nature neuroscience

Article

https://doi.org/10.1038/s41593-026-02496-6

Single-neuron encoding of rapidly learned visual information reshapes human perception

Received: 8 October 2025

Accepted: 17 September 2026

Published online: 04 November 2026

Check for updates

Marcelo Armendariz[1,2,10] ✉, Julie Blumberg[3,10], Jed Singer[1], Franz Aiple[4], Jiye Kim[1], Nuria Domínguez-Iturza[5,6], Armin Brandt[7], Peter Reinacher[8,9], Andreas Schulze-Bonhage[7] & Gabriel Kreiman[1,2] ✉

Humans can swiftly learn to recognize visual objects after minimal exposure. Integrating new information with existing knowledge requires forming enduring neuronal representations to enable future recognition. Yet, the neuronal mechanisms in the human brain underlying such rapid perceptual changes remain unclear. Here we recorded single-neuron activity in occipital cortex (OC) and medial temporal lobe (MTL) regions as participants rapidly learned to recognize degraded images. OC and MTL neurons modulated their activity to encode newly learned visual information and reshape perception. Population decoding revealed that OC neurons required additional processing time to resolve the identity of learned images, with delayed neuronal responses in the MTL. Our findings suggest that OC supports recognition after rapid learning via extensive recurrent processing, potentially involving higher-order neocortical areas, without relying on feedback from MTL. These results provide mechanistic constraints for biologically plausible models of visual recognition and few-shot visual learning.

Humans and other primates display a remarkable ability to visually recognize objects. The main family of models to describe visual recognition capabilities is based on hierarchical deep neural networks[1–6]. Current instantiations of such networks are typically trained via back-propagation through millions of examples when learning from scratch, but also when fine tuning an existing pre-trained network[7–10].

In stark contrast to such long training processes, humans can swiftly learn to recognize visual objects with just one or a few exposures. A striking example of rapid learning is the sudden recognition of a degraded black-and-white image of an object (Mooney image)[11] (Fig. 1a). These degraded Mooney images tend to be unrecognizable initially. However, Mooney images become easily interpretable after a brief exposure to the original intact version of the image[12–14] (Fig. 1b). This perceptual ability for rapid learning develops gradually in late childhood (7–9 years) and is crucial for integrating new information with prior knowledge[15,16]. Rapid learning necessitates the formation of enduring neuronal signatures to enable subsequent recognition. Thus, newly acquired experience embedded in neural circuits dramatically shapes the perception of future incoming visual stimuli[15,17–19]. Despite extensive behavioral characterization, the neuronal mechanisms

[1]Boston Children's Hospital, Harvard Medical School, Boston, MA, USA. [2]Center for Brains, Minds and Machines, MIT, Cambridge, MA, USA. [3]Department of Neuropediatrics and Muscle Disorders, University Medical Center Freiburg, University of Freiburg, Freiburg, Germany. [4]Department of Neurosurgery, University Medical Center Freiburg, University of Freiburg, Freiburg, Germany. [5]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA. [6]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [7]Epilepsy Center, University Medical Center Freiburg, University of Freiburg, Freiburg, Germany. [8]Department of Stereotactic and Functional Neurosurgery, University Medical Center Freiburg, University of Freiburg, Freiburg, Germany. [9]Fraunhofer Institute for Laser Technology, Aachen, Germany. [10]These authors contributed equally: Marcelo Armendariz, Julie Blumberg. ✉e-mail: marcelo.armendarizgil@childrens.harvard.edu; gkreiman@gmail.com

underlying perceptual changes induced by rapid learning in the human brain are not well understood.

Previous neurophysiological recordings in monkeys showed that neurons in the temporal visual cortex can adapt their selectivity to represent recently learned degraded versions of objects such as Mooney images[20]. On the basis of noninvasive methods, rapid learning in humans is thought to modulate activity in a broad range of neocortical regions, including visual areas and frontoparietal regions[21–24]. However, little is known about the mechanisms underlying how individual neurons modify their response patterns after learning and how rapid and all-or-none perceptual changes are orchestrated by the neuronal dynamics across different regions.

Here, we recorded the spiking activity of 1,104 single neurons in medial occipital and temporal regions of the human brain in patients with pharmacologically resistant epilepsy, while 13 participants learned to rapidly recognize novel Mooney images. Participants demonstrated the hallmark characteristics of essentially all-or-none perceptual changes triggered by one or a few exposures to grayscale counterparts to the Mooney images. Neurons in both the occipital cortex (OC) and medial temporal lobe (MTL) modulated their responses to encode newly learned Mooney images. OC neurons required additional time to process the identity of learned Mooney images compared with intact grayscale images, with MTL neurons showing delayed responses. The learning-induced dynamics in OC may indicate the need for extensive recurrent processing, potentially involving top-down feedback from higher cortical areas, before information reaches MTL.

Results

Rapid learning during image recognition and single-neuron recordings

We designed an image recognition task to investigate changes in perception following rapid learning. In total, 13 participants were shown sequences of images and were instructed to report whether they recognized the identity of the depicted objects. Images consisted of two-tone black-and-white pictures (Mooney images) (Fig. 1a) and their corresponding grayscale counterparts (Fig. 1b). In each trial, an image was shown for 500 ms, followed by a blank screen for 1,300 ms (Fig. 1c). Participants were first exposed to the Mooney versions of the images, which they did not recognize (Fig. 1d, red). After exposure to the Mooney images, brief learning sequences were introduced. These sequences involved presenting participants with the original grayscale image followed by the corresponding Mooney version of that image (Fig. 1d, turquoise). After exposure to these learning sequences, participants were shown the same Mooney images again (Fig. 1d, blue). Additional grayscale images were also presented, though no longer shown sequentially with the corresponding Mooney images. Different Mooney and grayscale images were interleaved throughout the task. Using this design, we defined three trial categories: 'pre-learning', where Mooney images were presented before learning occurred; 'grayscale', where the original grayscale images were presented; and 'post-learning', where Mooney images were presented after the learning process. Mooney image recognition rapidly and drastically grew from almost none before learning to high levels after learning (Fig. 1e), requiring only one or a few learning sequences (on average, 1.8 ± 1). Overall, grayscale images were highly recognizable, with participants performing on average at 93% ± 5% accuracy (mean ± s.d. across sessions, $n = 34$ sessions) (Fig. 1f). For the Mooney images, recognition was poor in the pre-learning condition, with an accuracy of 4% ± 2% and greatly improved for the post-learning condition, reaching 83% ± 15%. Performance was statistically different between conditions (Gray versus Pre, $P = 1.4 \times 10^{-12}$; Gray versus Post, $P = 5.5 \times 10^{-5}$; and Pre versus Post, $P = 1.4 \times 10^{-12}$; two-sided Wilcoxon signed-rank test). In addition, we measured reaction times from stimulus onset to participants' responses (Fig. 1g). Response times were longer for recognized Mooney images (post-learning: median 0.53 s, IQR 0.44–0.57 s, mean 0.52 s) than for grayscale images (median 0.48 s, IQR 0.41–0.52 s, mean 0.48 s) (Post versus Gray, $P = 2 \times 10^{-5}$; two-sided Wilcoxon signed-rank test), suggesting that interpreting the learned Mooney stimuli required additional cognitive effort. Unrecognized Mooney images (pre-learning) elicited the slowest responses overall (median 0.60 s, IQR 0.49–0.72 s, mean 0.63 s) (Pre versus Gray, $P = 3.6 \times 10^{-7}$; Pre versus Post, $P = 4.3 \times 10^{-7}$, two-sided Wilcoxon signed-rank test).

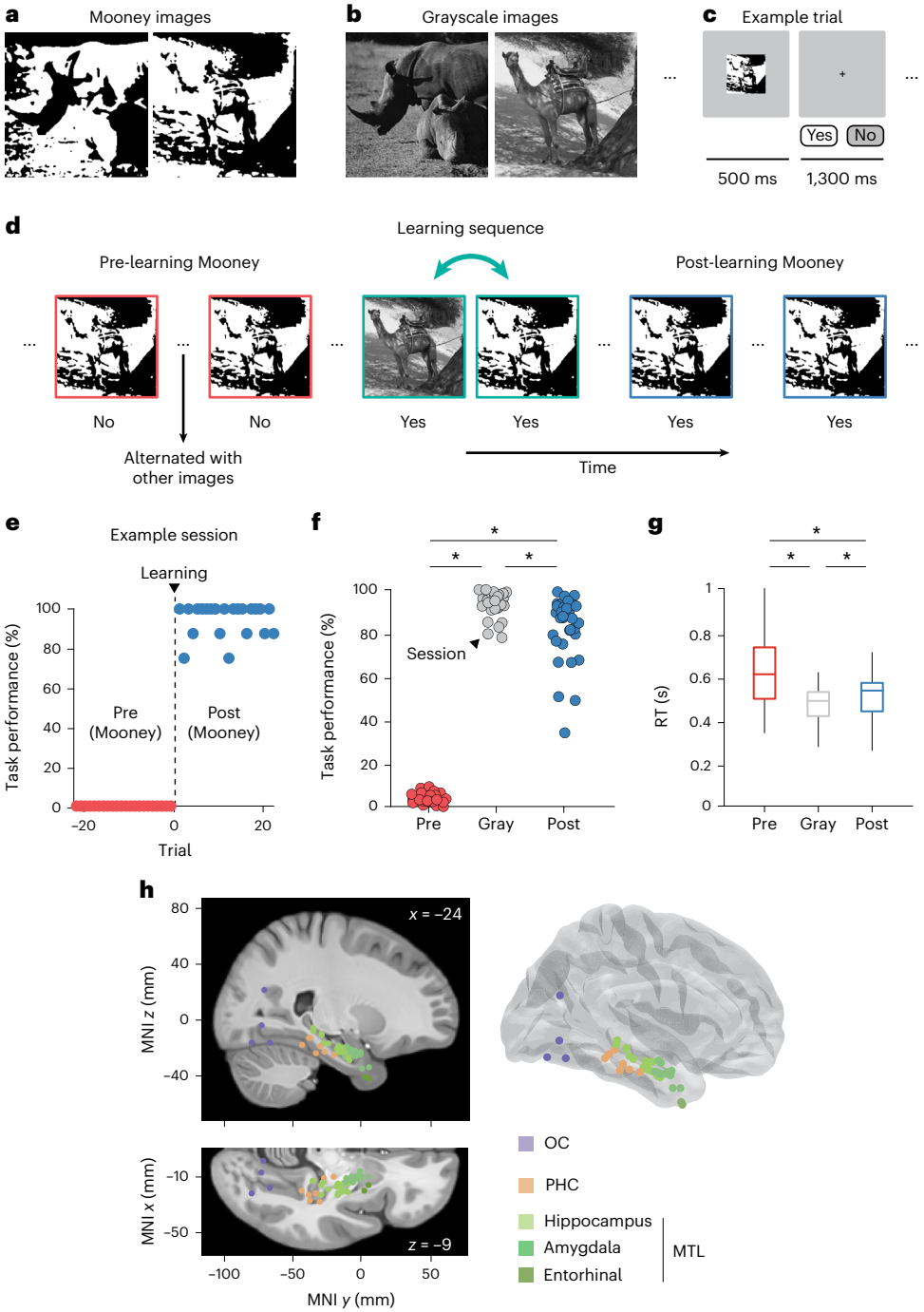
During this task, we recorded spiking activity from a total of 1,104 single neurons in medial OC and medial temporal brain areas in 34 sessions from 13 patients with pharmacologically resistant epilepsy who were chronically implanted with hybrid macro- and microelectrodes for seizure monitoring (Fig. 1h, Supplementary Figs. 1 and 2, and Supplementary Table 1). The recordings included 163 neurons from the medial OC, 212 in the parahippocampal cortex (PHC), 49 in the entorhinal cortex, 343 in the hippocampus and 337 in the amygdala (Supplementary Table 2).

We characterized eye movements during the task in 13 sessions across ten healthy participants (Extended Data Fig. 1). These

Fig. 1 | Experimental design, behavior and recording sites. a, Examples of Mooney images. **b**, Examples of grayscale images. **c**, An example trial displaying stimulus and blank screen durations. Participants reported whether they recognized the object on the screen via button press ('yes' or 'no'). **d**, An illustration of the learning process for Mooney images. The sequence depicts Mooney images before learning (pre-learning, red frame) and after learning (post-learning, blue frame), with the learning phase marked by a double-headed turquoise arrow. Pre-learning Mooney images are unrecognizable, as indicated by the response 'no' shown beneath the image. After multiple presentations of pre-learning Mooney images, a grayscale and a Mooney image depicting the same object are shown in immediate succession to facilitate learning of the Mooney image. During the learning sequence, both the grayscale and Mooney images are recognized ('yes' beneath the images). Following the learning sequence, Mooney images can be recognized. The colored frames were added for illustration purposes and were not visible during the task. Note that this panel illustrates a sequence for a single image (camel); the full sequence included multiple images, either grayscale or Mooney (pre- or post-learning), interleaved throughout the session, as indicated by the ellipses between stimuli (Methods). **e**, The proportion of Mooney images recognized across trials during an example session. Trials for each image are aligned with respect to its learning event (trial 0), indicated by a downward arrow. The session included multiple images, and each dot represents the mean accuracy across the different images. Mooney image recognition rapidly and drastically grew from almost none to almost perfect after learning. Post-learning trials (positive numbers, blue dots) show the proportion of Mooney images recognized after learning. Pre-learning trials (negative numbers, red dots) show that later-recognized Mooney images were not recognized before learning, even though the images were identical. **f**, Behavioral performance during the recognition task across all sessions. Before learning ('Pre', red), participants exhibited low performance (4%). Grayscale images ('Gray', gray) and Mooney images after learning ('Post', blue) were highly recognizable (93% and 83%, respectively). Each dot represents a session ($n = 34$). Asterisks indicate statistically significant differences between conditions (Post versus Gray, $P = 5.5 \times 10^{-5}$; Pre versus Gray, $P = 1.4 \times 10^{-12}$; and Pre versus Post, $P = 1.4 \times 10^{-12}$; two-sided Wilcoxon signed-rank test, Bonferroni-corrected). **g**, Response times (RTs) relative to image onset. Participants responded faster to grayscale images than to post-learning Mooney images, and slowest to pre-learning Mooney images. Box plots show the median (center line), 25th and 75th percentiles (box bounds), and whiskers extending to the minimum and maximum values of the response time distributions across sessions ($n = 34$). Asterisks show statistically significant differences between conditions (Post versus Gray, $P = 2 \times 10^{-5}$; Pre versus Gray, $P = 3.6 \times 10^{-7}$; and Pre versus Post, $P = 4.3 \times 10^{-7}$; two-sided Wilcoxon signed-rank test, Bonferroni-corrected). **h**, Electrode locations across all participants. Each dot represents a microwire bundle location, mapped onto a standardized MNI152 brain template (left) and a 3D brain model (right). All dots are displayed in the same hemisphere for visualization purposes only. The color of each dot denotes the microwire bundle location.

https://doi.org/10.1038/s41593-026-02496-6

participants exhibited similar recognition patterns to the 13 patients with epilepsy, with high recognition of Mooney images only after learning, and longer reaction times for recognized Mooney images compared with grayscale images (compare Extended Data Fig. 1 and Fig. 1). In the majority of trials (~83%), participants made no saccades during image presentation (0–0.5 s). The low incidence of saccades may be explained by the brief stimulus presentation and small image size, allowing participants to perceive the entire image in a single glance. In the subset of trials with saccades (~17%), the latency to the first saccade from image onset was 267 ± 33 ms. There were no significant differences across conditions in either the proportion of trials with saccades (Gray versus Pre, $P = 0.17$; Gray versus Post, $P = 0.52$; and Pre versus Post, $P = 0.38$; two-sided Wilcoxon signed-rank test) or the timing of the first saccade (Gray versus Pre, $P = 0.49$; Gray versus Post, $P = 0.99$; and Pre versus Post, $P = 0.49$; two-sided Wilcoxon signed-rank test). After image offset (0.5–1 s), saccades were rare, occurring in fewer than ~7% of trials, with no differences across conditions (Gray versus Pre, $P = 0.5$; Gray versus Post, $P = 0.09$; and Pre versus Post, $P = 0.59$; two-sided Wilcoxon signed-rank test). We also analyzed blink behavior and found that blinks were negligible during stimulus presentation (<3% of trials), occurring mostly after stimulus offset. No differences in blink rates were observed across conditions. In addition, to test whether performance could be explained by associative learning, the task included control trials in which a grayscale image was sequentially paired with a mismatched Mooney image. Participants readily recognized the grayscale images but were unable to identify the following unrelated Mooney images (mean accuracy 7%, comparable to pre-learning Mooney recognition) and did not misattribute the identity of the grayscale image to the Mooney image.



https://doi.org/10.1038/s41593-026-02496-6

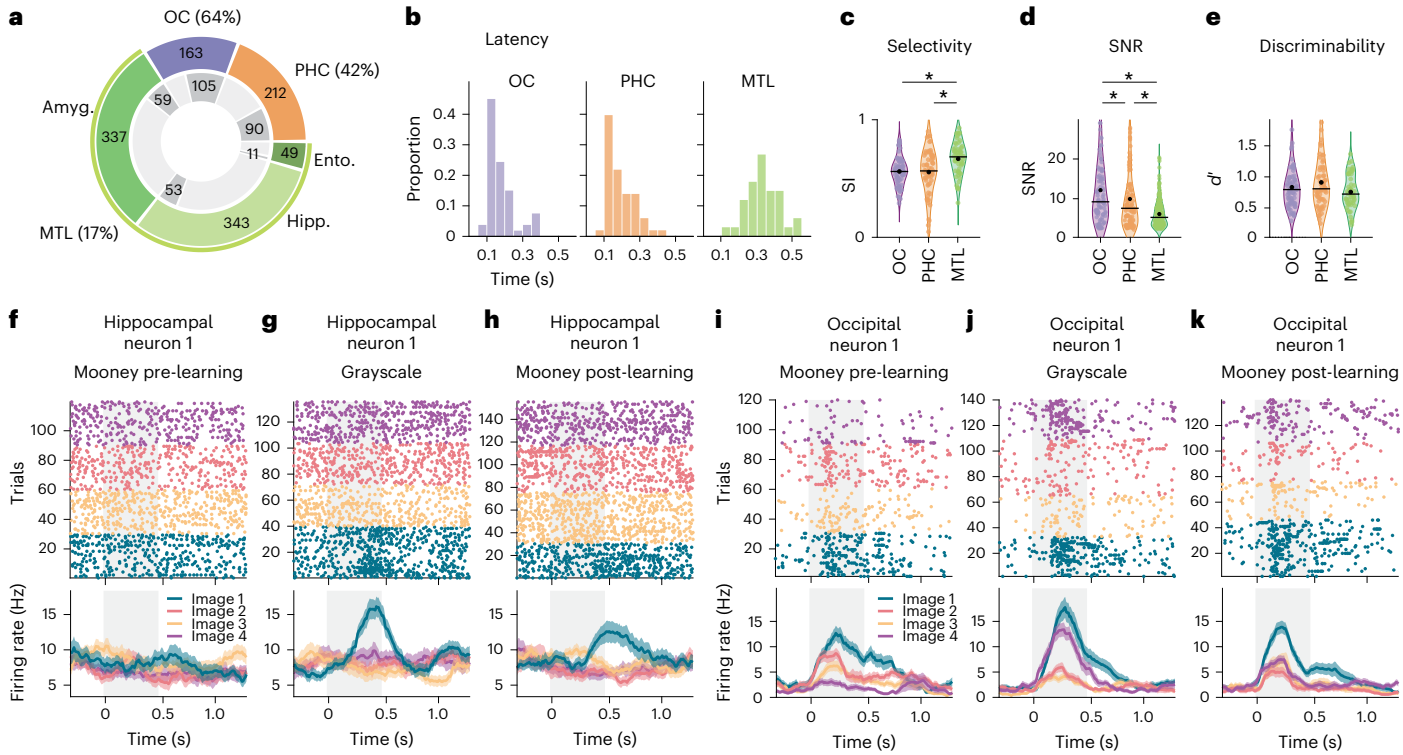


Fig. 2 | Neurons show different response latencies and visual selectivity across brain regions. a, The proportion and count of visually responsive neurons across human brain regions. Neurons in the hippocampus (Hipp.), amygdala (Amyg.) and entorhinal cortex (Ento.) are collectively referred to as the MTL from here on. Visually responsive neurons showed statistically significant modulation of their firing rates during image presentation (grayscale or Mooney) compared with the baseline period (two-sided permutation test, 1,000 iterations, $P < 0.05$, FDR corrected). The ring chart shows the number of visually responsive neurons (dark-gray segments; OC, $n = 105$; PHC, $n = 90$; and MTL, $n = 123$) relative to the total number of recorded neurons (colored segments). Percentages indicate the proportion of responsive neurons within each brain region (reported in aggregate for the MTL). **b**, The distribution of neuronal response latencies across brain regions. Response latency was defined as the time from grayscale stimulus onset to the first significant change in firing rate relative to baseline (Methods). Latencies differed significantly across regions ($P = 8.7 \times 10^{-11}$, Kruskal–Wallis test). **c**, The stimulus SI, computed from responses to grayscale images. **d**, The SNR of neuronal responses across regions. **e**, The stimulus discriminability (d'), computed by comparing responses with preferred versus least-preferred stimuli. Dots in **c**–**e** show the distribution of values across neurons, with kernel density outlines, the horizontal line marking the median and the black dot marking the mean. Asterisks in **c** and **d** indicate statistically significant differences between regions (two-sided Wilcoxon rank-sum test; **c**: OC versus MTL, $P = 1.6 \times 10^{-5}$; PHC versus MTL, $P = 1.2 \times 10^{-3}$; **d**: OC versus MTL, $P = 1.4 \times 10^{-10}$; OC versus PHC, $P = 0.02$; and PHC versus MTL, $P = 6.6 \times 10^{-4}$); the absence of an asterisk in **c**–**e** indicates no significant difference. **f**,**g**,**h**, An example of a highly selective neuron in the hippocampus. Panels show the spiking activity in response to different images (Images 1–4) in the pre-learning Mooney (**f**), grayscale (**g**) and post-learning Mooney (**h**) conditions. This neuron responded only to Image 1 (turquoise) for the grayscale (**g**) and post-learning Mooney (**h**) images, but not in the pre-learning condition (**f**). Raster plots (top) show spiking activity (each dot represents a spike event) across trials (rows) and over time. Colors correspond to different images. Responses are shown for a subset of four images that were presented during the task. Trials are grouped by image for visualization. The light-gray rectangle in the background indicates the image presentation interval. Mean firing rates (bottom) are shown separately for each image as post-stimulus time histograms (solid colored lines; the bin size is 250 ms and the step size is 10 ms). Shaded areas represent s.e.m. across trials. Firing rates for the Mooney images increased after learning (**h**) compared with before learning (**f**). **i**,**j**,**k**, The spiking activity of an example neuron in the OC, shown with the same conventions as in **f**–**h**. This neuron responded to multiple stimuli (Images 1–4) and exhibited firing rate modulation depending on the image and condition.

Neuronal latency and selectivity reflect hierarchical processing across brain regions

We first identified visually responsive neurons. A neuron was considered responsive if its firing rate showed significant modulation to any of the presented images, during a 500-ms interval (matching stimulus duration) from 200 to 700 ms after image onset, compared with the baseline period (−500 to 0 ms before image onset). Significance was determined using a permutation test (1,000 iterations) with a threshold $P < 0.05$, corrected for false discovery rate (FDR). We identified 318 visually responsive neurons (29% of the total). Of these, 105 were located in the OC (64% of the total number of neurons in this area), 90 in the PHC (42%), 53 in the hippocampus (15%), 59 in the amygdala (17%) and 11 in the entorhinal cortex (22%) (Fig. 2a). We will refer to the neurons in the hippocampus, amygdala and entorhinal cortex together as the MTL, comprising 123 visually responsive neurons. Representative examples are shown in Extended Data Fig. 2a–d. For instance, Extended Data Fig. 2a illustrates the responses of a hippocampal neuron to a specific image in its pre-Mooney, grayscale and post-Mooney versions. This neuron exhibited distinct response patterns across conditions, with the highest responses observed in the grayscale condition. Subsequent analyses were based only on visually responsive neurons.

We next estimated the onset latencies for responsive neurons. Response latency was defined as the first significant change in the neuron's firing rate relative to the baseline period after stimulus onset for the grayscale images, taking the earliest latency across the images for each neuron (Methods). Response latencies significantly differed between areas with the fastest responses in the OC (median 155 ms, interquartile range IQR 129–204 ms), followed by the PHC (median 161 ms, IQR 138–248 ms) and the MTL (median 322 ms, IQR 272–392 ms) ($P = 8.7 \times 10^{-11}$, Kruskal–Wallis test) (Fig. 2b).

Next, we characterized the response properties of the recorded neurons. We quantified neuronal selectivity using a selectivity index (SI)[25] (Methods), computed from responses to the grayscale images. Selectivity differed across regions (Kruskal–Wallis, $P = 0.0001$). MTL

https://doi.org/10.1038/s41593-026-02496-6

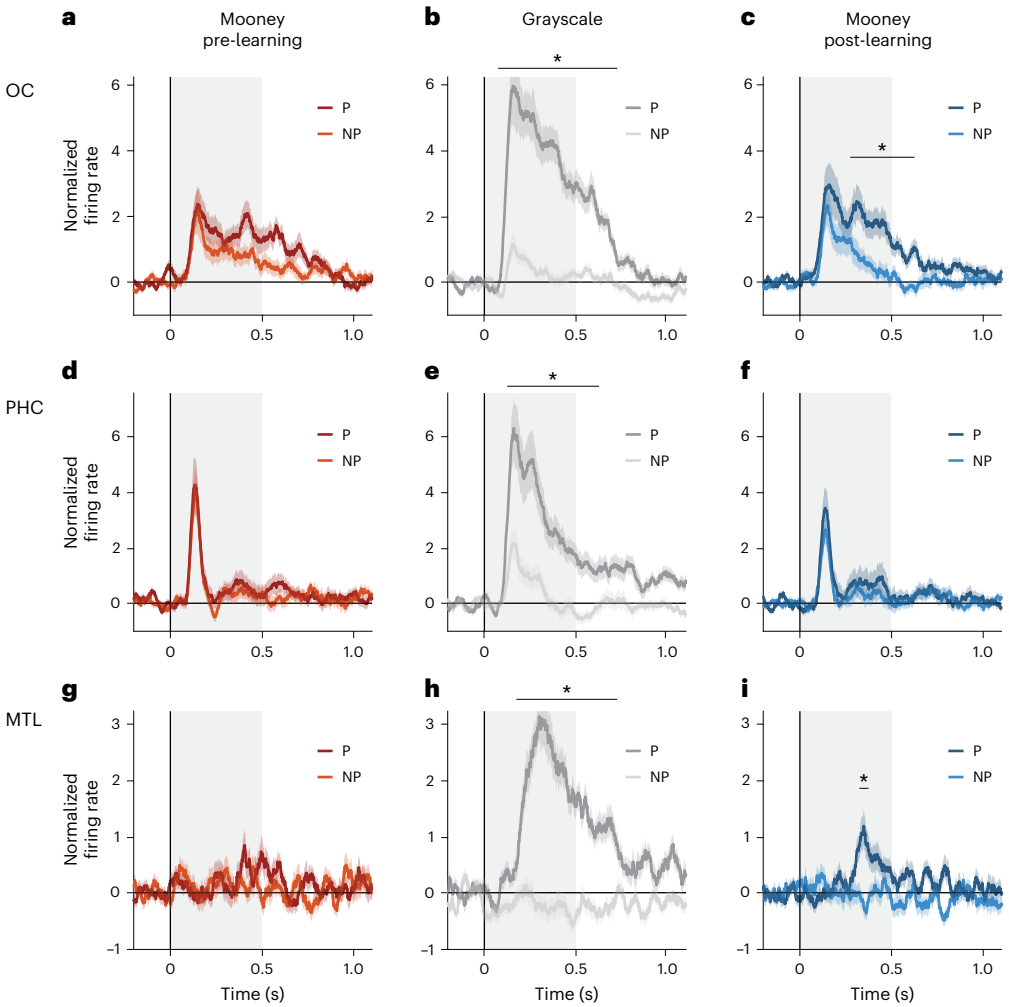


Fig. 3 | Stimulus preference emerged with learning. **a**–**c**, Mean normalized firing rates averaged across neurons in OC for grayscale-defined preferred (P) and nonpreferred (NP) stimuli in the pre-learning Mooney (**a**), grayscale (**b**) and post-learning Mooney (**c**) conditions. **d**–**f**, The same as **a**–**c**, but for neurons in the PHC. **g**–**i**, The same as **a**–**c**, but for neurons in the MTL. At each time point (nonoverlapping 50-ms bins), a two-sided Wilcoxon signed-rank test was performed between preferred and nonpreferred responses, with significant time points ($P < 0.05$, Bonferroni-corrected. Peak significance: **b**, $P = 1.6 \times 10^{-11}$; **c**, $P = 7.7 \times 10^{-6}$; **e**, $P = 9.1 \times 10^{-8}$; **h**, $P = 2.9 \times 10^{-11}$; and **i**, $P = 3.6 \times 10^{-5}$) indicated by a black line and asterisk above the curves. Shaded areas represent s.e.m. across neurons. Time zero marks the image onset, and the shaded gray area (0–0.5 s) indicates the stimulus presentation period.

neurons exhibited the highest selectivity compared with OC and PHC (SI 0.67) (Fig. 2c). PHC and OC neurons showed similar selectivity indices (SI 0.56). We next quantified the signal-to-noise ratio (SNR) of neuronal responses (Fig. 2d), which differed across regions (Kruskal–Wallis, $P = 2.6 \times 10^{-5}$) with OC neurons showing the highest SNR (8.9), followed by PHC (7.3) and MTL (4.3). Finally, we assessed the ability of neurons to discriminate between stimuli by computing a *d*-prime (*d*′) index comparing responses to preferred versus least-preferred stimuli (Fig. 2e; Methods). Despite differences in SNR, *d*′ discriminability was comparable across regions (Kruskal–Wallis, $P = 0.21$). SNR and *d*′ patterns remained consistent across time (Supplementary Fig. 3).

Figure 2f–k shows the spiking activity (top, raster plots; bottom, mean firing rates) of two example neurons in response to different images (color-coded, Images 1–4) (Supplementary Fig. 4) and across conditions. Figure 2f–h illustrates a highly selective neuron in the hippocampus. In the grayscale condition, this neuron responded only to Image 1 (Fig. 2g, turquoise). A second example, Fig. 2i–k, shows the spiking activity of a neuron in the OC that responded to multiple images. This neuron exhibited firing rate modulation depending on the image and condition. For instance, responses to Images 1 (turquoise) and 4 (purple) were higher in the grayscale condition (Fig. 2j) compared with the Mooney conditions (Fig. 2i,k). However, for Image 3 (yellow), firing rates were slightly lower in the grayscale condition compared with the Mooney conditions. Additional example neurons are shown in Extended Data Figs. 2 and 3. Overall, these results are consistent with a hierarchical organization in visual processing, where both selectivity and latencies increase as neural signals flow from the occipital region to the MTL.

Learning enhances neuronal stimulus selectivity

Modulation in the responses to grayscale images compared with their Mooney counterparts (Fig. 2g versus Fig. 2f,h; Fig. 2j versus Fig. 2i,k; Extended Data Fig. 3b versus Extended Data Fig. 3a,c; and Extended Data Fig. 3e versus Extended Data Fig. 3d,f) could be expected and may reflect the large differences between the two types of images at the pixel level. Remarkably, firing rates were also modulated between the identical post- and pre-learning Mooney images. For instance, the hippocampal neuron in Fig. 2 remained unresponsive to all Mooney images before learning (Fig. 2f), but its firing rate was selectively modulated in response to Mooney Image 1 after learning (Fig. 2h). See Extended Data Fig. 2 for additional examples showing a direct comparison between the responses to specific images across conditions and Extended Data Fig. 3 for further examples in the same format as Fig. 2.

We examined whether our population of neurons exhibited enhanced stimulus selectivity after learning. Specifically, we tested if

https://doi.org/10.1038/s41593-026-02496-6

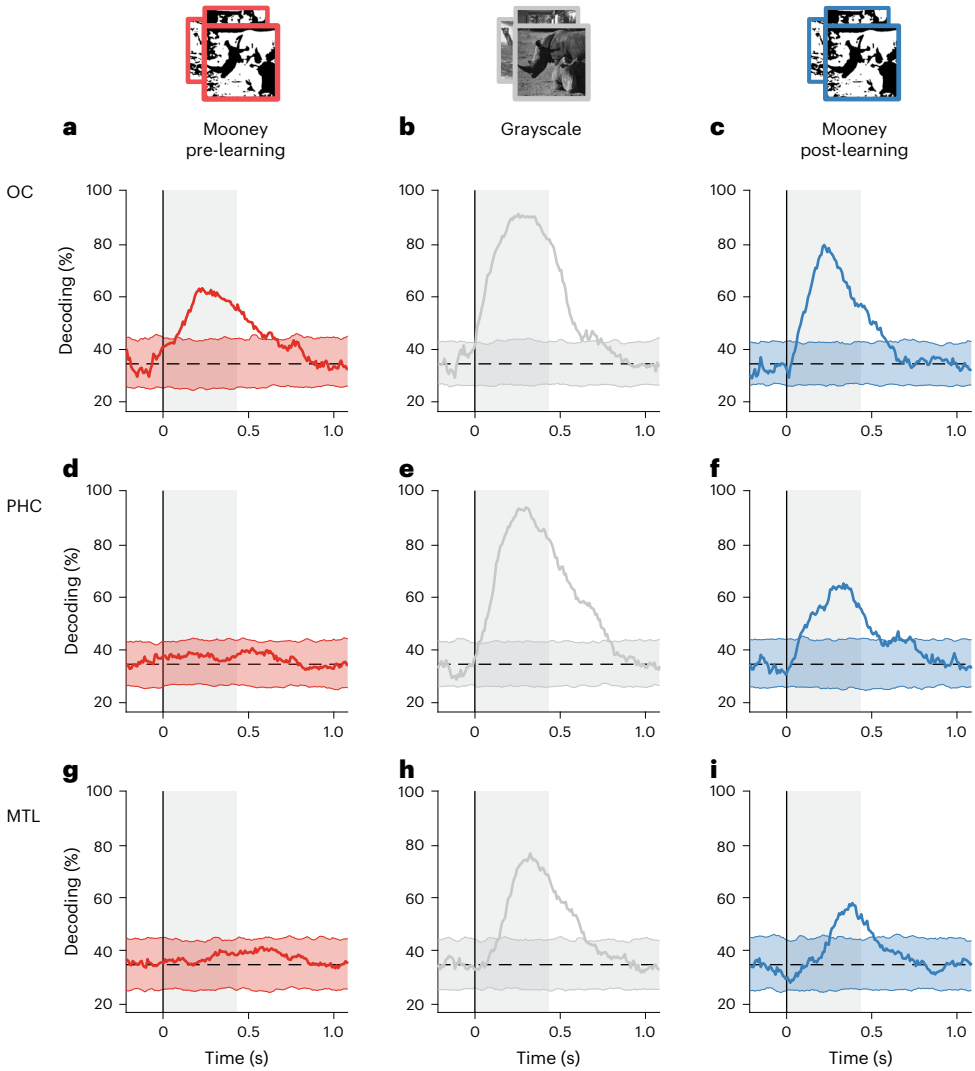


Fig. 4 | Learning enhances the neuronal discriminability of Mooney images. **a**–**c**, Decoding accuracy over time for image identity in the pre-learning Mooney (**a**), grayscale (**b**) and post-learning Mooney (**c**) conditions for OC neurons. **d**–**f**, The same as in **a**–**c**, but for PHC neurons. **g**–**i**, The same as in **a**–**c**, but for MTL neurons. Decoding was performed using pseudopopulations of neurons pooled across participants, with image identity predicted using an SVM with a linear kernel (bin size 250 ms, step size 25 ms). Three-class decoders were constructed using the three images that elicited the most neuronal responses per session (chance 33%; Methods). The horizontal dashed line indicates chance level. The shaded area represents the 95% confidence interval centered on the mean of the empirical null distribution obtained by shuffling the labels (500 iterations). Time zero marks the image onset, and the shaded gray area (0–0.5 s) indicates the stimulus presentation period.

stimulus preferences defined by grayscale image responses transferred to Mooney images following learning. For each neuron, we defined the most and least preferred images based solely on responses in the grayscale condition. We then computed the average normalized firing rates across neurons for the preferred versus nonpreferred images in the pre- and post-learning Mooney conditions (Fig. 3).

As expected for the grayscale condition, responses were stronger for the preferred versus the nonpreferred stimulus in all three regions (Fig. 3b,e,h). For Mooney images, OC neurons showed little selectivity before learning (Fig. 3a). After learning, however, responses to preferred images were significantly greater than responses to nonpreferred images ($P < 0.05$, two-sided Wilcoxon signed-rank test, corrected; Fig. 3c). This selectivity emerged during stimulus presentation, with significant differences arising at ~300–350 ms after image onset, markedly later than for the grayscale images (~100–150 ms; Fig. 3b). Similarly, neurons in the MTL did not encode stimulus preference before learning (Fig. 3g), but selectivity emerged after learning at ~350–400 ms (Fig. 3i; $P < 0.05$, corrected), again later than for the grayscale condition (~250 ms). The PHC did not show a reliable increase in preference selectivity either before or after learning (Fig. 3d,f).

Overall, these results indicate that stimulus preference emerges with learning in neurons of the OC and MTL. After learning, Mooney images evoke response patterns consistent with neurons' grayscale-defined preferences, demonstrating experience-dependent refinement of neuronal selectivity.

Neuronal representations of Mooney images become more discriminable after learning

To further investigate how learning shapes neuronal representations, we asked whether learning enhances overall discriminability of activity patterns for Mooney images. Both pre- and post-learning conditions included the same stimuli, providing identical visual input. However, only post-learning Mooney images could be readily recognized (Fig. 1d–f). We hypothesized that this perceptual difference would sharpen neuronal representations at the population level.

To test this, we trained linear decoders to discriminate between response patterns for different image identities separately for each

https://doi.org/10.1038/s41593-026-02496-6

condition. Decoding was assessed using pseudopopulations constructed across participants[26,27] (Methods). For each region, we pooled responsive neurons and used their responses as features to predict image identity across time (support vector machine (SVM) with a linear kernel; bin size 250 ms, step size 25 ms; Methods). Because different sessions used distinct subsets of images, we selected the top three images from each session on the basis of responses in the grayscale condition. Images were ranked by the number of neurons that responded to them, and the three highest-ranked images were labeled Images 1–3 to construct three-class decoders (Methods). All three regions showed the highest image discriminability in the grayscale condition (OC 90%, PHC 93%, MTL 76%; $P < 0.05$, permutation test versus empirical null distribution; chance 33%) (Fig. 4b,e,h), but performance diverged for the Mooney conditions. Neurons in OC significantly decoded Mooney images both before and after learning, with decoding accuracy increasing after learning (pre-Mooney 62%, post-Mooney 80%; $P < 0.05$, permutation test) (Fig. 4a,c). By contrast, neurons in PHC and MTL did not show significant decoding before learning (Fig. 4d,g) but did so after learning (PHC 63%, MTL 57%) (Fig. 4f,i). In all cases, reliable decoding emerged during image presentation and briefly persisted after stimulus offset before returning to chance level. Together, these findings indicate that learning reshapes neuronal responses to identical stimuli, enhancing the discriminability of Mooney representations. Complementary analyses corroborated these learning-induced changes, showing reliable decoding of pre- versus post-learning Mooney responses at both the single-neuron and population levels ($P < 0.05$, permutation test) (Extended Data Fig. 4a,d–f). Moreover, discrimination between grayscale and pre-learning Mooney responses was stronger than between grayscale and post-learning Mooney responses ($P < 0.05$; Extended Data Fig. 4b,c,g–i and Supplementary Fig. 5), suggesting that learning makes neuronal responses to Mooney images more similar to those evoked by grayscale images.

We conducted additional control analyses to assess the potential impact of time-dependent changes in neuronal excitability[28–32]. We assessed whether neurons systematically changed their firing rate over time (Supplementary Fig. 6a; Methods). Only a small proportion of neurons (~3%, permutation test) exhibited significant slope trends. To determine whether these trends were specific to the responsive population, we compared the distributions of slopes between responsive and nonresponsive neurons (Supplementary Fig. 6b). Nonresponsive neurons showed a comparable distribution of firing rate trends to responsive neurons ($P = 0.1$, Kolmogorov–Smirnov test), with ~3% exhibiting significant slopes, consistent with the proportion observed in the responsive population. Because the nonresponsive population shared similar trend distributions but lacked stimulus selectivity, we repeated the decoding analyses using these neurons to assess whether such time-dependent trends could account for the observed effects. Decoding on the basis of nonresponsive neurons did not exceed chance in any region (Supplementary Figs. 7 and 8), suggesting that firing rate drifts do not account for our findings. We next performed two additional conservative control analyses focusing on neurons with the most extreme firing rate trends. First, we repeated the decoding analyses using the responsive neuron population after removing neurons with the largest firing rate trends. Specifically, we excluded neurons with slopes deviating by more than 1 s.d. from zero (trend-controlled population; Supplementary Fig. 6c). Our findings remained robust after excluding these neurons (Supplementary Fig. 9), suggesting that changes in overall excitability do not account for the observed effects. We then constructed a control population by selecting nonresponsive neurons with the largest firing-rate trends, with slopes deviating by more than 1 s.d. from zero (trend-dominated population; Supplementary Fig. 6d), and repeated the decoding analyses. In this trend-dominated population, decoding did not yield significant image discriminability in any region (Supplementary Fig. 10), indicating that firing-rate trends alone were insufficient to produce reliable decoding.

Overall, these results show that learning enhanced neuronal representations, making Mooney images more discriminable across occipital and medial temporal regions, with effects emerging during stimulus presentation.

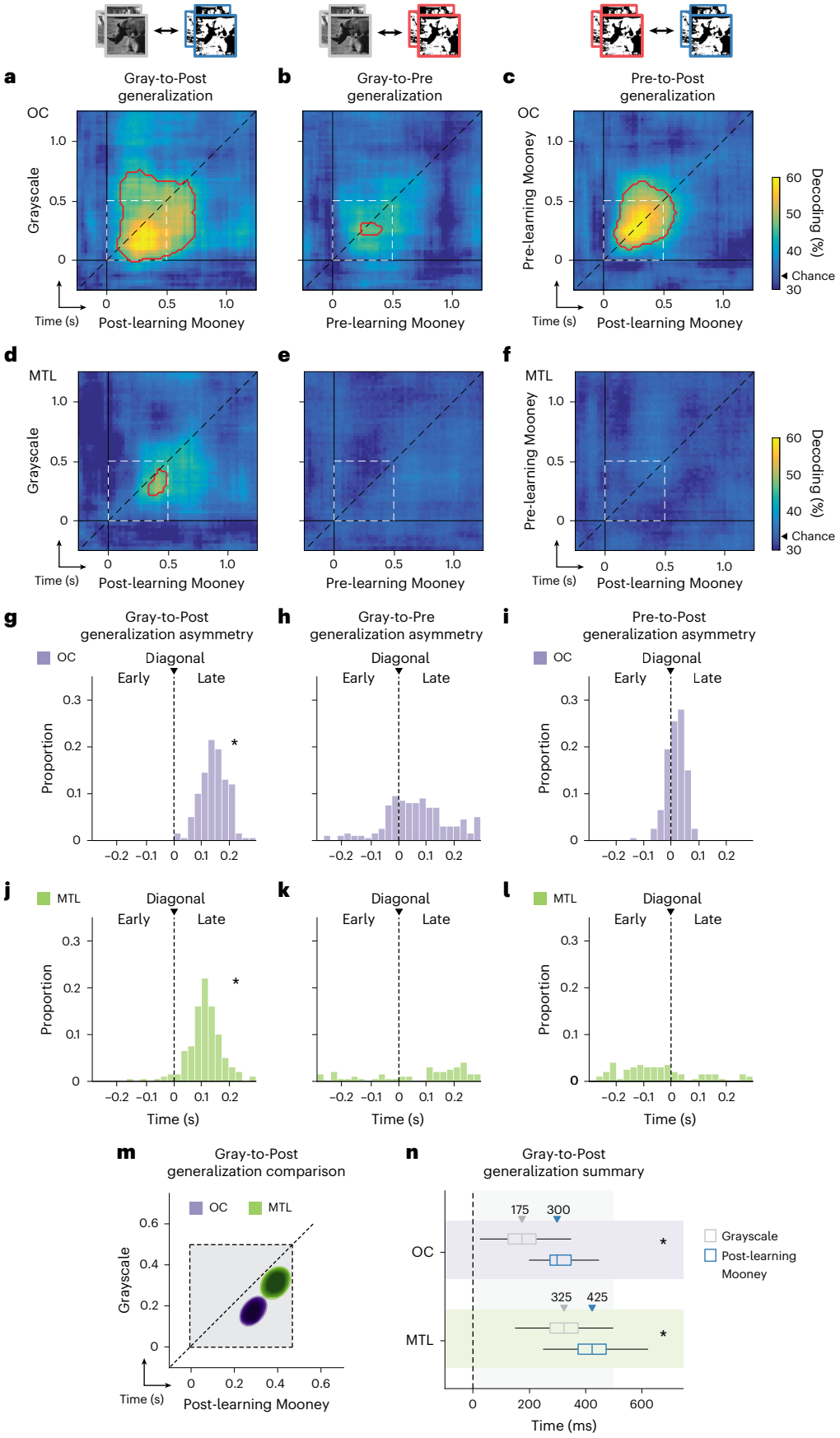
Neuronal patterns for grayscale images generalize to Mooney images after learning

To further understand the content encoded by neuronal populations before versus after learning, we assessed whether activity patterns elicited by grayscale images generalized to two-tone Mooney images.

Fig. 5 | Generalization to learned Mooney images is delayed in OC and MTL. a–**c**, Generalization matrices in OC for: grayscale to post-learning Mooney (Gray-to-Post) (**a**), grayscale to pre-learning Mooney (Gray-to-Pre) (**b**) and pre-learning to post-learning Mooney (Pre-to-Post) (**c**). Generalization was assessed by training a decoder with one condition and testing with the other, including the responses of neurons as features. Each matrix entry *x*,*y* indicates the mean decoding accuracy of image identity for the two train–test configurations across conditions as follows: first, we computed decoding accuracy by training with the grayscale at time *y* and testing with the Mooney condition at time *x*; second, we computed decoding accuracy by training with the Mooney at time *x* and testing with the grayscale condition at time *y*; third, the mean of these two values corresponds to entry *x*,*y*. Accuracies correspond to three-class decoders (chance of 33%) of the top three images chosen on the basis of their responses in the grayscale condition (Methods). The *x* and *y* axes indicate time points of neuronal activity for the corresponding conditions. Time zero marks image onset. The small pictures above the matrices indicate the conditions included in the generalization matrix (gray frame for grayscale, blue frame for post-learning Mooney and red frame for pre-learning Mooney). The dashed black line represents the matrix diagonal, where conditions were trained and tested at the same time points. The dashed line square represents the image presentation interval. Solid red lines delineate statistically significant decoding, assessed with a one-sided permutation test against an empirical null distribution generated by shuffling condition labels (200 iterations, cluster-corrected; $P < 0.05$). **d**–**f**, Generalization matrices in MTL. Panels and data follow the same conventions as in **a**–**c**. **g**–**i**, Generalization asymmetry corresponding to decoding profiles in **a**, **b** and **c**, respectively. We quantified the deviation of peak decoding from the diagonal to measure generalization asymmetry. Deviation of the distribution to the right indicates that neuronal patterns for grayscale images generalize to later activity for the post-learning Mooney images. Asterisks indicate a statistically significant deviation of the decoding peak from the diagonal, assessed with a resampling-based test (200 iterations; Methods). Deviation was considered significant if at least 95% of the values fell on the same side of zero as the observed deviation (OC: $P < 0.005$, 125-ms deviation in **g**). The absence of asterisks indicates no statistically significant deviation from the diagonal. **j**–**l**, Generalization asymmetry in MTL corresponding to the decoding profiles in **d**, **e** and **f**, respectively. Data use the same conventions as in **g**–**i** (MTL: $P = 0.04$, 100-ms deviation in **j**). Results for the OC are color-coded in purple, and MTL in green. **m**, Comparison of Gray-to-Post generalization in the OC and MTL. The ellipses illustrate generalization asymmetry in the OC (purple) and MTL (green), corresponding to **a** and **d**, respectively. In both cases, peak generalization falls below the diagonal, occurring earlier in the OC than in the MTL. The ellipses span the interquartile range of the probability density functions of peak generalization (Extended Data Fig. 6a,d). The shaded gray area with dashed lines delineates the stimulus presentation period. **n**, A summary of generalization times for Gray-to-Post. Downward arrows and numbers indicate the median peak generalization times for grayscale (gray box) and post-learning Mooney (blue box) images in the OC (top, purple background) and MTL (bottom, green background). Box plots show the median (center line), 25th and 75th percentiles (box bounds), and whiskers extending to the minimum and maximum values of the marginals of the probability density functions (Extended Data Fig. 6a,d). Statistical significance was assessed with the same resampling-based test as in **g** and **j** (200 iterations; OC: $P < 0.005$, MTL: $P = 0.04$). The vertical dashed line at time zero indicates image presentation onset, and the shaded gray area in the background represents the presentation period.

https://doi.org/10.1038/s41593-026-02496-6

If neurons changed their firing patterns after learning to reflect image identity, then patterns encoding grayscale images would be expected to generalize to Mooney images after learning, but not (or to a lesser extent) before learning. To test this hypothesis, we employed a cross-condition decoding approach in which classifiers were trained on one condition and tested on another across time (Fig. 5a–f). As in the previous within-condition decoding analyses, we selected the top three images from each session on the basis of grayscale responses to



https://doi.org/10.1038/s41593-026-02496-6

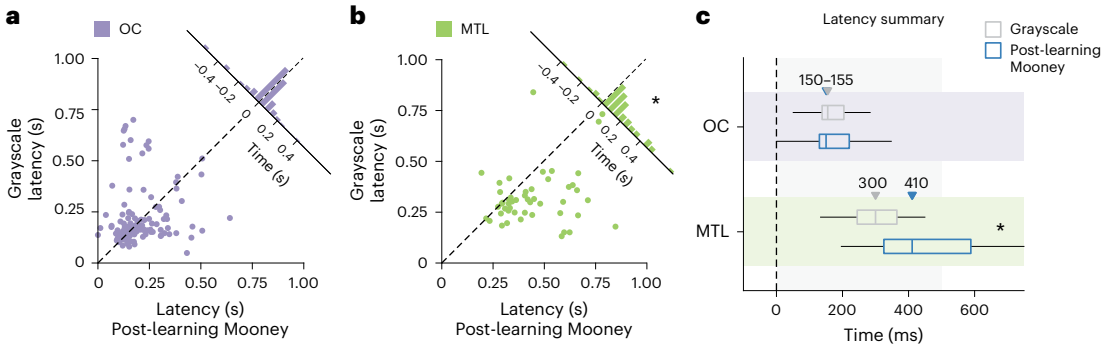


Fig. 6 | Neurons in MTL respond later to learned Mooney images compared with grayscale images. a,**b**, Scatter plots of neuronal response latencies for grayscale (*y* axis) versus post-learning Mooney images (*x* axis) in the OC (**a**) and MTL (**b**). Each dot represents the response latency of a neuron to an individual image. The tilted distribution represents the differences between latencies for grayscale and post-learning Mooney images. The distribution centered at zero indicates no latency difference between conditions. The asterisk in **b** denotes significant latency differences ($P = 0.0001$, two-sided Wilcoxon signed-rank test) between grayscale and post-learning Mooney images in MTL. Thus, dots largely falling below the diagonal (dashed line) indicate that neuronal responses to post-learning Mooney images are delayed with respect to responses to grayscale images. **c**, A summary of neuronal response latencies. Downward arrows and accompanying numbers indicate the median latency for grayscale (gray box) and post-learning Mooney (blue box) images in the OC (top, purple background) and MTL (bottom, green background). In the OC, latencies did not differ between the two conditions (150 ms for post-Mooney and 155 ms for grayscale), and arrows overlapped. Box plots show the median (center line), 25th and 75th percentiles (box bounds), and whiskers extending to the minimum and maximum values of the latency distributions in **a** and **b**. The asterisk denotes significant latency differences ($P = 0.0001$, two-sided Wilcoxon signed-rank test). The vertical dashed line at time zero indicates image presentation onset, and the shaded gray area in the background represents the presentation period.

construct three-class decoders (Methods; chance 33%). At each time point and for every pair of conditions, decoders were used to predict image identity from neuronal activity patterns.

We examined generalization of neuronal responses for grayscale to Mooney images ('Gray-to-Post' and 'Gray-to-Pre') separately for each region. Figure 5a shows Gray-to-Post generalization for neurons in the OC. Each entry at time *x*,*y* indicates the mean decoding accuracy for two train–test configurations: first, the classifier was trained with the responses to grayscale images at time *y* and tested with the responses to post-learning Mooney images at time *x*; second, the classifier was trained with the responses to post-learning Mooney images at time *x* and tested with the responses to grayscale images at time *y*. OC neurons showed significant decoding for Gray-to-Post ($P < 0.05$, versus empirical null distribution; chance 33%; Fig. 5a) and Gray-to-Pre (Fig. 5b) generalization. Decoding accuracies were higher for Gray-to-Post (up to 62%) than for Gray-to-Pre (49%), and remained significant for ~200 ms after image offset for the post-learning Mooney images. We next evaluated generalization in MTL neurons, which showed significant transfer of image identity information only after learning, with decoding accuracy reaching 50% for Gray-to-Post ($P < 0.05$; Fig. 5d), and chance-level performance for Gray-to-Pre (Fig. 5e).

In addition, we asked whether neural responses would generalize from pre- to post-learning Mooney images, given that the visual input was identical despite the change in perception, from unrecognized to recognized. We found that pre-to-post generalization was significant in OC neurons (up to 65%) (Fig. 5c), with decoding mostly confined to the image presentation interval. By contrast, MTL neurons showed no significant pre-to-post generalization, with decoding remaining at chance level (Fig. 5f). Generalization analyses in the PHC revealed no significant decoding across conditions (Extended Data Fig. 5).

As with the within-condition decoding analyses, we conducted control analyses to assess the potential influence of firing-rate trends on the generalization results. Repeating the generalization analyses using the nonresponsive neuron population yielded no generalization in any region (Supplementary Fig. 11). Results remained robust when analyses were restricted to the trend-controlled population (Supplementary Fig. 12). Finally, decoding in the trend-dominated population showed no evidence of generalization in any region (Supplementary Fig. 13), confirming that firing-rate trends alone cannot account for the observed generalization effects.

Consistent with the within-condition decoding results, neuronal response patterns evoked by grayscale images generalized to Mooney images predominantly after learning. This generalization emerged during stimulus presentation and was observed across OC and MTL, further suggesting that learning reshapes neuronal representations such that Mooney responses converge toward those evoked by grayscale images.

Generalization of neuronal responses for grayscale to learned Mooney images is delayed in OC and MTL

We further analyzed the temporal dynamics of the cross-condition generalization matrices (Fig. 5a–f). Entries along the diagonal reflect generalization from grayscale images to post-learning Mooney images at matched time points, while entries above or below the diagonal indicate cross-time generalization.

In the OC, the Gray-to-Post generalization matrix exhibited an asymmetric profile with a prominent off-diagonal peak (Fig. 5a). We quantified this asymmetry by estimating the probability density function of peak decoding times via resampling (Methods). The decoding peak deviated 125 ms from the diagonal ($P < 0.005$, empirical distribution versus the diagonal) (Fig. 5g), indicating that neuronal patterns in response to grayscale images at 175 ms most closely resembled those for post-learning Mooney images at 300 ms (Extended Data Fig. 6a). To evaluate whether generalization favored early or late neuronal responses, we compared decoding accuracies above and below the diagonal. Decoding accuracy was significantly higher at later compared with earlier time points in the Gray-to-Post case (Supplementary Fig. 14a,d). By contrast, the generalization profiles for Gray-to-Pre and Pre-to-Post were centered on the diagonal and no difference was observed between early and late decoding stages (Fig. 5h,i, Extended Data Fig. 6b,c and Supplementary Fig. 14b,c,e,f).

As in the occipital region, the Gray-to-Post generalization profile in MTL was also asymmetric, with higher accuracies falling below the diagonal and the decoding peak 100 ms off-diagonal (Fig. 5j). Thus, neuronal patterns in response to grayscale images at 325 ms were most similar to those for learned Mooney images at 425 ms (Extended Data Fig. 6d), and generalization was again higher at later stages (Supplementary Fig. 15). Gray-to-Pre and Pre-to-Post generalization profiles in the MTL showed no significant asymmetry with respect to the diagonal (Fig. 5k,l and Extended Data Fig. 6e,f).

Together, these findings uncovered key neuronal dynamics induced by learning. Consistent with the behavioral findings, grayscale-to-Mooney generalization emerged only after learning. The more pronounced below-diagonal decoding indicates that neuronal patterns for grayscale images predominantly generalize to post-learning Mooney images at a late processing stage. Thus, with additional processing time, responses to learned Mooney images evolve to resemble those elicited by grayscale images. For pre-to-post Mooney generalization, where images are identical, the observed on-diagonal decoding in OC might reflect shared low-level visual representations.

Finally, we summarized the timing of Gray-to-Post generalization in OC and MTL (Fig. 5m,n). Consistent with hierarchical processing, peak generalization occurred earlier in OC (175 ms and 300 ms, in grayscale and post-learning Mooney conditions, respectively) than in MTL (325 ms and 425 ms). The generalization delays were comparable across regions (125 ms in OC and 100 ms in MTL), suggesting that the temporal lag observed in OC may propagate to MTL, with MTL not contributing additional delays.

Neurons in MTL respond later to learned Mooney images compared with grayscale images

Both OC and MTL neuronal populations showed delayed generalization from grayscale to Mooney images after learning. To determine whether this generalization delay was reflected in the timing of individual neuronal responses, we computed response latencies to grayscale and post-learning Mooney images in OC and MTL (Fig. 6a,b). In OC, response latencies did not differ between conditions (Fig. 6a,c). By contrast, MTL neurons responded significantly later to post-learning Mooney images than to grayscale images (time difference of 110 ms; $P = 0.0001$, two-sided Wilcoxon signed-rank test; Fig. 6b,c). These results suggest that MTL neurons fire only after earlier stages have devoted additional processing time to resolving the identity of the learned Mooney images.

Discussion

We investigated how the neuronal dynamics induced by rapid learning in the human brain reshape perception using degraded two-tone Mooney images[11]. Mooney images are unrecognizable before learning. However, after brief exposure to the corresponding grayscale images, recognition of Mooney images increases dramatically[11,12,14,16,21]. Participants in our study showed essentially all-or-none recognition of Mooney images mediated by rapid learning. In addition, recognizing Mooney images after learning required additional processing time compared with grayscale images.

Medial occipital neurons showed lower selectivity and faster responses, while MTL neurons showed higher selectivity and slower responses (Fig. 2), reflecting hierarchical processing from the ventral visual pathway to downstream areas in the MTL[33–36]. The time required for visual signals to reach the human MTL was longer compared with other primates[37–39]. Rapid learning led to changes in firing patterns in neurons in medial occipital and temporal regions. Despite identical visual input, neurons in these areas showed increased stimulus selectivity, accompanied by enhanced discriminability of Mooney images after learning (Figs. 2–4). Previous studies in monkeys showed that neurons in the inferior temporal cortex, known for their role in representing objects and faces, change their activity in response to Mooney images after exposure to their intact versions[20], following learned associations[40–42], and while building size-invariant object representations[43]. Although we did not have neurons in inferior temporal cortex areas, the occipital and medial temporal regions examined here—which are generally less selective for complex shapes such as faces and objects than the inferior temporal cortex—exhibited clear experience-dependent modulation, suggesting that rapid learning extends beyond object-selective areas to engage broader visual circuits. While we did not measure receptive fields, prior receptive field mapping studies in humans indicate that neurons in ventromedial and dorsomedial OC at the recorded locations typically have relatively large receptive fields (>4°)[28,44,45], comparable to the image size used here (5°). Moreover, human imaging[46] and monkey invasive studies[47,48] suggest that the ventral visual regions sampled here process mid-level visual features such as curvature, shape and color relevant for object recognition. In humans, neuroimaging studies have revealed learning-related effects on the representation of Mooney face and object images in high-level visual cortex, including the occipital and fusiform areas, as well as in the precuneus and adjacent parietal regions implicated in top-down attention[21,24,49]. Moreover, single-shot associative memories of unrelated concepts are represented by neurons in the human MTL[50]. Linking the work in monkeys and humans, our results demonstrate that the effects of recently acquired experience are reflected in neuronal activity in both visual cortex and MTL regions.

Following learning, neuronal responses to Mooney images became more similar to those evoked by their grayscale counterparts (Figs. 2, 3 and 5). Consistently, neuroimaging studies suggested that the representation of Mooney images is more similar to the representation of intact images after learning[22,23]. We observed generalization in both OC and MTL, although it was weaker in MTL, which may partly be explained by the lower SNR of neuronal responses in MTL compared with OC. Decoding Mooney images after learning required additional processing time compared with the processing of intact grayscale images (Fig. 5), consistent with behavioral results showing slower response times for post-learning Mooney images (Fig. 1g). The generalization delays for degraded Mooney images (~125 ms) are reminiscent of those previously reported in human visual areas (110–160 ms) when participants had to recognize highly occluded objects compared with whole objects[27]. Furthermore, studies in monkeys have shown that objects that are harder to recognize take longer to be reliably decoded from neurons in the primate ventral stream[51]. These delays may indicate the need for recurrent processing, involving signals from other visual areas as well as top-down feedback from areas in the frontal and temporal lobes[52–55]. Unlike these studies, our work used learning to facilitate recognition. Learning transformed an essentially impossible recognition task (pre-Mooney) into a challenging but solvable task (post-Mooney), enabling the brain to reconstruct visual objects from impoverished images. In contrast to learning new images—which may require neocortical neurons to expand their tuning to encode novel objects[56]—recognizing Mooney images may rely on learning figure-ground segregation and image completion processes. The effects of learning in our study may manifest through extensive recurrent processing to solve object identity in the human brain, which might involve top-down feedback from higher-order cortical areas[19,57,58].

Our analyses reveal that key neuronal dynamics that support recognition after learning occur faster in OC than in MTL, highlighting their distinct roles in learning-induced visual recognition across the processing hierarchy. The generalization delay (~100 ms) and increased response latency (~110 ms) in MTL are consistent with the delayed emergence of generalizable patterns in OC, which appear ~125 ms later for Mooney images than for grayscale images. A plausible interpretation is that the delayed generalization in MTL may reflect the relatively late arrival of visual information from earlier neocortical processing stages, including OC. Although direct causality is not established in the current work, these findings are consistent with a model in which the additional processing required in OC—and potentially other neocortical regions—after learning constrains the temporal dynamics of MTL responses.

Overall, learning-related changes at the neuronal level emerged during image presentation (Figs. 3c,i, 4c,i and 5a,d). The brief stimulus duration and small image size likely minimized saccades and allowed the entire stimulus to be perceived within a single glance. In OC, Pre-to-Post generalization (Fig. 5c) and above-chance image discriminability before learning (Fig. 4a) suggest that Mooney image processing engages shared visual feature representations independent of recognition. However, once the Mooney images are learned, neuronal representations evolve to integrate the visual input with newly

acquired knowledge, enabling decoding of object identity in OC within ~300 ms (Fig. 5a). By contrast, chance-level Pre-to-Post generalization in MTL (Fig. 5f) suggests limited or no involvement in the processing of low-level visual features of the images. Instead, MTL appears to encode abstract representations related to image identity, resolving object identity of learned Mooney images within ~425 ms (Fig. 5d). Decoding persisting after stimulus offset (Figs. 4 and 5, and Extended Data Fig. 4) may partially reflect short-term visual memory of recently perceived stimuli, as well as noncontent-specific post-recognition processes, including attentional and decisional mechanisms[23].

Experience-driven synaptic changes in neuronal circuits are fundamental to learning[59] and engage multiple processes and brain regions[60]. Rapidly storing memory traces is often attributed to MTL structures[18,34,50,61], which facilitate memory reinstatement in neocortical regions via feedback[18,40]. However, whether all forms of rapid learning require the participation of the MTL remains unclear. The behavioral effects observed here are reminiscent of the phenomenon of priming, in which perceptual identification of objects can be facilitated by single exposure to related stimuli[17]. Yet, whereas classical priming is typically fast and short-lived, the rapid learning of Mooney images persists for weeks or longer[12]. By contrast, perceptual learning is long-lasting, but gradual and generally involves training[60]. Recent work suggests that priming and perceptual learning may co-exist to support perceptual improvements while remaining distinct mechanisms[62]. A recent study[14] argued that rapid one-shot perceptual learning may be conceptualized as a form of priming with long-term effects[63]. This study[14] found that memory-impaired patients with lesions in the MTL performed similarly to healthy control participants on a rapid one-trial perceptual learning task, suggesting cortical changes may occur independently of the MTL. Although the perceptual effects persisted for several days, the patients' memory of the test was impaired, arguing against the involvement of associative memory[13,14]. The spatiotemporal dynamics presented here, with generalizable representations emerging earlier in OC and delayed neuronal responses in MTL, suggest that recognition after rapid perceptual learning may rely on neocortical processing without requiring feedback from MTL.

The last decade has seen remarkable progress in the development of successful artificial intelligence algorithms. Such algorithms are typically trained on large datasets in a supervised fashion (however, see examples demonstrating few-shot and zero-shot performance capabilities[64,65]). Humans can use few-shot and unsupervised cues to rapidly learn, as strikingly demonstrated by recognition of impoverished two-tone Mooney images. Our single-neuron data characterize how learning-induced spatiotemporal dynamics encode new visual information reshaping perception. Unlike other forms of fast learning that require recruiting the MTL, rapid perceptual changes may primarily rely on neocortical processing. The current study provides initial steps toward understanding the dynamic interplay between visual cortex and MTL memory structures to instantiate rapid and unsupervised learning and provides constraints for developing computational models that include recurrent computations for rapid learning.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-026-02496-6.

References

1. Fukushima, K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.* **36**, 193–202 (1980).
2. Wallis, G. & Rolls, E. T. Invariant face and object recognition in the visual system. *Prog. Neurobiol.* **51**, 167–194 (1997).
3. Riesenhuber, M. & Poggio, T. Hierarchical models of object recognition in cortex. *Nat. Neurosci.* **2**, 1019–1025 (1999).
4. Serre, T. Deep learning: the good, the bad, and the ugly. *Annu. Rev. Vis. Sci.* **5**, 399–426 (2019).
5. Rolls, E. T. Learning invariant object and spatial view representations in the brain using slow unsupervised learning. *Front. Comput. Neurosci.* **15**, 686239 (2021).
6. Rolls, E. T. *Brain Computations and Connectivity* (Oxford Univ. Press, 2023).
7. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems* Vol. 25 (eds Pereira, F., Burges, C. J., Bottou, L. & Weinberger, K. Q.) 1097–1105 (Curran Associates, Inc., 2012).
8. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
9. Szegedy, C., Ioffe, S., Vanhoucke, V. & Alemi, A. Inception-v4, Inception-ResNet and the impact of residual connections on learning. *In Proc. Conf. AAAI Artif. Intell.* **31**, 4278–4284 (2017).
10. Srivastava, S. & Sharma, G. OmniVec: learning robust representations with cross modal sharing. In *Proc. IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* 1236–1248 (IEEE, 2024).
11. Mooney, C. M. Age in the development of closure ability in children. *Can. J. Psychol.* **11**, 219–226 (1957).
12. Ludmer, R., Dudai, Y. & Rubin, N. Uncovering camouflage: amygdala activation predicts long-term memory of induced perceptual insight. *Neuron* **69**, 1002–1014 (2011).
13. Chang, R., Baria, A. T., Flounders, M. W. & He, B. J. Unconsciously elicited perceptual prior. *Neurosci. Conscious.* **2016**, niw008 (2016).
14. Squire, L. R., Frascino, J. C., Rivera, C. S., Heyworth, N. C. & He, B. J. One-trial perceptual learning in the absence of conscious remembering and independent of the medial temporal lobe. *Proc. Natl Acad. Sci. USA* **118**, e2104072118 (2021).
15. Albright, T. D. On the perception of probable things: neural substrates of associative memory, imagery, and perception. *Neuron* **74**, 227–245 (2012).
16. Milne, G. A. et al. Perceptual reorganization from prior knowledge emerges late in childhood. *iScience* **27**, 108787 (2024).
17. Tulving, E. & Schacter, D. L. Priming and human memory systems. *Science* **247**, 301–306 (1990).
18. McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* **102**, 419–457 (1995).
19. Hardstone, R. et al. Long-term priors influence visual perception through recruitment of long-range feedback. *Nat. Commun.* **12**, 6288 (2021).
20. Tovee, M. J., Rolls, E. T. & Ramachandran, V. S. Rapid visual learning in neurones of the primate temporal visual cortex. *Neuroreport* **7**, 2757–2760 (1996).
21. Dolan, R. J. et al. How the brain learns to see objects and faces in an impoverished context. *Nature* **389**, 596–599 (1997).
22. González-García, C., Flounders, M. W., Chang, R., Baria, A. T. & He, B. J. Content-specific activity in frontoparietal and default-mode networks during prior-guided visual perception. *eLife* **7**, e36068 (2018).
23. Flounders, M. W., González-García, C., Hardstone, R. & He, B. J. Neural dynamics of visual ambiguity resolution by perceptual prior. *eLife* **8**, e41861 (2019).
24. Gorlin, S. et al. Imaging prior information in the brain. *Proc. Natl Acad. Sci. USA* **109**, 7935–7940 (2012).
25. Quiroga, R. Q., Reddy, L., Koch, C. & Fried, I. Decoding visual inputs from multiple neurons in the human temporal lobe. *J. Neurophysiol.* **98**, 1997–2007 (2007).

26. Hung, C. P., Kreiman, G., Poggio, T. & DiCarlo, J. J. Fast readout of object identity from macaque inferior temporal cortex. *Science* **310**, 863–866 (2005).
27. Tang, H. et al. Spatiotemporal dynamics underlying object completion in human ventral visual cortex. *Neuron* **83**, 736–748 (2014).
28. Kreiman, G. *Biological and Computer Vision* (Cambridge Univ. Press, 2021).
29. Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M. & Egner, T. Neural repetition suppression reflects fulfilled perceptual expectations. *Nat. Neurosci.* **11**, 1004–1006 (2008).
30. Gotts, S. J., Chow, C. C. & Martin, A. Repetition priming and repetition suppression: a case for enhanced efficiency through neural synchronization. *Cogn. Neurosci.* **3**, 227–237 (2012).
31. Larsson, J. & Smith, A. T. fMRI repetition suppression: neuronal adaptation or stimulus expectation? *Cereb. Cortex* **22**, 567–576 (2012).
32. Vinken, K., Boix, X. & Kreiman, G. Incorporating intrinsic suppression in deep neural networks captures dynamics of adaptation in neurophysiology and perception. *Sci. Adv.* **6**, eabd4205 (2020).
33. Mormann, F. et al. Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *J. Neurosci.* **28**, 8865–8872 (2008).
34. Quiroga, R. Q. Concept cells: the building blocks of declarative memory functions. *Nat. Rev. Neurosci.* **13**, 587–597 (2012).
35. Quian Quiroga, R. et al. Single neuron responses underlying face recognition in the human midfusiform face-selective cortex. *Nat. Commun.* **14**, 5661 (2023).
36. Connor, C. E., Brincat, S. L. & Pasupathy, A. Transformation of shape information in the ventral pathway. *Curr. Opin. Neurobiol.* **17**, 140–147 (2007).
37. Jutras, M. J. & Buffalo, E. A. Recognition memory signals in the macaque hippocampus. *Proc. Natl Acad. Sci. USA* **107**, 401–406 (2010).
38. Quian Quiroga, R. An integrative view of human hippocampal function: differences with other species and capacity considerations. *Hippocampus* **33**, 616–634 (2023).
39. Rolls, E. T., Xiang, J. & Franco, L. Object, space, and object-space representations in the primate hippocampus. *J. Neurophysiol.* **94**, 833–844 (2005).
40. Naya, Y., Yoshida, M. & Miyashita, Y. Backward spreading of memory-retrieval signal in the primate temporal cortex. *Science* **291**, 661–664 (2001).
41. Sakai, K. & Miyashita, Y. Neural organization for the long-term memory of paired associates. *Nature* **354**, 152–155 (1991).
42. Messinger, A., Squire, L. R., Zola, S. M. & Albright, T. D. Neuronal representations of stimulus associations develop in the temporal lobe during learning. *Proc. Natl Acad. Sci. USA* **98**, 12239–12244 (2001).
43. Li, N. & DiCarlo, J. J. Unsupervised natural visual experience rapidly reshapes size-invariant object representation in inferior temporal cortex. *Neuron* **67**, 1062–1075 (2010).
44. Dumoulin, S. O. & Wandell, B. A. Population receptive field estimates in human visual cortex. *NeuroImage* **39**, 647–660 (2008).
45. Wandell, B. A. & Winawer, J. Computational neuroimaging and population receptive fields. *Trends Cogn. Sci.* **19**, 349–357 (2015).
46. Brewer, A. A., Liu, J., Wade, A. R. & Wandell, B. A. Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nat. Neurosci.* **8**, 1102–1109 (2005).
47. Bashivan, P., Kar, K. & DiCarlo, J. J. Neural population control via deep image synthesis. *Science* **364**, eaav9436 (2019).
48. Pasupathy, A. & Connor, C. E. Population coding of shape in area V4. *Nat. Neurosci.* **5**, 1332–1338 (2002).
49. Andrews, T. J. & Schluppeck, D. Neural responses to Mooney images reveal a modular representation of faces in human visual cortex. *NeuroImage* **21**, 91–98 (2004).
50. Ison, M. J., Quian Quiroga, R. & Fried, I. Rapid encoding of new memories by individual neurons in the human brain. *Neuron* **87**, 220–230 (2015).
51. Kar, K., Kubilius, J., Schmidt, K., Issa, E. B. & DiCarlo, J. J. Evidence that recurrent circuits are critical to the ventral stream's execution of core object recognition behavior. *Nat. Neurosci.* **22**, 974–983 (2019).
52. Fyall, A. M., El-Shamayleh, Y., Choi, H., Shea-Brown, E. & Pasupathy, A. Dynamic representation of partially occluded objects in primate prefrontal and visual cortex. *eLife* **6**, e25784 (2017).
53. Tang, H. et al. Recurrent computations for visual pattern completion. *Proc. Natl Acad. Sci. USA* **115**, 8835–8840 (2018).
54. Kreiman, G. & Serre, T. Beyond the feedforward sweep: feedback computations in the visual cortex. *Ann. N.Y. Acad. Sci.* **1464**, 222–241 (2020).
55. Kar, K. & DiCarlo, J. J. Fast recurrent processing via ventrolateral prefrontal cortex is needed by the primate ventral stream for robust core visual object recognition. *Neuron* **109**, 164–176 (2021).
56. Rolls, E. T., Baylis, G. C., Hasselmo, M. E. & Nalwa, V. The effect of learning on the face selective responses of neurons in the cortex in the superior temporal sulcus of the monkey. *Exp. Brain Res.* **76**, 153–164 (1989).
57. Bar, M. et al. Top-down facilitation of visual recognition. *Proc. Natl Acad. Sci. USA* **103**, 449–454 (2006).
58. Gilbert, C. D. & Sigman, M. Brain states: top-down influences in sensory processing. *Neuron* **54**, 677–696 (2007).
59. Feldman, D. E. Synaptic mechanisms for plasticity in neocortex. *Annu. Rev. Neurosci.* **32**, 33–55 (2009).
60. Dosher, B. & Lu, Z.-L. Visual perceptual learning and models. *Annu. Rev. Vis. Sci.* **3**, 343–363 (2017).
61. Squire, L. R. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* **99**, 195–231 (1992).
62. Lin, Z., Dosher, B. A. & Lu, Z.-L. Mixture of easy trials enables transient and sustained perceptual improvements through priming and perceptual learning. *Sci. Rep.* **7**, 7421 (2017).
63. Cave, C. B. Very long-lasting priming in picture naming. *Psychol. Sci.* **8**, 322–325 (1997).
64. Sorscher, B., Ganguli, S. & Sompolinsky, H. Neural representational geometry underlies few-shot concept learning. *Proc. Natl Acad. Sci. USA* **119**, e2200800119 (2022).
65. Zhang, M. et al. Finding any Waldo with zero-shot invariant and efficient visual search. *Nat. Commun.* **9**, 3730 (2018).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026

Methods

Participants

The participants undergoing neuronal recordings were 13 adult patients (34 sessions; 5 female; aged 25–50 years; Supplementary Table 1) with epilepsy who were implanted with hybrid macro- and microelectrodes for epilepsy monitoring and evaluation for treatment of drug-resistant epilepsy, at the Freiburg Epilepsy Center in Freiburg im Breisgau, Germany. Electrode placement was solely determined by the need to localize epileptogenic regions. The electrode locations were determined using post-implantation computed tomography co-registered with preoperative MRI. These locations were then transformed into standardized MNI152 space and plotted on a template brain for visualization purposes only (Fig. 1h). All participants volunteered for this study and gave informed consent. The experimental procedures were approved by the Ethics Committee of the University of Freiburg, Freiburg im Breisgau, Germany.

Experimental task

Participants performed an image recognition task that included grayscale and two-tone Mooney images (Fig. 1). Stimuli depicted animals, objects and people, presented as grayscale pictures (Fig. 1b and Supplementary Fig. 4) and their corresponding two-tone Mooney counterparts[11] (Fig. 1a). They viewed sequences of images and reported recognition with a button press in a two-alternative forced-choice manner ('yes' or 'no'). The first time a given Mooney image was recognized, participants were asked to verbally report the identity of the image to confirm successful recognition. No further verbal reports were requested for subsequent presentations of recognized images. Responses were marked as incorrect if the object was not correctly identified, and participants received accuracy feedback. At the beginning of each session, participants received basic instructions for the task and were shown examples of Mooney and the corresponding grayscale pictures to illustrate the image sequence. Participants did not receive directions regarding how or where to fixate. Sessions were divided into blocks solely to allow participants to take brief breaks and reduce fatigue, and block boundaries had no task-related significance. Images were presented in blocks of 60 trials, and participants completed on average 16.6 ± 3.3 blocks (mean ± s.d.) per session (Supplementary Table 1 shows the number of sessions for each participant). Sessions ended when a maximum of 25 blocks was reached, or earlier if a participant wanted to stop. In each trial, a grayscale or Mooney image (5° × 5° visual angle) was presented at the center of the screen for 500 ms followed by a blank screen with a fixation cross for 1,300 ms (Fig. 1c). All trials in the task had identical temporal structure, differing only in the image presented.

At the start of the session, participants were exposed to the Mooney versions of four different images, intermixed during presentation, which were largely unrecognized (pre-learning Mooney images) (Fig. 1d, left, red frame; Fig. 1e,f). The set of images presented increased over the task based on participants' behavior. For instance, if participants recognized the identity of a Mooney image before learning, that image was no longer presented, the learning sequence for that image was omitted, and a new Mooney image was introduced in the sequence. This resulted in an interleaved presentation structure, with blocks containing balanced presentations of four to five distinct images. Participants were exposed to up to ten different images over the course of the task. After multiple presentations of unrecognized Mooney images (30 trials), participants were presented with a learning sequence. Each learning sequence consisted of two trials in which a Mooney image was immediately preceded by its corresponding grayscale image to facilitate learning (Fig. 1d, middle, turquoise frame). Learning sequences represented a small fraction of the total trials in a session (~2–3%). Participants reported recognition of the depicted object for each trial within the learning sequence, as in all other trials. An image was marked as learned once the participant successfully recognized both the grayscale and Mooney versions. The learning sequence was repeated until both versions were correctly identified. In most cases, images required only a single learning sequence within a session, reflecting rapid learning (median of 1). Once an image was learned, no further learning sequences for that image were presented. After learning, the Mooney images were presented again (now as post-learning Mooney images) (Fig. 1d, right, blue frame). As the experiment progressed and Mooney images transitioned from the pre-learning to post-learning condition, new previously unseen Mooney images were introduced into the sequence, with pre- and post-learning images presented intermixed across the session. Grayscale images were also interleaved with Mooney images throughout the task, with the proportion of trials balanced across the three conditions (one-third each). Participants who completed multiple sessions (Supplementary Table 1) were shown different images in each session. Behavioral performance was measured as the proportion of recognized images per condition (Fig. 1e,f), and response times were defined as the interval between stimulus onset and the participant's button press indicating recognition.

Eye tracking

To assess eye movements during visual recognition, we conducted the image recognition task described above in an independent cohort of healthy participants (13 sessions across ten individuals, five female, aged 24–65 years). Eye position was recorded using an infrared-based eye tracker (EyeLink 1000, SR Research) sampling at 500 Hz. Calibration was performed at the beginning of each session using a standard nine-point grid and was accepted only when the validation error was less than 1° of visual angle (mean validation error: 0.35°). The Eyelink system's default settings were used to define fixation and saccade onsets. Visual stimuli and trial structure matched those used in the patient recordings, and participants received the same instructions as patients for performing the task, without any additional directions regarding eye movements or fixation.

Neurophysiological recordings and spike sorting

We conducted extracellular electrophysiological recordings in patients using Behnke–Fried electrodes (Ad-Tech), which contained eight microwires (40 µm in diameter) located at the tip of the electrode shaft[66–68]. Microelectrode coverage included the OC, PHC, entorhinal cortex, hippocampus and amygdala (Supplementary Table 2). Electrodes implanted in the OC were primarily located in the ventromedial portion of the occipital lobe (Supplementary Fig. 1), within the anterior part of the lingual gyrus, and falling within regions corresponding to visual areas VO, V8, V4v and V3v based on multimodal cortical parcellation atlases[69,70]. One exception was an electrode located in the dorsomedial occipital lobe, in the superior portion of the cuneus, bordering the parieto-occipital sulcus and falling within the region corresponding to area V3d. We recorded broadband signals (0.1–9,000 Hz) at a sampling rate of 30 kHz using the NeuroPort system (Blackrock). Subsequently, these signals were filtered offline between 300 and 3,000 Hz using a zero-phase digital filter. We performed spike detection and sorting for each microwire using the semiautomated template-matching OSort algorithm[71].

Single-neuron analysis

To identify visually responsive neurons, we compared firing rates during the interval from 200 to 700 ms after image onset with the baseline period, defined as the −500 to 0 ms interval before image onset. This interval maximized the detection of visually responsive neurons and followed similar windows used in previous work[72,73]. We also evaluated alternative earlier windows (100–600 ms and 150–650 ms), which yielded similar results. Neurons were considered visually responsive if their firing rates significantly changed (increased or decreased) in response to at least one image (either grayscale or Mooney). Significance was assessed using a two-sided permutation test (1,000

iterations) with $P < 0.05$, corrected for FDR. FDR correction involved shuffling the labels of firing rates and repeating the analysis 1,000 times to generate a null distribution of results. The P threshold for visual responsiveness was then adjusted to ensure that fewer than 5% of neurons were expected to be identified as responsive by chance. Neurons in the hippocampus, amygdala and entorhinal cortex were considered together as the MTL. Subsequent analyses were based only on the visually responsive neurons (Fig. 2a).

We measured neuronal response latency as the time from stimulus onset to the first significant change in firing rate relative to the baseline period (Fig. 2b). For each responsive neuron, we estimated the instantaneous firing rate by convolving its spike train with an asymmetric kernel function[74]

$$R(t) = (1 - e^{-t/\tau_g})e^{-t/\tau_d},$$

where $\tau_g = 1$ ms and $\tau_d = 25$ ms are the growth and decay constants, respectively. Latency was determined as the time point where the instantaneous firing rate exceeded 3 s.d. from the baseline for at least 100 ms. Latency distributions were compared across regions using the Kruskal–Wallis test.

Stimulus selectivity (Fig. 2c) was assessed using an SI based on Quiroga et al.[25], bounded between 0 and 1, with higher values indicating greater selectivity. The SI is defined as

$$\mathrm{SI} = 1 - \frac{1}{M}\sum_{j=1}^{M} R(T_j),$$

with $T_j = f_{\min} + j(f_{\max} - f_{\min})/M$ being equidistant thresholds between the minimum and maximum firing rates $(f_{\min}, f_{\max})$, $j = 1, \ldots, M$; and M being the number of thresholds used (here, $M = 1{,}000$). $R(T)$ denotes the relative number of stimuli eliciting responses larger than T, given by

$$R(T) = \frac{1}{N}\sum_{i=1}^{N}\theta(f_i - T), \quad \theta(x) = \begin{cases} 1, x > 0 \\ 0, x \le 0 \end{cases},$$

where f_i is the firing rate of a neuron to stimulus i ($i = 1, \ldots, N$), and N is the number of stimuli presented.

For each neuron, we quantified the SNR as the number of standard deviations by which the stimulus-evoked firing rate exceeded the mean baseline activity (Fig. 2d). Regional differences in SNR and their consistency over time were assessed by computing SNR separately per block within the session (Supplementary Fig. 3a).

Discriminability (d') between preferred and least-preferred stimuli (Fig. 2e) was computed for each neuron as

$$d' = \frac{\mu_{\mathrm{pref}} - \mu_{\mathrm{least}}}{\sqrt{\frac{1}{2}\left(\sigma^2_{\mathrm{pref}} + \sigma^2_{\mathrm{least}}\right)}}$$

where μ_{pref} and μ_{least} are the mean responses of a neuron to its preferred and least-preferred stimuli, and σ^2_{pref} and $\sigma^2_{\mathrm{least}}$ are the corresponding response variances across trials. Higher d' values indicate a greater discriminability. Discriminability consistency over time was assessed by computing d' separately for each block in the session (Supplementary Fig. 3b).

Stimulus preference

To evaluate the emergence of stimulus preference with learning (Fig. 3), we defined, for each neuron, the most and least preferred images on the basis of responses to the grayscale condition alone, independently of the Mooney responses. We then computed the average normalized firing rate across neurons for these preferred versus nonpreferred images in the pre- and post-learning Mooney conditions separately for each brain region. Responses were analyzed in nonoverlapping 50-ms time bins. At each time bin, a two-sided Wilcoxon signed-rank test was performed across conditions, and time points were considered significant if $P < 0.05$ after correction for multiple comparisons.

Decoding analysis

Single-trial decoding analyses were conducted using neuronal responses as features at either the pseudopopulation (Fig. 4) or individual neuron level (Extended Data Fig. 4a–c). Pseudopopulations were constructed by aggregating neurons recorded across sessions[26,27,75].

For each neuron, baseline firing rate was subtracted on a trial-by-trial basis to reduce slow drifts in neural activity. Responses were then z-scored using normalization parameters estimated from the training set (mean subtraction and division by standard deviation across trials), and the same parameters were applied to the test set[72]. The decoders used SVMs with a linear kernel, as they provide a well-established method for assessing the performance of feature representations of stimulus data in discrete classification tasks and are particularly suitable for high-dimensional neural datasets[26,76]. Each binary classifier learned a decision function

$$f(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b,$$

with feature vector $\mathbf{x}$, weight vector $\mathbf{w}$ and bias term b. The parameters $\mathbf{w}$ and b were estimated by solving the optimization problem

$$\min_{\mathbf{w},b,\xi} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\xi_i$$

subject to

$$y_i\left(\mathbf{w}^T\mathbf{x} + b\right) \ge 1 - \xi_i, \quad \xi_i \ge 0,$$

where $(\mathbf{x}_i, y_i)$ are the training samples with labels $y_i \in \{-1, 1\}$, ξ_i are slack variables allowing misclassification, and C is the regularization parameter balancing margin size and error penalty. Performance was estimated using five-fold cross-validation. Decoding across time was assessed using a 250-ms sliding window, with steps of 25 ms.

For within-condition image decoding, we used three-class decoders (chance of 33%) (Fig. 4). To account for differences in the number of images across sessions, we selected the top three images per session, ranked by the number of neurons responding in the grayscale condition. Multiclass classification was implemented using the fitcecoc function in MATLAB (R2022b, MathWorks) with a one-versus-all scheme, which constructs a set of binary SVM learners and combines their outputs to assign class labels[73,77]. To pool images across sessions, we arbitrarily labeled images in each session as Images 1–3, and then grouped those with the same labels. This approach is suitable because the SVM algorithm requires that each class elicits distinct neuronal responses to separate patterns in a high-dimensional space, independent of image identity[73]. Using this procedure, cross-time image decoding was computed separately for each condition.

For condition decoding, we used binary classifiers (chance of 50%) trained to discriminate between pre- versus post-learning conditions (Extended Data Fig. 4d–f), grayscale versus pre-learning and grayscale versus post-learning (Extended Data Fig. 4g–i). Each classifier label included different images. For instance, the pre-learning label contained only pre-learning Mooney images, but these depicted different objects, such as a camel, a tool or people. Because different sessions contained varying numbers of images, we balanced the number of trials across sessions by randomly subsampling those with more images.

We determined significance by comparing decoding performance to an empirical null distribution. The null distribution was generated by shuffling the condition labels and running decoding for 500 iterations. Time points where decoding exceeded the 95th percentile of the null distribution for at least 250 ms were considered significantly different from chance.

Generalization analysis

We assessed generalization following a cross-condition population decoding approach, training decoders on one condition and testing them on another condition[72,78]. SVMs with a linear kernel were employed using a 250-ms sliding window with 25-ms steps. Pseudopopulations of neurons, pooled across sessions, were used to build classifiers. Classifiers were first trained to predict image identity from neuronal patterns in one condition. We then froze the classifier weights and tested predictions on neuronal patterns from a second condition. Following the same procedure as the within-condition decoding analyses, we selected the top three images from each session to train and test the decoders (chance of 33%). Using this procedure, we computed cross-time generalization profiles for Gray-to-Post (Fig. 5a,d), Gray-to-Pre (Fig. 5b,e) and Pre-to-Post (Fig. 5c,f). For each case, we computed both possible train–test combinations. For instance, in the Gray-to-Post generalization, we first trained on the grayscale condition and tested on the post-learning Mooney condition to estimate decoding accuracy at each time point. Then, we repeated this process in reverse order, training on the post-learning Mooney condition and testing on the grayscale condition. Each entry at time x,y in the generalization matrix (for example, Fig. 5a) represents the mean decoding of these two train–test combinations. Significance was assessed by comparing decoding performance to an empirical null distribution generated by shuffling condition labels (200 iterations). Significant decoding was defined at $P < 0.05$ (cluster-size correction for multiple comparisons)[73].

We quantified deviation of the generalization profiles from the diagonal by computing the probability density function of peak decoding (Extended Data Fig. 6). For each generalization profile, we resampled our data 200 times and ran our decoder as described above. Each iteration produced an x,y coordinate for peak decoding. Thus, each entry in the probability density function represents the frequency of that coordinate being identified as the maximum peak decoding. We then calculated the deviation from the diagonal (in milliseconds) as the difference between x and y. We compared the distribution of deviations from the diagonal against the null hypothesis that these deviations are zero (Fig. 5g–l). If 95% of the values in the distribution are greater (or lower) than zero, the deviation is considered significant. A positive deviation indicated that the patterns for the condition on the y axis generalized to later times for the condition on the x axis. A negative deviation indicated that the patterns for the y axis condition generalized to earlier times for the x axis condition. We then statistically compared the mean decoding accuracies below and above the diagonal. For each generalization profile, we selected one cluster from each side of the diagonal and compared their mean decoding (Supplementary Figs. 14 and 15). These clusters included all significant time points on both sides of the diagonal and their corresponding symmetric points. This process was repeated 200 times by resampling our data. We then employed a permutation test to compare the two distributions, determining significance at $P < 0.05$.

Control analysis

To assess the impact of excitability changes on our decoding analyses, we quantified firing rate changes over time for each neuron (Supplementary Fig. 6). For each neuron, we computed the mean firing rate during the post-stimulus interval (0–1 s) on every trial and fit a linear model $y_i = \alpha + \beta t_i$, where y_i denotes the firing rate across trials, t_i corresponds to the time of trial i and α is the intercept. To determine whether firing rates changed systematically over the course of the experiment, we tested whether the slope parameter β differed significantly from zero using a permutation test, in which trial order was shuffled 1,000 times. Neurons were considered significant at $P < 0.05$ (permutation test, corrected for multiple comparisons). We computed β slopes for both responsive and nonresponsive neuronal populations and compared their distributions using a Kolmogorov–Smirnov test.

Statistics and reproducibility

No statistical method was used to predetermine sample size. Our sample size is comparable to those reported in previous human single-neuron intracranial studies[72,73]. No data were excluded from the analyses. Of the 1,104 recorded neurons, 318 were identified as visually responsive and used in subsequent analyses (see above). The experiments were not randomized. The investigators were not blinded to the experimental conditions during data collection or analysis. Comparisons between conditions or groups were assessed using two-sided Wilcoxon signed-rank or rank-sum tests. Comparisons across more than two groups were performed using the Kruskal–Wallis test. Visual responsiveness of neurons was assessed using a permutation test. Decoding and generalization analyses were evaluated using permutation tests, in which significance was determined by comparing observed values to empirical null distributions generated by shuffling condition labels.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data supporting the main findings in this study are available at https://kreimanlab.com/code/rapid-visual-learning-in-the-human-brain/. The grayscale and Mooney images used in the experiments are also available at this link.

Code availability

Code supporting the main findings in this study is available at https://kreimanlab.com/code/rapid-visual-learning-in-the-human-brain/ and via Zenodo at https://doi.org/10.5281/zenodo.21708343 (ref. 79).

References

66. Fried, I., MacDonald, K. A. & Wilson, C. L. Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron* **18**, 753–765 (1997).
67. Kreiman, G., Koch, C. & Fried, I. Imagery neurons in the human brain. *Nature* **408**, 357–361 (2000).
68. Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C. & Fried, I. Invariant visual representation by single neurons in the human brain. *Nature* **435**, 1102–1107 (2005).
69. Amunts, K., Mohlberg, H., Bludau, S. & Zilles, K. Julich-Brain: a 3D probabilistic atlas of the human brain's cytoarchitecture. *Science* **369**, 988–992 (2020).
70. Glasser, M. F. et al. A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).
71. Rutishauser, U., Schuman, E. M. & Mamelak, A. N. Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo. *J. Neurosci. Methods* **154**, 204–224 (2006).
72. Minxha, J., Adolphs, R., Fusi, S., Mamelak, A. N. & Rutishauser, U. Flexible recruitment of memory-based choice representations by the human medial frontal cortex. *Science* **368**, eaba3313 (2020).
73. Kamiński, J. et al. Persistently active neurons in human medial frontal and medial temporal lobe support working memory. *Nat. Neurosci.* **20**, 590–601 (2017).
74. Lee, J., Kim, H. R. & Lee, C. Trial-to-trial variability of spike response of V1 and saccadic response time. *J. Neurophysiol.* **104**, 2556–2572 (2010).
75. Khanna, A. R. et al. Single-neuronal elements of speech production in humans. *Nature* **626**, 603–610 (2024).
76. Hong, H., Yamins, D. L. K., Majaj, N. J. & DiCarlo, J. J. Explicit information for category-orthogonal object properties increases along the ventral stream. *Nat. Neurosci.* **19**, 613–622 (2016).
77. Bishop, C. M. *Pattern Recognition and Machine Learning* (Springer, 2006).

78. Meyers, E. M., Freedman, D. J., Kreiman, G., Miller, E. K. & Poggio, T. Dynamic population coding of category information in inferior temporal and prefrontal cortex. *J. Neurophysiol.* **100**, 1407–1419 (2008).
79. Armendariz, M. Code for: single-neuron encoding of rapidly learned visual information reshapes human perception. *Zenodo* https://doi.org/10.5281/zenodo.21708343 (2026).

Acknowledgements

We thank the patients who participated in the study, as well as the participants who took part in the behavioral study.

Author contributions

M.A., J.B. and G.K. developed the hypothesis and research approach. J.B., F.A., P.R., A.B. and A.S.-B. recruited patients. M.A. and N.D.-I. recruited healthy participants. P.R. implanted electrodes. J.B. and M.A. collected data with support from F.A., A.S.-B., A.B. and N.D.-I. M.A., J.B., J.K. and J.S. performed the data curation. M.A. performed the analyses. M.A. developed methodology and software. M.A. and G.K. interpreted the findings and wrote the manuscript.

Funding

This work was supported by NIH grant R01EY026025 and NSF grant CCF-1231216. M.A. received support from the Research Foundation Flanders (FWO 1299924N). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41593-026-02496-6.

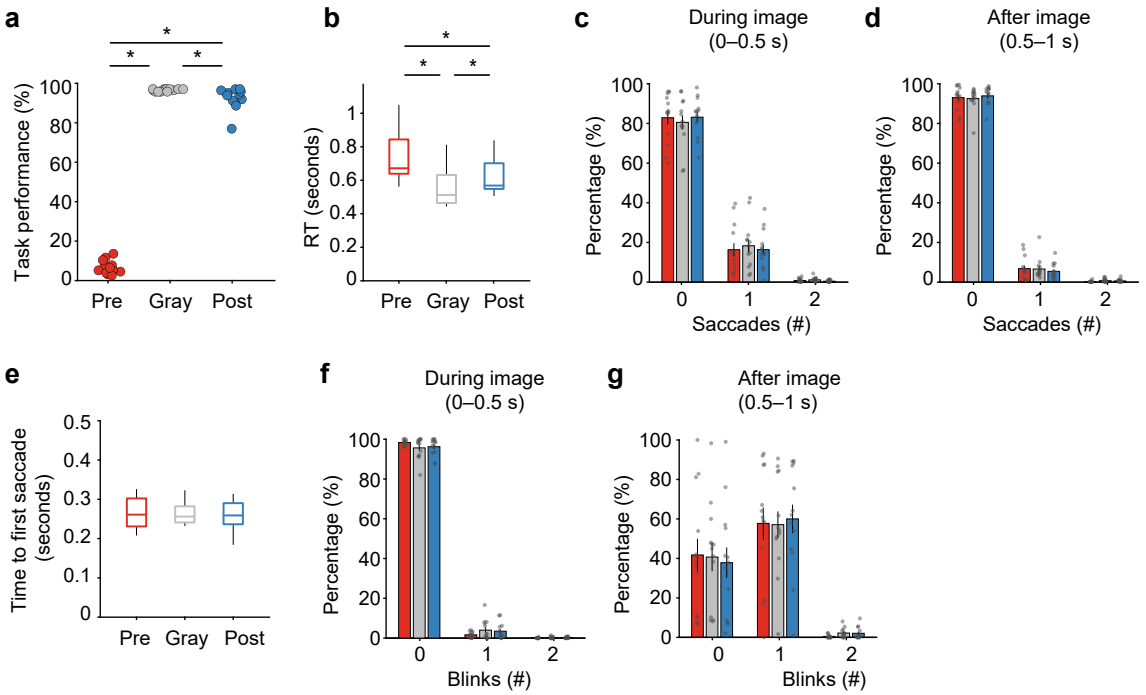
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41593-026-02496-6.

Correspondence and requests for materials should be addressed to Marcelo Armendariz or Gabriel Kreiman.

Peer review information *Nature Neuroscience* thanks Christopher Pack, Pieter Roelfsema and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

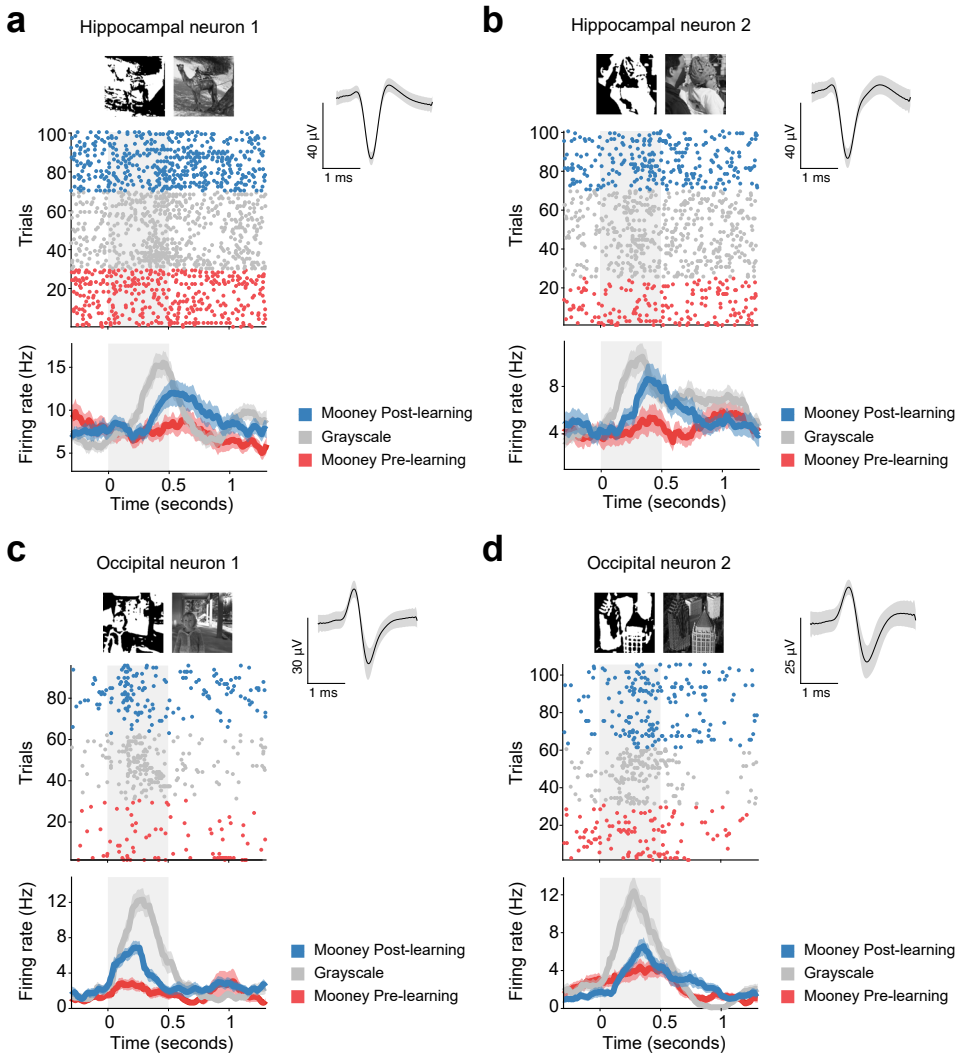
Reprints and permissions information is available at www.nature.com/reprints.

https://doi.org/10.1038/s41593-026-02496-6



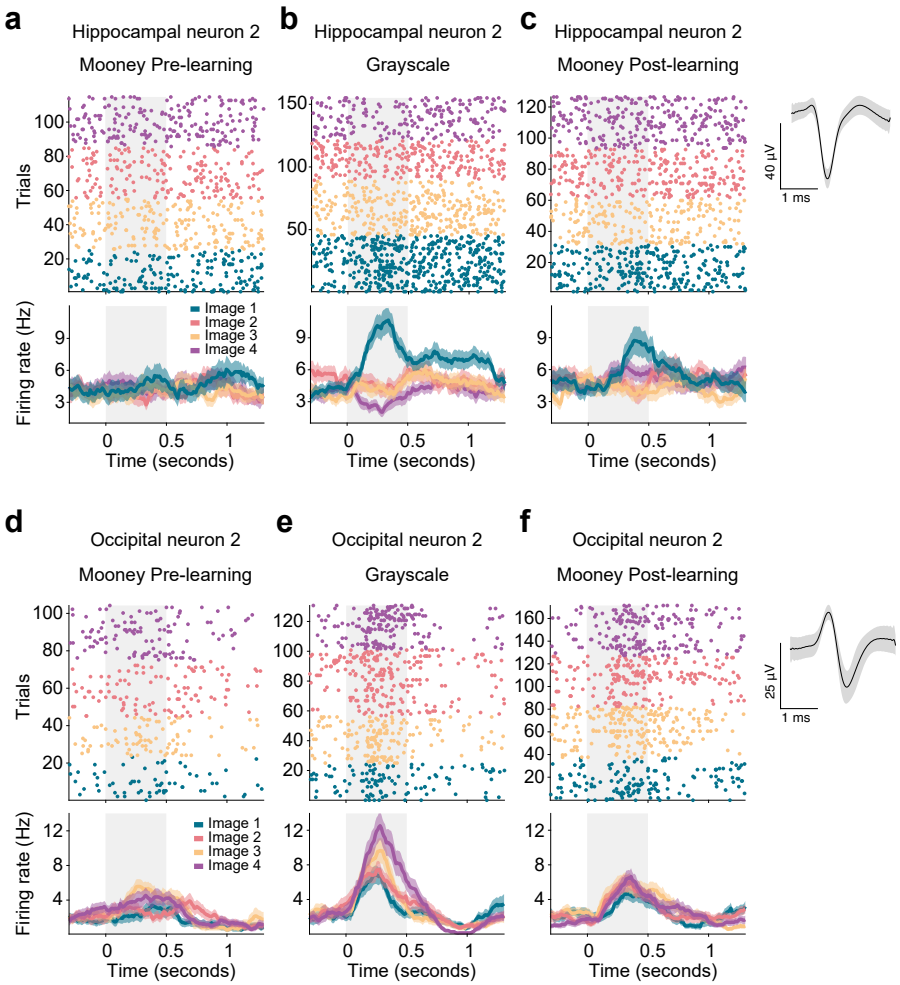
Extended Data Fig. 1 | Eye movements were similar across conditions. **a**, Performance during the recognition task with eye tracking. Participants exhibited poor image recognition for Mooney images before learning (Pre, red; 6%) and nearly perfect recognition for Grayscale images (Gray, gray; 99%). After learning, recognition for Mooney images improved to 95% (Post, blue). Each dot represents one session ($n = 13$; compare to Fig. 1f). Asterisks indicate statistically significant differences between conditions (Post versus Gray, $P = 0.004$; Pre versus Gray, $P = 0.0002$; Pre vs Post, $P = 0.0002$, two-sided Wilcoxon signed-rank test, Bonferroni-corrected). **b**, Response times (RT) relative to image onset. Participants were slower to recognize Mooney images compared to Grayscale images. Box plots show the median (center line), 25th and 75th percentiles (box bounds), and whiskers extending to the minimum and maximum values of the response time distributions across sessions ($n = 13$). Asterisks show statistically significant differences between conditions (Post vs Gray, $P = 0.0007$; Pre vs Gray, $P = 0.0007$; Pre vs Post, $P = 0.005$, two-sided Wilcoxon signed-rank test, Bonferroni corrected; compare to Fig. 1g). **c**, Percentage of trials with saccades during the image presentation interval (0–0.5 s). Participants made no saccades in most trials (~83%) and the rates of saccades were comparable across conditions ($P > 0.05$, two-sided Wilcoxon signed-rank test). **d**, Percentage of trials with saccades after the image presentation interval (0.5–1 s). Saccades were rare, occurring in fewer than ~7% of trials, with no differences across conditions ($P > 0.05$, two-sided Wilcoxon signed-rank test). Bars in **c** and **d** show the mean, error bars represent s.e.m., and dots represent individual sessions. **e**, Latency to first saccade from image onset. When saccades occurred, they happened on average at 267 ms, and there were no significant differences across conditions ($P > 0.05$, two-sided Wilcoxon signed-rank test). Box plots show the median (center line), 25th and 75th percentiles (box bounds), and whiskers extending to the minimum and maximum values of the distribution ($n = 13$). **f**, Percentage of trials with eye blinks during the image presentation interval (0–0.5 s). Blinks were negligible during stimulus presentation (<3% of trials). **g**, Percentage of trials with blinks after the image presentation interval (0.5–1 s). Participants blinked predominantly after stimulus offset (~60%), with no differences across conditions ($P > 0.05$, two-sided Wilcoxon signed-rank test). Bars in **f** and **g** show the mean, error bars represent s.e.m., and dots represent individual sessions.

https://doi.org/10.1038/s41593-026-02496-6



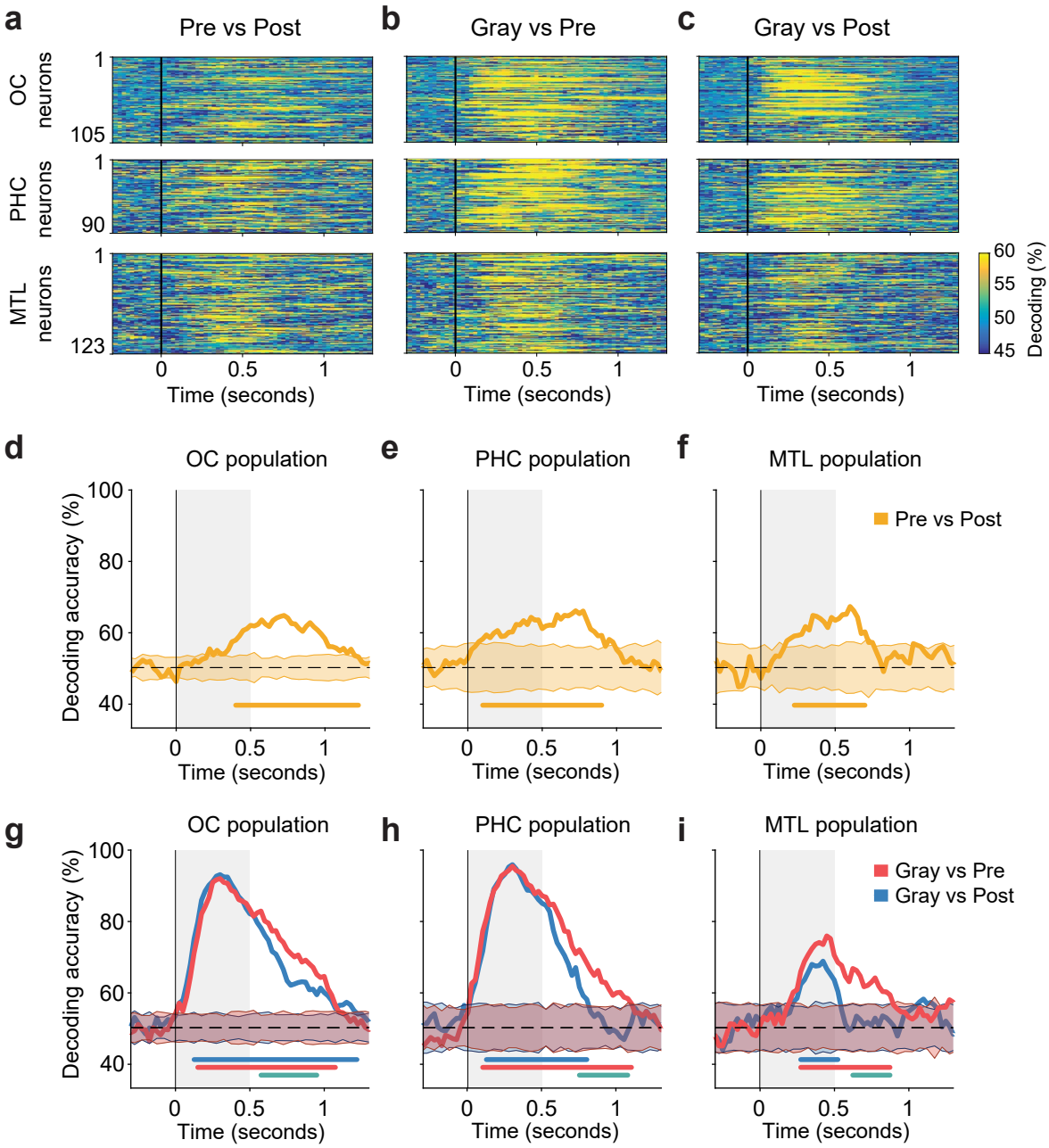
Extended Data Fig. 2 | Examples of visually responsive neurons. a–d, Spiking activity of 4 example neurons in response to individual images. Raster plots (top) show spiking activity (each dot represents a spike) for an image across trials (rows) and over time. Insets show the mean extracellular spike waveform for each neuron. Colors correspond to spiking activity of an image across the three different conditions: Pre-learning Mooney (red), Grayscale (gray), and Post-learning Mooney (blue). The corresponding Mooney and Grayscale images are shown above the raster plots. Trials are grouped by condition for visualization. The light gray rectangle indicates the image presentation interval. Mean firing rates (bottom) are shown separately for each condition (solid colored lines; bin size is 250 ms and step size is 10 ms). Shaded areas represent s.e.m. across trials.

https://doi.org/10.1038/s41593-026-02496-6



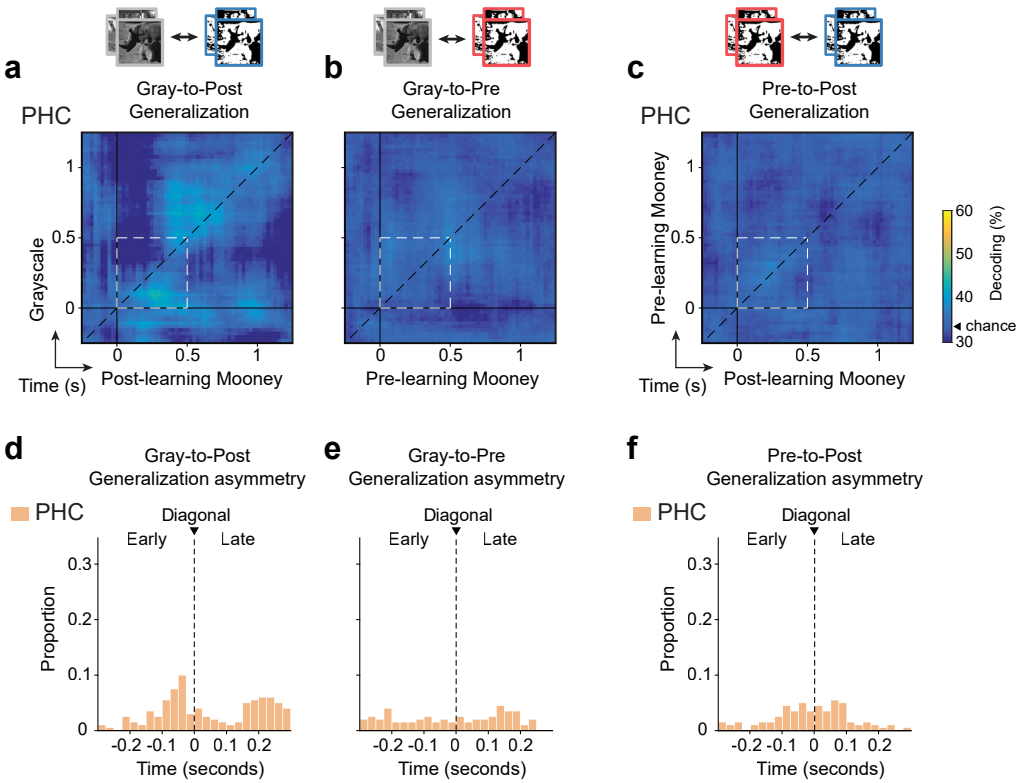
Extended Data Fig. 3 | Spiking activity for example neurons across images and conditions. a–**c**, Spiking activity of an example neuron in the hippocampus in response to different images in the Pre-learning Mooney (**a**), Grayscale (**b**), and Post-learning Mooney (**c**) conditions. **d**–**f**, Spiking activity of an example neuron in the occipital cortex in the Pre-learning Mooney (**d**), Grayscale (**e**), and Post-learning Mooney (**f**) conditions. Insets in panels **c** and **f** show the mean extracellular spike waveform for each neuron. Data use the same conventions as Fig. 2 in the main text.

https://doi.org/10.1038/s41593-026-02496-6



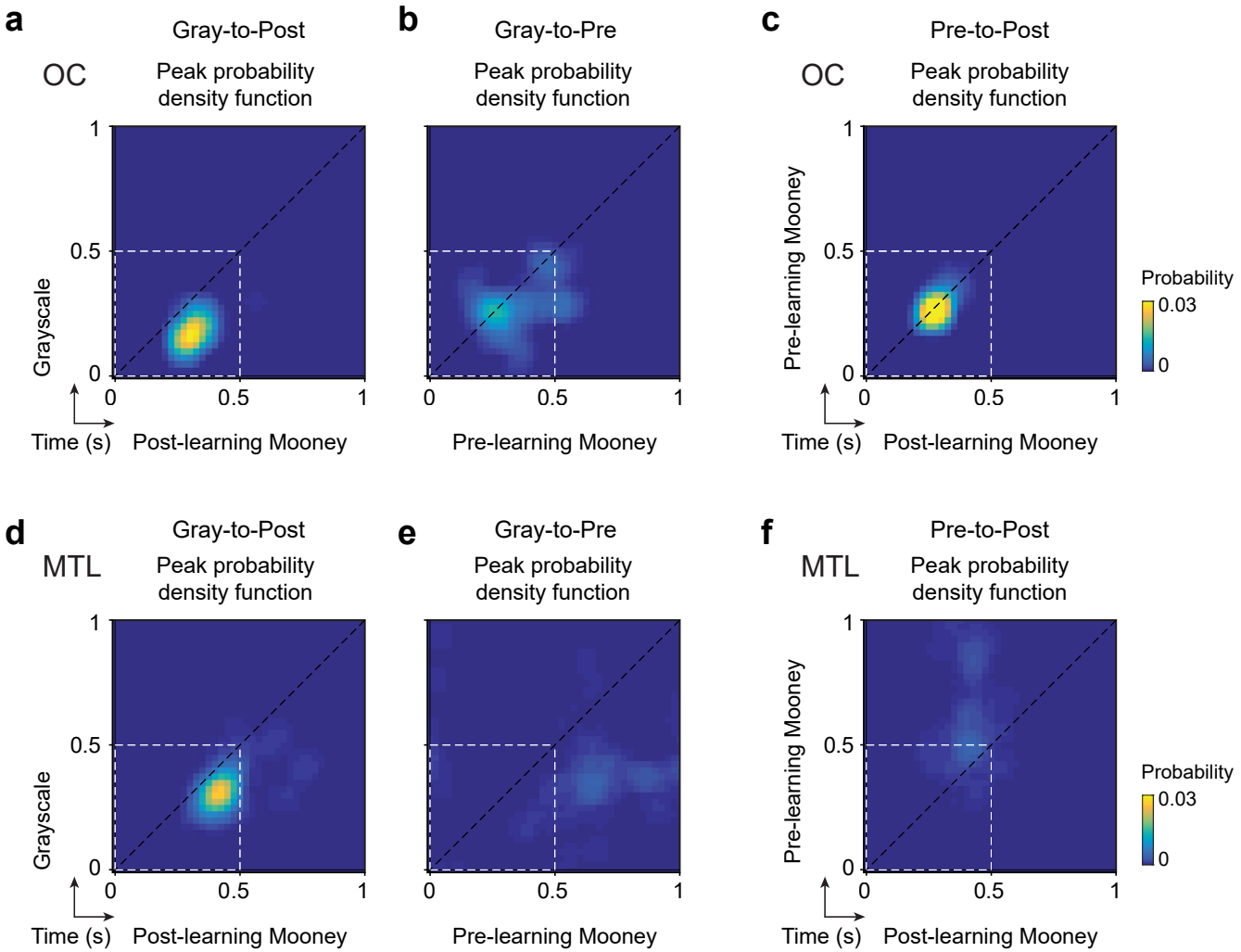
Extended Data Fig. 4 | Neuronal activity is modulated by learning. a–c, Single-neuron decoding accuracy for Pre- versus Post-learning Mooney images (**a**), Grayscale versus Pre-learning Mooney images (**b**), and Grayscale versus Post-learning Mooney images (**c**). Each row corresponds to a single neuron in OC (top), PHC (middle), and MTL (bottom). See color scale on bottom right. **d–f**, Population decoding accuracy for Pre- versus Post-learning Mooney images for OC (**d**), PHC (**e**), and MTL (**f**). **g–i**, Population decoding accuracy for Grayscale versus Pre-learning Mooney images (red) and Grayscale versus Post-learning Mooney images (blue) for OC (**g**), PHC (**h**), and MTL (**i**). In **d–i**, decoders were binary linear classifiers (support vector machine; chance = 50%) using sliding windows of 250 ms and a step size of 25 ms. The horizontal dashed lines indicate chance level. Colored shaded areas around the chance line represent the 95% confidence interval centered on the mean of the empirical null distribution obtained by shuffling the labels (500 iterations). Solid lines below the curves indicate time points with significantly above-chance decoding, assessed with a one-sided test against the empirical null distribution ($P < 0.05$; corrected for multiple comparisons across time; see Methods). The solid turquoise line below the curves in **g–i** represents time points with significant differences between the two curves, assessed with a two-sided permutation test ($P < 0.05$). Time zero indicates image onset and the shaded gray area represents the stimulus presentation period.

https://doi.org/10.1038/s41593-026-02496-6



Extended Data Fig. 5 | Generalization in PHC. a–c, Generalization profiles for Gray-to-Post (**a**), Gray-to-Pre (**b**), and Pre-to-Post (**c**) in PHC. **d–f**, Generalization asymmetry corresponding to the decoding profiles in **a–c**, respectively. Data follow the same conventions as Fig. 5a–c and Fig. 5g–i. No significant deviation from the diagonal was observed.

https://doi.org/10.1038/s41593-026-02496-6



Extended Data Fig. 6 | Peak decoding probability reveals asymmetric generalization dynamics in OC and MTL. a–c, Probability density functions of peak decoding in OC corresponding to the generalization profiles shown in Fig. 5a–c. The matrices were computed using a resampling method to identify the maximum decoding (Methods). Each matrix entry represents the probability of observing the decoding maximum at that position. The sum of all entries in each matrix equals 1. Peak decoding below the diagonal suggests delayed generalization from Grayscale to Post-learning Mooney images. Time zero indicates image onset, the black dashed line represents the matrix diagonal, and the square with white dashed lines represents the image presentation interval. For Gray-to-Post generalization in OC (**a**), peak decoding probability is highly concentrated below the diagonal around coordinates [175, 300] ms (Grayscale and Post-learning Mooney coordinates, respectively), suggesting delayed generalization from Grayscale to Post-learning Mooney images in OC. Conversely, for Pre-to-Post (**c**), peak decoding probability falls largely on the diagonal. **d–f**, Probability density functions of peak decoding in MTL corresponding to the generalization profiles shown in Fig. 5d–f. For Gray-to-Post generalization in MTL (**d**), peak decoding probability is highly concentrated below the diagonal around coordinates [325, 425] ms (Grayscale and Post-learning Mooney coordinates, respectively), indicating delayed generalization in MTL.

nature portfolio

Corresponding author(s): Gabriel Kreiman, Marcelo Armendariz

Last updated by author(s): 2026/07/28

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Neurophysiological data were collected using the NeuroPort system (Blackrock Microsystems, Salt Lake City, UT, USA) and hybrid depth electrodes (Ad-Tech, Oak Creek, Wisconsin, USA). See the Methods section for a detailed description.
Data analysis	Data analyses were performed with MATLAB 2022b (The MathWorks, Inc., Natick, MA, USA), using Statistics and Machine Learning Toolbox toolboxes and custom MATLAB code.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data and code supporting the main findings of this study are available at the following link:
https://kreimanlab.com/rapid-visual-learning-in-the-human-brain

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	The study patients were 5 female and 8 male. The Methods section contains the self-reported gender for all patients. Sex and gender were not part of the study design.
Reporting on race, ethnicity, or other socially relevant groupings	Race, ethnicity, or other socially relevant groupings were not part of the study design.
Population characteristics	The study patients were 13 adult patients (5 female, 8 male; aged 25 to 50 years) with epilepsy who were implanted with hybrid macro- and microelectrodes for epilepsy monitoring and evaluation for treatment of drug-resistant epilepsy, at the Freiburg Epilepsy Center in Freiburg im Breisgau, Germany. Supplementary Table 1 provides information on each patient
Recruitment	The study participants were adult patients undergoing surgery for epilepsy monitoring and evaluation for treatment of drug-resistant epilepsy. Patients who were capable of and willing to perform the task were included in the study. All participants volunteered for this study and gave informed consent.
Ethics oversight	The experimental procedures were approved by the Ethics Committee of the University of Freiburg, Freiburg im Breisgau, Germany.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We recorded spiking activity from a total of 1,104 single neurons in medial occipital and temporal brain areas in 34 sessions from 13 patients. No statistical method was used to predetermine sample size. Our sample size is comparable to those reported in previous human single-neuron intracranial studies.
Data exclusions	No participants were excluded from the analyses. Of the 1,104 neurons, we identified 318 visually responsive neurons which were used in the analyses.
Replication	Analyses were performed at the population level, separately for each anatomical region and at the single-neuron level. Permutation tests and corrections for multiple comparisons were used to assess significant effects.
Randomization	All participants and sessions were included in the analyses and recorded neurons were not assigned to experimental groups. Neurons were grouped based on their anatomical region to perform region-specific analyses.
Blinding	Participants were not informed about the research questions of the study. Participants were only given basic instructions to perform the image recognition task. The investigators were not blinded to the experimental conditions during data collection or analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.