Supporting Information

Meyer and Olson 10.1073/pnas.1112895108

SI Results

Neuronal Discrimination Is Improved for Unpredicted Images. To determine how strongly neuronal activity represented the identity of predicted and unpredicted trailing images, we identified, for each neuron, the trailing images eliciting the strongest and weakest responses, giving equal weight to trials in which the image was predicted and unpredicted. Then, independently for trials in which the images were predicted and unpredicted, we constructed plots representing the population response as a function of time. The spread in firing rates between responses to the best and worst images (indicated by filling) was greater for unpredicted images (Fig. S1*B*) than for predicted images (Fig. S1*A*).

To assess whether the greater spread in firing rates translated into more information about image identity, we carried out a signal detection analysis on data from each neuron (1). For each image, we randomly selected five trials out of eight in which it followed its training partner (predicted condition) and all five trials in which it followed another leading image (unpredicted condition). Independently for predicted and unpredicted conditions, we ranked the trailing images in ascending order by response strength. Then we asked how well the activity of the neuron discriminated between the best image (rank 6) and each other image (ranks 1–5). This investigation involved the following steps for each image pair. Twenty-five times, we omitted one response from each distribution of five firing rates, computed the mean of each residual distribution, identified the more effective image as the one for which the mean was greater, and then guessed that the higher omitted firing rate was elicited by the more effective image, so defined. The outcome could be correct, incorrect, or indeterminate. It was indeterminate if the firing rates on the two omitted trials were identical or if the average firing rates of the two residual distributions were identical. We computed the percent correct as $C/(C+I)$ where C and I were the counts of correct and incorrect guesses. Because the design incorporated a cross-validation step (identifying the image on a given trial with a discriminator based on the other trials), the outcome expected in the absence of a signal was 50%.

Deciding, on the basis of the neuronal firing rate, whether the best image or another image had been presented yielded a percent correct score that ranged from an average of around 90%, when the alternative was the worst image, to slightly above 50%, when the alternative was the next-best image (Fig. S2). Signal detection was consistently better for unpredicted trials (red squares) than for predicted trials (blue diamonds), with the difference achieving significance in four of five comparisons (two-tailed paired t test, $\alpha = 0.05$, $n = 81$). We conclude that when an image violated prediction, the gain of the visual response was enhanced, and neuronal activity carried more information about image identity.

Neurons Are Sensitive to Transitional Statistics as Distinct from Joint Statistics. Neuronal responses in the inferotemporal cortex (ITC) might truly be sensitive to the transitional statistics of the training displays (with A_n unidirectionally predicting B_n) or, alternatively, might hinge on joint statistics (with A_n predicting B_n and vice versa). To distinguish between these possibilities, we recorded from 17 neurons (five neurons at five sites in monkey 1 and 12 neurons at nine sites in monkey 2) while presenting the images both in forward order as during training (A then B) and in reverse order (B then A). At the level of neuronal population activity, there was a moderate transitional surprise signal during forward but not reverse presentation (Fig. S3). An example from a single neuron is shown also (Fig. S4). Local field potential (LFP) data from the same sessions are displayed in Fig. 3. We conclude that the surprise signal reflects a violation of transitional statistics and not just of joint statistics.

Transitional Surprise Effect Develops Earlier in Spikes than in the LFP. It is conceivable that the transitional surprise effect develops in visual areas antecedent to the ITC and is relayed by afferents from these areas to the ITC. If so, and if, as is widely thought, the LFP consists of summed synaptic currents, then the LFP should exhibit a transitional surprise effect, based on synaptic events induced by ascending afferents, even before such an effect is evident at the level of spikes generated by ITC neurons. To investigate this possibility, we compared the timing of the surprise signals measured at the level of the LFP and the timing of spiking activity. When we took the time to half height as a measure of signal latency, the LFP surprise signal clearly lagged rather than led the spike-based surprise signal (Fig. S5). This observation provides no support for the notion that the surprise effect is relayed to the ITC by ascending afferents. Too little is known about the neural origins of the LFP for this observation, in itself, to be taken as ruling the notion out.

Transitional Surprise Was Consistent Across Monkeys. The population histogram of Fig. 2*A*, which demonstrates a robust transitional surprise signal, is based on combined data from two monkeys. To demonstrate that the effect was present in both monkeys and of comparable magnitude, we broke down the results by monkey (Fig. S6).

1. Newsome WT, Britten KH, Movshon JA (1989) Neuronal correlates of a perceptual decision. *Nature* 341:52–54.

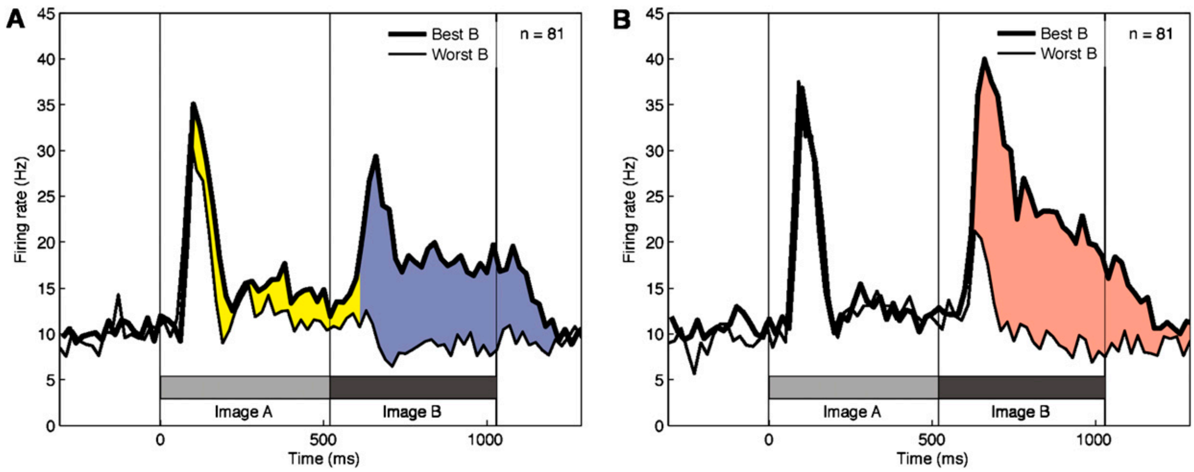

Fig. S1. Population activity discriminated better between unpredicted than between predicted images. (*A*) Trials in which the trailing image confirmed prediction. Mean firing rate across all 81 neurons for trials in which the trailing image was the one to which the neuron responded most strongly (thick curve) or least strongly (thin curve). The strength of the discriminative signal is indicated by the blue area. The difference in firing rate following onset of the leading image (yellow area) reflects experience-induced pair coding (the leading image paired during training with the best trailing image tended to elicit a stronger response than the leading image paired during training with the worst trailing image). (*B*) Trials in which the trailing image violated prediction. The strength of the discriminative signal, indicated by the red area, is greater than in *A*.

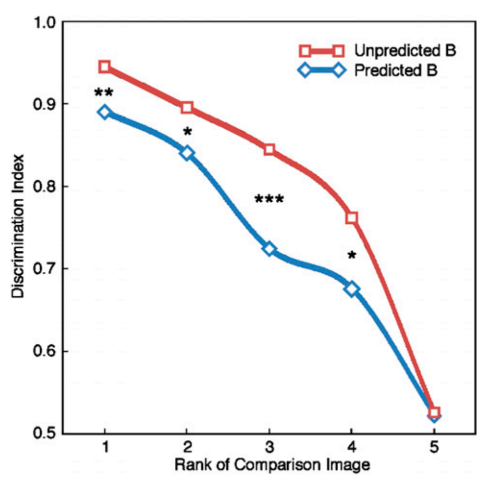

Fig. S2. The ability of neuronal activity to discriminate between the best trailing image (rank 6) and other images (ranks 1–5) was greater in trials when the image was unpredicted (red squares) than in trials when it was predicted (blue diamonds). The mean discrimination index (fraction of trials in which the identity of the image was inferred correctly from the neuronal firing rate) is plotted against the rank of the comparison image. The difference between the values for predicted and unpredicted conditions was significant in four of five cases (paired *t* test; $^*P < 0.05$, $^{**}P < 0.005$, $^{***}P < 0.001$).

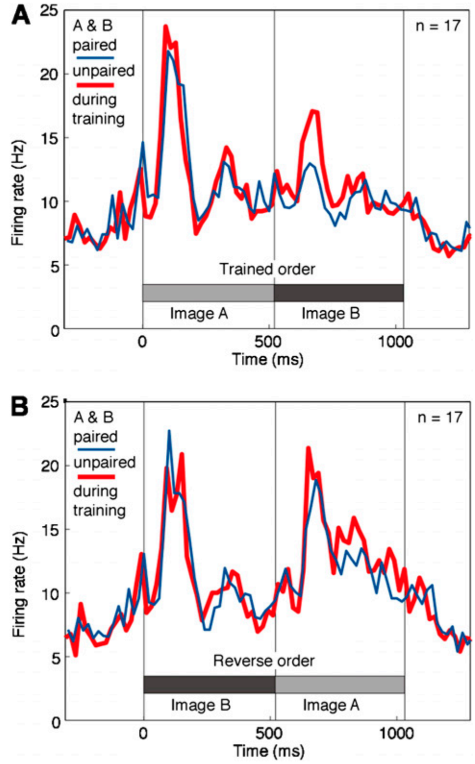

Fig. S3. Population activity of 17 neurons tested with displays in both forward and reverse order. (*A*) Neuronal population activity under the forward condition (A_m as leading image and B_n as trailing image). Just as in the larger population, there was a transitional surprise signal: The response to the trailing image was stronger when it was unpredicted ($m \neq n$, thick red curve) than when it was predicted ($m = n$, thin blue curve). (*B*) Neuronal population activity under the reverse condition (B_n as leading image and A_m as trailing image). The trailing-image response was stronger, but the transitional surprise signal (elevation of thick red above thin blue curve) was attenuated.

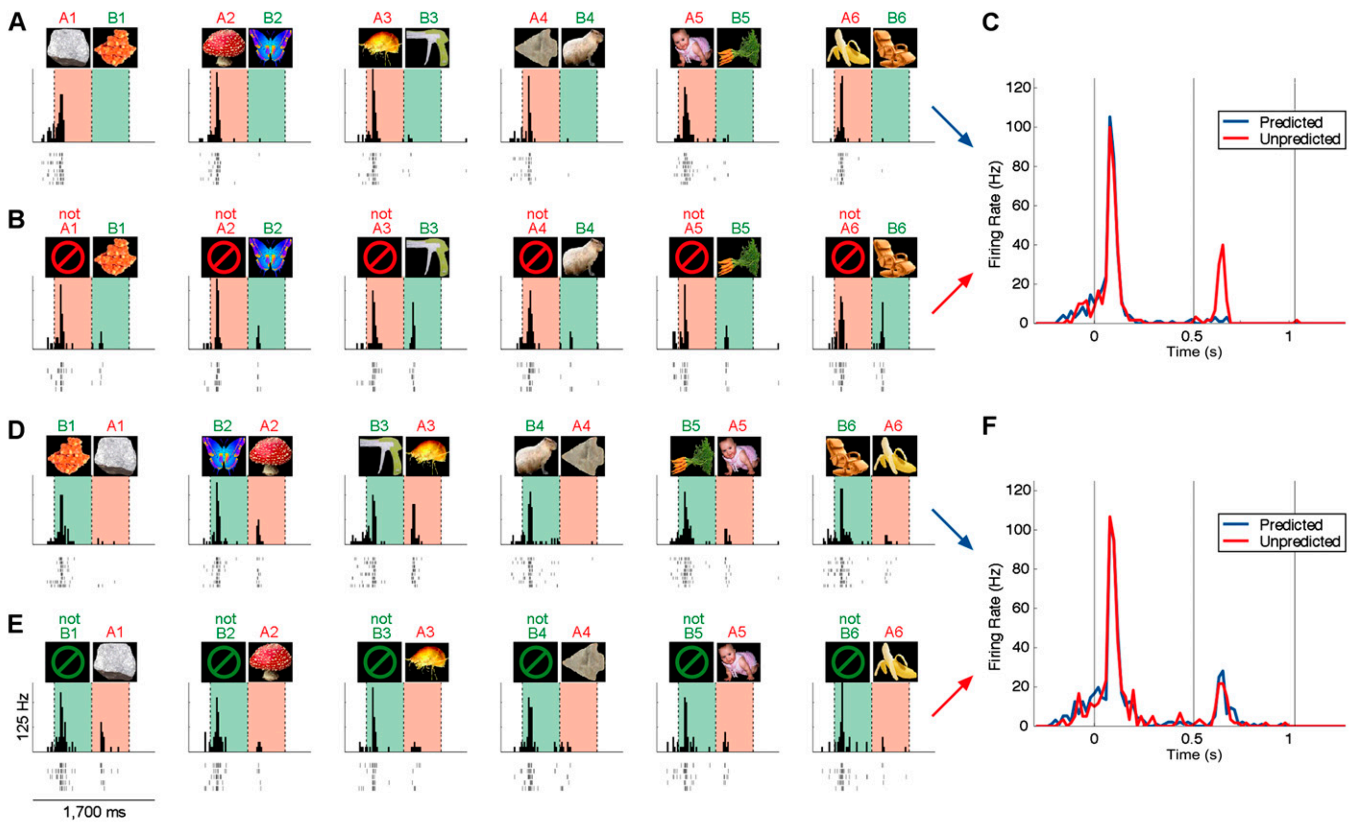

Fig. S4. Responses of a single neuron tested with displays in both forward and reverse order. (*A*) Forward sequence and trained pairing. Each histogram represents mean firing rate as a function of time across eight trials in which the leading image and trailing image were paired as they had been during training (A_mB_n, m = n). The leading and trailing images were visible during periods demarcated by dashed lines and shaded with red and green, respectively. (*B*) Forward sequence and untrained pairing. Each histogram represents firing in five trials in which the trailing image followed the five leading images not paired with it during training (A_mB_n, m ≠ n). (*C*) Mean firing rate across all trial conditions in *A* (predicted: blue) and all trial conditions in *B* (unpredicted: red). Note the presence of a transitional surprise signal (red > blue) in the response to the trailing image. (*D*) Reverse sequence and trained pairing. Each histogram represents mean firing rate as a function of time across eight reverse-sequence trials in which the leading image and trailing image were paired as they had been during training (B_nA_m, m = n). (*E*) Reverse sequence and untrained pairing. Each histogram represents firing in five reverse-sequence trials in which the trailing image followed the five leading images not paired with it during training (B_nA_m, m ≠ n). (*F*) Mean firing rate across all trial conditions in *D* (paired as during training: blue) and all trial conditions in *E* (not paired as during training: red). Note the absence of a transitional surprise signal (red = blue) in the response to the trailing image.

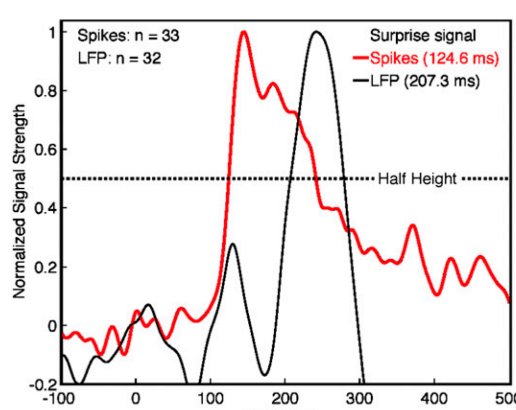

Fig. S5. Timing of the transitional surprise signal at the level of neuronal population activity (thick red curve) based on all 33 neurons in which the signal was significant and in the LFP (thin black curve) based on all 32 sites at which those neurons were recorded. The transitional surprise signal was computed for neurons by taking the mean response to the trailing image when it was unpredicted minus the mean response when it was predicted. The transitional surprise signal was computed for the LFP by taking the mean voltage to the trailing image when it was predicted minus the mean voltage when it was unpredicted. Each curve was smoothed (by convolution with a Gaussian function with 10-ms SD) and normalized to bring the preresponse baseline (0–50 ms after image onset) to 0 and peak response to 1. Each time to half height is indicated in parentheses.

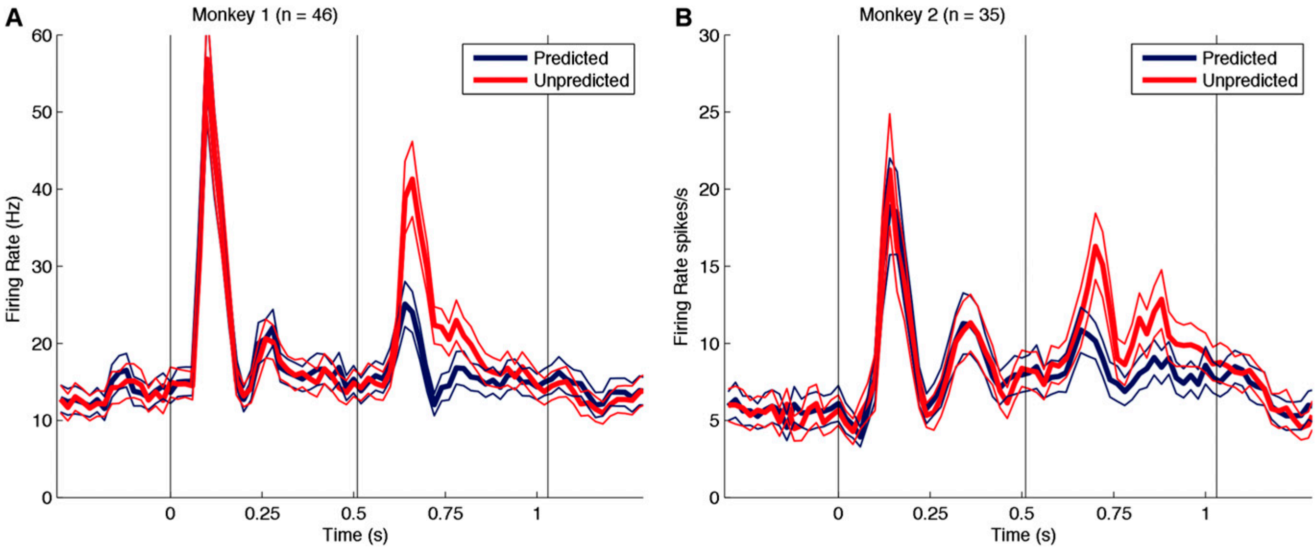

Fig. S6. Consistency of the transitional surprise signal between monkeys. (*A*) In monkey 1, the population firing rate elicited by the trailing image was greater when it followed another image's predictor (red curve) than when it followed its own predictor (blue curve). Data show the mean across all 46 neurons in 20-ms bins. The bounds on each curve represent ±1 SEM. (*B*) This effect also was present in monkey 2. Data show the mean across all 35 neurons.