R. analyzed data; D.C.B., M.C.M., G.R. and A.B. wrote the paper.

This work was supported by European Research Council Grant 832813. Italian Ministry of Education, University and Research (MIUR) PRIN 2017.

The authors declare no competing financial interests.

Correspondence should be addressed to David C. Burr at davidcharles.burr@unifi.it.

<https://doi.org/10.1523/JNEUROSCI.1879-21.2022>

Copyright © 2022 the authors

level origin of perceptual priors (Kim et al., 2020; Ceylan et al., 2021) and assume that the priors also act at a high level (Bliss et al., 2017; Feigin et al., 2021). Others claim, with supporting evidence, that the effects take place at low-level sensory areas (St John-Saaltink et al., 2016; Cicchini et al., 2021). However, both models require the existence of memory of the previous stimuli, either in the form of decaying neuronal activity or as silent activity reactivated on the specific task. Whether task relevance is necessary to elicit the dependence is still an open question (Manassi et al., 2018; Bae and Luck, 2020; Kim et al., 2020; Murai and Whitney, 2021).

EEG-based decoding is a new statistical approach that attempts to classify properties of the stimulus from distributions of signals across the scalp (Foster et al., 2016; Wolff et al., 2017; Bae and Luck, 2018; Noah et al., 2020). Bae and Luck (2019) demonstrated that information from previous trials may be reactivated in future trials; by activity elicited by the current trial they decode with reasonable accuracy the orientation of the previous stimulus. Interestingly, classification was not possible before trial onset, suggesting that prior stimuli act through activity-silent synaptic traces that generate an active representation when a new stimulus is presented. Similarly, Barbosa et al. (2020) illustrated the reactivation of an activity-silent trace in monkey prefrontal cortex and in human brain, analyzing a signal that influences spatial location in a classic memory saccade task.

In the current study, we looked for neural representations of priming effects in EEG recordings using support vector machine classifiers. We chose the serial dependence paradigm as this leads to response biases, which should affect classification. We show that the response to previous stimulus could be decoded from the response to the current one. Importantly, models trained on current responses decoded the previous trial responses robustly, and decoding accuracy correlated with serial dependence effects in individual participants, establishing a link between behavioral outcome and classification metrics of recent experience. These findings demonstrate that prior stimulation manifests as reactivation of a neural trace and that this reactivation is crucially linked to serial dependence effects in behavior.

Materials and Methods

Participants

Eighteen healthy adults (10 females, age range, 25–34 years; mean, 27.1 years; SD = 2.5 years), all with normal or corrected-to-normal vision, voluntarily participated in the experiment. All participants provided written informed consent. Sample size estimation was based on similar decoding studies (Foster et al., 2016). One participant was excluded from analysis from preprocessing because of irreconcilable noise in the EEG data. The experimental design was approved by the local regional ethics committee (Comitato Etico Pediatrico Regionale—Azienda Ospedaliero-Universitaria Meyer—Firenze, FI) and is in line with the *Declaration of Helsinki* for ethical principles for medical research involving human subjects.

Stimuli and apparatus

The stimuli appeared on a Display++ LCD Monitor (120 Hz, 1920 × 1080 resolution, Cambridge Research Systems), gamma corrected, placed 70 cm from the eyes. The stimulus was a briefly presented (8.3 ms) grating patch consisting of two vertical gratings, one on the upper part of the screen, the other on the lower part. The gratings appeared randomly on the left or right half of the screen, inner bound 3° from the fixation dot (32° height, 16° width, 50% contrast, random phase). The gratings had a fixed spatial frequency of 1 c/degree and 1.1 c/degree (10% difference), randomly presented either on the upper or lower half of the screen.

Procedure

A white fixation dot was present in the center of the display during the whole experiment. Each trial began with a voluntary button press, which initiated stimulus presentation after a pseudo-random time interval, ranging from 16 to 833 ms. The delay minimized stereotyping of attention. The highly visible grating patches were presented either on the left half or right half of the screen, at random. Participants gave a verbal response after at least 1 s after stimulus presentation, recorded by the experimenter. The long delay was introduced to strengthen the memory trace and to have a large interval of stable EEG recording with no task related interference. After responding, participants could initiate the following trial by a voluntary button press (average trial length, 3.32 s ± 0.58 s STD). The task was a two-alternative forced-choice discrimination, where participants chose which grating had higher spatial frequency (Fig. 1a). Each participant completed 648 trials.

EEG acquisition

EEG data were collected on a Nautilus Research headset (g.tec) at a sample rate of 500 Hz with no online filtering. The ground electrode was placed on the center of the head. The data were referenced online to a unilateral electrode placed behind the left ear. Activity was measured from 32 gel-based active electrodes (g.LADYbird technology) arranged according to the 10/20 system. Impedance was kept below 50 kΩ.

EEG preprocessing

Offline EEG preprocessing was performed in MATLAB (MathWorks) with custom code. EEG data were referenced to the common average reference and filtered with a finite-impulse response lowpass filter (Chebyshev window, 128th order, stopband 25 Hz, side lobe magnitude factor 30 dB). Epochs were extracted aligned to stimulus presentation composed of a segment of data from −500 ms to 1800 ms after the stimulus. Epochs were visually inspected for motor artifacts and wireless failure of signal transmission (manual rejection of 1.2 ± 0.2% STD of data across subjects). Ocular artifacts were removed through blind source separation with independent components analysis (ICA) decomposition (Jung et al., 2000).

Data analysis

Psychophysics. We examined the behavioral data for the individual participants. We applied signal detection theory (SDT) to discern sensitivity and criterion (Fig. 1). Sensitivity was measured with d' , given by the z -transform of hit rate minus the z -transform of false alarms. Criterion (c) was given by the sum of the z -transform of hit rate and the z -transform of false alarms. The psychophysical measure of serial dependence was defined as the difference in criterion when the previous response was up from when it was down. Receiver operating characteristics (ROCs) curves plot the values of the proportion of correct responses against the proportion of false alarms for individual participants. We show the best-fitting (Least Square Estimation) sensitivity curve for the average participant.

Decoding. The major analysis of this article was the decoding of the EEG signal, largely based on the methods in Bae and Luck (2019). We decoded multiple conditions of trials, the horizontal position of the stimulus (current hemifield), the vertical position of the current stimulus (current target), the horizontal and vertical position of the previous target (previous hemifield and previous target), the response of the participant to the current stimulus and to the previous stimulus. For each time point taken from the −500 to +1800 ms interval from stimulus onset, the average potential across the 32 electrodes was decoded independently.

Before decoding, the signal was down sampled to 50 Hz, resulting in 150 time points per trial. In each decoding pipeline, trials were divided into two conditions (such as stimulus left and stimulus right; response up and response down; previous target up and previous target down, etc.), regardless of other stimulus characteristics. Trials were randomly assigned to 10 sets per condition (with an average of 15 trials per set, leaving out from 1 to 9 random trials per condition as the number of trials were not always multiples of 10). Trials in each set were averaged (from here on, this average is referred to as sample) to increase the signal-to-noise ratio, a common method used by previous studies (Foster et al., 2016). Although this means averaging the EEG signals from very

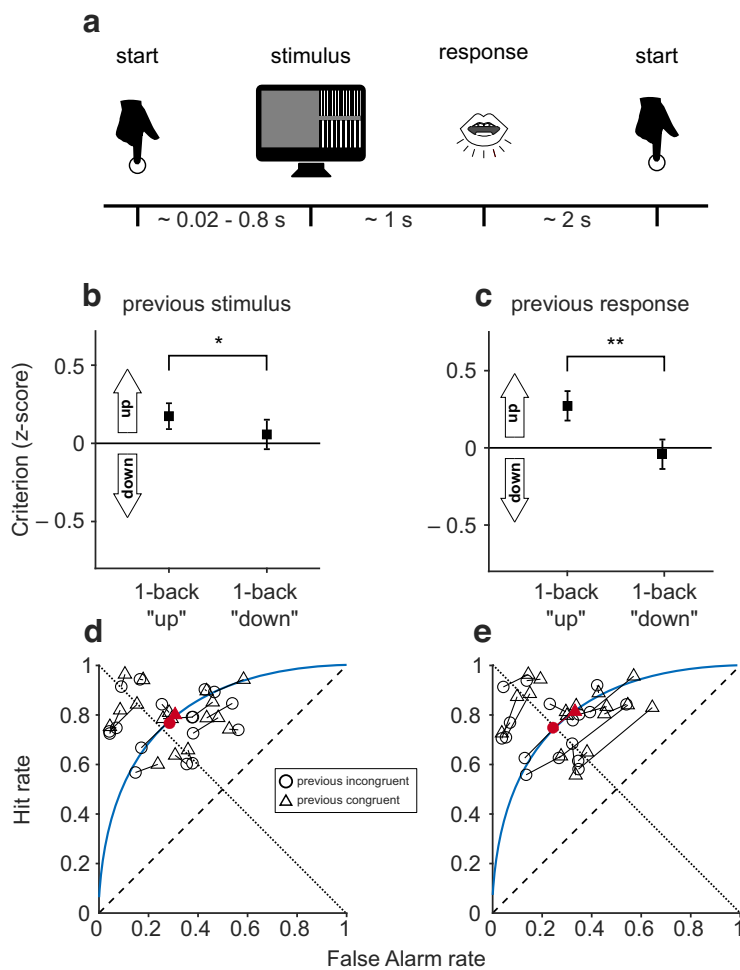


Figure 1. Responses are biased toward recent stimuli and responses. **a**, Schematics of the experimental paradigm. Voluntary button press starts the trial, then two vertically stacked vertical gratings appear on the left or right side of the screen at random. Participants are prompted to respond verbally after waiting at least 1 s, deciding which of the two gratings had higher spatial frequency. After a pause of at least 2 s participants could initiate the following trial. **b**, Serial dependence effect based on previous stimulus. Squares indicate mean criterion (z-scored, positive values indicate higher tendency toward responding up vs down). Criterion for previous stimuli up is statistically different from criterion for previous stimuli down ($p = 0.04$, $BF = 1.8$). Error bars indicate Standard Error of the Mean (SEM) across the 17 participants. The asterisk indicates a p value lower than 0.05. **c**, Serial dependence effect based on previous response. Criterion for previous responses of up is statistically different from criterion for previous responses of down ($p = 0.003$, $BF = 16$). The 2 asterisks indicate a p value lower than 0.01. **d**, Individual results, plotted as a proportion of hits against false alarms; red points show average. Sensitivity is indicated by the distance from the dashed line, criterion by the angle from the dotted line. The circles (red and black) show performance when the previous target was different from the current trial, triangles when it was the same. The blue curve indicates the best-fitting ROC curve of the average participant. **e**, Proportion of hits plotted against false alarms as in **d**, with respect to the response rather than the stimulus congruency of the previous trial.

different conditions and sources, it allows sufficient numerosity to address the specific decoding. We partitioned the data into training and test sets; eight samples per condition were assigned to the training set and two samples to the test set (where each sample is the average of trials as explained above). The decoding for each time point was performed with a five-fold cross-validation procedure, whereby four samples were used for testing in each fold. This ensured that all samples were used both in the training set and in the test set as standard procedure for datasets with low sample size. Each time a new model was trained and tested with no memory of previous models. We repeated the classification procedure generating samples by averaging trials with the same spatial properties (e.g., only up left trials). We reproduced the method with two strategies. First, we decoded labels by training and testing on spatially homogeneous samples, reproducing very comparable results to the main classification procedure; second, we performed separate classifications for each spatial location (e.g., classifying previous hemifield only for up

left trials), again obtaining similar results. However, given that this latter procedure did not give stable results for the low numerosity of the sets, we report here only the analysis with samples from averages of heterogeneous stimuli.

The classification used binary Support Vector Machines, implemented with a linear kernel via the MATLAB function `fitcsvm`. The classifier found the best margin for classification based on voltage values at 32 electrode locations. The model was then used to predict the classes of samples in the test set. This procedure was repeated for 1000 iterations. In each iteration the trials were assigned randomly to the 10 buckets so that samples were not biased because of lucky splits of the data. Given the 10-fold cross-validation procedure where four samples were tested, the 100 reiterations lead to a total of 40,000 tests for each final decoding values. The whole procedure was performed for each independent time bin of 20 ms duration, producing a measure of decoding accuracy in time. Predictions were compared with true labels, yielding percentage classification accuracy by averaging the number of correct predictions across iterations and cross-validations. Data from participants were averaged to obtain a mean time course of decoding accuracy with the associated SE.

We computed topographical maps associated with each decoding instance. The relevance of each electrode in the classification (referred in the literature as activation pattern; Min et al., 2016) was calculated by multiplying the signed weights of the trained model by the voltage of the samples of the corresponding test set. The activation pattern was then averaged through all iterations of the decoding instance. This procedure produced an $N \times M$ matrix, where N is the 32 electrodes, and M is the 150 time points. For visualization, we averaged activation patterns across ± 2 timepoints (100 ms windows) and displayed the topographies with consistent color mapping and no normalization.

Generalization matrices. To verify the similitude in time of the signal that led to decoding, we computed a temporal generalization matrix for each decoding instance (King and Dehaene, 2014). The procedure is an extension of the method described above. The classifier trained on a particular time point was tested on all 150 time points. The result is a measure of decoding accuracy as a 150×150 matrix.

The statistical significance of decoding was assessed in the 2D matrices by a two-tailed t test against the null hypothesis of chance classification independently for each time point (50% for the case of binary classification). Each classification procedure at each time point is conceptually independent. However, given that low-pass filters at 25 Hz induce correlations between 40 ms time bins, we also introduced a threshold for cluster size of time bins when at least 3×3 points had classification accuracy with a p value lower than 0.05. Clusters of 3×3 points were identified by simulating the null hypothesis with a label shuffle ($n = 2000$). This produced a distribution of clusters of contiguous significant time points in the null hypothesis, which allowed us to define a threshold for cluster size as the 95th percentile of that distribution. To visualize cross-participant decoding significance, we highlighted the perimeter of the areas that surpassed the t test. The statistical annotation for classification accuracy plots across time (Figs. 2, 3) correspond to the diagonal of the respective generalization matrix. For graphical visualization only, we smoothed the decoding accuracy with a Gaussian temporal filter of 72 ms time constant.

We also computed additional matrices to verify the cross-conditional generalizability of decoding, ensuring that the training and test sets were independent, and no signal participated in the individual averages of both sets. We trained the classifier on data split according to previous

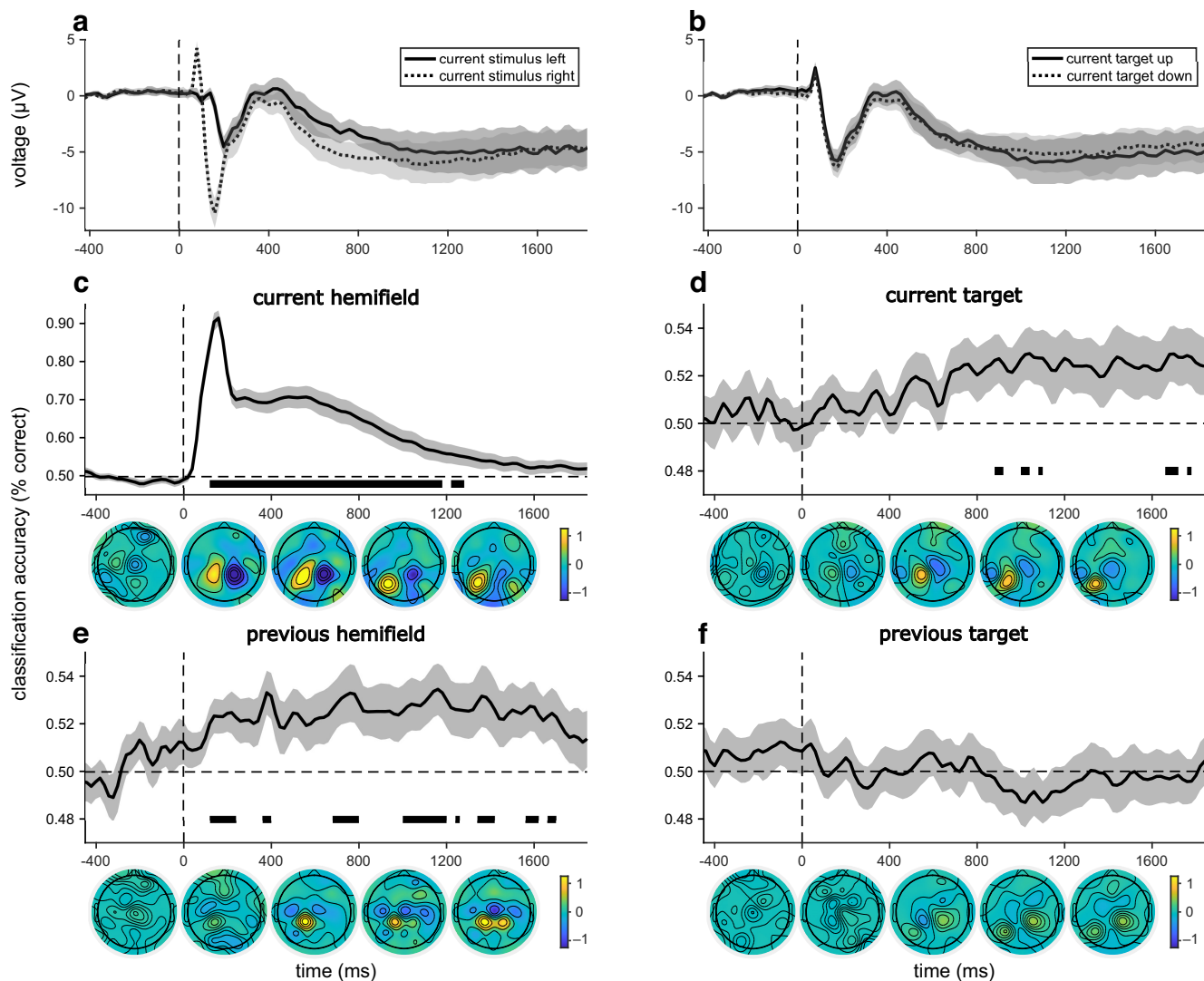


Figure 2. EEG and decoding of current and previous stimulus features. *a, b*, ERPs for current stimulus conditions in electrode P03. Shaded areas show ± 1 SEM between participants. Different gray tone for different stimuli. ERP for stimuli presented on the left or on the right hemifield (*a*). ERP for targets on the top or on the bottom side of the screen (*b*). *c–f*, Decoding accuracy plots for current and previous stimulus labels. Horizontal dashed lines indicate random classifier performance (50%). Time bins with significant decoding are marked by the black bar (see above, Materials and Methods). Topographies indicate relevance of electrodes for classifications (tested samples weighted based on model coefficients), averaged across 100 ms windows (showing representative activation patterns at -200 , 200 , 600 , 1000 , and 1400 ms). Topographies share consistent color mapping, both bright yellow and deep blue indicating reliable electrodes. Decoding accuracy for current hemifield (*c*). Decoding accuracy for current target (*d*). Decoding accuracy for previous hemifield (*e*). Decoding accuracy for previous target (*f*).

response labels and tested the model on current responses. We also did the inverse classification, training on current labels and testing on previous labels. The resulting accuracy is a measure of similarity between the current perceptive signal and the memory signal of the previous trial.

Correlations. Average classification accuracy of multiple decoding conditions (current response, previous response, current response trained on previous response, previous response trained on current response) was correlated against the magnitude of individual serial dependence. Serial dependence was calculated as the difference in criterion calculated on trials with a previous response up and trials with a previous response down. We computed Pearson's coefficients and p values against the null hypothesis of no significant covariance between serial dependence and decoding of previous response, as well as Bayes factors (BFs) for the likelihood ratio of the alternative hypothesis.

Results

Behavioral results

Participants performed a two-alternative forced-choice visual target discrimination task (Fig. 1*a*) while we recorded EEG scalp

potentials. We first demonstrated psychophysical serial dependence, confirming that behavior at trial n was influenced by trial $n-1$. We calculated criterion using SDT, by summing hits and false alarms, resulting in a measure of bias toward responding up. We calculated criterion separately for when the previous target was congruent or incongruent (Fig. 1*b*), verifying that serial dependence occurred in this experiment. Despite variability among participants, there was a significant positive effect of previous target on criterion, as upper targets led to increased tendency of responding up in the following trial and vice versa (mean difference = 0.12 , $SE = 0.05$, $t = 2.25$, $p = 0.037$, $BF = 1.8$). Because the task used a two-alternative forced-choice paradigm, with targets at discriminative threshold, we expected the effect to be stronger when considering previous responses instead of previous targets (Fig. 1*c*). Indeed, responses influenced future behavior more than targets (mean difference = 0.29 , $SE = 0.08$, $t = 3.51$, $p = 0.003$, $BF = 16$).

Figures 1, *d* and *e*, show individual results illustrating the effect in typical ROC curves. They plot the proportion of hits (correct responses) against false alarms (incorrect responses)

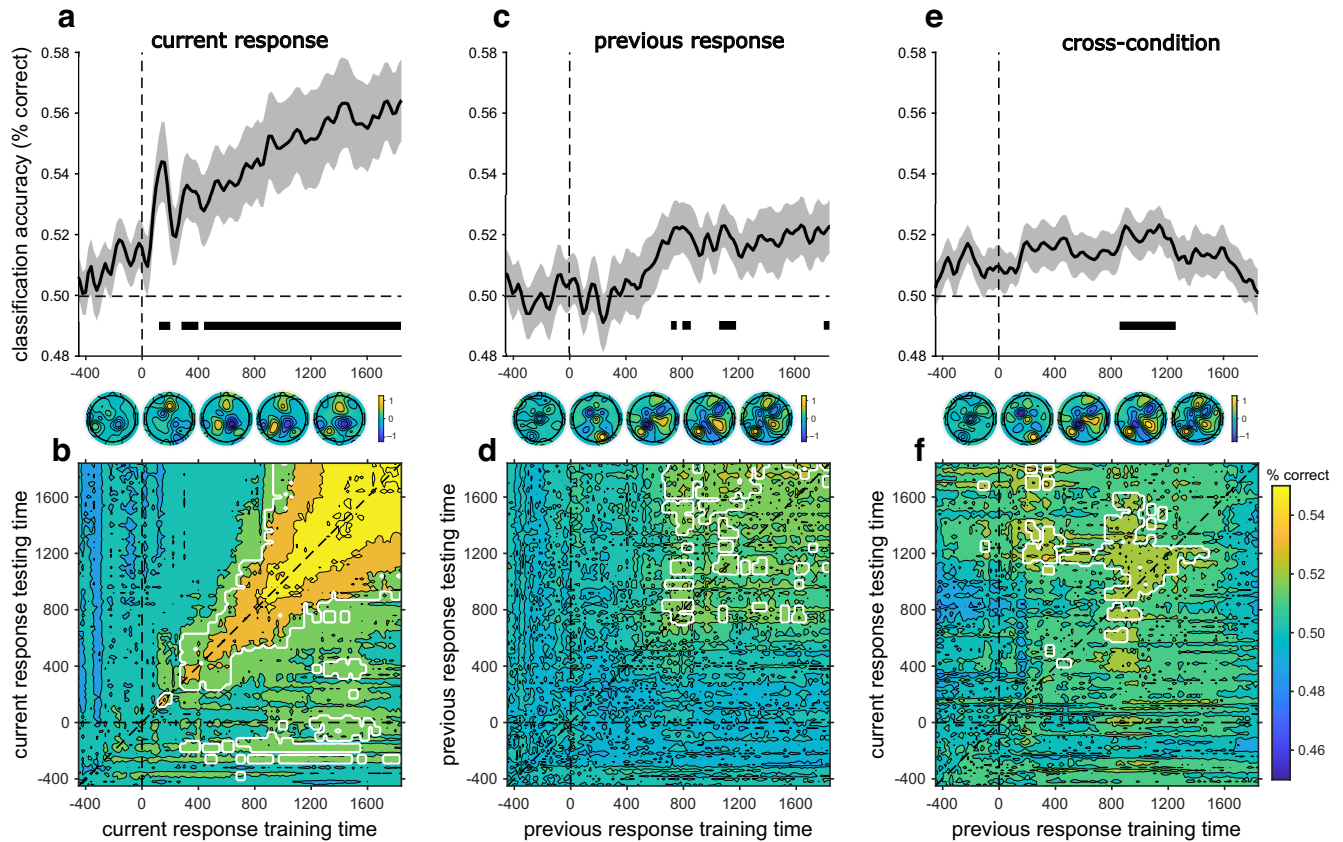


Figure 3. Current EEG activity contains integrated memory of previous responses. **a**, Decoding accuracy plots for current response. Shaded areas show ± 1 SEM. Horizontal dashed lines indicate random classifier performance (50%). Time bins with significant decoding are marked by black bars. Activation patterns are averaged across 100 ms windows and show consistent color mapping. **b**, Temporal generalization matrix for decoding of current response, assessing the stability of the response signal. Classifiers trained at each time point in the *x*-axis are tested on all time points in the *y*-axis. Classification accuracy is mapped with color, with yellow indicating higher accuracy. Significant regions of decoding, cluster corrected, are circumscribed by white perimeters. **c**, Decoding accuracy plots for previous response. **d**, Temporal generalization matrix for decoding of previous response. **e**, Cross-conditional decoding accuracy plots for classifiers trained on previous responses and tested on current responses for corresponding testing and training time. **f**, Conditional generalization matrix for classifiers trained on previous responses and tested on current responses.

separately for each participant (red points show average). Sensitivity is given by the distance from the dashed line, criterion by the rotation across the best-fitting ROC curve of the point. Points sitting on the negative diagonal (dotted line) show unbiased responses, with criterion of 0, while points angled clockwise from the diagonal refer to a tendency to respond up. The circles refer to incongruent trials when the previous trial stimulus (Fig. 1*d*) or response (Fig. 1*e*) was different from the current trial, and triangles refer to congruent trials when it was the same. The general effect of similar past stimuli or response was to rotate the responses in criterion with no change in sensitivity. This is particularly evident when the plots are based on previous response rather than stimulus (Fig. 1*e*, red arrow). As the target position varied randomly from trial to trial, no serial effect was expected for sensitivity (distance from the dashed line), and none was found (mean difference = 0.03, SE = 0.04, $t = 0.76$, $p = 0.46$, BF = 0.31).

To control for potential confounds, we verified that future stimuli and responses did not modulate participants' criterion in the current trial. There was no effect of future target (mean difference = -0.09 , SE = 0.06, $t = 1.70$, $p = 0.11$) and no effect of future response (mean difference = -0.04 , SE = 0.06, $t = 0.62$, $p = 0.54$) on criterion, meaning that response biases based on previous responses cannot be explained by autocorrelation in the decision.

Event-related potentials

The event-related potential (ERP) was strong and reliable at all electrodes, with the dynamics depending on electrode position. For illustrative purposes we show the ERPs recorded at PO3, where the difference between conditions was strongest. The ERPs for stimuli presented to the right and to the left visual hemifield (Fig. 2*a*) show major differences at intervals of 40–200 ms and 320–1500 ms (N1 amplitude effect, window = 130–170 ms, electrodes = PO3, mean of the difference = $5.29 \mu\text{V}$, SE = 0.25, $p < 0.001$), as to be expected given the lateralized visual field representation. ERPs to different target vertical positions diverged at late epochs (~ 600 –1800 ms, Fig. 2*b*), with smaller differences. This is to be expected given the smaller difference in dipole (dorsoventral representation of calcarine sulcus) and the task being near discrimination threshold.

Decoding results

We used decoding techniques to search for neurophysiological traces in both the current and the previous stimuli. We first decoded the hemifield (task-irrelevant feature) of the current stimulus from stimulus activity, a relatively simple task given the retinotopic or spatiotopic representation of much of visual cortex. We also decoded the target of the current presentation, a more difficult task, as the target was defined by subtle differences in spatial frequency at threshold detection. We then attempted to

decode the previous stimuli from the current response, both the hemifield and the target, to search for memory traces of previous stimulation related to serial dependence.

We trained the classifier to classify trials labeled by hemifield of stimulus presentation (regardless of target position), for consecutive intervals of 20 ms. These classifiers were used to predict the hemifield of left-out averaged trials (see above, Materials and Methods). The hemifield of the grating patch was decoded with high accuracy, contingent on stimulus presentation [significant intervals = (80–1160, 1200–1320 ms), peak = 0.93, peak latency = 160 ms]. Decoding accuracy was strongly significant when the ERP difference for the two sets of stimuli was strongest (Fig. 2*a*). The classifier was never above chance before stimulus presentation.

We computed activation maps related to current-hemifield decoding by projecting classification weights onto the scalp distribution (Fig. 2*c*, bottom images). The resulting topographic maps (saturated yellow or blue colors) indicate how relevant the single electrodes were for accurate discrimination of the current hemifield. These maps show that the decoder relies on occipitoparietal activity to label left and right trials.

After confirming that the algorithm decoded well the strong signals associated with the hemifield of presentation, we repeated the procedure for the vertical position of the target, the feature related to the perceptual task. The dynamics of the ERP differences on vertical target position (illustrated in Fig. 2*b* for the PO3 electrode) were slightly different, as expected from dipole orientation (Di Russo et al., 2003). Although the ERPs were more similar for the two target positions at all electrode positions, the classifiers were able to decode the position from scalp topography. Decoding was not as strong as for hemifield, which was to be expected as stimuli had equal high contrast and were close to 75% discrimination performance, but it was significant [significant intervals = (840–920, 960–1080, 1100–1160, 1640–1700, 1720–1760 ms), peak = 0.53, peak latency = 1620 ms], with accuracies in line with those of previous studies (Foster et al., 2016; Bae and Luck, 2019). Interestingly the significant classification occurs quite late, in correspondence again to the greatest difference of the ERPs.

We then attempted to classify stimuli of the previous trial from ERP distributions of the current trial. We first decoded the hemifield of the previous stimulus presentation, labeling the current EEG activity with the hemifield of the previous trial. Decoding was successful after ~100 ms from stimulus presentation, for quite long periods [Fig. 2*e*; significant intervals = (100–280, 340–420, 680–820, 1000–1200, 1220–1260, 1300–1440, 1560–1660, 1680–1720 ms), peak = 0.53, peak latency = 1160 ms]. This implies that a memory trace of the previous presentation, possibly related to an expectation of where the next trial will appear, remains for at least one trial. However, it was not present before stimulus onset or during the major visual responses. We further divided the data on the basis of hemifield position but found no significant change in decoding for consecutive stimuli that were in the same or different hemifields. The scalp maps at the bottom of Figure 2*c* highlight a more centroparietal activation compared with maps of the current target or hemifield.

We then attempted to decode the vertical position of the previous target from the current EEG (Fig. 2*f*). Classification was not possible, never reaching significance. This is perhaps not surprising, given that classifying the current target was itself very weak, consistent with the near threshold performance. We therefore used the same technique to decode the responses to the

stimuli, Figure 3*a* shows that labeling trials with current participant responses rather than with the target position led to strong decoding after stimulus presentation [Fig. 3*a*; significant intervals = (100–220, 280–440, 460–1800 ms), peak = 0.56 at 1800 ms].

Figure 3*b* shows the temporal generalization matrices for the classification of current responses, showing how training at various times generalizes to the different testing times. Temporal generalization is a measure of how well models built at a certain time interval can accurately classify voltage distributions at all time, assessing the evolution of the signal and its relative stability (King and Dehaene, 2014). Maximum coding clearly occurred along the diagonal (when training and testing times coincide) and fell off symmetrically and quite rapidly away from the diagonal, consistent with a dynamic model. No decoding was significant before zero, but good response decoding was apparent soon after stimulus appearance. The activation map here shows an early frontal activation and a later occipitoparietal activation. The later occipitoparietal activation may be driven by feedback from frontal areas, modulating the neural representation of the stimulus.

Encouraged by the successful decoding of current responses, we attempted to classify the responses of the previous trials from current activity (Fig. 3*c*). Decoding was reliably >50% after 700 ms and reached statistical significance at several points [significant intervals = (700–740, 820–880, 1040–1180, 1760–1800 ms), peak = 0.52, latency = 1100 ms]. Scalp maps in Figure 3 show an earlier activation of occipitoparietal electrode locations compared with the current response and a weaker contribution from frontal sites. Figure 3*d* shows the temporal generalization matrix. Decoding shows a clear square pattern, pointing to generalization across time, consistent with static decoding, which may be expected from a memory signal. Interestingly, the previous response decoded from the current response is significant only very late, again consistent with decoding of a memory signal.

We measured decoding separately for stimuli that were spatially congruent or incongruent (in horizontal position). Decoding was similar in both cases, consistent with a general memory trace. In addition, as a standard sanity check (Maljkovic and Nakayama, 1994), we attempted to classify future responses from current activity. This was completely impossible, with decoding accuracy always below 51%, discarding possible confounds from response autocorrelation or other artifacts.

Cross-condition coding

The previous section showed that previous response labels can be decoded from scalp potentials of the current trial, suggesting that a trace of the previous response is represented in the current one. We examined further the traces common to both current and previous responses by measuring cross-condition coding, training the support vector machine on previous responses and testing on the current (and vice versa). This cross-condition measures how well the previous response model predicts the current response.

Figure 3*e* shows the results for training on the previous response (using current ERP distributions) and testing the current responses. The decoding is not significant before stimulus onset and rises to significance only late after stimulus presentation [significant intervals = (880–1260 ms), peak = 0.52, latency = 920 ms]. Figure 4*f* shows the generalization matrix using the model trained on the previous response to

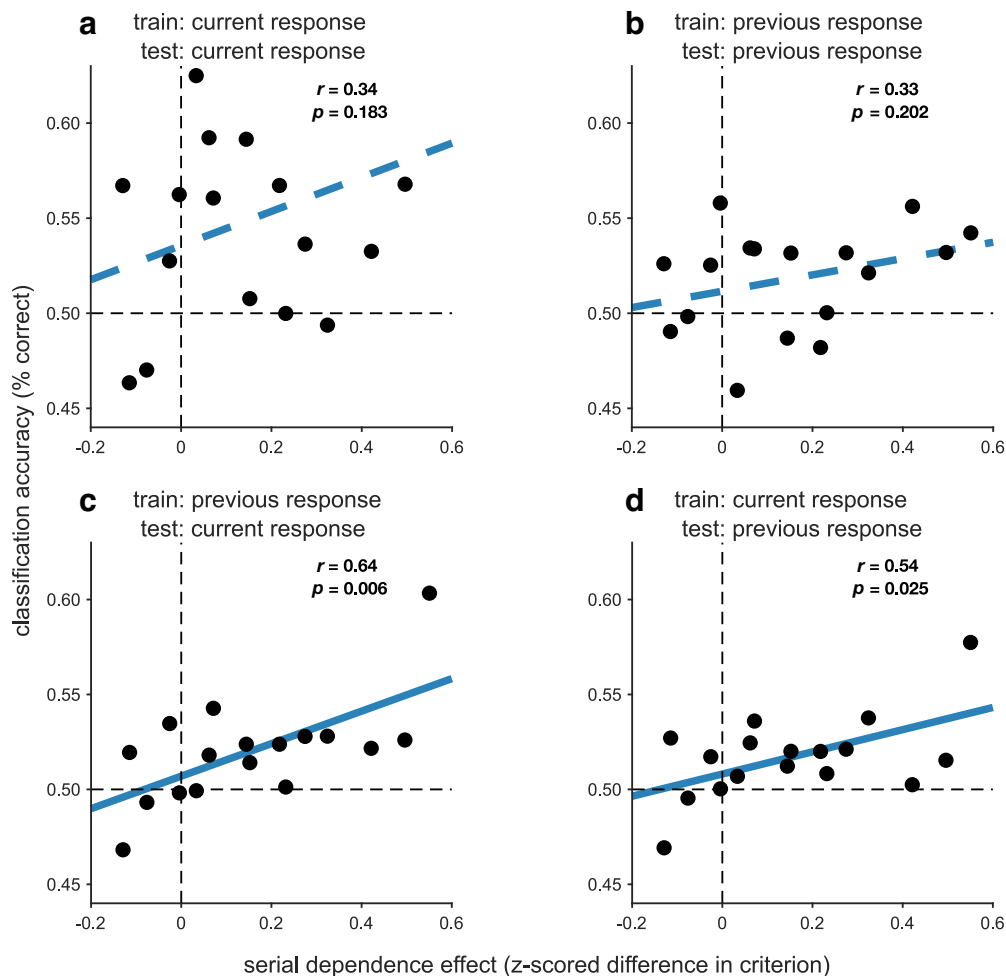


Figure 4. Shared EEG activity between the past and the present activation predicts serial dependence effects. Correlations across participants ($n = 17$) of serial dependence effects. The abscissa plots serial dependence of each participant, calculated as the difference in criterion between trials where the previous response was the same as the current one and those when it was different. The ordinates show the average decoding accuracy over the poststimulus window (0–1800 ms) for different decoding conditions. Statistics shown are two-tailed Pearson's correlation coefficient r and p value against the null hypothesis of no correlation. Line of best fit in blue (dashed for correlations with $p < 0.05$). **a–d**, The ordinates show decoding accuracy of current responses trained on current responses (**a**), decoding of previous responses trained on previous responses (**b**), cross-condition decoding of current responses trained on previous responses (**c**), and cross-condition decoding of previous responses trained on current responses (**d**).

decode the current response. The matrix again shows good decoding not only along the diagonal but a square-like generalization pattern consistent with static decoding, as may be expected from a memory signal.

The scalp maps of cross-condition decoding share many similarities with those for previous response temporal generalization, as to be expected given that activation maps are highly dependent on training model coefficients. We also tested the reciprocal condition, training on current response and testing on previous response, finding very similar results both for the qualitative assessment of the generalization matrix and for the statistical results of the diagonal [significant intervals = (760–800, 840–1280 ms), peak = 0.52, latency = 1140 ms].

Linking decoding to behavior

The successful decoding of previous responses, particularly in the cross-condition, suggests that a consistent component of the actual response is driven by the previous response, along the lines suggested by serial dependence studies. However, there is to date no evidence that the decoding represents a functionally useful signal. We investigated the possibility that the decoding of the memory signal shown in Figure 3 could have a perceptual

consequence by correlating the decoding accuracy of previous stimuli with the magnitude of serial dependence of individual participants.

Figure 4 plots the average poststimulus decoding (from 0 to 1800 ms) against the serial dependence measure for all participants for four different conditions. We first checked whether accuracy of decoding current responses trained on current responses correlated with serial dependence (Fig. 4a). There was a tendency toward correlation, but this was not significant ($r = 0.33$, $p = 0.19$, $BF = 0.3$). We then checked whether decoding of previous responses, trained on previous responses, correlates with serial dependence; again, this measure did not correlate significantly ($r = 0.29$, $p = 0.26$, $BF = 0.3$). However, the two cross-conditions of decoding produced significant correlations for training on previous and testing on current responses Fig. 5c, $r = 0.64$, $p = 0.006$, $BF = 7.6$) and training on current and testing on previous responses (Fig. 5d, $r = 0.52$, $p = 0.025$, $BF = 2.2$). As decoding accuracy for current stimuli also showed a tendency to correlate with serial dependency, potentially reflecting an intervening variable driving both decoding and serial dependence (e.g., such as attention), we remeasured the correlations with decoding of past responses after regressing out

the dependence on current decoding. This did not change the pattern of results, training on previous response and testing on current response (Fig. 4c), and vice versa (Fig. 4d); both remained significant ($p = 0.020$ and 0.050 , respectively), whereas training and testing on previous responses remained insignificant ($p = 0.22$).

Finally, we verified that for the prestimulus period (from -500 to 0 ms) there was no correlation between serial dependence and either of the cross-condition decoding conditions (training on previous response and testing on current response, $r = 0.36$, $p = 0.16$, $BF = 0.5$; decoding training on current response and testing on previous response, $r = 0.34$, $p = 0.19$, $BF = 0.6$).

Discussion

This study used classification techniques to characterize the neurophysiological substrate of serial dependence of response biases. From the scalp distribution of EEG signals in response to visual stimuli, we trained classifiers to decode the target stimuli, both their horizontal position (task irrelevant) and their vertical position (task relevant), as well as responses to the targets. The classifiers decoded the horizontal position well, both of the current and the previous stimuli. We were also able to classify the current target and the current and previous responses from the current EEG distribution, suggesting that the EEG included a representation of the previous stimulus leading to the response. The most successful classification used a cross-condition paradigm, training the classifier with labels from the previous response and testing on the current response (and vice versa). Importantly, the accuracy of this cross-condition classification correlated strongly with the psychophysically measured magnitude of serial dependence of participants, strongly supporting the possibility that the neural traces decoded by the classifiers correspond to a neural representation of an expectation associated with serial dependence.

Our research is consistent with and expands on recent research on by Bae and Luck (2019), who showed that current-trial EEG contained information about the orientation of the previous trial. They concluded that serial dependence may be driven by the reactivation of memory traces but provided little evidence for a direct relationship between serial biases and the decoded traces. Similarly, Fornaciai and Park (2020) observed memory trace reactivations in a numerosity discrimination task subject to serial dependence, but again with no direct evidence that the memory traces were connected with serial dependence. The strong correlations we found between accuracy of decoding and magnitude of serial dependence (across participants) strongly suggest that the trace of the representation of the previous target modulates the response to the present stimulus and participates in serial dependence perceptual effects.

Previous research (Barbosa et al., 2020; Fornaciai and Park, 2020) has suggested that perceptual history is communicated through activity-silent signals, not observable before the current stimulus is presented. In our study, decoding the previous hemifield (task-irrelevant feature) was not possible before stimulus onset, consistent with activity-silent traces. Similarly, decoding of previous responses (training and testing on previous responses) was significant only after presentation of the current stimulus.

One concern for many of these studies may be that classification decoders may be influenced by stereotypical eye position signals in the orbits (e.g., toward the target), which generate an

electrical dipole that survives ICA oculomotor detection. In our case, it is unlikely that this potential signal made a major contribution as we obtained similar results using only parietal and occipital electrodes. Furthermore, we observed similar decoding dynamics for horizontal and vertical positions of the previous trial, whereas any hypothetical eye-movement pattern would be very different. Another point to mention is that participants waited for at least 1 s before responding verbally. We cannot know if this affected the results and if paradigms without this pause may prove less effectual, given that some evidence suggests that serial dependence requires time to build up (Bliss et al., 2017).

Under the conditions of our experiment, we were able to decode previous responses from current EEG activation, but not previous targets. Similarly, decoding of current responses was stronger than the decoding of current targets. Several reasons could contribute to this result. First, the stimuli were particularly weak at response threshold (75% correct). Second, it is likely that the response to the target represents the neural representation of the target better than the target itself does, even when it is perceived incorrectly (on 25% of trials). This would be consistent with Ress and Heeger's (2003) observation that the BOLD response in early visual cortex is better predicted by the participant response than by the stimulus itself. In addition, perceptual decisional mechanisms that lead to the responses could contribute to the decoding. Finally, the actual motor response may also contribute to the decoding, although using a verbal rather than a button-press response should have minimized this motor contribution. Unfortunately, our data do not allow us to irrefutably distinguish between all these signals in forming the memory trace that promotes serial dependence; more complex experimental designs are required to address these points.

Activation maps give some information about the source of the memory signal. Although dipole localization was not possible with only 32 electrodes, the difference in dynamics in the activation patterns between current response decoding and cross-condition decoding is informative. Frontal activation became prominent only very late in cross-condition decoding, although it was very strong at 200 ms after stimulus onset for the current response decoding. This may suggest that the memory signal decoded from previous responses is mainly localized in the occipitoparietal cortex, consistent with studies that support that serial biases act directly on the current percept (Cicchini et al., 2017; Manassi et al., 2018; Alamia and VanRullen, 2019; Cicchini et al., 2021).

To highlight the shared component of decoding of previous and current representation, we tested cross-condition decoding, which is training on previous responses and testing on current ones, or vice versa. This led both to stronger decoding over longer intervals (compare Fig. 3c and e), and to strong and highly significant correlations with behavioral measures of serial dependence (Fig. 4c,d). Importantly, only the response after the stimulus correlated with psychophysical serial dependence, showing it is not artifactual.

A range of theoretical and empirical work in both auditory and visual perception suggests that the memory of the previous stimulus may be transmitted via an oscillatory signal (Alamia and VanRullen, 2019; Friston, 2019; Zhang et al., 2019; Bell et al., 2020). For example, Bell et al. (2020) showed that in judging the gender of faces, oscillations of specific frequencies in the low beta range were associated with serial dependence; faces preceded by a male face showed oscillations in reporting criteria ~ 17 Hz, and

those preceded by female faces showed oscillations at 14 Hz. Similarly, Ho et al. (2019) reported biases in auditory perception that oscillated at ~9 Hz only for trials preceded by a target tone to the same ear, strongly implicating neural oscillations in predictive perception. Both these studies showed that the memory signal is silent and reactivated only if the current stimulus is congruent with the past experience. This is consistent with our observation of reactivation of the memory signal only after stimulus presentation. Unfortunately, the current study was not designed to measure oscillations, but future studies could attempt to study the oscillatory dynamics of predictive signals by EEG decoding, monitoring the dynamics of memory traces.

In conclusion, our study provides evidence for a neurophysiological signal related to perceptual priors generated by the response to the previous stimuli. The ability of classifiers to generalize across previous and current conditions indicates that the two representations coexist in the EEG scalp potential in the current trial, consistent with a neural echo (Chang et al., 2017; Ho et al., 2019) of recent experience during stimulus processing. That the accuracy of decoding the signals of the previous responses was tightly related to serial dependence effects suggests that the intensity of prior signals drives the predictive processes. This elaborate mechanism presumably serves to enhance perceptual efficiency and to help preserve continuity of perceptual experience.

References

- Alamia A, VanRullen R (2019) Alpha oscillations and traveling waves: signatures of predictive coding? *PLoS Biol* 17:e3000487.
- Bae GY, Luck SJ (2018) Dissociable decoding of spatial attention and working memory from EEG oscillations and sustained potentials. *J Neurosci* 38:409–422.
- Bae G-Y, Luck SJ (2019) Reactivation of previous experiences in a working memory task. *Psychol Sci* 30:587–595.
- Bae G-Y, Luck SJ (2020) Serial dependence in vision: merely encoding the previous-trial target is not enough. *Psychon Bull Rev* 27:293–300.
- Barbosa J, Stein H, Martinez RL, Galan-Gadea A, Li S, Dalmau J, Adam KCS, Valls-Solé J, Constantinidis C, Compte A (2020) Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. *Nat Neurosci* 23:1016–1024.
- Bell J, Burr DC, Crookes K, Morrone MC (2020) Perceptual oscillations in gender classification of faces, contingent on stimulus history. *iScience* 23:101573.
- Bliss DP, Sun JJ, D'Esposito M (2017) Serial dependence is absent at the time of perception but increases in visual working memory. *Sci Rep* 7:14739.
- Ceylan G, Herzog MH, Pascucci D (2021) Serial dependence does not originate from low-level visual processing. *Cognition* 212:104709.
- Chang AY-C, Schwartzman DJ, VanRullen R, Kanai R, Seth AK (2017) Visual perceptual echo reflects learning of regularities in rapid luminance sequences. *J Neurosci* 37:8486–8497.
- Cicchini GM, Anobile G, Burr DC (2014) Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proc Natl Acad Sci U S A* 111:7867–7872.
- Cicchini GM, Mikellidou K, Burr D (2017) Serial dependencies act directly on perception. *J Vis* 17:6 1–9.
- Cicchini GM, Benedetto A, Burr DC (2021) Perceptual history propagates down to early levels of sensory analysis. *Curr Biol* 31:1245–1250.e2.
- Di Russo F, Martínez A, Hillyard SA (2003) Source analysis of event-related cortical activity during visuo-spatial attention. *Cereb Cortex* 13:486–499.
- Feigin H, Baror S, Bar M, Zaidel A (2021) Perceptual decisions are biased toward relevant prior choices. *Sci Rep* 11:648.
- Fischer J, Whitney D (2014) Serial dependence in visual perception. *Nat Neurosci* 17:738–743.
- Fornaciai M, Park J (2020) Neural dynamics of serial dependence in numerosity perception. *J Cogn Neurosci* 32:141–154.
- Foster JJ, Sutterer DW, Serences JT, Vogel EK, Awh E (2016) The topography of alpha-band activity tracks the content of spatial working memory. *J Neurophysiol* 115:168–177.
- Friston KJ (2019) Waves of prediction. *PLoS Biol* 17:e3000426.
- Fritsche M, Spaak E, de Lange FP (2020) A Bayesian and efficient observer model explains concurrent attractive and repulsive history biases in visual perception. *Elife* 9:e55389.
- Gregory RL (1968) Perceptual illusions and brain models. *Proc R Soc Lond B Biol Sci* 171:279–296.
- Helmholtz H (1867) *Handbuch der physiologischen Optik*. Leipzig, Germany: Voss.
- Ho HT, Burr DC, Alais D, Morrone MC (2019) Auditory perceptual history is propagated through alpha oscillations. *Curr Biol* 29:4208–4217.e3.
- Jung TP, Makeig S, Humphries C, Lee TW, McKeown MJ, Iragui V, Sejnowski TJ (2000) Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37:163–178.
- Kersten D, Mamassian P, Yuille A (2004) Object perception as Bayesian inference. *Annu Rev Psychol* 55:271–304.
- Kim S, Burr D, Cicchini GM, Alais D (2020) Serial dependence in perception requires conscious awareness. *Curr Biol* 30:R257–R258.
- King JR, Dehaene S (2014) Characterizing the dynamics of mental representations: the temporal generalization method. *Trends Cogn Sci* 18:203–210.
- Kiyonaga A, Scimeca JM, Bliss DP, Whitney D (2017) Serial dependence across perception, attention, and memory. *Trends Cogn Sci* 21:493–497.
- Kristjánsson Á, Ásgeirsson ÁG (2019) Attentional priming: recent insights and current controversies. *Curr Opin Psychol* 29:71–75.
- Lashley KS (1951) The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior; the Hixon Symposium* (pp. 112–146). Wiley.
- Liberman A, Fischer J, Whitney D (2014) Serial dependence in the perception of faces. *Curr Biol* 24:2569–2574.
- Maljkovic V, Nakayama K (1994) Priming of pop-out: I. Role of features. *Mem Cognit* 22:657–672.
- Manassi M, Liberman A, Chaney W, Whitney D (2017) The perceived stability of scenes: serial dependence in ensemble representations. *Sci Rep* 7:1971.
- Manassi M, Liberman A, Kosovicheva A, Zhang K, Whitney D (2018) Serial dependence in position occurs at the time of perception. *Psychon Bull Rev* 25:2245–2253.
- Min B-K, Dähne S, Ahn M-H, Noh Y-K, Müller K-R (2016) Decoding of top-down cognitive processing for SSVEP-controlled BMI. *Sci Rep* 6:36267.
- Murai Y, Whitney D (2021) Serial dependence revealed in history-dependent perceptual templates. *Curr Biol* 31:3185–3191.e3.
- Noah S, Powell T, Khodayari N, Oliván D, Ding M, Mangun GR (2020) Neural mechanisms of attentional control for objects: decoding EEG alpha when anticipating faces, scenes, and tools. *J Neurosci* 40:4913–4924.
- Ress D, Heeger DJ (2003) Neural correlates of perception in early visual cortex. *Nat Neurosci* 6:414–420.
- St John-Saaltink E, Kok P, Lau HC, de Lange FP (2016) Serial dependence in perceptual decisions is reflected in activity patterns in primary visual cortex. *J Neurosci* 36:6186–6192.
- Weingarten E, Chen Q, McAdams M, Yi J, Hepler J, Albarracín D (2016) From primed concepts to action: a meta-analysis of the behavioral effects of incidentally presented words. *Psychol Bull* 142:472–497.
- Wolff MJ, Jochim J, Akyürek EG, Stokes MG (2017) Dynamic hidden states underlying working-memory-guided behavior. *Nat Neurosci* 20:864–871.
- Wolpert DM, Ghahramani Z, Jordan MI (1995) An internal model for sensorimotor integration. *Science* 269:1880–1882.
- Zhang H, Morrone MC, Alais D (2019) Behavioural oscillations in visual orientation discrimination reveal distinct modulation rates for both sensitivity and response bias. *Sci Rep* 9:1115.