Familiarity Increases Processing Speed in the Visual System

Mariya E. Manahova, Eelke Spaak, and Floris P. de Lange

Abstract

■ Familiarity with a stimulus leads to an attenuated neural response to the stimulus. Alongside this attenuation, recent studies have also observed a truncation of stimulus-evoked activity for familiar visual input. One proposed function of this truncation is to rapidly put neurons in a state of readiness to respond to new input. Here, we examined this hypothesis by presenting human participants with target stimuli that were embedded in rapid streams of familiar or novel distractor stimuli at different speeds of presentation, while recording brain activity using magnetoencephalography and measuring behavioral performance. We investigated the temporal and spatial dynamics of signal truncation and whether this phenomenon bears relationship to participants' ability to categorize target items within a visual stream.

Behaviorally, target categorization performance was markedly better when the target was embedded within familiar distractors, and this benefit became more pronounced with increasing speed of presentation. Familiar distractors showed a truncation of neural activity in the visual system. This truncation was strongest for the fastest presentation speeds and peaked in progressively more anterior cortical regions as presentation speeds became slower. Moreover, the neural response evoked by the target was stronger when this target was preceded by familiar distractors. Taken together, these findings demonstrate that item familiarity results in a truncated neural response, is associated with stronger processing of relevant target information, and leads to superior perceptual performance. ■

INTRODUCTION

The brain rapidly processes an immense amount of sensory information and is constantly engaged in learning, adapting its responses based on new knowledge of the environment. A large body of evidence (e.g., Huang, Ramachandran, Lee, & Olson, 2018; Woloszyn & Sheinberg, 2012; Anderson, Mruczek, Kawasaki, & Sheinberg, 2008; Mruczek & Sheinberg, 2007; Freedman, Riesenhuber, Poggio, & Miller, 2006; Xiang & Brown, 1998; Fahy, Riches, & Brown, 1993; Li, Miller, & Desimone, 1993; Sobotka & Ringo, 1993) has demonstrated that familiarity with a stimulus modulates neural processing in various ways.

One recently described consequence of familiarity is temporal sharpening, or "truncation," of the signal evoked by a familiar stimulus (Meyer, Walker, Cho, & Olson, 2014). When the visual system is familiar with a stimulus, it is able to process the stimulus more quickly, leading to a truncated response for familiar compared with novel images. Put differently, activity in the visual system returns to baseline levels more rapidly for familiar than for novel images, thus leaving the visual system more aptly poised to process future input. This may lead to faster, more efficient processing of new information, and a behavioral consequence may be that familiar images have lower

saliency than novel ones. Indeed, monkeys as well as humans spend less time looking at familiar than novel images (Ghazizadeh, Griggs, & Hikosaka, 2016; Jutras & Buffalo, 2010), and familiar distractors interfere with behavioral tasks more than novel ones (Mruczek & Sheinberg, 2007).

When images are presented in rapid succession, image truncation leads to a larger dynamic range (i.e., baseline-to-peak difference) for a stream of familiar than novel images (Meyer et al., 2014). In such a case, neural processing has not yet terminated for one novel image when the next one is presented, effectively reducing the ability of the visual system to respond to new input. In a previous magnetoencephalography (MEG) study in humans (Manahova, Mostert, Kok, Schoffelen, & de Lange, 2018), we observed a larger dynamic range in the visual response to familiar object items compared with novel ones localized in the lateral occipital complex (LOC), thus demonstrating that signal truncation also occurs in the human brain.

Although signal truncation appears a robust and potentially useful phenomenon, there are still several outstanding questions. First, it is not clear to what extent signal truncation is present at different stages in the visual system. It may be a mechanism that underlies familiarity in the whole visual system, or it could be specific to object-selective cortex such as human area LOC and macaque area inferotemporal cortex (IT; Manahova et al., 2018; Meyer et al., 2014). Second, the process's temporal boundary conditions are

unclear. Signal truncation has been observed for streams where stimuli were presented at a rate of 120–180 msec, but it is unknown whether the phenomenon extends to other temporal scales. Multiple studies show that processing within earlier cortical regions takes place at a faster intrinsic time scale than for later regions (Murray et al., 2014). Therefore, short image durations may potentially elicit the strongest signal truncation in early visual regions, whereas long image durations may do so in downstream visual areas. Third, if signal truncation aids in faster neural processing, then this phenomenon may be linked to behavioral improvements such as improved perceptual speed and accuracy.

To address these open questions, we examined the neural dynamics of stimulus processing along the visual hierarchy and their relationship to behavior. We embedded target stimuli in a stream of familiar or novel distractors while measuring human neural activity using MEG. To preview our findings, we found robust truncation of neural activity for familiar input, which was localized to early visual cortex for the shortest image duration and to later, more anterior, visual areas for medium-length image durations. Moreover, familiarity with the distractor stream led to stronger neural processing of the target and to superior behavioral performance, which was most pronounced for rapid visual streams. Together, our results suggest that neural truncation helps to rapidly put the visual system in a state of readiness to respond to new input.

METHODS

Data and Software Availability

Data and code used for stimulus presentation are available online at the Donders Institute for Brain, Cognition, and Behavior repository at hdl.handle.net/11633/aacjmbax. Videos displaying how different trial types appear to a participant are also available there.

Participants

Thirty-seven healthy human volunteers (25 women, 11 men; mean age = 26.4 years, SD = 6.6 years) with normal or corrected-to-normal vision, recruited from the university's participant pool, completed the experiment and received monetary compensation. The sample size, which was defined a priori, ensured at least 80% power to detect within-participant experimental effects with an effect size of Cohen's $d > 0.5$. The study was approved by the local ethics committee (CMO Arnhem-Nijmegen, Radboud University Medical Center) under the general ethics approval ("Imaging Human Cognition", CMO 2014/ 288), and the experiment was conducted in compliance with these guidelines. Written informed consent was obtained from each individual.

Stimuli

Stimuli were chosen from the image sets provided at cvcl. mit.edu/MM/uniqueObjects.html and cvcl.mit.edu/MM/ objectCategories.html (see also Konkle, Brady, Alvarez, & Oliva, 2010). A different object was represented in each image, and all objects were shown against a gray background. A total of 554 images were used (20 target images, six familiar distractors, and 528 novel distractors). Familiar images were randomly selected for each participant and were shown during the behavioral training session as well as during the MEG testing session, whereas novel images were only shown during the MEG testing session. In both the behavioral and MEG sessions, the images subtended 4° × 4° of visual angle.

Apparatus

MATLAB (The Mathworks, Inc.) and the Psychophysics Toolbox extensions (Brainard, 1997) were used to show the stimuli on a monitor with a resolution of 1920 × 1080 pixels and a refresh rate of 120 Hz for the behavioral session. For the MEG session, a PROpixx projector (VPixx Technologies) was used to project the images on the screen, with a resolution of 1920 × 1080 pixels and a refresh rate of 120 Hz.

Trial Structure

During the behavioral session as well as the MEG session, each trial began with a fixation dot (see Figure 1A for the trial structure). The fixation dot was presented for a randomly selected period between 500 and 750 msec. Then, a stream of images was presented, lasting for 2400 msec. The duration of the trial was kept the same, whereas the number of images per trial varied depending on the image duration. For images lasting 50 msec, there were 48 images per trial; for images lasting 100 msec, there were 24 images per trial; for images lasting 150 msec, there were 16 images per trial; for images lasting 200 msec, there were 12 images per trial; and for images lasting 300 msec, there were eight images per trial. At the end of the trial, if a correct response had been given (i.e., the target image had been categorized correctly as animate or inanimate), the fixation dot turned green for 500 msec. If the response was incorrect, the fixation dot turned red for 500 msec. Two types of responses were considered incorrect: incorrect classification as animate or inanimate or no response (i.e., response failure). Participants were instructed to respond on every trial, and they did so on almost all trials (M = 99.56%, SD = 0.64%). Afterward, a blank screen was presented for 1250 msec, and participants were encouraged to blink during this period.

Experimental Design

This experiment consisted of two sessions, conducted on separate days: a behavioral training session on the first

Figure 1. (A) Experimental setup. The behavioral training session, taking place on Day 1, included learning the target images and practicing the experimental task. The MEG testing session, taking place on Day 2, included performing the experimental task and demonstrating the knowledge participants had acquired of the familiar images. (B) Trial structure. First, participants saw a bull's-eye fixation point for a jittered period between 500 and 750 msec. Then, a stream of distractors was presented (here, highlighted in yellow), with each image lasting for the allotted amount of time for that trial (50, 100, 150, 200, or 300 msec) and no gaps between images. Somewhere in the distractor stream, a target image was embedded (here, highlighted in orange). Participants had to respond as quickly as possible after

detecting the target, categorizing the target image as animate or inanimate. Finally, a fixation dot was presented at the end of the trial, and participants received feedback on their performance. (C) Accuracy and RTs per image duration. Blue lines denote familiar distractors, and red lines denote novel distractors. Error bars show the within-participant *SEM*.
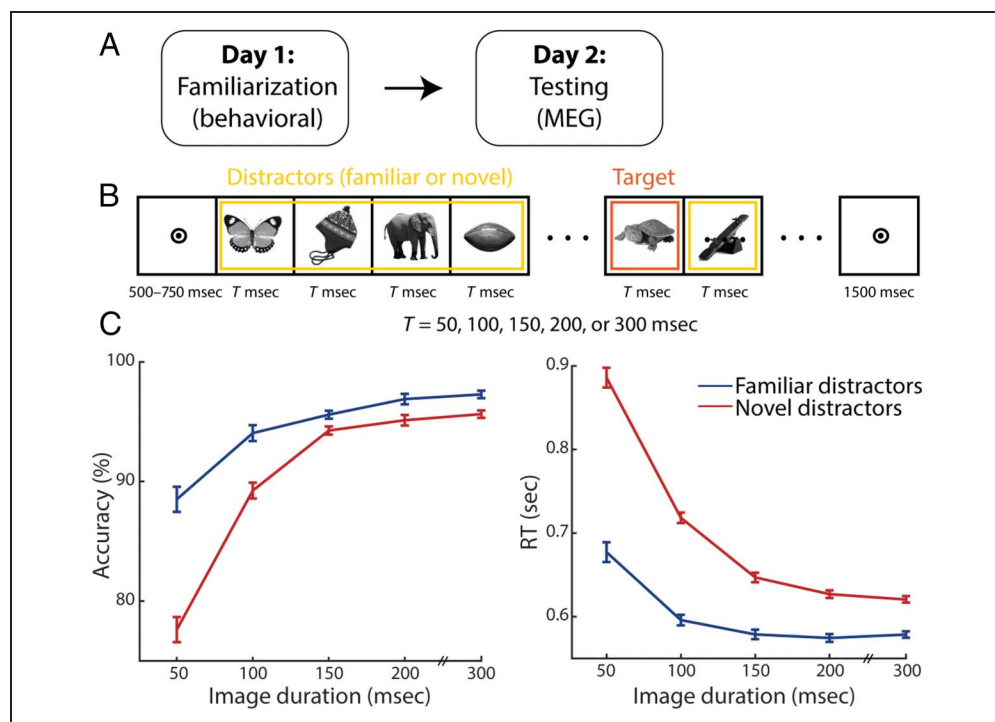
day and an MEG testing session on the second day. The behavioral training session consisted of two parts: a target learning part and a distractor familiarization part. Specifically, in the first part of the behavioral training session, participants were taught to categorize the target images into animate or inanimate. There were 20 target images, and they saw them one at a time. The images were presented on the screen until a response was made. The participants pressed the left and right arrow keys to indicate their answer, with the response mapping randomized across participants. There were six blocks, with each block showing all 20 target images. Participants were given feedback after each image about whether they categorized that image correctly. They were also shown their average accuracy at the end of each block.

Then, in the second part of the behavioral training session, participants proceeded to practice the experimental task (see Figure 1A). They observed streams of images of distractors, and they had to detect the target image, which was 1 of the 20 images they had just learned. Once they detected the target among the distractors, they had to immediately press a button to categorize it as animate or inanimate. The target could occur at different time points in the trial. The trial lasted for 2400 msec, and on 90% of trials, the target was presented sometime between 1200 and 2100 msec. On 10% of trials, the target was presented before the 1200-msec mark of the trial to make sure that participants were paying attention throughout the trial and not only during the second

half. These trials were discarded from further analysis. Notably, during the behavioral training session, the distractors were always the same six images. Because participants saw these repeatedly on every trial, they became highly familiar. Participants completed five blocks of 100 trials each, for a total of 500 trials.

One or two days later, participants completed the MEG testing session in which they saw familiar (i.e., those presented during the behavioral session) and novel (not seen before) images. Half of the trials in the experiment were familiar, and the other half were novel. On each trial, six images were chosen as distractors and repeated to match the number of images needed for that image duration (see Trial structure). If the trial was familiar, then these six images were the six distractors participants had seen repeatedly during the behavioral training session. If the trial was novel, the six images were randomly drawn (without replacement) from the set of novel images. Each novel image was shown four times during the course of the experiment; in contrast, each familiar image was displayed 350 times during the MEG session and 250 times during the behavioral session. The task was the same as during the behavioral training session: detect the target and categorize it as animate or inanimate immediately after seeing it. Participants completed seven blocks of 100 trials each, for a total of 700 trials. At the end of the MEG testing session, participants' knowledge of the familiar images was assessed. Participants saw 50 images, the six familiar ones and 44 selected at

random from the novel images participants had been shown, and participants had to indicate whether the image was familiar or novel. "Familiar" referred to images seen repeatedly during the behavioral training session as well as during the MEG testing session, whereas "novel" referred to images seen only during the MEG testing session.

Data Acquisition

MEG Recordings

Brain activity was recorded using a 275-channel MEG system with axial gradiometers (VSM/CTF Systems) in a magnetically shielded room. During the experiment, head position was monitored online and corrected if necessary (Stolk, Todorovic, Schoffelen, & Oostenveld, 2013). Head position monitoring was done using three coils: one placed on the nasion, one in an earplug in the left ear, and one in an earplug in the right ear. MEG signals were sampled at 1200 Hz. A projector outside the magnetically shielded room projected the visual stimuli onto a screen in front of the participant via mirrors. Participants gave their behavioral responses via an MEG-compatible button box. Participants' eye movements and blinks were also monitored using an eye-tracker system (EyeLink, SR Research Ltd.).

MRI Recordings

To allow for source reconstruction, anatomical MRI scans were acquired using a 3-T MRI system (Siemens) and a T1-weighted magnetization prepared rapid gradient echo sequence with a GRAPPA acceleration factor of 2 (repetition time = 2300 msec, echo time = 3.03 msec, voxel size = $1 \times 1 \times 1$ mm, 192 transversal slices, 8° flip angle).

Data Analysis

Preprocessing of MEG Data

The MEG data were preprocessed offline using the FieldTrip software (Oostenveld, Fries, Maris, & Schoffelen, 2011). Trials where the target was presented earlier than 1200 msec into the trial (10% of trials) were removed from analysis. The data were demeaned, and noise was removed based on third-order gradiometers. Then, trials with high variance were manually inspected and removed if they contained excessive and irregular artifacts. This resulted in retaining, on average, 93% of trials per participant (range = 86–99%). The number of trials that remained after preprocessing did not vary between conditions (familiar trials: $M = 292.70$, $SD = 9.62$; novel trials: $M = 295.76$, $SD = 11.36$), $t(36) = -1.25$, $p = .22$. Afterward, independent component analysis was applied to identify regular artifacts such as heartbeat and eye blinks and remove the respective components.

ERFs

The data were filtered using a sixth-order Butterworth low-pass filter with a cutoff frequency of 40 Hz. Before calculating ERFs, the data were baseline-corrected on the interval starting at 200 msec before stimulus onset until stimulus onset (0 msec), that is, the onset of the first distractor. For the ERF analysis of target-related data, the data were baseline-corrected on the interval starting at 100 msec before target onset until target onset (0 msec). Subsequently, the data were split into conditions based on the familiarity and the image duration for the trial. Then, the data were transformed to planar gradiometer representation to facilitate interpretation as well as averaging over participants. The familiar and novel conditions had an equal number of trials by design. The resulting ERFs were averaged over participants for visualization purposes. Standard error of the mean (*SEM*) was computed within participants (Cousineau, 2005) and with bias correction (Morey, 2008).

Power Analysis

To assess the dynamic range (i.e., the baseline-to-trough difference) of the signal, we computed the power in the ERF as in Manahova et al. (2018). First, we averaged the data over trials, time-locking to stimulus onset. The time window of interest was from 200 to 1200 msec. Next, we applied the planar transformation to the data. Then, we conducted a spectral analysis for all frequencies between 1 and 30 Hz with a step size of 1 Hz. We applied the fast Fourier transform to the planar-transformed time domain data, after tapering with a Hanning window. The power analysis was carried out separately per condition, where a condition was defined by familiarity and image duration. Afterward, the horizontal and vertical components of the planar gradient were combined by summing. The resulting power per frequency was averaged over participants.

Coherence Analysis

Complementary to the power analysis on the evoked fields, we computed the coherence between neural activity for each image duration and a synthetic stimulus signal at the frequency of interest. We selected data from 500 to 1000 msec after stimulus presentation because the visual response was strongly rhythmic during that period. Next, we applied the planar transformation and conducted a spectral analysis for the frequency of interest for each condition, applying the fast Fourier transform after tapering with a Hanning window. Then, we computed coherence with the synthetic stimulus signal for each condition (i.e., image duration), and afterward, the horizontal and vertical components of the planar gradient were combined by summing. The resulting coherence values per frequency were averaged over participants.

Source Reconstruction

For maximal sensitivity, we carried out source reconstruction analysis. We used each participant's anatomical MRI scan to create a volume conduction model based on a single-shell model of the inner surface of the skull (Nolte & Dassios, 2005). We computed participant-specific dipole grids, which were based on a regularly spaced 6-mm grid in normalized Montreal Neurological Institute space. Then, the sensor-level axial gradiometer data were split into conditions determined by familiarity and image duration. For each condition, we carried out source analysis with the DICS method (Gross et al., 2001), quantifying coherence with a synthetic stimulus signal at the frequency of interest, and selected the time window between 500 and 1000 msec poststimulus because the visual response was strongly rhythmic during that period. We used coherence as a measure of dynamic range and thus signal truncation. Finally, we averaged the data over participants.

To investigate the topographical spread of coherence for different image durations, we compared source-level activity during stimulus presentation to baseline activity. We computed the source-level stimulus versus baseline coherence difference for each of the five image durations. We defined the brain area of interest for each image duration as including all source locations that had an activity value of 50% or higher of the peak activity value for that condition. We also extracted the y coordinate of the peak of the coherence map for each image duration, indicating the position of this peak on the anterior–posterior axis.

Statistical Analysis

Behavioral Results

Mean RT and accuracy were first calculated within participant per condition. RT values were log-transformed and accuracy values were arcsine-transformed to reduce the skewness and more closely approach a normal distribution of the data. Then, a two-way repeated-measures ANOVA assessed the differences in RT, and another two-way repeated-measures ANOVA evaluated the differences in accuracy. There were two independent variables: Familiarity (two levels: familiar and novel) and Image Duration (five levels: 50, 100, 150, 200, and 300 msec).

Overall Amplitude in ERFs

To statistically assess the MEG activity difference between familiar and novel trials and control for multiple comparisons, we applied cluster-based permutation tests (Maris & Oostenveld, 2007), as implemented by FieldTrip (Oostenveld et al., 2011). The tests were carried out on the period from 0 msec (the onset of the first distractor) to 1200 msec, over all sensors, and with 1000 permutations. For each sensor and time point, the MEG signal was compared univariately between two conditions, using a paired t test. Positive and negative clusters were then formed separately by grouping spatially and temporally adjacent data points whose corresponding p values were lower than .05 (two tailed). Cluster-level statistics were calculated by summing the t values within a cluster, and a permutation distribution of this cluster-level test statistic was computed. The null hypothesis was rejected if the largest cluster in the considered data was found to be significant, which was the case if the cluster's p value was smaller than .05 as referenced to the permutation distribution.

Influence of Familiarity and Image Duration on Coherence

To assess how image duration affected the topographical spread of coherence (see above for how this was computed), we conducted a one-way ANOVA with five levels, which were the five image durations. To assess the influence of familiarity and image duration on coherence with the frequency of interest (quantified as source-level coherence, as explained above), we performed a two-way repeated-measures ANOVA with two factors, Familiarity (two levels) and Image Duration (five levels). Post hoc t tests were used to assess each pairwise comparison. Coherence values for each condition were taken from the general visual system ROI, which was constructed by taking the union of the five condition-specific areas of entrainment of the stimulus (see Results). Because coherence data are skewed, we log-transformed them to facilitate statistical comparisons between conditions. Results remain qualitatively unaltered if this transformation is not applied.

Next, we computed coherence values for each image duration within ROIs defined separately for each image duration (in contrast to the general visual system ROI described above). We computed the difference in coherence between familiar and novel trials for each image duration in each ROI. Each ROI was based on a stimulus versus baseline comparison for that image duration, meaning that the contrast for selection was independent from the tested contrast; in addition, the two conditions did not differ in terms of the number of trials (see Preprocessing of MEG data). Furthermore, each successive area did not include the previous one (e.g., the ROI for the 150-msec condition did not include any locations belonging to the 50- or 100-msec conditions). To test how image duration and ROI influenced the magnitude of difference in coherence, we ran a two-way repeated-measures ANOVA with two factors, Image Duration (three levels: 50, 100, and 150 msec) and ROI (five levels). Moreover, we fitted regression lines for each image duration across the five ROIs, and we conducted an ANCOVA to compare the slopes of the regression coefficients for each image duration. For each of these analyses, we did not test the 200- and 300-msec conditions because they did not show a significant difference in coherence between

familiar and novel trials when tested with the general visual system ROI approach described above (see Results).

Correlation between Dynamic Range and RTs

To test for the presence of a relationship between coherence (and thus dynamic range) and RT, we computed the correlation between the difference in coherence between familiar and novel trials (familiar–novel) and the difference in log-transformed RT (novel–familiar). We computed five Spearman's rank correlation coefficients, one for each image duration.

Influence of Familiarity and Image Duration on Target-related Amplitude

To determine whether familiarity and image duration influenced ERF amplitude for the target stimulus, we performed a two-way repeated-measures ANOVA with two factors, Familiarity (two levels) and Image Duration (five levels), on the time window from 0 msec (target onset) to 300 msec and including all occipital sensors. Post hoc t tests were conducted to assess each pairwise comparison. Because combined planar-transformed ERF data are skewed, we log-transformed them to ensure the data approximately conformed to a normal distribution and facilitate statistical comparisons between conditions. Results remained qualitatively unaltered when this transformation was not applied. Note that this analysis included both correct and incorrect trials; we conducted the same analysis on correct trials only, and the results were nearly identical and qualitatively unaltered.

Correlation between Target-related ERF Amplitude and RTs

To quantify the potential relationship between target-related amplitude and RT, we computed the correlation between the difference in target-related ERF amplitude (from 200 to 1200 msec after the onset of stimulus presentation) between familiar and novel trials (familiar–novel) and the difference in log-transformed RT (novel–familiar). We computed five Spearman's rank correlation coefficients, one for each image duration.

Correlation between Target-related ERF Amplitude and Dynamic Range

To test for the presence of a relationship between target-related amplitude and coherence (and thus dynamic range), we computed the correlation between the difference in target-related ERF amplitude (from 0 to 300 msec after target presentation) between familiar and novel trials (familiar–novel) and the difference in log-transformed coherence values (familiar–novel). We computed five Spearman's rank correlation coefficients, one for each image duration.

RESULTS

We measured MEG activity while participants viewed target stimuli in a stream of familiar or novel distractors. Image duration varied per trial; images were shown for 50, 100, 150, 200, or 300 msec. The participants' task was to categorize a familiar target image as animate or inanimate as quickly as possible.

Behavioral Results

Behavioral Performance Improved with Distractor Familiarity and Image Duration

Target categorization accuracy was higher when distractors were familiar compared with novel images, $F(1, 36) = 61.56$, $p = 2.64\text{e-}09$, and higher for longer image durations, $F(4, 144) = 116.41$, $p < 1\text{e-}09$ (see Figure 1B). Moreover, the effect of familiarity was most pronounced for the most challenging rapid visual streams, as indicated by a Familiarity × Duration interaction, $F(4, 144) = 8.37$, $p = 4.22\text{e-}06$.

Target categorization speed was also faster when distractors were familiar compared with novel, $F(1, 36) = 439.35$, $p < 10\text{e-}09$, and faster for longer image durations, $F(4, 144) = 166.59$, $p < 10\text{e-}09$ (see Figure 1B). Again, the effect of familiarity was most pronounced for the most challenging rapid visual streams, as indicated by a Familiarity × Duration interaction, $F(4, 144) = 48.27$, $p < 10\text{e-}09$.

At the end of the MEG session, participants' knowledge of the distractor image familiarity was assessed. On average, participants correctly identified the familiar images in 91.6% of trials ($SD = 8.9\%$), showing that they were clearly aware of the familiarity manipulation.

MEG Results

Novel Stimuli Led to Higher Overall Activity than Familiar Ones

To investigate the time courses of each trial type, we computed ERFs (Figure 2A). There was a marked difference in the overall amplitude of the signal, with novel items leading to higher activity than familiar ones ($p < 1\text{e-}03$ for all image durations), as found previously in Manahova et al. (2018). Interestingly, the MEG amplitude difference between familiar and novel images appeared to decrease over time, especially for short image durations. This may be the case because, although the visual system is better able to process familiar images, it is eventually also overwhelmed by quick succession of images, reducing the difference between conditions. The power spectra for the evoked responses and corresponding topographies for each condition are illustrated in Figure 2B. The stimulus-evoked power at the driving frequency approximates the dynamic range of the response (Manahova et al., 2018). We obtained highly similar sensor topographies
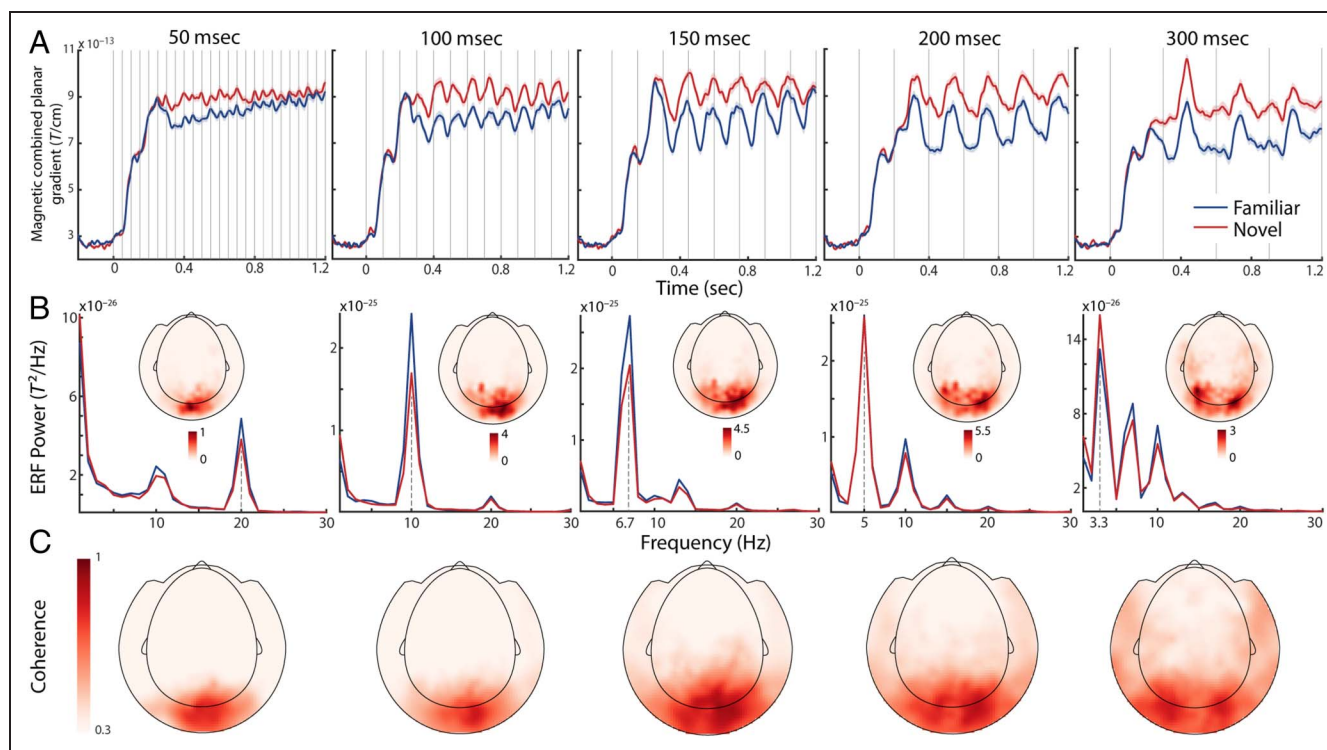
Figure 2. Familiarity effect in sensor-level data. (A) MEG activity (magnetic fields with planar combination) over time for familiar (blue) and novel (red) conditions, separately for each image duration (50, 100, 150, 200, and 300 msec). Activity was averaged over all occipital sensors. Shaded areas are error bars illustrating within-participant *SEM* for familiar (light blue) and novel (light red) conditions. Vertical lines denote the onset of each image. (B) Power spectra for the ERFs for familiar (blue) and novel (red) conditions, separately for each image duration (50, 100, 150, 200, and 300 msec). Dotted vertical lines denote the frequency of interest for each image duration. Activity was averaged over all occipital sensors. Topography plots show the spatial distribution of the mean activity of familiar and novel trials for the corresponding image duration. The color bars of the topography plots are of the same units as the *y* axis of the power spectra. (C) Topographies of coherence. The topography plots show the spatial distribution of the mean activity of familiar and novel trials for the corresponding image duration. The same color bar applies to all five plots.

when analyzing coherence with the stimulus, as shown in Figure 2C.

Anatomical Location of Stimulus Tracking

Next, we aimed to identify the anatomical location of stimulus tracking. We therefore shifted our focus to the source level and computed source-level coherence with the stimulus using a beamformer approach. We contrasted the resulting coherence values between stimulus and baseline periods. The resulting five areas that track stimuli are depicted in Figure 3A, and the observed areas are anatomically compatible with early visual cortex.

Strong stimulus tracking was observed within the ventral visual stream. The topography of this effect generally spread more anteriorly as the image duration became longer. As image durations increased, the anatomical peak location of the highest stimulus tracking moved to progressively more downstream regions of the visual hierarchy. Indeed, the peak *y* coordinate shifted more anteriorly as the image durations lengthened (Figure 3A), $F(4, 184) = 5.98, p = .0002$, as shown by a one-way ANOVA. Note that this is also in line with the sensor-level topographies of

stimulus-evoked power (Figure 2B) and coherence (Figure 2C).

Familiar Images Led to Higher Dynamic Range than Novel Ones

Next, we aimed to determine whether familiar and novel stimuli led to significantly different coherence, and thus dynamic range, at each stimulus frequency. We constrained this analysis to an ROI corresponding to the visual system as identified by our data. Namely, we combined the five areas that track stimuli shown in Figure 3A into one ROI.

First, we focused on the coherence averaged over this single ROI encompassing a large part of visual cortex. Familiar distractors were associated with stronger coherence than novel ones, $F(1, 36) = 22.09, p = 3.75e-05$, and coherence also differed as a function of image duration, $F(4, 144) = 7.07, p = 3.16e-05$, but the interaction was not significant, $F(4, 144) = 1.10, p = .36$ (see Figure 3B). We found that coherence was significantly higher for familiar than for novel trials when images were presented for 50 msec, $t(36) = 4.23, p = 1.50e-04$; 100 msec, $t(36) = 2.39, p = .02$; and 150 msec, $t(36) = 3.08, p = .004$. There
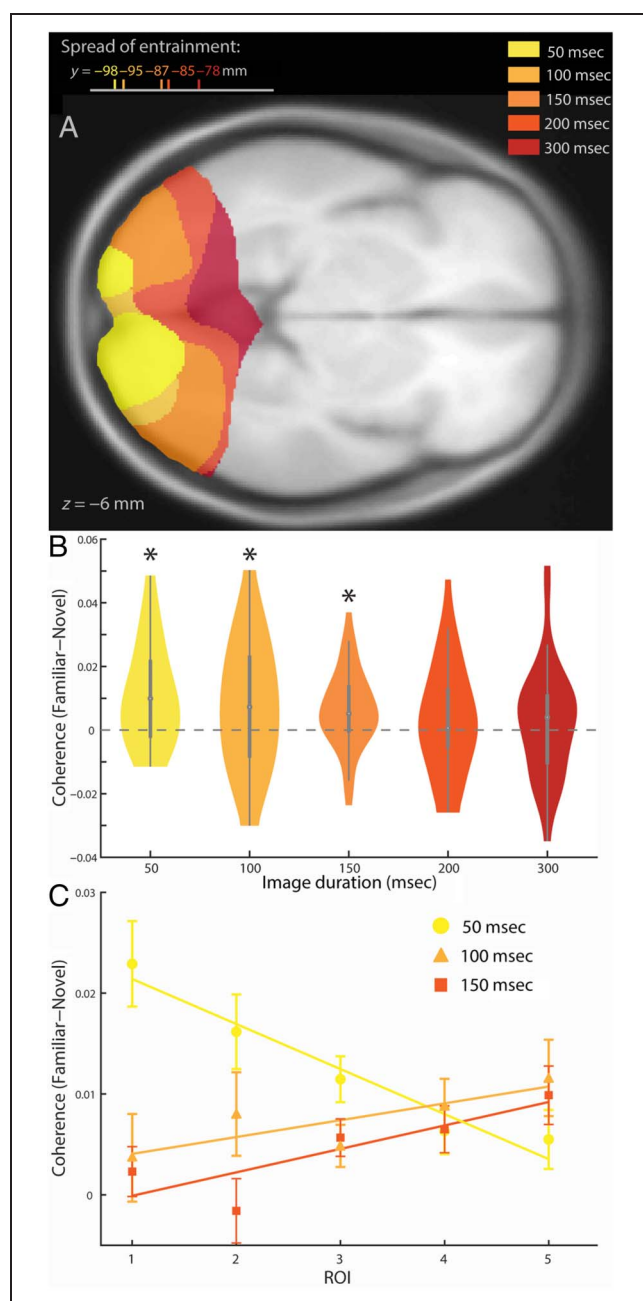
Figure 3. (A) Anatomical location of stimulus tracking in the brain for activity resulting from source reconstruction for stimulus versus baseline comparison. The anatomical location of activity was estimated separately for each image duration (50, 100, 150, 200, and 300 msec) and, for visual representation, thresholded at 50% of peak value for each condition. The peak y coordinates for each image duration are shown in the top left corner. (B) Difference in coherence between familiar and novel trials, separately per image duration. Stars denote significance level at alpha = .05. Dashed horizontal line shows the level of no difference. (C) Difference in coherence between familiar and novel trials for different image durations per ROI. Each ROI is based on the stimulus versus baseline comparison for the corresponding image duration; that is, ROI 1 corresponds to the 50-msec condition, ROI 2 corresponds to the 100-msec condition, and so forth. An ROI does not contain any of the locations belonging to the previous ROI; for example, ROI 3 does not contain any locations belonging to ROI 1 or 2. Error bars denote within-participant *SEM*. Lines show regression fits for each image duration.

was no significant difference between familiar and novel for images lasting 200 msec, $t(36) = 1.14$, $p = .26$, or 300 msec, $t(36) = 0.89$, $p = .38$. Thus, the dynamic range was higher for familiar than for novel stimuli. The data suggest that this difference was particularly prominent for the short and medium image durations (50, 100, and 150 msec), although the lack of a significant interaction indicates there is no statistical support for significantly stronger stimulus tracking for short compared with long (200 and 300 msec, respectively) image durations.

Familiarity-induced Truncation Shifted Topographically Depending on Image Duration

Next, we shifted our attention from the single visual system ROI to individual ROIs defined for each image duration. To assess how the topographical distribution of familiarity-induced truncation changed with image duration, we quantified the difference in coherence between familiar and novel trials in the five ROIs, each based on the area of stimulus tracking for an image duration (see Figure 3C). Importantly, each successive area did not include the previous one (e.g., for the ROI for the 150-msec condition, we excluded any locations that were included in the ROIs for the 50- or 100-msec conditions). For this analysis, we examined activity for the 50-, 100-, and 150-msec conditions because they showed significant differences in coherence between novel and familiar distractors when tested within the general visual system ROI (see Figure 3B). The different conditions exhibited different patterns of signal truncation across ROIs, as evidenced by a significant interaction, $F(4, 144) = 4.17$, $p = 9.83\text{e-}05$. To understand this interaction effect better, we conducted an ANCOVA, fitting regression lines across ROIs for each image duration (see Figure 3C) and then comparing the regression coefficients with each other. We found that signal truncation decreased significantly with anteriority for the 50-msec condition ($t = -4.24$, $p < .01$), whereas we observed a trend in the opposite direction for the 100-msec condition (positive slope but not significantly different from zero; $t = 1.80$, $p = .07$). For the 150-msec condition, signal truncation increased significantly with anteriority ($t = 2.44$, $p = .02$). Thus, for the 50-msec condition, signal truncation peaked in the earliest ROI and decreased in successive ones, whereas for the 100- and 150-msec conditions, signal truncation was low in early ROIs and increased in successive ones.

No Relationship between Dynamic Range and RTs

We found a clear RT benefit for familiar stimuli compared with novel ones, as well as a clear difference in signal truncation. This raises the question of whether these two effects are correlated across participants. To quantify this, we computed the correlation between the difference in coherence on familiar and novel trials (familiar–novel) and the difference in RT on familiar and novel

trials (novel–familiar) across participants. We computed five Spearman's rank correlation coefficients, one for each image duration. None of these correlations was significant (50 msec: $r = .02, p = .90$; 100 msec: $r = −.02, p = .90$; 150 msec: $r = .02, p = .91$; 200 msec: $r = .05, p = .78$; 300 msec: $r = .10, p = .57$).

Stronger Processing of Target Image when Embedded in Familiar Distractors

We assessed how strongly a target was processed when it was embedded in familiar or novel distractors by examining the overall amplitude of target-evoked activity in the time window from 0 msec (target onset) to 300 msec over occipital sensors, as seen in the target-related ERFs (see Figure 4). The target evoked a stronger response when it was embedded in familiar distractors rather than novel ones, $F(1, 36) = 233.72, p < 10e-15$. Although this effect was not equally strong for all image durations (Familiarity × Duration interaction: $F(4, 144) = 5.27, p = 5.46e-04$), post hoc tests confirmed that it was robustly present for all image durations (50 msec: $t(36) = 11.46, p = 1.43e-13$; 100 msec: $t(36) = 10.08, p = 4.99e-12$; 150 msec: $t(36) = 11.32, p = 2.04e-13$;

200 msec: $t(36) = 8.14, p = 1.11e-09$; 300 msec: $t(36) = 5.78, p = 1.35e-06$). Therefore, the target stimulus resulted in stronger evoked responses for familiar than for novel trials, regardless of image duration. Note that this pattern was absent when we applied this analysis to a distractor in the middle of the visual stream instead of on the target stimulus (data not shown). This control analysis demonstrates that the ERF differences reported here are truly because of target processing and not simply to the preexisting differences in neural activity because of the surrounding distractor streams. Note also that, for target-related activity, the MEG activity continues to rise as the trial progresses because more images are shown on the screen after the end of the time window depicted here.

No Relationship Between Target-related Amplitude and RTs

We found significant differences in target-related ERF amplitude as well as in RTs, and we wanted to determine whether these two were related. Thus, we computed the correlation between the difference in target-related amplitude on familiar and novel trials (familiar–novel) and the
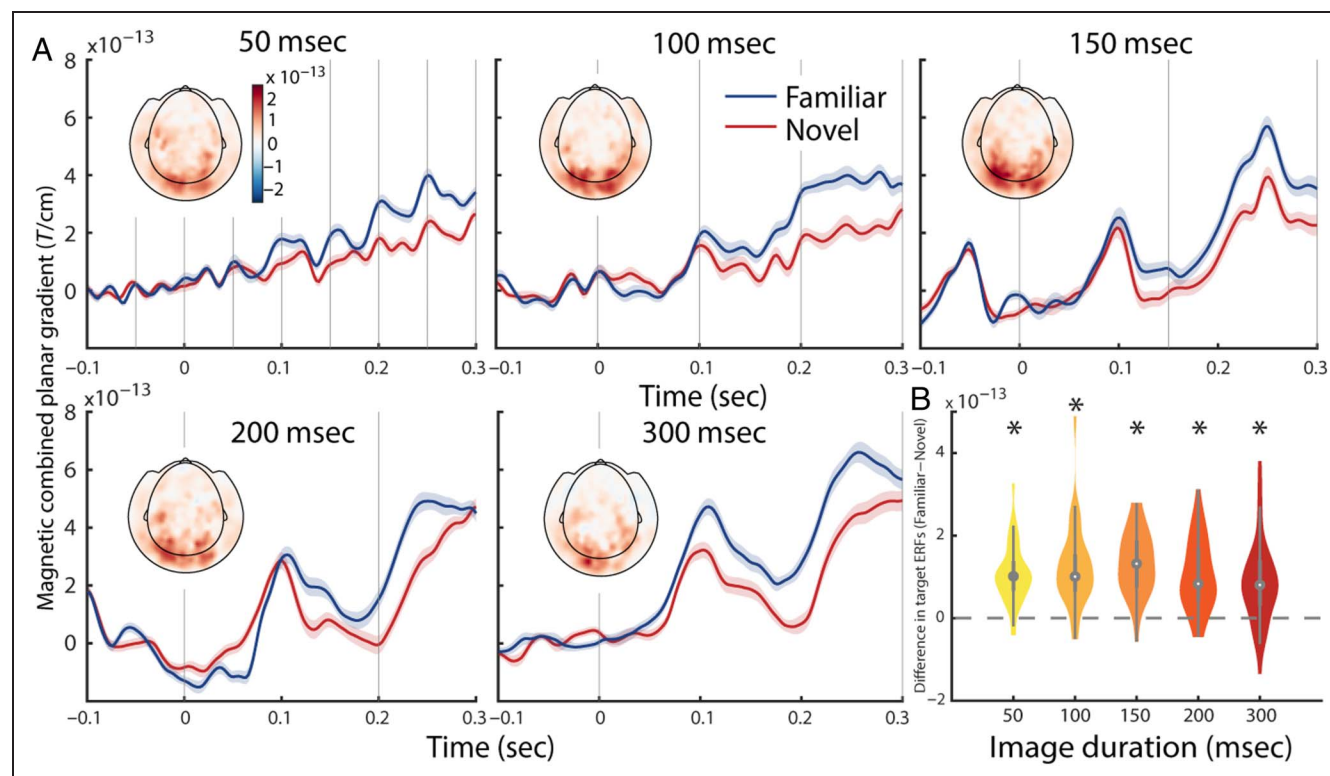
Figure 4. Target-related activity in sensor-level data. (A) MEG activity (combined planar gradient) over time for familiar (blue) and novel (red) conditions, separately for each image duration (50, 100, 150, 200, and 300 msec). Activity was averaged over all occipital sensors. Shaded areas are error bars illustrating within-participant *SEM* for familiar (light blue) and novel (light red) conditions. Vertical lines denote the onset of each image. Topography plots show the spatial distribution of the difference in target-related activity between familiar and novel streams (familiar–novel), separately for each image duration. Activity was averaged over the time window from 0 msec (target onset) to 300 msec. The color bars of the topography plots are of the same units as the *y* axis of the ERFs, and the scale of the color bars is the same for all topography plots. (B) Difference in target-related activity between familiar and novel trials (familiar–novel), separately per image duration. Stars denote significance level at alpha = .05. Dashed horizontal line shows the level of no difference.

difference in RT on familiar and novel trials (novel–familiar) across participants. We computed five Spearman's rank correlation coefficients, one for each image duration. None of these correlations was significant (50 msec: $r = .07$, $p = .67$; 100 msec: $r = -.06$, $p = .72$; 150 msec: $r = .13$, $p = .43$; 200 msec: $r = .02$, $p = .91$; 300 msec: $r = .05$, $p = .76$).

No Relationship between Target-related Amplitude and Dynamic Range

We found significant differences in target-related ERF amplitude as well as in coherence (and thus dynamic range), so we wanted to determine whether there is an association between these two variables. Therefore, we computed the correlation between the difference in target-related amplitude on familiar and novel trials (familiar–novel) and the difference in coherence on familiar and novel trials (familiar–novel) across participants. We computed five Spearman's rank correlation coefficients, one for each image duration. None of these correlations was significant (50 msec: $r = .07$, $p = .69$; 100 msec: $r = -.18$, $p = .29$; 150 msec: $r = -.09$, $p = .58$; 200 msec: $r = -.15$, $p = .39$; 300 msec: $r = -.19$, $p = .27$).

DISCUSSION

In this study, we investigated the temporal and spatial dynamics of signal truncation and whether this phenomenon bears relationship to participants' ability to detect target items within a stream of visual distractors. In short, we found truncation of neural activity for familiar input to be the strongest when image streams were presented rapidly (images last between 50 and 150 msec each). For the shortest image duration (50 msec), signal truncation was localized to early visual cortex, whereas successively longer image durations were linked to progressively later visual areas. Furthermore, the neural processing of the target stimulus was stronger when targets were embedded in familiar streams compared with novel ones. Behavioral performance was also better for targets in familiar streams, especially for short image durations.

A wealth of research has demonstrated that familiarity affects neural processing, with the most commonly observed effect being familiarity suppression (Huang et al., 2018; Woloszyn & Sheinberg, 2012; Anderson et al., 2008; Mruczek & Sheinberg, 2007; Freedman et al., 2006; Xiang & Brown, 1998; Fahy et al., 1993; Li et al., 1993; Sobotka & Ringo, 1993). Viewing an object image repeatedly and becoming familiar with it result in reduced spiking activity in IT in monkeys (Miller, Li, & Desimone, 1991) and reduced hemodynamic activity in human LOC as measured with fMRI (Grill-Spector, Henson, & Martin, 2006). This reduction of the population response may indicate sharper neuronal tuning and a sparser population representation for familiar than for novel images (Woloszyn & Sheinberg, 2012; Freedman et al., 2006). Our data are in accordance

with this reduction of activity for familiar stimuli: We found a sustained higher amplitude response for novel than for familiar items in the ERFs (see Figure 2A).

Image familiarity not only influences the magnitude of the neural response but also truncates the neural response (Meyer et al., 2014), such that neural activity returns to baseline levels more quickly for a familiar image than a novel image. This puts neurons in a state of readiness to respond to new input more rapidly. We observed signal truncation throughout the ventral visual system (Figure 3A). Interestingly, the topographical distribution of the effect shifted anteriorly and laterally as image duration increased. Perhaps surprisingly, signal truncation was robustly present in early visual cortex, particularly for very rapid streams (50 msec per image). In line with this, it has recently been observed that correlates of visual familiarity are observed as early as macaque area V2 (Huang et al., 2018).

Because signal truncation of distractors may benefit target processing, we explored whether signal truncation was related to target processing and behavioral performance. Indeed, we found that, although novel distractors showed a higher amplitude than familiar distractors (Figure 2A), target-related activity was higher when the targets were embedded in familiar images than in streams of novel images (Figure 4). In terms of behavioral performance, participants were markedly faster and more accurate in categorizing the target when the target was embedded within familiar distractors, suggesting a direct relationship between signal truncation of the distractors and behavioral performance. However, a between-participant correlation analysis showed no reliable correlation between truncation and behavior. Nevertheless, both the stronger evoked responses and the enhanced behavioral performance for targets embedded in familiar streams suggest a benefit for the processing of visual input, possibly stemming from the signal truncation of the distractors. This is in accordance with findings showing that familiar distractors are less salient (Ghazizadeh et al., 2016; Jutras & Buffalo, 2010) and less disruptive (Mruczek & Sheinberg, 2007) than novel ones. Although it is tempting to conclude that there is a direct link between signal truncation, enhanced target processing, and improved behavior, our data, unfortunately, do not offer clear evidence for or against this idea. It would be interesting for future work to address this possible relationship directly.

We hypothesized that signal truncation would be prevalent in early visual cortex for short image durations and occur in later visual areas as images are shown longer. We based this idea on the fact that early regions reach their peak activity relatively soon after stimulus onset, whereas more downstream regions show peak spiking later (Nowak & Bullier, 1997; Dinse & Krüger, 1994), and that higher-order cortical regions have a slower intrinsic time scale than early sensory regions (Murray et al., 2014). Moreover, regions lower in the visual hierarchy accumulate input for shorter image durations than

downstream areas (Honey et al., 2012; Hasson, Yang, Vallines, Heeger, & Rubin, 2008). Our results are in accordance with these notions. First, we found overall stimulus tracking (from a stimulus vs. baseline contrast) that showed an anterior and lateral topographical shift with increasing image duration. Second, the spatial localization of the difference in signal truncation between familiar and novel stimuli also shifted to more downstream areas as images were shown longer, with the 50-msec condition showing the strongest signal truncation difference for early ROIs, whereas the 100- and 150-msec conditions demonstrated stronger signal truncation difference for more downstream ROIs. Therefore, both overall tracking and the signal truncation difference showed a topographical distribution that shifted up the brain's visual processing hierarchy along with increasing image duration.

It is intriguing to speculate how signal truncation can be observed for such short image durations as 50 msec and in early visual cortex. This implies that familiarity may be encoded in early visual cortex somehow, which may sound implausible because familiarity suppression effects are commonly established in IT (e.g., Li et al., 1993). However, Huang et al. (2018) showed that a type of familiarity suppression is also found in macaque V2, which suggests that early visual cortex may encode familiarity. Alternatively, the signal truncation effect we observe in early visual areas may be the result of feedback from local recurrent connections or from higher-order areas. In our current analysis, and in accordance with previous work (Manahova et al., 2018), we quantify signal truncation by computing coherence over a 1000-msec time window, so we are unable to detect at which time points in the image stream truncation is present. It is possible, however, that the visual evoked response may be truncated only after the first few stimuli in a stream are presented, which may indicate the involvement of other areas. After all, the processing of an image continues after other images are presented even when images are presented as briefly as 17 msec (Mohsenzadeh, Qin, Cichy, & Pantazis, 2018), suggesting that an effect such as signal truncation may build up in magnitude as a challenging visual stream is being processed.

In the current study, image durations of 200 and 300 msec did not elicit a significant difference in signal truncation. Future research could use more complex or naturalistic (Kayser, Körding, & König, 2004) images because such stimuli, which require more extensive and demanding processing, may engage regions higher up in the visual hierarchy (Murray et al., 2014), thus leading to signal truncation differences even if images are shown for 200 msec or longer.

Although we compared familiar and novel images, our familiarity manipulation involved both familiarity and recency. Xiang and Brown (1998) define familiarity as whether an object was seen during a previous session, usually on a previous day, and recency as whether an object has been seen during the same recording session. Although these effects are clearly related, as they refer to whether the system has been exposed to this stimulus before, they denote different time scales (Fahy et al., 1993) and may rely on different mechanisms. Familiarity, related to a longer time scale, requires plasticity changes in the neural network that encodes the stimulus. Recency, related to a shorter time scale, is more similar to repetition suppression or adaptation effects and may result from synaptic depression or adapted input from other neurons in the same network (Vogels, 2016). In the current study, we manipulated both familiarity and recency in our familiar condition, meaning that we are unable to distinguish the effects of these two phenomena. It would be interesting to further disentangle the respective contributions of recency and familiarity to signal truncation.

In conclusion, familiarity with a visual stimulus leads to a truncation of neural activity in the visual system, and this truncation was strongest for the fastest presentation speeds. Moreover, this truncation for familiar distractor stimuli is associated with stronger target processing and behavioral improvements in perceptual categorization, suggesting a functional role of this phenomenon.

REFERENCES

Anderson, B., Mruczek, R. E., Kawasaki, K., & Sheinberg, D. (2008). Effects of familiarity on neural activity in monkey inferior temporal lobe. *Cerebral Cortex*, 18, 2540–2552.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Massons method. *Tutorials in Quantitative Methods for Psychology*, 1, 42–45.

Dinse, H. R., & Krüger, K. (1994). The timing of processing along the visual pathway in the cat. *NeuroReport*, 5, 893–897.

Fahy, F. L., Riches, I. P., & Brown, M. W. (1993). Neuronal activity related to visual recognition memory: Long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Experimental Brain Research*, 96, 457–472.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2006). Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cerebral Cortex*, 16, 1631–1644.

Ghazizadeh, A., Griggs, W., & Hikosaka, O. (2016). Ecological origins of object salience: Reward, uncertainty, aversiveness, and novelty. *Frontiers in Neuroscience*, 10, 378.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10,* 14–23.

Gross, J., Kujala, J., Hämäläinen, M., Timmermann, L., Schnitzler, A., & Salmelin, R. (2001). Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences, U.S.A., 98,* 694–699.

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *Journal of Neuroscience, 28,* 2539–2550.

Honey, C. J., Thesen, T., Donner, T. H., Silbert, L. J., Carlson, C. E., Devinsky, O., et al. (2012). Slow cortical dynamics and the accumulation of information over long timescales. *Neuron, 76,* 423–434.

Huang, G., Ramachandran, S., Lee, T. S., & Olson, C. R. (2018). Neural correlate of visual familiarity in macaque area V2. *Journal of Neuroscience, 38,* 8967–8975.

Jutras, M. J., & Buffalo, E. A. (2010). Recognition memory signals in the macaque hippocampus. *Proceedings of the National Academy of Sciences, U.S.A., 107,* 401–406.

Kayser, C., Körding, K. P., & König, P. (2004). Processing of complex stimuli and natural scenes in the visual cortex. *Current Opinion in Neurobiology, 14,* 468–473.

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General, 139,* 558–578.

Li, L., Miller, E. K., & Desimone, R. (1993). The representation of stimulus familiarity in anterior inferior temporal cortex. *Journal of Neurophysiology, 69,* 1918–1929.

Manahova, M. E., Mostert, P., Kok, P., Schoffelen, J. M., & de Lange, F. P. (2018). Stimulus familiarity and expectation jointly modulate neural activity in the visual ventral stream. *Journal of Cognitive Neuroscience, 30,* 1366–1377.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164,* 177–190.

Meyer, T., Walker, C., Cho, R. Y., & Olson, C. R. (2014). Image familiarization sharpens response dynamics of neurons in inferotemporal cortex. *Nature Neuroscience, 17,* 1388–1394.

Miller, E. K., Li, L., & Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science, 254,* 1377–1379.

Mohsenzadeh, Y., Qin, S., Cichy, R. M., & Pantazis, D. (2018). Ultra-rapid serial visual presentation reveals dynamics of feedforward and feedback processes in the ventral visual pathway. *ELife, 7,* e36329.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4,* 61–64.

Mruczek, R. E., & Sheinberg, D. L. (2007). Context familiarity enhances target processing by inferior temporal cortex neurons. *Journal of Neuroscience, 27,* 8533–8545.

Murray, J. D., Bernacchia, A., Freedman, D. J., Romo, R., Wallis, J. D., Cai, X., et al. (2014). A hierarchy of intrinsic timescales across primate cortex. *Nature Neuroscience, 17,* 1661–1663.

Nolte, G., & Dassios, G. (2005). Analytic expansion of the EEG lead field for realistic volume conductors. *Physics in Medicine and Biology, 50,* 3807–3823.

Nowak, L. G., & Bullier, J. (1997). The timing of information transfer in the visual system. In K. S. Rockland, J. H. Kaas, & A. Peters (Eds.), *Extrastriate cortex in primates* (pp. 205–241). Boston, MA: Springer.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011,* 156869.

Sobotka, S., & Ringo, J. L. (1993). Investigation of long term recognition and association memory in unit responses from inferotemporal cortex. *Experimental Brain Research, 96,* 28–38.

Stolk, A., Todorovic, A., Schoffelen, J. M., & Oostenveld, R. (2013). Online and offline tools for head movement compensation in MEG. *Neuroimage, 68,* 39–48.

Vogels, R. (2016). Sources of adaptation of inferior temporal cortical responses. *Cortex, 80,* 185–195.

Woloszyn, L., & Sheinberg, D. L. (2012). Effects of long-term visual experience on responses of distinct classes of single units in inferior temporal cortex. *Neuron, 74,* 193–205.

Xiang, J. Z., & Brown, M. W. (1998). Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. *Neuropharmacology, 37,* 657–676.