Image familiarization sharpens response dynamics of neurons in inferotemporal cortex

Travis Meyer[1], Christopher Walker[2], Raymond Y Cho[2] & Carl R Olson[1,3]

Repeated viewing of an image over days and weeks induces a marked reduction in the strength with which neurons in monkey inferotemporal cortex respond to it. The processing advantage that attaches to this reduction is unknown. One possibility is that truncation of the response to a familiar image leaves neurons in a state of readiness to respond to ensuing images and thereby enhances their ability to track rapidly changing displays. We explored this possibility by assessing neuronal responses to familiar and novel images in rapid serial visual displays. Inferotemporal neurons responded more strongly to familiar than to novel images in such displays. The effect was stronger among putative inhibitory neurons than among putative excitatory neurons. A comparable effect occurred at the level of the scalp potential in humans. We conclude that long-term familiarization sharpens the response dynamics of neurons in both monkey and human extrastriate visual cortex.

Inferotemporal cortex, the terminus of the ventral stream of visual areas[1], has a critical role in object vision. This role is thought to depend on the fact that neurons in inferotemporal cortex, including its largest division, area TE, respond selectively to complex images and on the fact that their responsiveness can be modified by visual experience. Changes induced by visual experience include the familiarity effect, in which repeated viewing of an image leads to a reduction in the strength with which neurons in TE respond to it. This effect develops not only if monkeys actively process images[2–8], but also if they view them passively[5]. It is evident, on subsequent testing, not only during active discrimination[2–9], but also during passive exposure[2–8,10]. The effect is generally assumed to depend on prolonged experience. Previous studies have employed hundreds[2–8] or thousands[8] of exposures spanning a period of weeks[5] or months[2–8]. However, even one exposure to an image leads to a short-term reduction in response strength that may be a precursor to the long-term effect[3,4,11–26]. The nature of the visual processing advantage that attaches to the familiarity effect is unknown. On one hand, the reduction itself might serve some purpose. This could be to signal the familiarity of the image[3] or to reduce the salience of an expected and therefore uninformative event[27]. On the other hand, the reduction might be incidental to some primary beneficial change. For example, if familiarization induced a few neurons selective for the image to became more strongly responsive and led to a reduction of response strength among all other neurons, this would create a primary benefit, sharper tuning, in conjunction with an incidental effect, reduced population response strength[5,8].

We tested an alternative idea, namely that the weakening of the response induced by familiarization is secondary to an improvement in dynamic tracking ability. It is a little remarked fact that the reduction in response strength induced by familiarization is not uniform in time, but rather comes about in consequence of a sharp truncation occurring around 80 ms after response onset. Truncation of the response might put neurons in a state of readiness to respond to any ensuing stimulus and so might contribute to tracking events in a rapidly changing environment. To explore this idea, we recorded neuronal responses to dynamic displays. Under conditions of rapid serial visual presentation, which involves the display of images in a continuous stream with an interstimulus interval of around 100 ms, human subjects are able to perform perceptual tasks, but do so with greater difficulty than at lower presentation rates[28]. Likewise, neurons of monkey inferotemporal cortex continue to carry information about image identity[29–31], but do so less effectively than at lower rates[32]. We asked whether, under conditions of rapid serial visual presentation, neurons in TE track the content of the changing display more effectively if the images are familiar than if they are novel.

RESULTS

During the initial phase of the experiment, each monkey repeatedly viewed 16 images of background-free objects (**Fig. 1a**). Monkey 1 viewed each image 140 times during 28 runs spanning 4 weeks. Monkey 2 viewed each image 160 times during 32 runs spanning 2 weeks. Each run consisted of 80 trials in a sequence that was random with the exception that each of the 16 images had to appear five times. On each trial, while the monkey fixated the center of the screen, an image appeared at fixation for a second (**Fig. 1b**). The image sets used for training monkeys 1 and 2 contained no items in common. During subsequent microelectrode recording sessions, we assessed the responses of each neuron to two 'familiar' images selected arbitrarily from the training set and two 'novel' images selected arbitrarily from the large library of background-free objects from which the training images had been drawn. The novel images used in each session had never before been seen by the monkey and were used in no further sessions. Each familiar image was used in multiple sessions. No single familiar image was used in fewer than 7% or more than 14% of all

Figure 1 Training and testing procedures. (**a**) 16 images rendered familiar by repeated viewing over the course of a month in monkey 1. The exposure set for monkey 2 consisted of other background-free object images. (**b**) On each exposure trial, during the familiarization period, the monkey maintained fixation at the center of the screen for 1,600 ms. During fixation, a single image was presented for 1,000 ms at the fovea. Fluid reward was delivered at the end of each trial contingent solely on the monkey's having maintained fixation. (**c**) Procedure for rapid serial visual presentation during neuronal recording. On each trial, the monkey maintained central fixation while two images were presented in alternation at the fovea. The images could be both familiar (F_1 and F_2) or both novel (N_1 and N_2). (**d**–**i**) Data from a neuron giving strong periodic responses to rapidly alternating familiar images, but not to rapidly alternating novel images. Successive displays in the left column show the responses of a representative neuron to familiar images presented in the sequence F_1F_2 (**d**) or F_2F_1 (**e**) and in data combined from the two sequences (**f**). Successive displays in the right column show responses of the same neuron to novel images presented in the sequence N_1N_2 (**g**) or N_2N_1 (**h**) and in data combined from the two sequences (**i**).
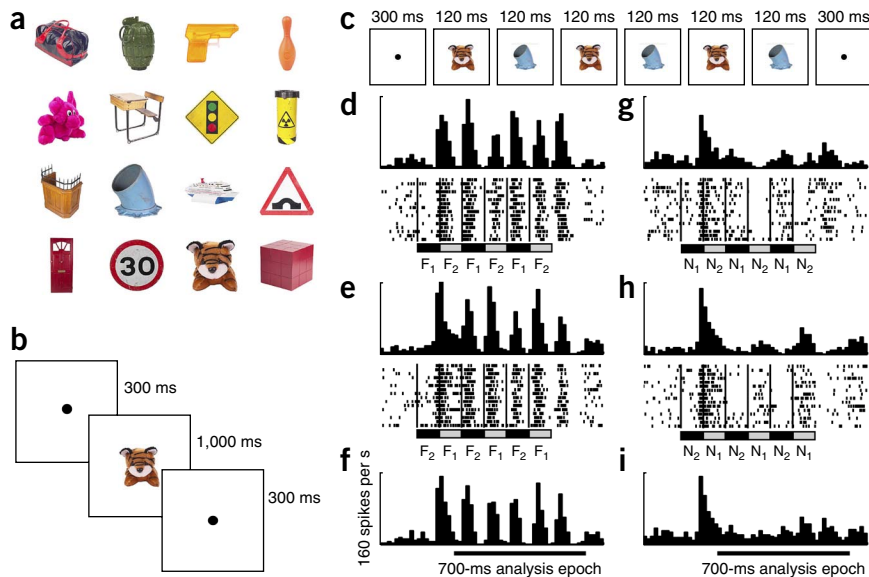
sessions. The selection procedure necessarily precluded finding the optimal stimulus for any neuron. We monitored responses elicited by familiar and novel images in 66 neurons (29 in monkey 1 and 37 in monkey 2) of anterior area TE. The data were combined between the two monkeys, but the key phenomena were present in both monkeys and were generally notable in each monkey considered individually (**Supplementary Figs. 1** and **2**).

To assess the ability of neurons to track rapidly changing displays, we presented two images in back-to-back alternation for 120 ms each until each had appeared three times (**Fig. 1c**). The representative neuron in **Figure 1d**–**i** responded with a strong burst to the onset of each successive familiar image regardless of whether one image (**Fig. 1d**) or the other (**Fig. 1e**) initiated the alternating string. In contrast, it gave a strong phasic response only to the first image in a novel string regardless of order (**Fig. 1g,h**). To quantify this effect, we created histograms on the basis of averaging across the two orders (**Fig. 1f,i**). We then used a fast Fourier transform to compute power in the histogram at the driving frequency of 8.33 Hz corresponding to the interstimulus interval of 120 ms. This analysis focused on a 700-ms window (**Fig. 1f,i**), excluding the response to the first image in the string. We converted power to peak-to-peak amplitude so as to render the results in units of spikes per second. Then, for the entire population of 66 neurons, we plotted amplitude for familiar images against amplitude for novel images (**Fig. 2a**). There was a highly significant tendency for amplitude to be greater when periodic activity was driven by familiar images than when it was driven by novel images ($P = 0.000026$, Mann-Whitney U test, $n = 66$). That driven periodic activity was stronger for familiar images than for novel images was also clear from histograms representing the population mean firing rate (**Fig. 2b**).

To be certain that the effect depended on the familiarity and novelty of the images rather than on some other factor correlated with familiarity and novelty, we carried out several further analyses. The phenomenon might have depended on differences between familiar and novel images with regard to response strength or selectivity as measured during individual presentation. However, when the dependence of driven periodic activity on these factors was eliminated through a regression procedure, the significance of the effect actually increased ($P = 1.2 \times 10^{-10}$, Mann-Whitney U test, $n = 66$; **Supplementary Fig. 3**).

Figure 2 Rapid sequences of familiar images elicit stronger periodic responses than rapid sequences of novel images. (**a**) For each of 66 neurons, the amplitude of the periodic response elicited by a familiar string is plotted against the amplitude of the periodic response elicited by a novel string. Amplitude was measured during the 700-ms window indicated by the black bars in **Figure 1f,i**. The circled point is from the example neuron shown in **Figure 1d**–**i**. The P value indicates the outcome of a paired t test on the 66 pairs of values. (**b**) Mean across 66 neurons of the firing rate elicited under familiar-image (red) and novel-image (blue) conditions. (**c**) For each of 55 LFP sites, the amplitude of the periodic response elicited by a familiar string is plotted against the amplitude of the periodic response elicited by a novel string. (**d**) Mean across 55 LFP sites of the voltage elicited under familiar-image (red) and novel-image (blue) conditions. Ribbons in **b** and **d** indicate s.e.m.
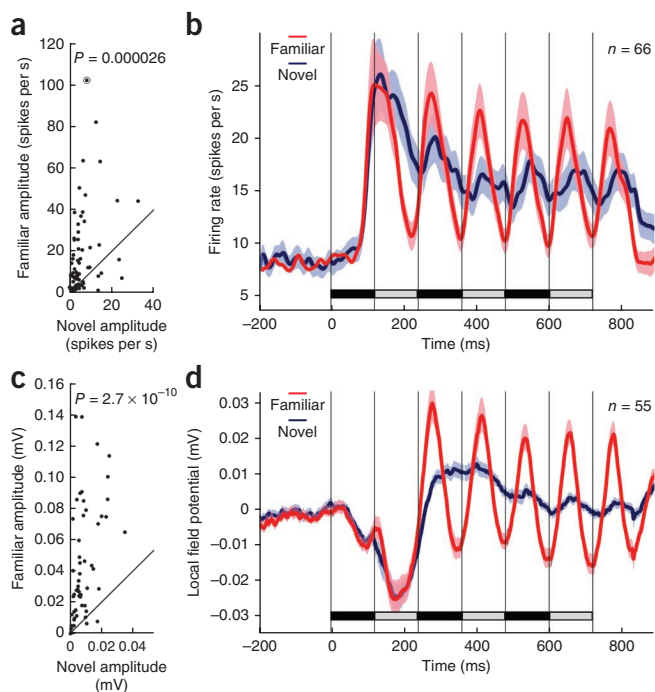
Figure 3 The tendency for familiar images to elicit strong periodic responses is more prominent among fast-spiking (putative inhibitory) than among regular-spiking (putative excitatory) neurons. (**a**,**b**) Data from 32 fast-spiking neurons. (**c**,**d**) Data from 34 regular-spiking neurons. Data are presented as in **Figure 2a**. For criteria on the basis of which neurons were classified, see **Supplementary Figure 4**.

The phenomenon might have depended on the monkeys' adopting different gaze strategies when viewing familiar and novel displays. However, saccade rate and horizontal gaze angle did not differ between conditions. The mean vertical gaze angle did differ (being lower by 0.5° during viewing of familiar displays), but this could not explain the effect, which persisted and was significant ($P = 0.0063$, Mann-Whitney U test, $n = 66$) in a subset of trials selected to reverse the trend. Finally, the phenomenon might have depended on short-term rather than long-term familiarity. If so, then novel images should have begun to elicit strong driven periodic activity over the course of a session. Periodic activity evoked by novel image strings did not, however, differ in strength between the first 12 and the final 12 alternation trials in the session ($P = 0.99$, Mann-Whitney U test, $n = 66$).

Familiar images elicited strong driven periodic responses not only at the level of neuronal spiking activity, but also at the level of the local field potential (LFP) as recorded at all 55 sites (26 in monkey 1 and 29 in monkey 2) from which neuronal data were collected. At each site, we computed the average of voltage as a function of time across all familiar-image trials and all novel-image trials. Then, independently for familiar and novel images, we measured the peak-to-peak amplitude of the periodic component at 8.33 Hz in the same 700-ms time window employed for neuronal data analysis. On comparing the amplitude of periodic activity driven by alternating familiar and novel images (**Fig. 2c**), we observed a strong and significant tendency for the amplitude to be greater for familiar images ($P = 2.7 \times 10^{-10}$, Mann-Whitney U test, $n = 55$; Monkey 1, $P = 0.0000017$; Money 2, $P = 0.000052$). Plots representing the average of voltage as a function of time across all sites supported this observation (**Fig. 2d**).
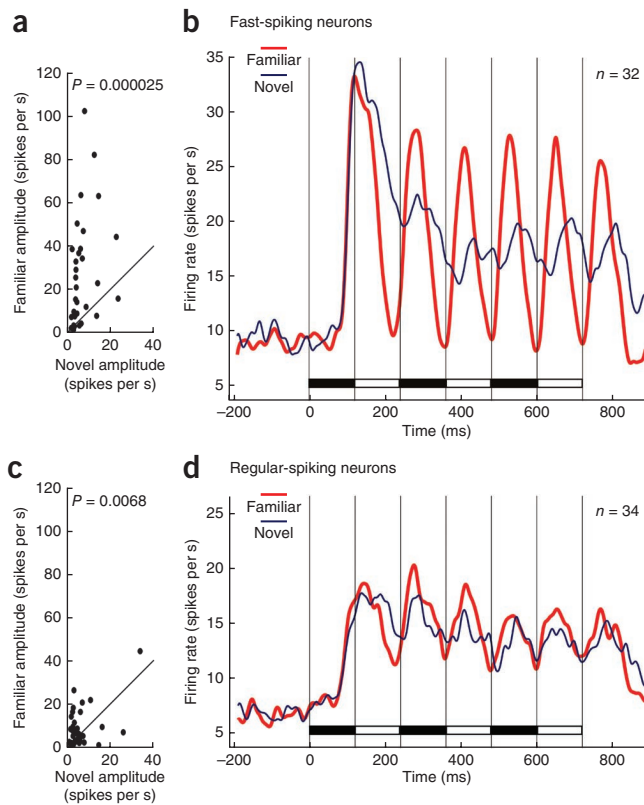
To determine whether the effect depended on the inhibitory or excitatory status of the neuron, we measured the duration of the interval between the most prominent peak and trough in each neuron's action potential waveform[8]. The distribution of the measured durations was clearly bimodal, with a cut at around 400 µs (**Supplementary Fig. 4**). We categorized neurons with durations below and above the cut as fast spiking (putative inhibitory) and regular spiking (putative excitatory), respectively. The effect was very prominent among 32 putative inhibitory neurons (**Fig. 3a,b**). In this population, the difference in peak-to-peak amplitude between familiar and novel displays (mean effect size = 21.2 spikes per s) was highly significant ($P = 0.000025$, paired t test, $n = 32$). The difference achieved significance ($\alpha = 0.05$, Mann-Whitney U test) in 21 of 32 of these neurons considered individually, with 16 exhibiting stronger periodic driving in response to familiar displays. The effect was less evident among 34 putative excitatory neurons (**Fig. 3c,d**). In this population, the difference in peak-to-peak amplitude between familiar and novel displays (mean
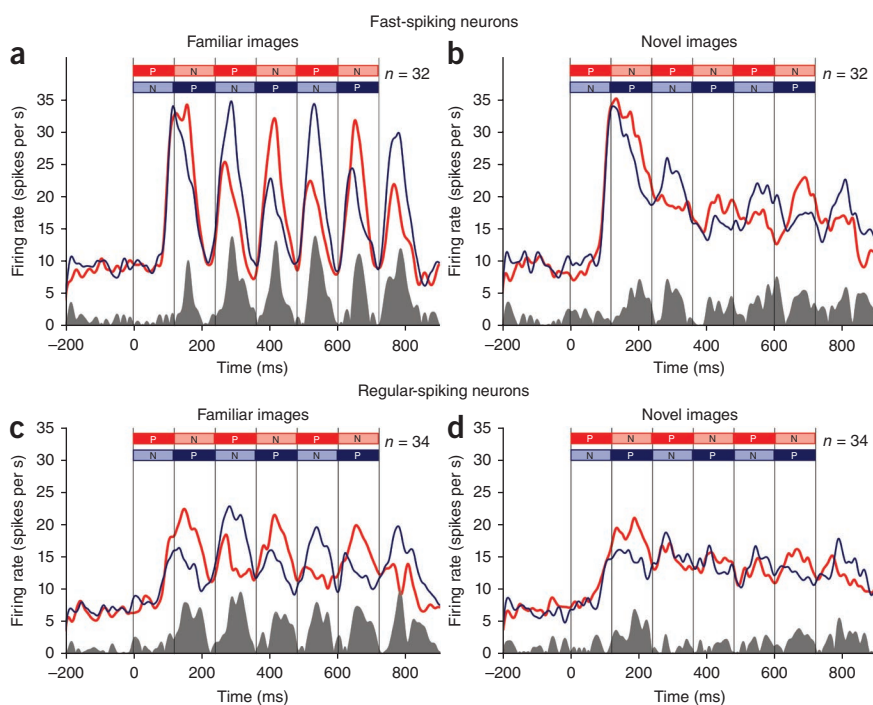
Figure 4 During rapid sequential presentation, neurons represent familiar-image identity more strongly than novel-image identity. (**a**–**d**) The red curve represents the mean firing rate elicited by strings in which the preferred image (P) led and the blue curve represents the mean of the firing rate on trials in which the non-preferred image (N) led. The shaded histogram at the base of each panel represents the absolute difference between the red and blue curves and is a measure of image-selective activity.
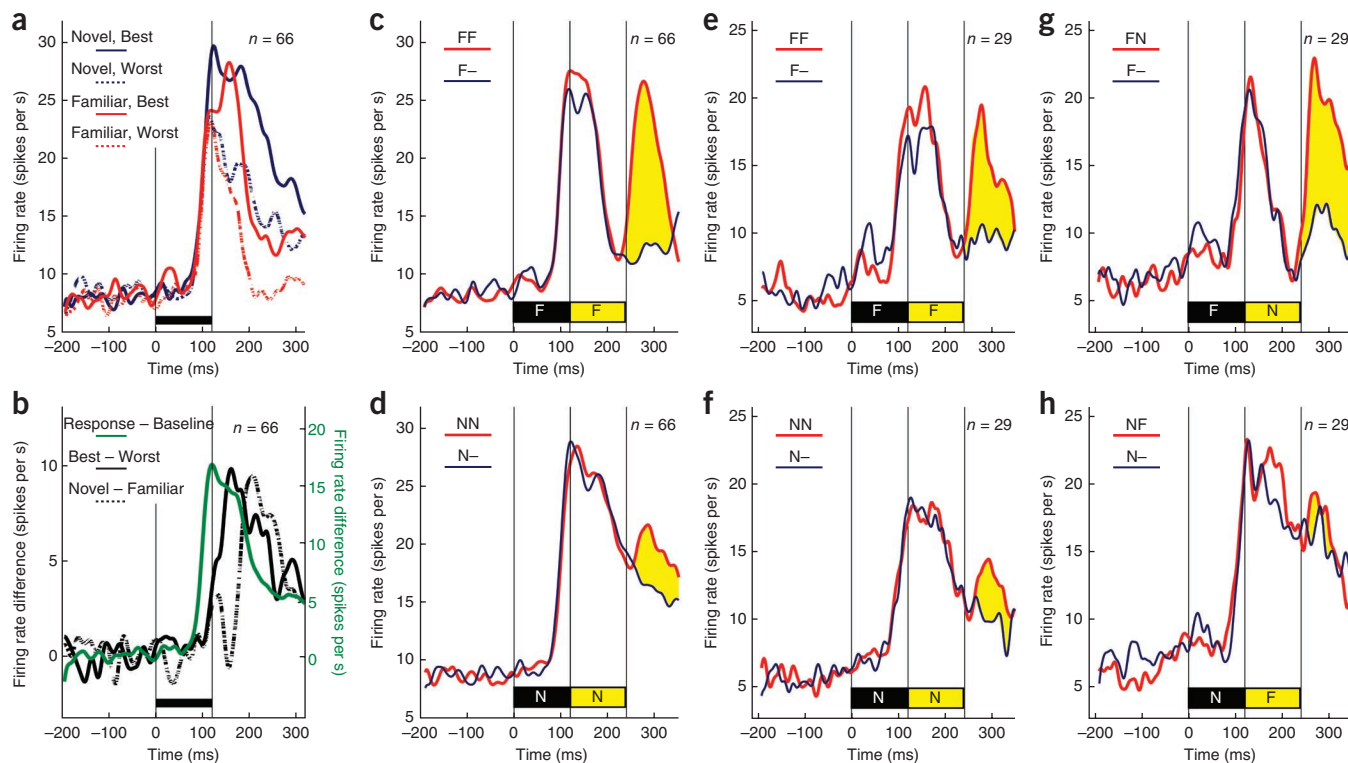
Figure 5 Truncation of the response to a familiar image is accompanied by enhanced responsiveness to an immediately ensuing image. (**a**) The population response to a familiar image was truncated shortly after it begins. Firing rate of 66 recorded neurons is represented as a function of time following image onset. The black bar above the time line indicates the 120-ms duration of the image. Firing is represented independently for trials involving the preferred novel image (solid blue), the non-preferred novel image (dashed blue), the preferred familiar image (solid red) and the non-preferred familiar image (dashed red). (**b**) Difference curves representing the following signals. Visual response: mean firing rate for all images minus time-zero baseline (solid green, scale to right). Image discrimination: firing rate for preferred images minus firing rate for non-preferred images (solid black, scale to left). Novelty detection: firing rate for novel images minus firing rate for familiar images (dashed black, scale to left). (**c**) Strong population response to a familiar probe image immediately following a familiar leading image. The thin blue curve represents the mean across all 66 tested neurons of the firing rate elicited by the leading image followed by nothing and the thick red curve represents the mean firing rate elicited by the leading image followed by the probe image. The difference, indicated in yellow, is the response to the probe. The duration of each image was 120 ms. (**d**) Weak population response to a novel probe image immediately following a novel leading image. (**e,f**) The subset of 29 neurons selected for further testing showed the same pattern as the full population under conditions depicted in **c** and **d**. (**g**) Neurons responded strongly to a novel probe immediately following a familiar leading image. (**h**) Neurons responded weakly to a familiar probe following a novel leading image.

effect size = 4.1 spikes per s) was moderately significant ($P = 0.0068$, paired t test, $n = 34$). The difference achieved significance ($\alpha = 0.05$, Mann-Whitney U test) in 14 of 34 of these neurons considered individually, with 11 exhibiting stronger periodic driving in response to familiar displays. The two neuronal populations were significantly different from each other with regard to the strength of the population effect ($P = 0.00055$, unpaired t test, n1 = 32, n2 = 34). The difference between them with regard to the percentage of neurons exhibiting a significant effect approached significance ($P = 0.081$, χ^2 test, 1 degree of freedom, Yates correction). We conclude that neurons with presumed inhibitory status made the major contribution to the population effect visible in **Figure 2a**.

To determine whether the identity of images in familiar strings was better represented than the identity of images in novel strings, we classified the two familiar images used for testing each neuron as relatively preferred (P) and non-preferred (N) and did likewise for the two novel images. The classification was based on responses to the images presented in isolation on trials otherwise excluded from the analysis. This approach was possible because neurons in TE commonly fire differentially in response to arbitrarily selected images, even though these may not elicit extreme levels of firing[31]. We then constructed plots representing population firing rate as a function of

time during trials with the sequence PNPNPN and NPNPNP (**Fig. 4**). Neurons exhibited image selectivity by responding more strongly to the sequence containing the preferred image at a given phase of the trial. In responding to images after the first one in the string, both putative inhibitory neurons (**Fig. 4a,b**) and putative excitatory neurons (**Fig. 4c,d**) appeared to be more strongly selective for image identity in familiar displays than in novel displays. To determine whether this effect was significant, we computed for each neuron, independently for familiar and novel images, a measure of image selectivity at each position in the sequence: $S_x = P_x - N_x$, where x covered the range 1 to 6, P_x was the mean firing rate $x * 120$ ms to $(x + 1) * 120$ ms after the onset of the string on trials in which the preferred image occurred at sequential position x, and N_x was the equivalent measure on trials in which the non-preferred image occurred at sequential position x. Neurons were equally selective for familiar and novel images when they occupied the leading position in the string ($P = 0.84$, paired t test on S_1, $n = 66$). In contrast, they were more selective for familiar images than for novel images when these appeared later in the string (mean familiar $S_{2-6} = 4.41$ spikes per s, mean novel $S_{2-6} = 1.52$ spikes per s, $P = 0.014$, paired t test, $n = 66$). Among putative inhibitory neurons, this effect approached significance (mean familiar $S_{2-6} = 4.93$ spikes per s, mean novel $S_{2-6} = 1.97$ spikes per s, $P = 0.12$, paired t test, $n = 32$).
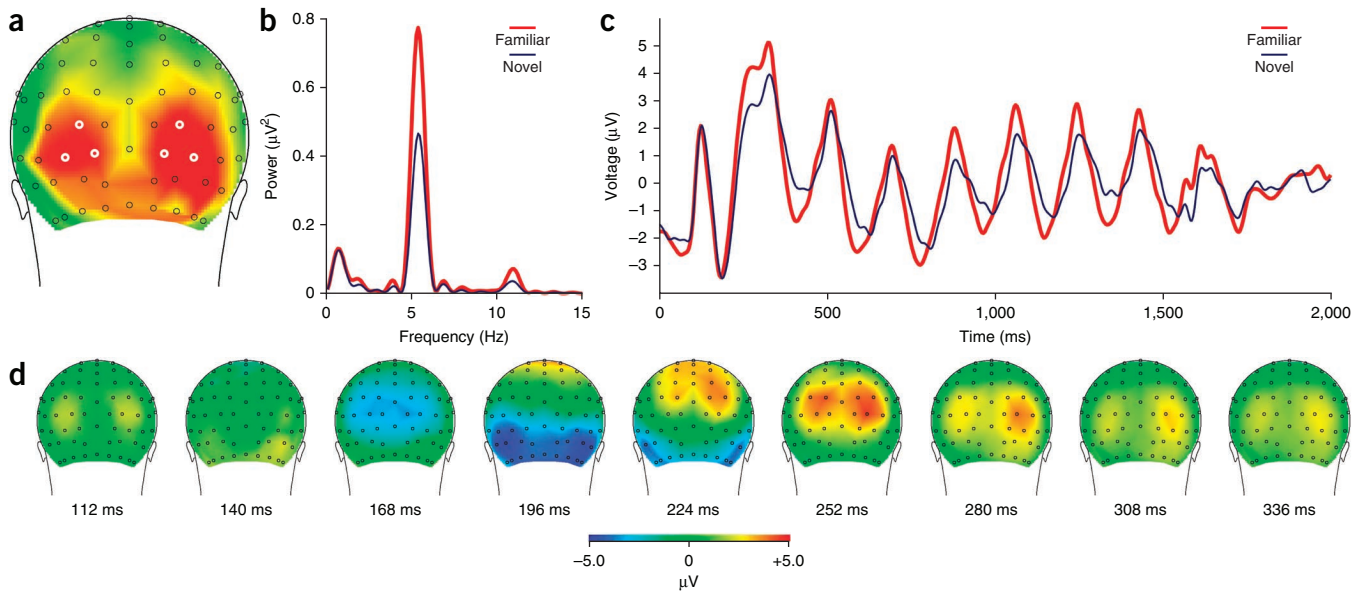
Figure 6 In human subjects, rapidly alternating familiar images elicit stronger periodic scalp responses than rapidly alternating novel images. (**a**) Saturated red indicates scalp locations at which, in the group data, the difference in strength achieved significance at a level of $P < 0.001$ with Bonferroni correction for multiple comparisons. (**b**) Frequency domain analysis. Power at the driving frequency of 5.45 Hz was greater for familiar (red) than for novel (blue) image strings. (**c**) Time domain analysis. The peak to trough excursion was greater on average for familiar (red) than for novel (blue) strings. Traces in **b** and **c** represent measures obtained by averaging across the six electrodes circled in white in **a**. (**d**) The scalp response elicited by an image that appeared at time zero and remained visible for 600 ms. The dorsal-to-ventral progress of the leading negative component (blue) presumably reflects the spread of activation from primary visual cortex to visual areas of higher order. The scalp region at which familiar image strings elicited significantly stronger driven periodic activity, visible as red in **a**, was displaced laterally and ventrally from the zone of earliest negativity, visible as blue at 168 ms in **d**.

Among putative excitatory neurons, it achieved significance (mean familiar $S_{2-6} = 3.93$ spikes per s, mean novel $S_{2-6} = 1.10$ spikes per s, $P = 0.037$, paired t test, $n = 34$). The difference between the two populations was not significant ($P = 0.63$, unpaired t test on familiar S_{2-6} minus novel S_{2-6}, n1 = 32, n2 = 34). We conclude that population activity represents the identity of familiar images more effectively than the identity of novel images during rapid serial visual presentation.

What could be the mechanistic underpinning of this effect? As a step toward answering this question, we measured the dynamics of neuronal responses to familiar and novel images presented in isolation. For the four conditions obtained by crossing preference status (preferred or non-preferred as defined on the basis of trials otherwise excluded from analysis) with training status (familiar or novel), we plotted the mean population firing rate as a function of time following image onset (**Fig. 5a**). The plots revealed a progression of events over time. The rate of firing was initially the same regardless of condition. Then, firing began to differentiate between preferred and non-preferred images. Finally, within each preference category, firing began to differentiate between novel and familiar images. The visual response, measured as the mean firing rate across all conditions, attained half-peak height at 99 ms, the difference between firing rates elicited by preferred and non-preferred images attained half-peak height at 129 ms and the difference between firing rates elicited by novel and familiar images attained half-peak height at 180 ms (**Fig. 5b**). The familiar-novel difference first achieved statistical significance at 180 ms ($\alpha = 0.05$, paired t tests on successive 20-ms bins, $n = 66$). The delay in onset of the familiarity effect was evident among both putative inhibitory neurons (**Supplementary Fig. 5a,b**) and putative excitatory neurons (**Supplementary Fig. 6a,b**). We conclude that neuronal responses to familiar images, although strong at the outset, undergo truncation at a delay of around 80 ms.

We next asked whether, after truncation of the response to a familiar image, the ability of neurons to respond to a new image was restored. To answer this question, we compared firing elicited by two familiar images presented in immediate succession for 120 ms each (combining data from sequences F_1F_2 and F_2F_1) to firing elicited by the leading image alone (combining data from $F_1\bullet$ and $F_2\bullet$, where the bullet indicates a blank screen). We took the difference between firing under the two conditions to represent the response to the trailing image. The trailing image in a familiar sequence indeed elicited a strong response (**Fig. 5c**). On carrying out an identical analysis on novel image sequences, we found that the response to the trailing image was much weaker (**Fig. 5d**). This effect was highly significant ($P = 0.000018$, paired t test based on firing rates 100–200 ms following offset of the leading image, $n = 66$). With consideration restricted to putative inhibitory neurons (**Supplementary Fig. 5c,d**), it remained significant ($P = 0.0010$, $n = 32$). Among putative excitatory neurons (**Supplementary Fig. 6c,d**), it also achieved significance ($P = 0.015$, $n = 34$). However, it was significantly stronger ($P = 0.047$, unpaired t test, n1 = 32, n2 = 34) among putative inhibitory neurons than among putative excitatory neurons. In the above test, the training status of the leading image was confounded with the training status of the trailing image. The strong response to the trailing image in the FF sequence might have depended on its familiarity rather than on the familiarity of the leading image. To rule out this possibility, we assessed the responses of a subset of 29 neurons (17 in monkey 1 and 12 in monkey 2) in an additional block of trials involving different images in FN and NF sequence. These neurons were representative of the population as a whole in that they showed stronger periodic driving for familiar sequences ($P = 0.024$, Mann-Whitney U test, $n = 29$) and responded more strongly to the trailing image in an FF than in an NN sequence ($P = 0.055$, paired t test on mean firing rate 100–200 ms

after offset of the leading image, n = 29; **Fig. 5e,f**). Upon testing with hybrid sequences, they responded strongly to the novel trailing image under the FN condition (**Fig. 5g**), but not to the familiar trailing image under the NF condition (**Fig. 5h**) with the difference achieving significance ($P = 0.0050$, paired t test on mean firing rate 100–200 ms after offset of the leading image, n = 29). The eight putative inhibitory neurons and 21 putative excitatory neurons making up the sample were statistically indistinguishable with regard to the strength of the effect ($P = 0.47$, unpaired t test, n1 = 21, n2 = 8). We conclude that visual responsiveness was greater during the trough of suppression following presentation of a familiar image than at the corresponding time following presentation of a novel image regardless of whether the trailing probe was familiar or novel. This phenomenon, occurring iteratively, presumably gives rise to the enhanced ability of familiar images to elicit periodic driven activity under conditions of rapid serial visual presentation (**Fig. 2a**).

To investigate whether a comparable effect occurs in humans, we exposed four observers to 16 digitized images under conditions identical to those imposed during the exposure phase of the monkey study. A participant viewed each image at least 360 times over the course of at least 4 weeks. At the end of the exposure period, using scalp arrays containing 128 electrodes, we recorded evoked potentials elicited in each individual by rapidly alternating sequences of familiar or novel images. On each trial, a pair of images was presented in back-to-back alternation for 1,440 ms. The images succeeded each other at an interstimulus interval of either 180 ms (corresponding to a frequency of 5.56 Hz) or 120 ms (corresponding to a frequency of 8.33 Hz). At the lower, but not at the higher, frequency, familiar images elicited significantly stronger periodic responses than did novel images. The effect achieved significance at multiple electrodes in each observer, with the number of significant electrodes ranging across individuals from a minimum of 16 to a maximum of 45 (t test with α = 0.05, Bonferroni-corrected for 129 comparisons, based on power at the driving frequency 500–1,500 ms after display onset). At no electrode did periodic driving by novel images significantly exceed periodic driving by familiar images. Electrodes exhibiting a significant effect of familiarity in the group data (**Fig. 6**) coincided with those exhibiting a comparatively late (196 ms) negativity (**Fig. 6d**). This pattern of scalp distribution suggests that the effect arose in occipital and temporal extrastriate visual areas. The effect was of substantial magnitude, as judged by reference to spectral (**Fig. 6b**) and time series (**Fig. 6c**) measures based on six electrodes, three in each hemisphere, centered in the zone showing the effect (**Fig. 6a**).

DISCUSSION

The central finding of this study is that image familiarization sharpened the dynamics of neuronal visual responses in monkey TE. Under conditions of rapid serial presentation, neurons gave stronger and more selective periodic responses to strings of familiar images than to strings of novel images. The enhancement of response strength was greater among putative inhibitory neurons than among putative excitatory neurons, whereas the enhancement of selectivity was of roughly the same magnitude. The fact that the effect was stronger in inhibitory than in excitatory neurons may cast light on the nature of the underlying circuitry, as we discuss below. The fact that excitatory neurons exhibited the effect, although in a weaker form than inhibitory neurons, is of interest because it indicates that the effect has the potential to propagate beyond TE via excitatory projections to other areas. These findings are, to the best of our knowledge, without precedent. Previous studies characterizing the responses of TE neurons to rapid serial displays did not include any manipulation of image

familiarity. Previous studies manipulating image familiarity did not employ dynamic displays[5,8] or did not characterize the effect of image familiarity on dynamic tracking[29]. At the level of mass potentials, familiarization is known to enhance the visual evoked response at the surface of the occipitotemporal junction in monkeys[33], the N200 and P300 components of the local field potential in TE of monkeys[7], and the N170 component of the human scalp potential[34,35]. However, no previous study has assessed the effect of familiarization on steady-state potentials evoked by sequential displays.

The site of the synaptic changes underlying the effect described here might lie in either TE itself or another area. Familiarity effects are known to occur in several areas outside of TE, including perirhinal cortex[3,4,36], entorhinal cortex[3,4], the hippocampus[3,4] and dorsolateral prefrontal cortex[37]. Regardless of the site of the change, our observations place constraints on the nature of the underlying mechanism. The key change is the sharp truncation of the response to a familiar image. This could be expected to reduce spike-frequency adaptation[38,39] and thereby to enhance the response to an immediately ensuing image. The response truncation itself might originate from any of three general mechanisms: weakening of the excitatory bottom-up pathway[40], strengthening of a recurrent inhibitory pathway[41,42] and weakening of a recurrent excitatory pathway. A mechanism based on weakening of the excitatory bottom-up pathway does not appear to be compatible with the observation that truncation began well after onset of the response. The delay of 80 ms measured in our study is in good agreement with estimates of 50, 80, 70, 90 and 70 ms derived from previous studies[5–8]. A mechanism based on strengthening of a recurrent inhibitory pathway does not, on its face, appear to be compatible with our observations. Because putative inhibitory neurons themselves exhibit strong response truncation, it seems implausible that they could impose response truncation on excitatory neurons. Paradoxical interactions conforming to this pattern do, however, occur in inhibition-stabilized networks[43,44]. A mechanism based on weakening of a recurrent excitatory pathway is the only obvious remaining alternative. Every image is thought to excite a unique subset of TE neurons responsive to the moderately complex features present in that image[45]. The late phase of the visual response might depend in part on recurrent excitatory connections among these neurons. If so, and if repeated exposure leads to reduced excitatory connectivity among the coactive neurons, then the upshot would be an image-specific weakening of the late phase of the familiar-image response. This effect would be contrary to the spirit of classic Hebbian models[40,41,46], in which coactivation enhances excitatory connectivity, but would be harmonious with the principle of efficient coding[47,48].

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

AUTHOR CONTRIBUTIONS
C.R.O. provided oversight and resources for the nonhuman primate study. T.M. and C.R.O. designed the experiment. T.M. trained the monkeys and collected the data. R.Y.C. provided oversight and resources for the human study.

ARTICLES

C.W. and R.Y.C. collected and analyzed the human data. All of the authors participated in the preparation of the manuscript.

1. Ungerleider, L.G. & Mishkin, M. Two cortical visual systems. in *Analysis of Visual Behavior* (eds. D.J. Ingle, M.A. Goodale & R.J.W. Mansfield) 549–586 (MIT Press, Cambridge, Massachusetts, 1982).
2. Sobotka, S. & Ringo, J. Investigation of long-term recognition and association memory in unit responses from inferotemporal cortex. *Exp. Brain Res.* **96**, 28–38 (1993).
3. Fahy, F.L., Riches, I.P. & Brown, M.W. Neuronal activity related to visual recognition memory: long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Exp. Brain Res.* **96**, 457–472 (1993).
4. Xiang, J.Z. & Brown, M.W. Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. *Neuropharmacology* **37**, 657–676 (1998).
5. Freedman, D.J., Riesenhuber, M., Poggio, T. & Miller, E.K. Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cereb. Cortex* **16**, 1631–1644 (2006).
6. Mruczek, R.E.B. & Sheinberg, D.L. Context familiarity enhances target processing by inferior temporal cortex neurons. *J. Neurosci.* **27**, 8533–8545 (2007).
7. Anderson, B., Mruczek, R.E.B., Kawasaki, K. & Sheinberg, D. Effects of familiarity on neural activity in monkey inferior temporal lobe. *Cereb. Cortex* **18**, 2540–2552 (2008).
8. Woloszyn, L. & Sheinberg, D.L. Effects of long-term visual experience on responses of distinct classes of single units in inferior temporal cortex. *Neuron* **74**, 193–205 (2012).
9. Anderson, B. & Sheinberg, D.L. Effects of temporal context and temporal expectancy on neural activity in inferior temporal cortex. *Neuropsychologia* **46**, 947–957 (2008).
10. Sobotka, S. & Ringo, J. Investigation of long term recognition and association memory in unit responses from inferotemporal cortex. *Exp. Brain Res.* **96**, 28–38 (1993).
11. De Baene, W. & Vogels, R. Effects of adaptation on the stimulus selectivity of macaque inferior temporal spiking activity and local field potentials. *Cereb. Cortex* **20**, 2145–2165 (2010).
12. Desimone, R. Neural mechanisms for visual memory and their role in attention. *Proc. Natl. Acad. Sci. USA* **93**, 13494–13499 (1996).
13. Li, L., Miller, E.K. & Desimone, R. The representation of stimulus familiarity in anterior inferior temporal cortex. *J. Neurophysiol.* **69**, 1918–1929 (1993).
14. Liu, Y., Murray, S.O. & Jagadeesh, B. Time course and stimulus dependence of repetition-induced response suppression in inferotemporal cortex. *J. Neurophysiol.* **101**, 418–436 (2009).
15. Lueschow, A., Miller, E.K. & Desimone, R. Inferior temporal mechanisms for invariant object recognition. *Cereb. Cortex* **4**, 523–531 (1994).
16. McMahon, D.B.T. & Olson, C.R. Repetition suppression in monkey inferotemporal cortex: relation to behavioral priming. *J. Neurophysiol.* **97**, 3532–3543 (2007).
17. Miller, E.K. & Desimone, R. Parallel neuronal mechanisms for short-term memory. *Science* **263**, 520–522 (1994).
18. Miller, E.K., Gochin, P.M. & Gross, C.G. Habituation-like decrease in the responses of neurons in inferior temporal cortex of the macaque. *Vis. Neurosci.* **7**, 357–362 (1991).
19. Miller, E.K., Li, L. & Desimone, R. A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* **254**, 1377–1379 (1991).
20. Miller, E.K., Li, L. & Desimone, R. Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *J. Neurosci.* **13**, 1460–1478 (1993).
21. Sawamura, H., Orban, G. & Vogels, R. Selectivity of neuronal adaptation does not match response selectivity: a single-cell study of the FMRI adaptation paradigm. *Neuron* **49**, 307–318 (2006).
22. Sobótka, S. Involvement of single unit activity in inferotemporal and perirhinal cortices in recognition memory of visual objects in the macaque. *Acta Neurobiol. Exp. (Warsz.)* **60**, 219–226 (2000).
23. Sobotka, S. & Ringo, J. Investigation of long term recognition and association memory in unit responses from inferotemporal cortex. *Exp. Brain Res.* **96**, 28–38 (1993).
24. Sobotka, S. & Ringo, J. Mnemonic responses of single units recorded from monkey inferotemporal cortex, accessed via transcommissural versus direct pathways: a dissociation between unit activity and behavior. *J. Neurosci.* **16**, 4222–4230 (1996).
25. Verhoef, B.-E., Kayaert, G., Franko, E., Vangeneugden, J. & Vogels, R. Stimulus similarity-contingent neural adaptation can be time and cortical area dependent. *J. Neurosci.* **28**, 10631–10640 (2008).
26. Vogels, R., Sary, G. & Orban, G.A. How task-related are the responses of inferior temporal neurons? *Vis. Neurosci.* **12**, 207–214 (1995).
27. Meyer, T. & Olson, C.R. Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc. Natl. Acad. Sci. USA* **108**, 19401–19406 (2011).
28. Intraub, H. Rapid conceptual identification of sequentially presented pictures. *J. Exp. Psychol. Hum. Percept. Perform.* **7**, 604–610 (1981).
29. Földiák, P., Xiao, D., Keysers, C., Edwards, R. & Perrett, D.I. Rapid serial visual presentation for the determination of neural selectivity in area STSa. *Prog. Brain Res.* **144**, 107–116 (2004).
30. Freiwald, W.A., Tsao, D.Y. & Livingstone, M.S. A face feature space in the macaque temporal lobe. *Nat. Neurosci.* **12**, 1187–1196 (2009).
31. Kiani, R., Esteky, H., Mirpour, K. & Tanaka, K. Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *J. Neurophysiol.* **97**, 4296–4309 (2007).
32. Allred, S.R. & Jagadeesh, B. Quantitative comparison between neural response in macaque inferotemporal cortex and behavioral discrimination of photographic images. *J. Neurophysiol.* **98**, 1263–1277 (2007).
33. Peissig, J.J., Singer, J., Kawasaki, K. & Sheinberg, D.L. Effects of long-term object familiarity on event-related potentials in the monkey. *Cereb. Cortex* **17**, 1323–1334 (2007).
34. Scott, L.S., Tanaka, J., Sheinberg, D. & Curran, T. The role of category learning in the acquisition and retention of perceptual expertise: a behavioral and neurophysiological study. *Brain Res.* **1210**, 204–215 (2008).
35. Scott, L.S., Tanaka, J., Sheinberg, D. & Curran, T. A reevaluation of the electrophysiological correlates of expert object processing. *J. Cogn. Neurosci.* **18**, 1453–1465 (2006).
36. Hölscher, C., Rolls, E.T. & Xiang, J. Perirhinal cortex neuronal activity related to long-term familiarity memory in the macaque. *Eur. J. Neurosci.* **18**, 2037–2046 (2003).
37. Rainer, G. & Miller, E. Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* **27**, 179–189 (2000).
38. Fuhrmann, G., Markram, H. & Tsodyks, M. Spike frequency adaptation and neocortical rhythms. *J. Neurophysiol.* **88**, 761–770 (2002).
39. Liu, Y.H. & Wang, X. Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron. *J. Comput. Neurosci.* **10**, 25–45 (2001).
40. Sohal, V.S. & Hasselmo, M. A model for experience-dependent changes in the responses of inferotemporal neurons. *Network* **11**, 169–190 (2000).
41. Bogacz, R., Brown, M. & Giraud-Carrier, C. Model of familiarity discrimination in the perirhinal cortex. *J. Comput. Neurosci.* **10**, 5–23 (2001).
42. Moldakarimov, S.B., McClelland, J.L. & Ermentrout, G.B. A homeostatic rule for inhibitory synapses promotes temporal sharpening and cortical reorganization. *Proc. Natl. Acad. Sci. USA* **103**, 16526–16531 (2006).
43. Ozeki, H., Finn, I.M., Schaffer, E.S., Miller, K.D. & Ferster, D. Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron* **62**, 578–592 (2009).
44. Tsodyks, M.V., Skaggs, W.E., Sejnowski, T.J. & McNaughton, B.L. Paradoxical effects of external modulation of inhibitory interneurons. *J. Neurosci.* **17**, 4382–4388 (1997).
45. Tanaka, K., Saito, H., Fukada, Y. & Moriya, M. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J. Neurophysiol.* **66**, 170–189 (1991).
46. Feldman, D.E. Synaptic mechanisms for plasticity in neocortex. *Annu. Rev. Neurosci.* **32**, 33–55 (2009).
47. Lewicki, M.S. Efficient coding of natural sounds. *Nat. Neurosci.* **5**, 356–363 (2002).
48. Olshausen, B.A. & Field, D.J. Sparse coding of sensory inputs. *Curr. Opin. Neurobiol.* **14**, 481–487 (2004).

ONLINE METHODS

Monkeys. Two healthy adult rhesus macaque monkeys participated in the experiments (monkey 1, male, laboratory designation Tu; monkey 2, female, laboratory designation Ec). All experimental procedures were approved by the Carnegie Mellon University Institutional Animal Care and Use Committee and were in compliance with the guidelines set forth in the United States Public Health Service Guide for the Care and Use of Laboratory Animals.

Exposure runs. Each monkey viewed 16 images repeatedly. The images represented natural and man-made objects against a blank background (**Fig. 1a**). Some were adapted from an online resource (A. Oliva, http://cvcl.mit.edu/datasets.html). On an LCD monitor at a viewing distance of 32 cm, the horizontal or vertical axis of each image, whichever was longer, subtended four degrees of visual angle (80 pixels along the vertical axis or 88 pixels along the horizontal axis). Each exposure trial consisted of the following events: fixation on a central spot (300 ms), image at screen center (1000 ms), fixation on a central spot (300 ms) and delivery of reward (**Fig. 1b**). During a single run, consisting of 80 trials, each image was presented five times. The sequence of images in a run was random subject to the constraint that each image had to be presented once in each block of 16 successfully completed trials. Monkey 1 viewed each image in its set 140 times over the course of 28 runs spanning 4 weeks. Monkey 2 viewed each image in its set 160 times over the course of 32 runs spanning 2 weeks.

Selection of images for neuronal testing. For each test carried out at a given recording site, we employed two novel images and two familiar images. The novel images were equivalent to the familiar images in that they represented natural and man-made objects against a blank background and were scaled so as to subtend 4° along the horizontal or vertical axis, whichever was longer. Each pair of novel images was discarded after use at one site. The familiar image pair was selected from the training set at the discretion of the experimenter. Inasmuch as this set contained only 16 items, it was necessary to re-use familiar images across sessions. The number of sessions during which each familiar image was used ranged from 5 to 9 and was 6.5 on average. We could not fully counterbalance image identity against training status across monkeys, by using images familiar to monkey 1 as novel images for monkey 2, because we needed far more than 16 novel images in each monkey. However, in recording from a subset of 12 neurons, we did achieve strict counterbalancing, as described in the Results.

Rapidly alternating stimulation: FF and NN sequences. All neurons were tested with rapidly alternating stimulation. Testing any given neuron required the use of two images selected arbitrarily from the set familiar to the monkey (F_1 and F_2) and two images never before seen by the monkey (N_1 and N_2). The novel images, like the familiar ones, subtended approximately 4° and represented objects extracted from their background. It was feasible to use a small number of arbitrarily selected images because most TE neurons are broadly tuned[49,50]. Trials conformed to eight conditions: four alternation conditions and four gap conditions. The sequence of images under the four alternation conditions was F_1F_2, F_2F_1, N_1N_2 or N_2N_1. Under each condition, the two images were presented centrally in back-to-back alternation three times, with each image being present for 120 ms (**Fig. 1c**). If the monkey maintained central fixation throughout the trial then it received fluid reward. The sequence of images under the four gap conditions was F_1•, F_2•, N_1• or N_2•, where the bullet represents an empty screen. The timing was the same as under the alternation conditions with the exception that the screen was empty during each 120-ms epoch when, in the alternation condition, the second image would have appeared. The gap conditions allowed us to identify visually responsive neurons and to compare responses to familiar and novel images presented in isolation (**Fig. 4a,b**). They also provided a first-image-plus-gap baseline to which to compare the first-image-plus-second-image condition in computing the response elicited by the second image (**Fig. 4c–h**). During a single run, over the course of 96 trials, each condition was imposed 12 times. The sequence of conditions in a run was random subject to the constraint that each condition had to be imposed once in each block of eight successfully completed trials. Because each pair of novel images, N_1 and N_2, was employed during only one run, the monkey's entire experience of N_1 (or N_2) occurred across 12 N_1N_2 trials, 12 N_2N_1 trials and 12 N_1• (or N_2•) trials. The total number of exposures to each novel image was therefore 36.

Rapidly alternating stimulation: FN and NF sequences. A subset of neurons was tested with a variant of the standard program in which the possible sequences were F_1N_1, F_2N_2, N_1F_1, N_2F_2, F_1•, F_2•, N_1• and N_2•. In recording from each neuron, this version was run after the standard version had been run. The images, F_1, F_2, N_1 and N_2, were different from those employed in the immediately preceding standard run. The purpose of this test was to allow analyzing the response to a novel image presented after a familiar image and vice versa. This analysis focused on the epoch encompassing responses to the first and second images in the string (**Fig. 4g,h**).

Recording sites. In each monkey, before recording, a cranial implant was surgically installed by use of standard procedures. The implant held a post for immobilization of the head during recording and a vertically oriented chamber through which the electrode could be introduced via a guide tube into TE along tracks forming a square grid with 1-mm spacing. Recording was carried out in the left hemisphere of monkey 1 and the right hemisphere of monkey 2. At the end of the data collection period, magnetic resonance imaging (MRI) was carried out in each monkey. The location of recording sites relative to gross morphological landmarks was determined by extrapolation from MRI-visible fiducial markers placed at known locations in the chamber. The recording sites occupied the ventral bank of the superior temporal sulcus and the inferior temporal gyrus lateral to the rhinal sulcus at levels A16–19 mm relative to the interaural plane in monkey 1 and A13–16 mm in monkey 2.

Data collection. All aspects of the behavioral experiment (stimulus presentation, eye position monitoring and delivery of water reward at the end of each trial) were under control of a computer running Cortex software (NIMH Cortex). Eye position was monitored by means of an infrared tracking system (model RK-826, ISCAN). The monkey was required during each trial to maintain fixation in a window commensurate in extent with the images being viewed. At the beginning of each day's session, a varnish-coated tungsten microelectrode with an initial impedance of ~1.0 MΩ at 1 kHz (FHC) was introduced into the temporal lobe through a transdural guide tube advanced to a depth such that its tip was ~1.5 cm above TE. The electrode was then advanced by use of a micromanipulator until phasic visual responses were observed. Action potentials of single neurons were isolated from the multi-neuronal trace by use of a commercially available spike-sorting system (Plexon). Action potential waveforms were recorded during the experiments and spike sorting was performed offline using commercially available software (Plexon). The raw signal was passed through a 4-pole filter with a high frequency cut-off of 170 Hz and stored continuously with 1-ms resolution for the analysis of LFP signals.

Neuronal data set. We monitored neuronal responses at 55 recording sites (26 in monkey 1 and 29 in monkey 2). We selected a neuron for inclusion in the data set if, for at least one of the gap conditions (F_1•, F_2•, N_1• and N_2•) in the standard version of the rapidly alternating stimulation program, it satisfied the following criterion. The mean firing rate during a response epoch 80–320 ms following onset of the first image exceeded the mean firing rate during a baseline epoch 160 ms before to 80 ms after the onset of the image at a level of $P < 0.05$ in a paired one-tailed t test. Out of 69 recorded neurons, 66 met this criterion (29 in monkey 1 and 37 in monkey 2).

Statistical analysis of stimulus-driven periodic activity. The aim of this procedure was to quantify, for each neuron and for the LFP at each recording site, the strength of the rapid periodic response elicited by the alternation of two familiar images or two novel images. After combining data from trials involving the same images in opposite order, we computed the average firing rate (for a neuron) or the average voltage (for an LFP site) as a function of time in 1-ms bins during a period selected to encompass responses to the second through the fifth image (200-900 ms from sequence onset). We then carried out a fast Fourier transform on the resulting distribution of values (Matlab spectrogram function) and extracted power and phase at the driving frequency of 8.33 Hz. We converted power (P) to peak-to-peak amplitude in spikes per seconds (A) by means of the formula $A = (2P)^{0.5}$ and converted phase (Ph) to latency in milliseconds (L) of the initial positive peak relative to onset of the second image by means of the formula $L = 80\ ms + K$, where $K = -Ph/3\ ms$ [if $Ph < 0$] else $K = (Ph/3 - 120)\ ms$.

To compare amplitude distributions obtained under distinct conditions, we employed a two-tailed Mann-Whitney U test.

Human electroencephalogram. The human experiment involved, as participants, two of the authors (T.M. and C.R.O.) and two other individuals (E.C. and S.R.) not otherwise involved in the study. Written informed consent was obtained in accordance with the policies of the Institutional Review Board at the University of Pittsburgh. Prior to the collection of electroencephalogram (EEG) data, each subject viewed each of 16 images in that subject's familiarization set at least 360 times over a period spanning at least 4 weeks. Each familiarization trial began with presentation of a fixation spot (300 ms) followed by presentation of the image (1,000 ms) after which the fixation spot reappeared (300 ms). The image was centered at fixation. Images forming the familiarization set for C.R.O. formed the novel set for T.M. and vice versa. Images forming the familiarization set for E.C. formed the novel set for S.R., but the reverse was not true.

EEG recording was carried out in each subject on a single day. Data were collected while each subject completed 240 trials involving alternate presentation at 5.56 Hz, 240 trials involving alternate presentation at 8.33 Hz and 256 trials involving continuous presentation for 600 ms. The 5.56-Hz and 8.33-Hz conditions were imposed in alternate 80-trial blocks until 240 trials had been completed under each condition. The run ended with 256 trials involving continuous presentation. Two subjects completed part or all of a second run because the experimenter judged on the basis of the signal-to-noise ratio that added trials were desirable. The following principles governed trial timing and sequencing.

5.56-Hz alternate presentation. On each trial, a pair of images appeared in back-to-back alternation four times, with an interstimulus interval of 180 ms, for a total duration of 1,440 ms. The intertrial interval, during which a fixation spot was continuously visible, varied randomly in the range 1,500–1,800 ms. Across 120 familiar-image trials, every possible pair of familiar images was presented once and likewise for novel images. Which member of a pair began each alternating sequence was determined arbitrarily. The order in which the conditions were imposed was random with the exception that within each 80-trial block half of the conditions were familiar and half novel.

8.33-Hz alternate presentation. On each trial, a pair of images appeared in back-to-back alternation six times, with an interstimulus interval of 120 ms, for a total duration of 1,440 ms. The principles of design were otherwise identical to those employed for the 5.56-Hz condition with the sole exception that the member of each image pair beginning the alternating sequence under this condition was not the one beginning it under the 5.56-Hz condition.

Continuous presentation. On each trial, a single image was presented for 600 ms. The intertrial interval, during which a fixation spot was continuously visible, varied randomly in the range 900–1,200 ms. The order of trials was random except for the constraint that each of the 32 images had to appear once in each block of 32 trials.

The EEG was monitored with a 128-channel Geodesic Sensor Net and with a sampling rate of 250 Hz. Signals were passed through a 0.01–100-Hz elliptical bandpass hardware filter and were referenced in common to a vertex electrode (Cz). The impedance of all channels was ≤50 kΩ. Collection and storage of the data were under the control of Netstation software (Electric Geodesics). Preprocessing was carried out offline by use of EEGLAB[51]. Data were first visually inspected for grossly aberrant channels, which were removed and estimated using spherical spline interpolation. The data stream from alternation blocks was broken into single-trial segments 2,500-ms-long extending from 500 ms before display onset to 560 ms after display offset. The data stream from continuous blocks was broken into single-trial segments 1,500 ms long extending from 500 ms before display onset to 400 ms after display offset. The voltage during each segment was adjusted by subtraction of the mean measured from 350 to 150 ms before display onset. A segment was rejected if 20 or more channels met any of the following criteria: (1) difference between the maximal and minimal voltages >200 μV, (2) difference between any pair of consecutive bins >60 μV, (3) flat voltage trace. Ocular and cardiographic artifacts were removed with the EEGLAB RUNICA. Then the data were passed through a 1–30-Hz bandpass, infinite impulse filter. On a segment-by-segment basis, channels that met any of the three rejection criteria noted above were removed and estimated using spherical spline interpolation. Segments with any channel still above threshold were visually inspected and removed or included at the discretion of the experimenter. Finally, data at each time point were referenced to the average across all electrodes. The number of alternate-presentation trials selected for analysis by these steps ranged across subjects from a minimum of 222 (out of 240 trials run) to a maximum of 435 (out of 480 trials run). The corresponding numbers of continuous-presentation trials were 247 of 256 and 493 of 512.

Quantitative and statistical analyses were carried out with Vision Analyzer software (Brain Products GmbH). Data from each trial were re-baseline corrected by subtraction of the mean measured from 350 to 150 ms before display onset. Power at the driving frequency during each trial was calculated by use of a Hamming windowed, fast Fourier transform applied to the voltage trace 500–1,500 ms after display onset. Statistical comparison of the power under familiar and novel conditions was based, in each subject, on an independent-samples t test applied across all electrodes, with Bonferroni correction for 129 comparisons. Electrodes selected for graphical summary presentation were GSN-128 electrodes 64, 65, 70, medially adjacent to TP9, Po7 and P7 in the 10-20 system, and 91, 95 and 96, medially adjacent to TP10, Po8 and P8.

A **Supplementary Methods Checklist** is available.

49. Lehky, S.R., Kiani, R., Esteky, H. & Tanaka, K. Statistics of visual responses in primate inferotemporal cortex to object stimuli. *J. Neurophysiol.* **106**, 1097–1117 (2011).
50. Rust, N.C. & Dicarlo, J. Selectivity and tolerance ("invariance") both increase as visual information propagates from cortical area V4 to IT. *J. Neurosci.* **30**, 12978–12995 (2010).
51. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004).