J Neurophysiol 118: 374–382, 2017.
First published April 19, 2017; doi:10.1152/jn.00136.2017.

RESEARCH ARTICLE | *Sensory Processing*

Prediction suppression and surprise enhancement in monkey inferotemporal cortex

Suchitra Ramachandran,[1,2,4] Travis Meyer,[1] and Carl R. Olson[1,2,3]

[1]*Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, Pennsylvania;* [2]*Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania;* [3]*Department of Neuroscience, University of Pittsburgh, Pittsburgh, Pennsylvania; and* [4]*Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland*

Submitted 24 February 2017; accepted in final form 6 April 2017

Ramachandran S, Meyer T, Olson CR. Prediction suppression and surprise enhancement in monkey inferotemporal cortex. *J Neurophysiol* 118: 374–382, 2017. First published April 19, 2017; doi: 10.1152/jn.00136.2017.—Exposing monkeys, over the course of days and weeks, to pairs of images presented in fixed sequence, so that each leading image becomes a predictor for the corresponding trailing image, affects neuronal visual responsiveness in area TE. At the end of the training period, neurons respond relatively weakly to a trailing image when it appears in a trained sequence and, thus, confirms prediction, whereas they respond relatively strongly to the same image when it appears in an untrained sequence and, thus, violates prediction. This effect could arise from prediction suppression (reduced firing in response to the occurrence of a probable event) or surprise enhancement (elevated firing in response to the omission of a probable event). To identify its cause, we compared firing under the prediction-confirming and prediction-violating conditions to firing under a prediction-neutral condition. The results provide strong evidence for prediction suppression and limited evidence for surprise enhancement.

NEW & NOTEWORTHY In predictive coding models of the visual system, neurons carry signed prediction error signals. We show here that monkey inferotemporal neurons exhibit prediction-modulated firing, as posited by these models, but that the signal is unsigned. The response to a prediction-confirming image is suppressed, and the response to a prediction-violating image may be enhanced. These results are better explained by a model in which the visual system emphasizes unpredicted events than by a predictive coding model.

inferotemporal; macaque; prediction

AREA TE of the macaque inferotemporal cortex, the terminus of the ventral visual stream (Ungerleider and Mishkin 1982), plays a critical role in visual object recognition. TE neurons respond to complex images with distinct individual patterns of selectivity (Kobatake and Tanaka 1994), in consequence of which, image identity can be decoded from the activity of the population as a whole (Lehky et al. 2014). The visual response properties of neurons in TE are subject to marked influence by visual experience. Training monkeys to discriminate between images (Baker et al. 2002; Jagadeesh et al. 2001; Kobatake et al. 1998), categorize them (Freedman et al. 2003; Sigala and Logothetis 2002), or form associations between them (Messinger et al. 2001; Sakai and Miyashita 1991) induces functional changes, which have the effect of strengthening the representation of image attributes relevant to task performance. Even passive visual experience induces dramatic changes in the functional properties of TE neurons. Repeated viewing of a single image leads to familiarity suppression: the experienced image elicits comparatively weak responses (Freedman et al. 2006; Meyer and Olson 2014; Mruczek and Sheinberg 2007). Repeated viewing of two images close together in time leads to pair coding: neurons responsive to one image tend to respond to the other (Erickson and Desimone 1999; Li and DiCarlo 2008; Miyashita 1988). Finally, repeatedly presenting pairs of images in fixed sequence leads to changes in TE that are the focus of the present study.

Exposing monkeys over days and weeks to displays in which a particular leading image is always followed by a particular trailing image (Fig. 1*A*) strongly affects the visual response properties of neurons in TE. The response to a trailing image is relatively weak if the image appears in a trained sequence and, thus, confirms a prediction based on experience during the training period (blue in Fig. 1*B*), but it is relatively strong if the image appears in an untrained sequence and, thus, violates a prediction based on experience during the training period (red in Fig. 1*B*) (Kaposvari et al. 2016; Meyer and Olson 2011; Meyer et al. 2014a; Ramachandran et al. 2016). This phenomenon could arise solely from prediction suppression (a reduction in response strength when the trailing image confirms a prediction), solely from surprise enhancement (an increase in response strength when the trailing image violates a prediction) or from a combination of the two effects. To decide among these possibilities, we compared responses elicited under the prediction-confirming and prediction-violating conditions to responses elicited under a prediction-neutral condition. Images used in the prediction-neutral condition were presented during training in all possible combinations with equal frequency, so that no leading image disproportionately predicted the occurrence of any particular trailing image (Fig. 1*C*). With the prediction-neutral condition as a basis for comparison, we asked whether firing was suppressed under the prediction-confirming condition (blue fill in Fig. 1*D*) or enhanced under the prediction-violating condition (red fill in Fig. 1*D*). We found that the response to a prediction-confirming image was robustly suppressed, whereas the response to a prediction-violating image showed slight but insignificant enhancement.

Address for reprint requests and other correspondence: S. Ramachandran, Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland (e-mail: suchitra.ramachandran@fmi.ch).

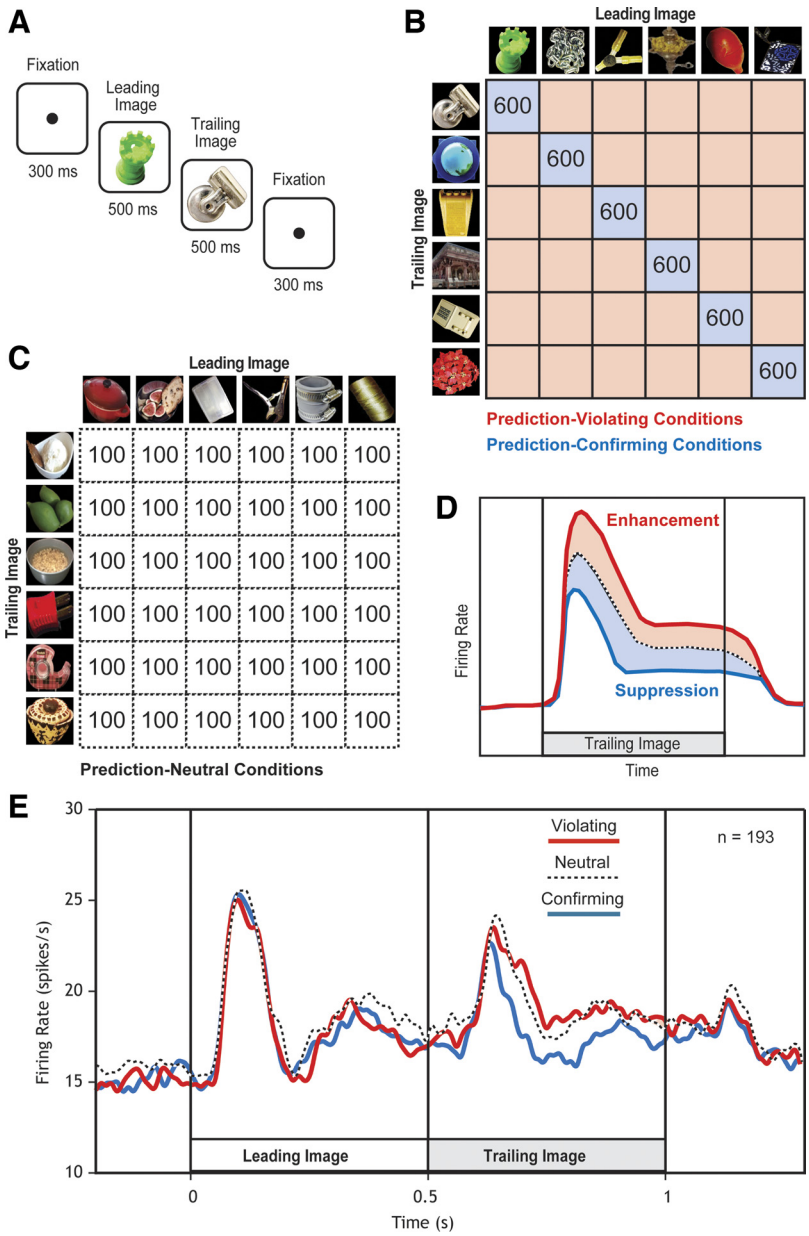

Fig. 1. *A*: sequence of events during a representative trial. Trial structure was identical during training and testing phases of the experiment. *B*: training in the standard paradigm involved presenting each leading image and a paired trailing image in fixed sequence ~600 times (blue cells). Testing involved recording neuronal responses not only to the trained sequences (blue cells) but also to the untrained sequences (red cells). *C*: training in the prediction-neutral paradigm involved presenting every possible sequence of leading image and trailing image ~100 times. The number of exposures to each image was the same as in the standard paradigm, but the conditional probabilities were different. Testing involved recording neuronal responses to all sequences. *D*: it is known from prior experiments that neurons respond to a trailing image more strongly when it violates a prediction (red curve) than when it confirms a prediction (blue curve). The present experiment compared firing under these conditions to firing under the prediction-neutral condition (dashed curve), so as to determine whether the previously reported effect was due to suppression (blue fill: reduced firing under the prediction-confirming condition) and/or enhancement (red fill: increased firing under the prediction-violating condition). *E*: mean population firing rate (193 neurons) as a function of time during trial under prediction-violating (red), prediction-confirming (blue), and prediction-neutral (dashed) conditions.

MATERIALS AND METHODS

Subjects. We studied two adult rhesus macaques: *monkey 1* (male; laboratory designation Tu) and *monkey 2* (female; laboratory designation Ec). Procedures were in accordance with guidelines set forth by the United States Public Health Service Guide for the Care and Use of Laboratory Animals and were approved by the Carnegie Mellon University Institutional Animal Care and Use Committee.

Task. The monkeys were required, during both the training and testing phases of the experiment, to engage in passive viewing of pairs of images presented in sequence. The succession of events in each trial was fixation spot (300 ms), leading image at screen center (503 ms), an 18-ms delay, trailing image at screen center (503 ms), an 18-ms delay, fixation spot (300 ms), and reward delivery (Fig. 1*A*). A trial was aborted without reward if at any point from onset of fixation at the beginning of the trial to offset of the fixation spot at the end of the trial, the monkey failed to maintain fixation within a 4° × 4° central window. The use of a relatively large fixation window was required because, during presentation of the images, which spanned 4°, no fixation spot was visible.

Training. On each training day, the monkey completed one or more training runs. Each run comprised a standard block and a prediction-neutral block. Which block was run first varied from day to day. During a standard block (Fig. 1*B*), each leading image was paired with a particular trailing image. The six sequences were presented six times each for a total of 36 trials. Ordering within the block was random, subject to the constraints *1*) that within each set of six successfully completed trials, each condition had to be imposed once and *2*) the same condition could not be imposed on two consecutive trials. During a prediction-neutral block (Fig. 1*C*), each leading image was paired with all six trailing images. The 36 possible sequences were presented once each for a total of 36 trials. Ordering within the block was random. *Monkey 1* completed 100 runs over the course of 34 days with the number of runs per day ranging from two to nine. *Monkey 2* completed 105 runs over the course of 59 days with the number of runs per day ranging from one to four.

Although the number of trials completed successfully was under precise control for each image sequence, the number of unsuccessful trials was not subject to control because the monkeys introduced

errors arbitrarily by breaking fixation. On rare trials it happened that a fixation break occurred after presentation of the second image. In such a case, although the trial was aborted, the monkey still experienced the two-image sequence. We did not record the frequency of these events in *monkey 2*. In *monkey 1*, the percentage increase in exposures due to these events was 2.7% in the standard paradigm and 2.0% in the prediction-neutral paradigm. Because these events were infrequent and because their frequency was comparable between the two conditions, we do not believe that their occurrence contributed significantly to the different neuronal effects observed after the two forms of training.

Testing. During neuronal data collection, the monkeys completed trials identical to training trials with regard to the timing of events (Fig. 1*A*). The status of the images as leading or trailing was the same as during training. However, during testing, all possible sequences were presented. Data were collected from each neuron over the course of a single run, in which each of six prediction-confirming sequences (blue cells in Fig. 1*B*) was presented five times for a total of 30 trials, each of 30 prediction-violating sequences (red cells in Fig. 1*B*) was presented once for a total of 30 trials, and each of 36 prediction-neutral sequences (Fig. 1*C*) was presented once for a total of 36 trials. Thus, the grand total of trials in a run was 96. Over the course of the run, the various possible sequences were interleaved rather than blocked. The order in which the sequences were imposed was fully random with replacement on error.

Images. All stimuli were digitized images of background-free objects. When presented on an LCD monitor 32 cm from the monkey's eyes, each image subtended 4° of visual angle along whichever axis, vertical or horizontal, was longer. The full stimulus set for *monkey 1* consisted of 12 leading images and 12 trailing images. Images used for the standard paradigm in *monkey 1* were used for the prediction-neutral paradigm in *monkey 2* and vice versa. Counterbalancing the image sets against the training paradigms across the two monkeys reduced the possibility of confounding activity dependent on the prediction status of an image with activity dependent on its identity.

Recording. An electrode was introduced through a vertical guide tube into the left (*monkey 1*) or right (*monkey 2*) temporal lobe. We determined the locations of the recording sites, by extrapolation from MRI-visible fiducial markers within the chamber. The identified locations, as judged by reference to standard maps of cytoarchitecturally defined areas in the temporal lobe, lay within area TE (von Bonin and Bailey 1947) in the ventral bank of the superior temporal sulcus, and on the inferior temporal gyrus and lateral to the perirhinal cortex (Suzuki and Naya 2014). They lay at levels anterior to the interaural plane by 16–19 mm in *monkey 1* and 13–16 mm in *monkey 2*.

During the search for neurons preceding each data collection run, the monkeys viewed image sequences under the same protocol employed during training. The advantage of this approach was twofold: it allowed detecting neurons responsive to images in the training set, and it counteracted any tendency for presentation of prediction-violating sequences during neuronal recording to wash out the effects of training. The number of trials completed during the search phase typically exceeded by a factor of around 10 the number of trials completed during the subsequent recording phase.

Database. We monitored neuronal activity at one site during each recording session. Following the recording session, neuronal spikes were classified offline by use of a hierarchical clustering algorithm (Plexon Offline Sorter) with the requirement that spikes, to be considered as deriving from separate neurons, form clearly distinct clusters. In total, we recorded from 227 neurons (184 from *monkey 1* and 43 from *monkey 2*). We classified a neuron as potentially visually responsive if, for any leading or trailing image, the mean firing rate in a window 50–300 ms following image onset exceeded the mean firing rate in a window extending from 50 ms before to 50 ms after image onset (one-tailed *t*-test, $\alpha = 0.05$). The aim of this procedure was to include any neuron with a hint of visual responsiveness. To establish the presence of statistically significant visual responsiveness would have required correcting for multiple comparisons. This, we did not attempt because the number of trials per image was too few to allow realistic testing. The population meeting this criterion and used in all subsequent analyses consisted of 193 neurons (155 in *monkey 1* and 38 in *monkey 2*), obtained from 95 sites (69 and 26 in *monkeys 1* and *2*, respectively). Traces from the same 95 sites subjected to low-pass filtering with a high-frequency cutoff of 170 Hz formed the local field potential (LFP) database.

Statistical analysis. Analyses of spiking data were based on the population mean firing rate of 100–500 ms after trailing image onset. The key questions concerned differences in this measure among prediction-confirming, prediction-violating, and prediction-neutral conditions. For each of these conditions, the distribution of values across the population of 193 neurons and 95 LFP sites deviated from normalcy (Jarque-Bera test). Before statistically comparing results obtained under any two conditions, we log-transformed the values. Following log transformation, no distribution deviated from normalcy. We then used a two-tailed paired *t*-test ($\alpha = 0.05$) to compare the distributions.

RESULTS

The experiment began with a training period extending over multiple weeks, during which the monkeys passively viewed pairs of images presented in fixed sequence for 500 ms each (Fig. 1*A*). In the standard paradigm, they viewed six sequences ~600 times each (Fig. 1*B*, blue squares). In the prediction-neutral paradigm, they viewed 36 sequences ~100 times each (Fig. 1*C*). The number of exposures to each image was identical between paradigms. This ensured that differences in neuronal response strength could not be explained in terms of image familiarity.

After completion of training, we measured the responses of 193 visually responsive neurons in anterior TE (155 in *monkey 1* and 38 in *monkey 2*) to trailing images presented in both trained and untrained sequences. On randomly interleaved trials, we presented all sequences that could be constructed from images used in the standard and prediction-neutral training sets. The six trained sequences from the standard set (Fig. 1*B*, blue squares) were presented five times each. In these sequences, the trailing image was prediction-confirming. The 30 untrained sequences from the standard set (Fig. 1*B*, red squares) were presented once each. In these sequences, the trailing image was prediction-violating. The 36 sequences from the prediction-neutral set (Fig. 1*C*) were presented once each. These sequences were prediction-neutral in that the trailing image neither confirmed nor violated a strong prediction. In analyzing the results, we focused on the mean population firing rate. This approach was necessary because visual responses in TE are image-selective and because the images used in the standard paradigm and the prediction-neutral paradigm were different. Averaging across neurons with varying patterns of selectivity reduced irrelevant variance in firing rate arising from image selectivity. This approach necessarily leaves open the possibility that an effect present at the population level might vary across neurons making up the population.

Population responses to the leading images employed in the standard paradigm (Fig. 1*E*, red and blue curves) and in the prediction-neutral paradigm (Fig. 1*E*, dashed curve) appeared similar and were statistically indistinguishable ($P = 0.65$, $t = 0.45$, two-tailed paired *t*-test on firing rate 100–500 ms after

leading-image onset, $n = 193$). This observation is important because it validates the assumption that using six images in each category rendered insignificant any differences in firing rate arising from minor image-to-image variations in efficacy. It allows confidence that differences in responses to the trailing images, of which there were also six in each category, depended on their prediction status rather than on their incidental properties. With regard to trailing-image responses, we confirmed prior reports that prediction-violating images elicited a stronger response than prediction-confirming images (Fig. 2*A* and difference plot in Fig. 2*B*). This effect was significant (mean = 1.7 spikes/s; $P = 1.0E\text{-}9$, $t = 6.42$, two-tailed paired *t*-test on firing rate 100–500 ms after trailing-image onset, $n = 193$). The unique contribution of the present study lay in allowing us to compare responses under the prediction-confirming and prediction-violating conditions to responses under the prediction-neutral condition. Firing in the prediction-confirming condition was lower than firing in the prediction-neutral condition, as expected from prediction suppression (blue fill in Fig. 2*A* and difference plot in Fig. 2*C*). This effect was significant (mean = 1.44 spikes/s, $P = 0.014$, $t = 2.47$, two-tailed paired *t*-test on firing rate 100–500 ms after trailing-image onset, $n = 193$). Firing in the prediction-violating condition was slightly higher than in the prediction-neutral condition, as expected from surprise enhancement (red fill in Fig. 2*A* and difference plot in Fig. 2*D*). However, this effect was not significant (mean = 0.29 spikes/s, $P = 0.28$, $t = 1.09$, two-tailed paired *t*-test on firing rate 100–500 ms after trailing-image onset; $n = 193$). We conclude that prediction suppression occurred robustly but that surprise enhancement existed only as a trend.

Having obtained these results from an analysis of the full data set, we checked to ensure that they were consistent across subsets of the data. First, we compared monkeys. On each of the key difference measures referenced in the preceding analysis—violating minus confirming (Fig. 2*B*), neutral minus confirming (Fig. 2*C*), and violating minus neutral (Fig. 2*D*)—the two animals were indistinguishable ($P = 0.42$, 0.40, and 0.69, respectively, using two-tailed unpaired *t*-test). Next, we checked whether the effects persisted in a database restricted to neurons giving a statistically significant visual response to at least one trailing image. The resulting subset consisted of 148 neurons (119 in *monkey 1* and 29 in *monkey 2*). Firing was higher under the violating condition than under the confirming condition ($P = 1.4E\text{-}7$, $t = 5.53$, two-tailed paired *t*-test). Firing was higher under the neutral condition than under the confirming condition ($P = 0.0073$, $t = 2.72$, two-tailed paired *t*-test). Firing was statistically indistinguishable under the vio-

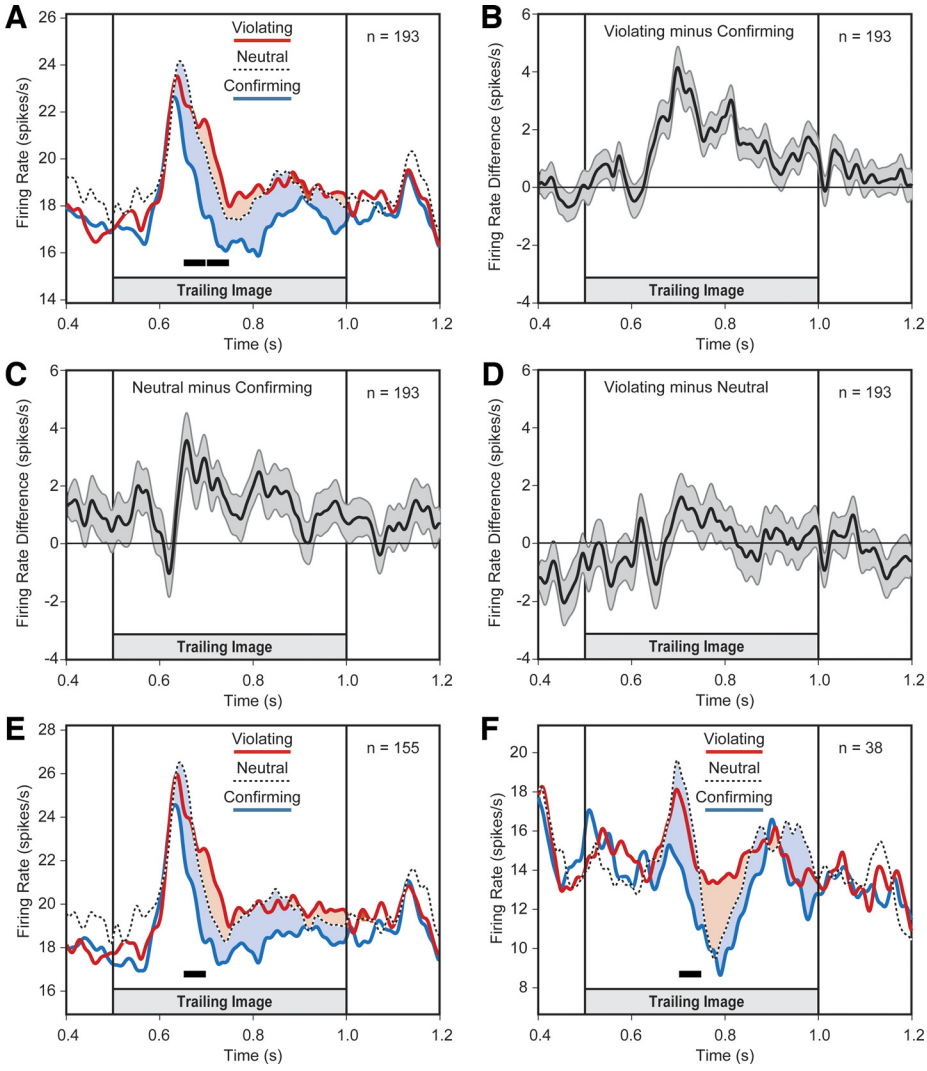

Fig. 2. The trailing-image response under the prediction-neutral condition was markedly greater than under the prediction-confirming condition and slightly less than under the prediction violating condition. *A*: mean firing rate of 193 neurons in response to prediction-confirming (blue), prediction-neutral (dashed), and prediction-violating (red) trailing images. Fill highlights the marked offset of the prediction-neutral firing rate from the prediction-confirming firing rate (blue fill) and the minor offset of the prediction-neutral firing rate from the prediction-violating firing rate (red fill). Black rectangles subjacent to curves indicate 50-ms epochs, during which the population firing rate was significantly reduced under the prediction-confirming condition, as compared with the prediction-neutral condition. In not one 50-ms window was firing significantly enhanced under the prediction-violating condition as compared with the prediction-neutral condition. *B*: mean difference between the prediction-violating and the prediction-confirming population firing rates. The positive displacement of the curve represents the standard prediction effect, as previously reported. *C*: mean difference between the prediction-neutral and the prediction-confirming population firing rates. *D:* mean difference between the prediction-violating and the prediction-neutral population firing rates. *E:* results for *monkey 1*. *F*: results for *monkey 2*. Ribbons in *B–D* represent means ± SE. Conventions in *E* and *F* are the same as in *A*.

lating and neutral conditions ($P = 0.95$, $t = 0.06$, two-tailed paired t-test). Thus, all three observations persisted in the reduced database.

It is evident from inspection of the population histograms that suppression (blue fill) and enhancement (red fill) may have varied over the course of the response (Fig. 2*A*) and that the dependence on time may have been different in the two monkeys (Fig. 2, *E* and *F*). To assess the time-dependent expression of each effect, we carried out post hoc tests on eight nonoverlapping 50-ms windows spanning the 100–500-ms analysis epoch, checking for the presence of statistically significant suppression and enhancement in each window (two-tailed paired t-test, $\alpha = 0.00625$, reflecting Bonferroni correction for eight comparisons). Despite the lack of power arising from the use of a narrow window and correction for multiple comparisons, this analysis revealed two windows with significant suppression in the combined data (Fig. 2*A*, black bars beneath histograms) and one window with significant suppression in each monkey (Fig. 2, *E* and *F*, black bars beneath histograms). It did not reveal any window with significant enhancement.

The firing rate before onset of the response to the trailing image was slightly higher under the prediction-neutral condition than under the other conditions (Fig. 2*A*) because, in one monkey (Fig. 2*E*), the leading images used in prediction-neutral sequences elicited stronger firing late in the response period than the leading images used in the other sequences. This observation raises the issue of whether, in computing response strength, we should have subtracted out the preonset firing rate. To do so would be appropriate if preonset firing represented a baseline on which the response to the trailing image was superimposed. It would be inappropriate if preonset firing were simply truncated and replaced by the trailing-image response. To distinguish between these possibilities, we considered trials involving the best and worst leading images for a given neuron, restricting consideration to cases in which the trailing image was predicted by neither. Elevated firing elicited by the best leading image appeared to carry over into the early phase of the trailing-image response (Fig. 3*A*: yellow fill). However, within the standard analysis window 100–500 ms after trailing-image onset, the resulting increase of firing rate (1.0 Hz) did not achieve significance ($P = 0.83$, $t = 0.22$, $n = 193$, paired t-test). We also considered a case in which the difference in strength between responses elicited by the leading images approximated more closely the small difference observed in our data (elevation of dashed curve over solid curves before onset of the response to the trailing image in Fig. 2*A*). In this analysis, we took advantage of the fact that neurons in TE develop pair-coding as a result of the training procedure used to induce prediction suppression: the leading image predicting the best trailing image elicits a stronger response than the leading image predicting the worst trailing image (Meyer and Olson 2011). We plotted firing rate as a function of time on trials involving the leading images predicting the best and worst trailing images for a given neuron, restricting consideration to trials in which the trailing image was neither of these.

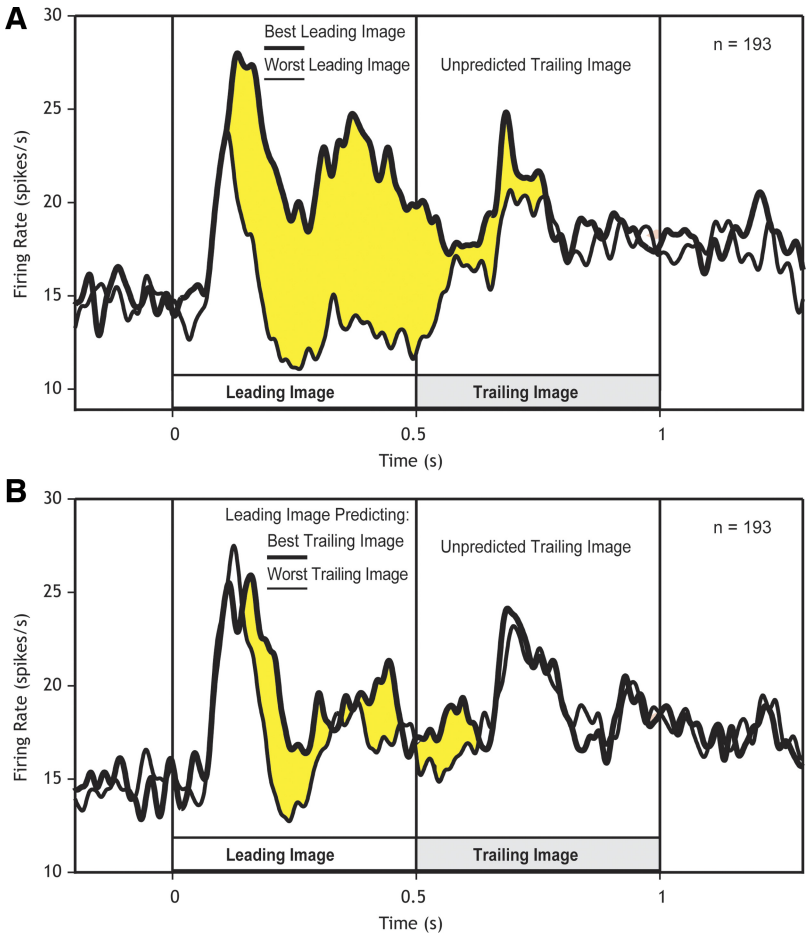

Fig. 3. The strength of the response to the trailing image was unaffected by moderate differences in the level of firing elicited by the leading image. *A*: mean firing rate computed across all 193 neurons on trials in which the leading image was the one to which the neuron responded most strongly (thick curve) or least strongly (thin curve). Consideration was restricted to trials involving the four trailing images associated with neither of those leading images. Differential activity reflecting selectivity for the best leading image (yellow shading) persisted only until the onset of the response to the trailing image. *B*: mean firing rate computed across all 193 neurons on trials in which the leading image predicted the trailing image to which the neuron responded most strongly (thick curve) or least strongly (thin curve). Consideration was restricted to trials involving the four other trailing images. Differential activity reflecting selectivity for the leading image paired during training with the best trailing image (pair coding) persisted only until the onset of the response to the trailing image.

As previously reported, the leading image predicting the best trailing image elicited stronger firing than the leading image predicting the worst trailing image (Fig. 3*B*). However the enhanced activity did not carry over into the succeeding visual response. On trials in which the leading image elicited the stronger response, firing elicited by the trailing image was actually weaker by 0.08 Hz, although the effect did not achieve significance ($P = 0.30$, $t = 1.03$, $n = 193$, paired t-test). We conclude that the response to the trailing image supplanted rather than rode on top of the response to the leading image and, accordingly, that baseline correction would be inappropriate.

The results described up to this point emerged from the analysis of data collapsed across all trials in a run. However, the effects might not have been stable within a run. Over the course of repeated presentations, the strength of the neuronal response to a given image might have changed. For instance, it might have declined due to repetition suppression. To assess response stability across the run, we analyzed response strength as a function of repetition number, independent of prediction status of the images, i.e., prediction-confirming, prediction-violating, or prediction-neutral. Over the course of a run, each of six trailing images from the standard paradigm was presented five times with prediction-confirming status and five times with prediction-violating status, and each of six trailing images from the prediction-neutral paradigm was presented six times. Thus, we could analyze firing rate as a function of sequential rank in the range 1–5 for six prediction-confirming and prediction-violating images and in the range 1–6 for six prediction-neutral images. The best-fit lines relating mean firing rate to repetition number are shown in Fig. 4*A*. The slopes were all slightly positive (0.024, 0.022, and 0.065 under the confirming, violating, and neutral conditions, respectively). This is contrary to expectation based on the occurrence of repetition suppression (Li et al. 1993; Liu et al. 2009; McMahon and Olson 2007; Vogels 2015). Repetition suppression must have been saturated before data collection by hundreds of exposures to the images occurring during training and by further exposures to the training sequences occurring during the search for neurons before each day's recording run. As expected from the near parallelism of the best-fit lines, linear regression analyses ($n = 965$) revealed no significant dependence on repetition number of any effect manifested as a difference in firing rate between conditions. This was true for the classic effect (violating minus confirming: $P = 0.63$), the suppression effect (neutral minus confirming: $P = 0.78$), and the enhancement effect (violating minus neutral: $P = 0.91$).

The results presented in preceding paragraphs were based on the analysis of data collapsed across an experimental period spanning several months in each monkey. However, the effects might not have been stable over the course of the experiment. For example, repeated exposure to prediction-violating sequences during neuronal data collection might have caused an attenuation of effects dependent on the prediction status of an image. To assess long-term stability, we divided the neurons into sextiles based on the phase of the experiment during which they were studied. The first sextile consisted of the first 25 neurons studied in *monkey 1* and the first 6 neurons studied in *monkey 2*; the second sextile consisted of the second 25 neurons studied in *monkey 1* and the second 6 neurons studied in *monkey 2*, and so on. We then plotted response strength for

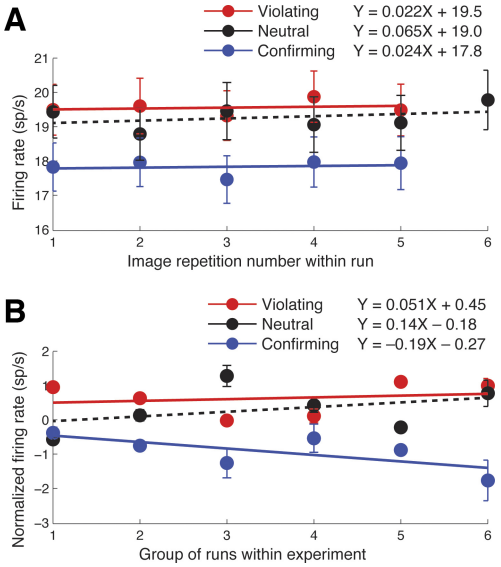

Fig. 4. Prediction-based effects persisted across trials in a run (*A*) and across runs in the full experiment (*B*). *A*: during a data collection run, each trailing image from the standard paradigm was presented five times in a prediction-confirming sequence and five times in a prediction-violating sequence, and each trailing image from the prediction-neutral paradigm was presented six times in a prediction-neutral sequence. Crossing repetition number with prediction status, thus, afforded 16 conditions corresponding to the 16 points in this plot. Each point with its error bars represents the mean and standard error of response strength 100–500 ms after image onset. Each of 193 neurons contributed to each of the 16 measures, and six firing rates were elicited by six distinct trailing images. *B*: we divided neurons into sextiles on the basis of the time during the experiment when they were recorded. The first sextile consisted of the first 25 neurons recorded in *monkey 1* and the first six neurons recorded in *monkey 2*, the second sextile consisted of the second 25 neurons recorded in *monkey 1* and the second six neurons recorded in *monkey 2*, and so on. The sextiles encompassed the first 150 out of 155 neurons recorded in *monkey 1* and the first 36 out of 38 neurons recorded in *monkey 2*. For each neuron, we computed the mean firing rate elicited by images in each category: prediction-confirming, prediction-violating, and prediction-neutral. We then subtracted from each measure the average of the three means. This step eliminated variance due to differences among neurons in mean firing rate. Each point in the plot represents the average across neurons in a given sextile of the mean-normalized response to images possessing a given prediction status. Error bars represent standard error of the mean. Analysis was based on firing 100–500 ms after image onset.

prediction-confirming, prediction-violating, and prediction-neutral images as a function of sextile number. We based the analysis on the firing rate for each condition minus the mean firing rate across all three conditions. This step was required to factor out nuisance variance arising from differences among neurons with regard to mean firing rate. The best-fit lines relating mean firing rate to sextile number are shown in Fig. 4*B*. The slope was negative for prediction-confirming images (slope = −0.19) and positive for prediction-violating images (slope = 0.05) and prediction-neutral images (slope = 0.14). As expected from the differences among the slopes, linear regression analyses ($n = 186$) revealed a significant dependence on sextile number of certain effects manifest as a difference in firing rate between conditions. The increase in the classic effect (violating minus confirming) approached significance ($P = 0.074$). The increase in the suppression effect (neutral minus confirming) achieved significance ($P = 0.041$). The decrease in the enhancement effect (violating minus neutral) also achieved significance ($P = 0.011$). These results did not depend on the particular measure used to characterize the

effects. Normalizing the firing rate under each condition to the mean of the firing rates observed under the three conditions left all trends and statistical outcomes the same with the sole exception that the decline in enhancement no longer achieved significance ($P = 0.25$). We conclude that the observed effects changed in strength, although not in sign over the course of the full experimental period.

Analysis of the local field potential (LFP) is a useful adjunct to the analysis of neuronal spiking activity. Arising from synaptic events spanning thousands of neurons, the LFP provides a low-noise window on events common to a local population. We previously reported that the LFP response to a prediction-confirming image is weaker than the LFP response to a prediction-violating image (Meyer and Olson 2011; Meyer et al. 2014a; Ramachandran et al. 2016). The response measure was the amplitude of the excursion from maximal negativity at around 200 ms to maximal positivity at around 300 ms following trailing image onset. On applying the same measure to LFP responses from the 95 sites at which neuronal data were collected in the present experiment (69 and 26 sites in *monkeys 1* and *2*, respectively), we found, in accordance with previous observations, that response amplitude was greater under the prediction-violating than under the prediction-confirming condition (Fig. 5, *A–C*, red vs. blue curves). This effect was significant (mean difference = 3.4 μV, $P = 1.6E\text{-}4$, $t = 3.93$, two-tailed paired *t*-test on excursion amplitude, $n = 95$). On comparing response amplitude in the prediction-confirming condition to response amplitude in the prediction-neutral condition, we observed marked suppression (Fig. 5, *A–C*, dashed vs. blue curves). This effect was significant (mean difference = 6.9 μV, $P = 9.4E\text{-}7$, $t = 6.5$, two-tailed paired *t*-test on excursion magnitude, $n = 95$). On comparing response amplitude in the prediction-violating condition to response amplitude in the prediction-neutral condition, we found no indication whatsoever of enhancement. In *monkey 1*, response amplitude under the prediction-neutral condition actually exceeded response amplitude under the prediction-violating condition (Fig. 5*B*), an effect for which we have no ready explanation. We conclude that the LFP response was suppressed under the prediction-confirming condition and was not enhanced under the prediction-violating condition.

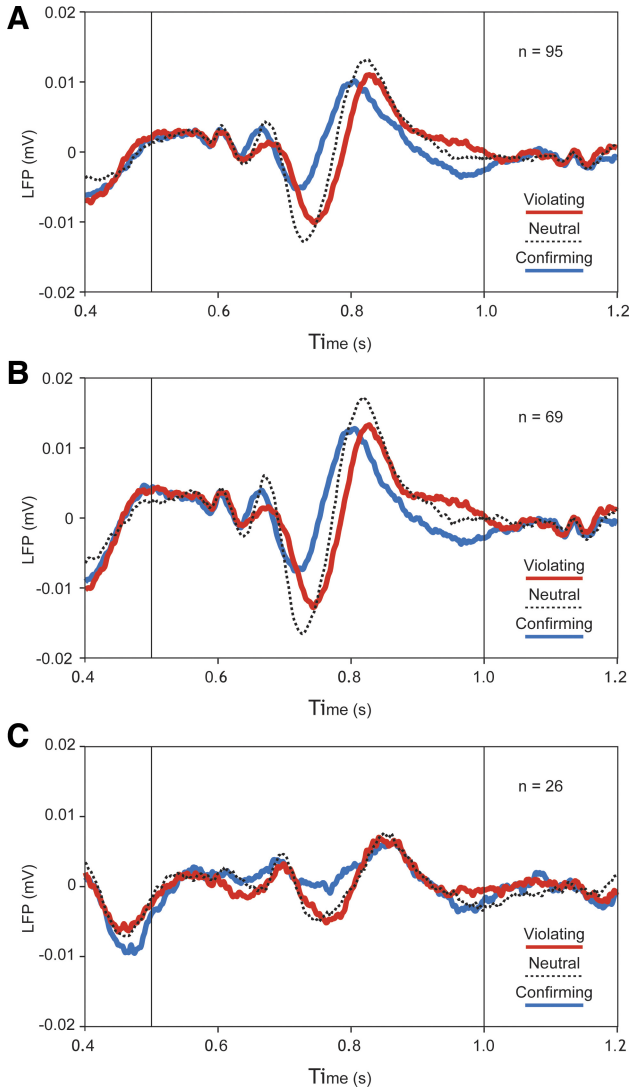

Fig. 5. The local field potential (LFP) response to the trailing image had a much greater peak-to-peak excursion under the prediction-neutral and prediction-violating conditions than under the prediction-confirming condition. *A*: average across all 95 recording sites. *B*: average across 69 sites in *monkey 1*. *C*: average across 26 sites in *monkey 2*.

DISCUSSION

A trailing image confirming a prediction elicits a relatively weak TE neuronal response, whereas a trailing image violating a prediction elicits a relatively strong response (Meyer and Olson 2011; Meyer et al. 2014a; Ramachandran et al. 2016). This pattern could arise from prediction suppression (reduced firing under the prediction-confirming condition) and/or surprise enhancement (elevated firing under the prediction-violating condition). To determine whether one or both of these phenomena were present, we compared response strength under the prediction-confirming and prediction-violating conditions to response strength under a prediction-neutral condition. We found clear indications of prediction suppression but obtained only weak evidence for surprise enhancement.

Evidence for prediction suppression was present at the level of both spiking activity and the LFP. Suppression occurred from the outset of the visual response onward. Prediction suppression as observed here may be related to other suppressive phenomena previously encountered in TE. These include repetition suppression, a reduction of response strength occurring when an image is repeated within a session (Li et al. 1993; Liu et al. 2009; McMahon and Olson 2007; Vogels 2015), and familiarity suppression, a reduction of response strength induced by repeated viewing of an image over many sessions (Anderson et al. 2008; Li et al. 1993; Meyer et al. 2014b; Mruczek and Sheinberg, 2007; Peissig et al. 2007; Woloszyn and Sheinberg, 2012). Repetition suppression and familiarity suppression are thought to result from fatigue-based adaptation at synapses or cell bodies in the pathway leading to TE (Grill-Spector et al. 2006; Kohn 2007; Vogels 2015). Adaptation could also give rise to prediction suppression if the leading image selectively excited and, thus, fatigued synapses or cell bodies responsive to the associated trailing image. This is not out of the question because the training regimen used in the present experiment leads to pair-coding at the level of TE: neurons responsive to the leading image tend also to respond to the associated trailing image, as demonstrated both here and in

a previous study (Meyer et al. 2014b). Cross-adaptation between the leading image and the trailing image, due to their sharing fatiguing neural elements at some point in the processing chain leading to TE, could, thus, give rise to prediction suppression. However, it is not immediately evident how the asymmetry of prediction suppression (image A suppresses the response to image B but not vice versa) (Kaposvari et al. 2016; Meyer et al. 2014b) could arise from symmetric pair-coding (neurons responsive to image A are responsive to image B and vice versa).

Before finally concluding that prediction suppression occurred, we must consider whether differential firing under the prediction-confirming and prediction-neutral conditions could possibly have arisen from physical differences between the images. One obvious possibility is that images presented under the prediction-neutral condition were physically more salient than those presented under the prediction-confirming condition. We can rule out this possibility because the pairing of image set with paradigm was reversed across animals and yet the results were comparable. A second possibility is that we might, by chance, have isolated in each monkey neurons selective for images in that monkey's prediction-neutral set. This is a technical possibility, but its probability is vanishingly small in light of the number of neurons from which we recorded. We note finally that both possibilities contemplated above depend on the existence of consistent differences between images in the two sets with regard to physical properties determining neuronal response strength. By including six images in each set, we intended to minimize the likelihood of there being any consistent differences. This approach had the intended effect, as evidenced by the fact that leading images in the two sets elicited responses of virtually identical strength (Fig. 1*E*).

Evidence for surprise enhancement was markedly weaker than for prediction suppression. On the positive side, there was a trend toward enhancement during the late phase of the visual response in both monkeys. Furthermore, there was a decline over the course of the experiment in the measured strength of enhancement (Fig. 4*B*), as if prediction-violating events were initially capable of eliciting enhanced responses but lost their potency over the course of multiple sessions. On the negative side, there was no hint of enhancement at the level of the LFP. Furthermore, because of an unavoidable feature of experimental paradigm, even the apparent trend toward enhancement at the level of spiking activity might have been a product of suppression. The conditional probability of the trailing image in a prediction-violating sequence was zero during training runs because the particular leading image was never followed by the particular trailing image. The conditional probability of the trailing image in a prediction-neutral sequence could not be set to zero in any practicable design. In our design, it was 0.17 because, during training runs, any particular leading image was followed by any particular trailing image on one-sixth of trials. Because the prediction-neutral image had a small non-zero conditional probability, the response to it might have been slightly suppressed.

Surprise enhancement, although marginal in our study, might still occur robustly in other experimental contexts, for example, ones utilizing other image-sequence protocols or requiring monkeys to engage in active processing of images. On one hand, prior studies have failed to demonstrate enhanced firing in TE under conditions in which an image violates a prediction set up by prior events within a trial or block (Kaliukhovich and Vogels 2014; 2011). On the other hand, TE neurons have recently been shown to respond with enhanced strength to images that violate predictions set up by long-term training on image sequences as in our study (Kaposvari et al. 2016). Using continuous sequences of images with fixed transitional statistics, the authors observed suppression early in the response to a prediction-confirming image and enhancement late in the response to a prediction-violating image, with enhancement being the more pronounced effect. The dominance of enhancement in that study may be related to the timing of the displays. Images were presented back-to-back for 300 ms each in unremitting succession. Consequently, the phasic response to each image was weak. In the already weak response, there may have been little room for further suppression.

The results described here place a definite limit on the degree to which TE neurons can be regarded as engaged in predictive coding of visual events. In predictive coding schemes, as events unfold, the brain forms predictions, monitors outcomes, and generates error signals representing prediction-outcome discrepancies (Bastos et al. 2012). The error signals serve as training signals for fine-tuning the predictive model. For such a scheme to work, the error signal must be signed. For example, the reward prediction error signal posited in classic learning theory (Rescorla and Wagner 1972) and observed in dopamine neurons of the ventral tegmental area (Waelti et al. 2001) represents the value of the delivered reward minus the value of the predicted reward. By direct analogy, TE neurons might carry a visual prediction error signal corresponding to the response associated with the presented image minus the response associated with the predicted image. In this case, the response to an intermediate trailing image should have been low following prediction of the best trailing image and high following prediction of the worst trailing image. No such effect occurred (Fig. 3*B*). We conclude that prediction-based modulation of response strength in TE falls into the category of a surprise signal rather than a prediction error signal. However, there are several formal measures of surprise. Some measures, notably Shannon surprise, are sensitive only to the occurrence of an improbable event (Itti and Baldi 2009). Prediction suppression, as observed in this study, can be regarded as an inverse function of Shannon surprise (Ramachandran et al. 2016). Other measures also register the omission of a probable event. These include Bayesian surprise (Itti and Baldi 2009; O'Reilly et al. 2013) and unsigned prediction error (Belova et al. 2007; Roesch et al. 2010; Roesch et al. 2012). The weakness of surprise enhancement can be taken as indicating that TE neurons are largely insensitive to this form of surprise in the context of our task. The main function of prediction-related activity in TE in this context apparently is to filter out events that are probable and, therefore, uninformative, not to signal the omission of probable events.

ACKNOWLEDGMENTS

We thank Karen McCracken for technical assistance.

GRANTS

This research received support from National Institutes of Health (NIH) RO1 EY-024912, NIH P50 MH-103204, NIH K08 MH-080329, the Pennsylvania Department of Health's Commonwealth Universal Research Enhancement Program, and technical support from NIH Grants P30 EY-008098 and P41RR-03631.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

S.R. and C.R.O. conceived and designed research; S.R., T.M., and C.R.O. performed experiments; S.R. and C.R.O. analyzed data; S.R., T.M., and C.R.O. interpreted results of experiments; S.R. and C.R.O. prepared figures; S.R. and C.R.O. drafted manuscript; S.R., T.M., and C.R.O. edited and revised manuscript; S.R., T.M., and C.R.O. approved final version of manuscript.

REFERENCES

Anderson B, Mruczek REB, Kawasaki K, Sheinberg DL. Effects of familiarity on neural activity in monkey inferior temporal lobe. *Cereb Cortex* 18: 2540–2552, 2008. doi:10.1093/cercor/bhn015.

Baker CI, Behrmann M, Olson CR. Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nat Neurosci* 5: 1210–1216, 2002. doi:10.1038/nn960.

Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. Canonical microcircuits for predictive coding. *Neuron* 76: 695–711, 2012. doi:10.1016/j.neuron.2012.10.038.

Belova MA, Paton JJ, Morrison SE, Salzman CD. Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron* 55: 970–984, 2007. doi:10.1016/j.neuron.2007.08.004.

Erickson CA, Desimone R. Responses of macaque perirhinal neurons during and after visual stimulus association learning. *J Neurosci* 19: 10404–10416, 1999.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J Neurosci* 23: 5235–5246, 2003.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK. Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cereb Cortex* 16: 1631–1644, 2006. doi:10.1093/cercor/bhj100.

Grill-Spector K, Henson R, Martin A. Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci* 10: 14–23, 2006. doi:10.1016/j.tics.2005.11.006.

Itti L, Baldi P. Bayesian surprise attracts human attention. *Vision Res* 49: 1295–1306, 2009. doi:10.1016/j.visres.2008.09.007.

Jagadeesh B, Chelazzi L, Mishkin M, Desimone R. Learning increases stimulus salience in anterior inferior temporal cortex of the macaque. *J Neurophysiol* 86: 290–303, 2001.

Kaliukhovich DA, Vogels R. Stimulus repetition probability does not affect repetition suppression in macaque inferior temporal cortex. *Cereb Cortex* 21: 1547–1558, 2011. doi:10.1093/cercor/bhq207.

Kaliukhovich DA, Vogels R. Neurons in macaque inferior temporal cortex show no surprise response to deviants in visual oddball sequences. *J Neurosci* 34: 12801–12815, 2014. doi:10.1523/JNEUROSCI.2154-14.2014.

Kaposvari P, Kumar S, Vogels R. Statistical learning signals in macaque inferior temporal cortex. *Cereb Cortex* 1–17, 2016. doi:10.1093/cercor/bhw374.

Kobatake E, Tanaka K. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J Neurophysiol* 71: 856–867, 1994.

Kobatake E, Wang G, Tanaka K. Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *J Neurophysiol* 80: 324–330, 1998.

Kohn A. Visual adaptation: physiology, mechanisms, and functional benefits. *J Neurophysiol* 97: 3155–3164, 2007. doi:10.1152/jn.00086.2007.

Lehky SR, Kiani R, Esteky H, Tanaka K. Dimensionality of object representations in monkey inferotemporal cortex. *Neural Comput* 26: 2135–2162, 2014. doi:10.1162/NECO_a_00648.

Li L, Miller EK, Desimone R. The representation of stimulus familiarity in anterior inferior temporal cortex. *J Neurophysiol* 69: 1918–1929, 1993.

Li N, DiCarlo JJ. Unsupervised natural experience rapidly alters invariant object representation in visual cortex. *Science* 321: 1502–1507, 2008. doi:10.1126/science.1160028.

Liu Y, Murray SO, Jagadeesh B. Time course and stimulus dependence of repetition-induced response suppression in inferotemporal cortex. *J Neurophysiol* 101: 418–436, 2009. doi:10.1152/jn.90960.2008.

McMahon DBT, Olson CR. Repetition suppression in monkey inferotemporal cortex: relation to behavioral priming. *J Neurophysiol* 97: 3532–3543, 2007. doi:10.1152/jn.01042.2006.

Messinger A, Squire LR, Zola SM, Albright TD. Neuronal representations of stimulus associations develop in the temporal lobe during learning. *Proc Natl Acad Sci USA* 98: 12,239–12,244, 2001. doi:10.1073/pnas.211431098.

Meyer T, Walker C, Cho, RY, Olson C. Image familiarization sharpens response dynamics of neurons in inferotemporal cortex. *Nat Neurosci* 17: 1388–1394, 2014. doi:10.1038/nn.3794.

Meyer T, Olson CR. Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc Natl Acad Sci USA* 108: 19,401–19,406, 2011. doi:10.1073/pnas.1112895108.

Meyer T, Ramachandran S, Olson CR. Statistical learning of serial visual transitions by neurons in monkey inferotemporal cortex. *J Neurosci* 34: 9332–9337, 2014a. doi:10.1523/JNEUROSCI.1215-14.2014.

Meyer T, Walker C, Cho RY, Olson CR. Image familiarization sharpens response dynamics of neurons in inferotemporal cortex. *Nat Neurosci* 17: 1388–1394, 2014b. doi:10.1038/nn.3794.

Miyashita Y. Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature* 335: 817–820, 1988. doi:10.1038/335817a0.

Mruczek REB, Sheinberg DL. Context familiarity enhances target processing by inferior temporal cortex neurons. *J Neurosci* 27: 8533–8545, 2007. doi:10.1523/JNEUROSCI.2106-07.2007.

O'Reilly JX, Schüffelgen U, Cuell SF, Behrens TEJ, Mars RB, Rushworth MFS. Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proc Natl Acad Sci USA* 110: E3660–E3669, 2013. doi:10.1073/pnas.1305373110.

Peissig JJ, Singer J, Kawasaki K, Sheinberg DL. Effects of long-term object familiarity on event-related potentials in the monkey. *Cereb Cortex* 17: 1323–1334, 2007. doi:10.1093/cercor/bhl043.

Ramachandran S, Meyer T, Olson CR. Prediction suppression in monkey inferotemporal cortex depends on the conditional probability between images. *J Neurophysiol* 115: 355–362, 2016. doi:10.1152/jn.00091.2015.

Rescorla RA, Wagner AR. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical Conditioning II: Current Research and Theory*, edited by Black AH and Prokasy WF. New York: Appleton-Century-Crofts, 1972, p. 64–99.

Roesch MR, Calu DJ, Esber GR, Schoenbaum G. Neural correlates of variations in event processing during learning in basolateral amygdala. *J Neurosci* 30: 2464–2471, 2010. doi:10.1523/JNEUROSCI.5781-09.2010.

Roesch MR, Esber GR, Li J, Daw ND, Schoenbaum G. Surprise! Neural correlates of Pearce-Hall and Rescorla-Wagner coexist within the brain. *Eur J Neurosci* 35: 1190–1200, 2012. doi:10.1111/j.1460-9568.2011.07986.x.

Sakai K, Miyashita Y. Neural organization for the long-term memory of paired associates. *Nature* 354: 152–155, 1991. doi:10.1038/354152a0.

Sigala N, Logothetis NK. Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415: 318–320, 2002. doi:10.1038/415318a.

Suzuki WA, Naya Y. The perirhinal cortex. *Annu Rev Neurosci* 37: 39–53, 2014. doi:10.1146/annurev-neuro-071013-014207.

Ungerleider LG, Mishkin M. Two cortical visual systems. In: *Analysis of Visual Behavior*, edited by Ingle DJ, Goodale MA, Mansfield RJW. Cambridge, MA: MIT Press, 1982, p. 549–586.

von Bonin G, Bailey P. *The Neocortex of Macaca Mulatta*. Urbana, IL: University of Illinois Press, 1947.

Vogels R. Sources of adaptation of inferior temporal cortical responses. *Cortex* 80: 185–195, 2016. doi:10.1016/j.cortex.2015.08.024.

Waelti P, Dickinson A, Schultz W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412: 43–48, 2001. doi:10.1038/35083500.

Woloszyn L, Sheinberg DL. Effects of long-term visual experience on responses of distinct classes of single units in inferior temporal cortex. *Neuron* 74: 193–205, 2012. doi:10.1016/j.neuron.2012.01.032.