Supporting Information

Zoccolan et al. 10.1073/pnas.0811583106

SI Methods

Timing of Stimulus Presentation. Rats initiated behavioral trials and reported the stimulus (object) identity as described in *Materials and Methods* (see Behavioral Rig and Task). Once presentation of a visual stimulus was prompted by the animal licking the central touch sensor, its duration depended on the animal's response. The default presentation time (in the event that the animal made no response after initiating a trial) was 3 s. However, if the animal responded correctly before these 3 s expired, the stimulus remained on the monitor for an additional 4 s from the time of the response (e.g., if the animal responded correctly after 500 ms from the stimulus onset, the stimulus was displayed on the monitor for a total of 4.5 s). In the event of an incorrect response, the stimulus was removed immediately and the time-out sequence started. If the animal did not make any response during the default presentation time of 3 s, it still had 1 s, after the offset of the stimulus presentation and before the end of the trial, to make a response.

To prevent rats from making very quick (presumably random) responses, a trial was aborted if the animal's reaction time was lower than 350 ms. In such a case, the animal's response was not evaluated (neither reward or time-out was administered), the stimulus was immediately turned off, and a brief tone was played.

Pseudorandom Stimulus Presentation. In each trial, each of the 2 target objects (shown in Fig. 1*A*) had a 50% chance to be randomly selected for presentation, but the same object (e.g., object 1) was allowed to be presented in a sequence of no more than n consecutive trials (n was set equal to 3 or 4, depending on the training session and the animal). Therefore, every time such a sequence occurred by chance, the other object (e.g., object 2) was forced to be presented in the next (the $n + 1$) trial. This pseudorandom presentation strategy was adopted to prevent the rats from developing a bias for a particular reward port based on the occurrence of a long sequence of consecutive trials with objects having the same identity. All results presented in this study were appreciably the same even if these "predictable" trials were removed (see below *Data Analysis*).

Data Analysis. Our behavioral rig allowed the collection of hundreds of behavioral trials per session (between 200 and 600). As a result, over the course of 5–20 sessions, we were able to collect the response to 50–90 presentations of each object appearance that was tested during phases III and IV of our study. This allowed assessing the significance of the recognition performance of each individual subject for any tested object condition, without the need to pool across animals and/or conditions (1-tailed Binomial test, under the null hypothesis that each animal response is a Bernoulli trial with 0.5 probability of being correct; see Fig. 2*B*, *Right*). To provide a more compact description of the data, we also assessed the significance of the animals' group mean performance over individual (Figs. 2*B*, *left*, and 4*C* and *D*) or pooled (Fig. 3) object conditions (1-tailed *t* test).

As explained above, target objects were presented in a pseudorandom order. As a consequence, in a fraction of trials, the identity of the presented objects could theoretically be predicted from the number of previous occurrences of consecutive trials with the same object identity. Although this fraction of predictable trials was small (12%), and although it seems very unlikely that rats could exploit them to boost their performance (they would need to constantly count the number of consecutive trials with the same object that happen during a session), we verified that the performances obtained by taking into account only the predictable trials were not significantly different from the performances obtained after removal of the predictable trials (combined χ^2 test; $P > 0.05$). Therefore, the rats' performances over the tested object transformation could be safely computed by taking into account all of the collected trials (as was done in Figs. 2–4).
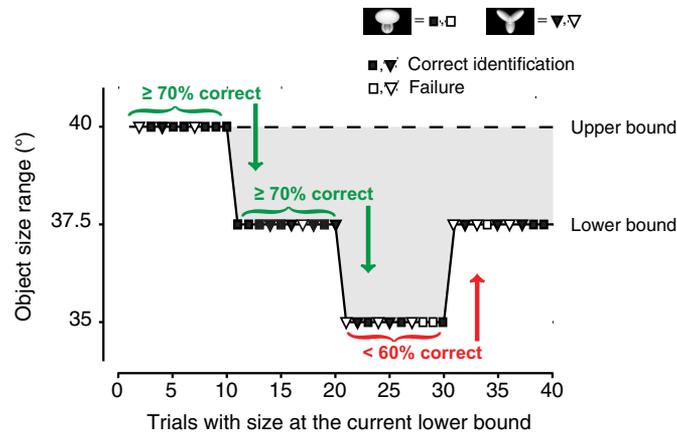
Fig. S1. Illustration of the staircase procedure used to update the range of object sizes that were presented to a subject during a training session in phase II of our study. At any given time during the session, sizes were sampled from a range (gray area) defined by a fixed upper bound of 40° (dashed line; this is the default object size used during phase I of the study) and a lower bound (solid line and symbols) that was determined by the staircase according to the animal's performance. Symbols (squares and triangles) show the identity of the object presented in a given trial (see key in the top of the figure) and the animal's response (filled symbols mean correct identification, while empty symbols mean failure). Note that, for clarity, only trials in which the target objects were presented at the current lower bound of the size range are shown in the figure (e.g., when the sizes' lower bound was 35°, objects were presented with size of 40°, 37.5°, and 35°, but only trials in which the object size was 35° are shown here). The arrows show how the sizes' lower bound was updated (in steps of 2.5°) according to the animal performance over the last 10 presentations of objects with size at the lower bound. The lower bound was decreased if the animal performance was equal or higher than 70% correct (green arrows), while was increased if the performance was equal or lower than 50% correct (red arrow). A similar staircase procedures was used to train the rats to tolerate variation along the other dimension tested in this study (i.e., in-depth azimuth rotation).

Fig. S2. Increasingly larger variation in objects' size and azimuth rotation that one of our subjects learned to tolerate during phase II of our study. Each plot on the left shows the range of image variation (colored areas) produced by one of the object transformations to which the animal was exposed across consecutive, staircase-controlled (see Fig. S1 and *Materials and Methods*) training sessions of the recognition task. The solid lines (and circles) show the minimal/maximal amount of image variation of the target objects over which the rat was able to maintain a performance equal or higher than 70% correct for at least 10 object presentations. That is, a solid line represents the minimal (maximal) lower (upper) bound reached by a staircase over the course (typically at the end) of a training session. The figure shows how, over the course of 10–15 days of training, the rat learned to recognize the target objects across sizes ranging from 40° to 10° (*A*) and azimuth rotations spanning ±60° (*B*). The dashed lines show the default values of the object properties (i.e., size and azimuth rotation) used during phase I. *Right Insets* show the rat's performance over the range of transformations that were tested during one of the training sessions (*n* is the total number of trials presented during the session).

Fig. S3. Comparison of the within-object with the between-object image differences, for the object conditions used in phases III, IVa, and IVb of our study. (*A*) This conceptual diagram helps understanding the rationale of our analysis. Our goal was to measure how much image variation was produced by either changing the appearance (e.g., size and azimuth rotation) of a given object (blue lines), or, instead, the identity of the object, while maintaining size and azimuth fixed (red lines). (*B*) This operational diagram shows how the within-object and the between-object image differences were computed for the object conditions used in phase III. Given an object (object 1, in the example) in a particular appearance (i.e., a size-azimuth conjunction; 40° size and -60° azimuth, in the example), we computed the following metrics: (*i*) the within-object image distance, i.e., the average of the pixel-wise Euclidean distances between this object appearance (image) and all other appearances of the same object that were presented to the subjects during phase III (blue lines); and (*ii*) the between-object image distance, i.e., the pixel-wise Euclidean distance between this object appearance and the appearance of the other object (object 2, in the example), when presented at the same size and azimuth (red line). Both metrics were computed for every object appearance used in phase III, so to obtain 2 sets of values that could be compared pair-wise. A similar procedure was used to compute the within-object and the between-object image differences for the novel lighting conditions used in phase IVa and the novel elevation conditions used in phase IVb. (*C*) The histograms show that, for the sets of object conditions used in phases III (*left*), IVa (*center*), and IVb (*right*), the average within-object distance is larger than the between-object image distance. For each set, this difference was highly significant according to a 2-tailed, paired *t* test.
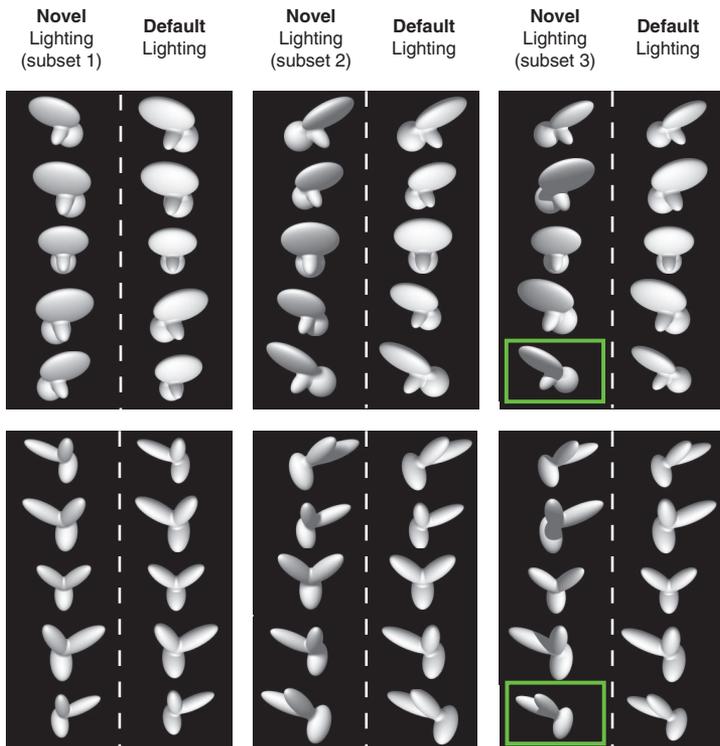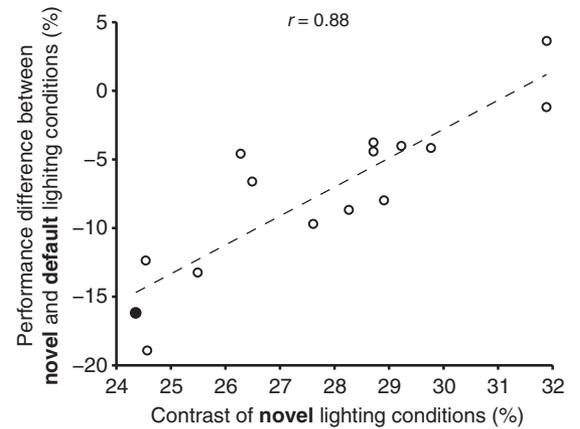
Fig. S4. The novel lighting conditions used during phase IVa of our study. (*A*) The full set of 15 arbitrary size-azimuth conjunctions of the target objects that, during phase IVa of our study, were presented to the rats both under novel lighting conditions and under default lighting conditions (i.e., the same used during phases I–III). As explained in *Materials and Methods*, these 15 novel lighting conditions were divided in 3 subsets of 5 (as shown in the figure), and each subset was presented, interleaved with the default lighting conditions from the previous phase, for 5–10 sessions. Note the large pixel-level image variation produced by the lighting manipulation and how the novel lighting condition images were overall substantially darker and lower contrast than their default lighting counterparts (the lowest-contrast condition for the 2 objects is indicated by the green frames). (*B*) The difference between performance over the novel lighting conditions and performance over their default counterparts (ordinate) is plotted against the contrast of the novel lighting conditions (abscissa). The contrast of an image was quantified by the ratio between the standard deviation of its pixel intensity values and the maximum of the pixel intensity scale (this ratio was multiplied by 100 to obtain a percentage value). For each novel lighting condition, the contrast of the resulting images of objects 1 and 2 was computed and then averaged (this is the value reported on the abscissa of the scatter plot). The performance difference and the contrast of the novel lighting conditions were strongly correlated ($r = 0.88$, $P < 10^{-4}$, 2-tailed *t* test). The black circle refers to the lowest-contrast novel lighting condition (see green frames in *A*), that is the only novel lighting condition for which the animals' performance was not significantly above change (corresponding to the black circle in Fig. 4*C*).

Fig. S5. The novel elevation conditions used during phase IVb of our study. As explained in *Materials and Methods*, 15 novel appearances of each target object were built by varying its elevation (i.e., the in-depth rotation about its horizontal axis) by either ±10° or ±20° and then combining this new in-depth rotation with 15 arbitrary size-azimuth conjunctions from phase III. All these new elevation conditions for both objects are shown in the figure, together with their default elevation counterparts. The 15 novel elevation conditions were divided in 3 subsets of 5 (as shown in the figure), and each subset was presented, interleaved with 45 default elevation conditions from phase III (i.e., the full matrix of size-azimuth conjunctions shown in Fig. 2*A*, with the exception of the 15° size), for 5–10 sessions. Note the substantial variation in the objects' silhouette produced by manipulating the objects' elevation.

Fig. S6. Within- and between-object image differences computed over the responses of a population of simulated V1-like simple cells to our image set. (*A*) This diagram shows how the simulated V1 population was constructed. The V1 simple cells were simulated using a bank of Gabor filters with orientations, spatial frequencies, and receptive field (RF) sizes matching those reported for rat primary visual cortex [Girman et al. (1999) *J Neurophysiol* 82:301–311] and RF centers tiling the visual field. More precisely, we built an array of Gabor filters resulting from all possible combinations of: 3 RF sizes (10°, 20°, and 30°; shown, respectively, in green, red, and cyan in *Left*); 11 orientations (evenly spaced around the clock; see *Right*); 2 phases (0 and π; shown, respectively, as the left and right sets of filters in *Right*); 10 spatial frequencies, ranging from 1 to 10 cycles per RF size and resulting in a 0.03–1 cycles per degree range (see examples in *Right*). This array of Gabor filters was replicated every 5° in both the vertical and horizontal direction over the 60° × 40° span of visual field occupied by our image stimuli (for sake of simplicity, only filters at one particular visual filed location are shown in the figure). The response of a Gabor filter to a given image was computed as the dot product of the filter and the image patch with the same visual field location and size. To simulate the nonlinear response properties of V1 simple cells (i.e., saturation, luminance and contrast normalization, and non-negative firing rates), both the filter and the image patch were normalized to 1 before computing their dot product and negative responses were clipped to 0. (*B*) For each image in our stimulus set, we computed its representation in the space of the simulated V1 population, and we obtained the within-object and the between-object image differences in this space, using the same rationale described in Fig. S3. As shown by the histograms, for the sets of object conditions used in phases III (*Left*), IVa (*Center*), and IVb (*Right*), the average within-object distance was larger than the between-object image distance, and this difference was highly significant according to a 2-tailed, paired *t* test.