

EFFECTS OF PERCEPTUAL LEARNING IN VISUAL BACKWARD MASKING ON THE RESPONSES OF MACAQUE INFERIOR TEMPORAL NEURONS

H. P. OP DE BEECK,^{a,b*} J. WAGEMANS^b
AND R. VOGELS^a

^aLaboratorium voor Neuro- en Psychofysiologie, K. U. Leuven Medical School, Herestraat 49, Leuven, Belgium

^bLaboratory of Experimental Psychology, K. U. Leuven, Tiensestraat 102, 3000 Leuven, Belgium

Abstract—Learning is critical for fast and efficient object recognition in primates. To understand the neuronal correlates of behavioral improvements due to training, we recorded the responses of single neurons in the inferior temporal (IT) cortex of monkeys that were trained to recognize briefly presented, backward-masked objects. First we investigated training effects that are specific to the objects shown during training and that do not transfer to untrained objects. Only one of two monkeys tested showed object-specific training effects at the behavioral level, and only this monkey showed a transient object-specific increase in object selectivity for trained compared with untrained backward-masked objects. However, in each monkey a substantial part of the training effect transferred to untrained objects. To investigate the neural correlates of these object-independent training effects, we compared the neural responses to masked objects in trained monkeys to the responses in untrained monkeys. Training was associated with a reduction of the responses to the irrelevant masking patterns. These findings suggest that extensive training in recognizing backward-masked objects results in neural changes that reduce IT responses to the interfering irrelevant masking patterns and enhance the processing of the relevant objects. © 2006 IBRO. Published by Elsevier Ltd. All rights reserved.

Key words: object recognition, extrastriate cortex, vision, pattern masking, shape perception, monkey.

Studies have shown that extensive exposure to or training with complex objects can result in strong behavioral improvements in discrimination and detection tasks (Goldstone, 1998; Gold et al., 1999; Gilbert et al., 2001). These learning effects reflect changes in perceptual processes, as shown in “perceptual learning paradigms” that involve a systematic degradation of a sensory signal while the overall task and stimulus-response associations have been learned previously (Hall, 1991). These paradigms include the addition of simultaneous luminance noise to stimuli

*Correspondence to: H. P. Op de Beeck, Department of Psychology, K. U. Leuven, Tiensestraat 102, 3000 Leuven, Belgium. Tel: +32-16-325930; fax: +32-16-326099.

E-mail address: hans.opdebeeck@psy.kuleuven.be (H. P. Op de Beeck).

Abbreviations: ANOVA, analysis of variance; DOS, depth of selectivity; IT, inferior temporal; MRI, magnetic resonance imaging; SOA, stimulus-mask onset asynchrony; S.E.M., standard error of the mean.

0306-4522/07\$30.00+0.00 © 2006 IBRO. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuroscience.2006.12.058

(e.g. Gold et al., 1999; Rainer and Miller, 2000) and presenting stimuli for a brief time period followed by mask patterns (backward masking, e.g. Furmanski and Engel, 2000; Grill-Spector et al., 2000), among others. Consistent findings in these paradigms are that signal degradation impairs performance, and that this effect of stimulus degradation on performance can be overcome partly with extensive training.

Monkey inferior temporal (IT) cortex is important for object recognition and categorization (Logothetis and Sheinberg, 1996; Tanaka, 1996; Vogels, 1999; Op de Beeck et al., 2001) and previous studies have suggested changes in IT responses as a consequence of exposure to complex objects (e.g. Li et al., 1993; Sakai and Miyashita, 1994; Ringo, 1996; Kobatake et al., 1998; Sigala and Logothetis, 2002; Baker et al., 2002; Freedman et al., 2006). However, none of these studies involved a perceptual learning paradigm as defined above. In the present study, we studied how training in recognizing backward-masked objects affects the responses of IT neurons. Although the primate visual system enables fast object recognition and categorization (Thorpe et al., 1996), presenting a mask after an object image can interfere strongly with object recognition when the stimulus-mask onset asynchrony (SOA) is less than 60 ms (Furmanski and Engel, 2000; Grill-Spector et al., 2000; Bacon-Macé et al., 2005). Single-unit studies have revealed a similar interference of backward masking with the amount of object selectivity and with the time window in which this object selectivity is found in the responses of IT neurons (Rolls and Tovee, 1994; Kovacs et al., 1995; Rolls et al., 1999). Nevertheless, multi-session training to recognize masked objects greatly improves the ability of human subjects to recognize masked objects, and it has been suggested that such training affects object processing in high-level visual cortex (Furmanski and Engel, 2000; Grill-Spector et al., 2000). Our study is the first to report the neural correlates of training in backward masking at the level of single neurons.

These human studies focused on object-specific training effects, that is, training effects that do not generalize to objects that have not been shown during training. We initially designed our study to investigate these object-specific training effects. As found for humans, we found that training monkeys in the recognition of masked objects improved object recognition at short SOA. However, only a small part of this effect was specific to trained objects (significant in only one of two monkeys), and a substantial part of this behavioral improvement generalized to objects that were not shown during the training period (object-

independent learning). In the only monkey that showed an object-specific training effect, it seemed to be related to a transient and object-specific increase in neuronal selectivity. Object-independent training effects were much more consistent across monkeys, and they were associated with a strong reduction of the responses to the masking patterns. Thus, training reduced the responses to interfering, irrelevant stimuli and in one monkey it enhanced the processing of discrimination-relevant stimuli.

EXPERIMENTAL PROCEDURES

Subjects

Four juvenile male monkeys (*Macaca mulatta*) were used. All procedures were approved by the K. U. Leuven Ethical Committee for animal experiments and followed U.S. National Institutes of Health guidelines. All efforts were made to minimize the number of animals used and their suffering. Separate surgeries were performed in order to implant a plastic head restraint, a scleral search coil, and a plastic recording chamber (Crist Instruments, Hagerstown, MD, USA). Monkeys were anesthetized with ketamine hydrochloride (10 mg/kg, i.m.), given atropine sulfate (0.04 mg/kg, i.m.). An i.v. line was established via the saphenous vein. The monkey was intubated and anesthesia was maintained with isoflu-

rane (1.2 MAC; 50% N₂O/50% O₂). Heart rate and blood oxygenation were monitored throughout the procedure. The monkey was positioned on a warmed blanket, and the head was mounted in a stereotaxic frame (David Kopf Instruments, Tujunga, CA, USA). The stereotactic implantation of the recording chamber was based on pre-operative anatomical magnetic resonance imaging (MRI) scanning. Surgeries were performed under aseptic condition. Antibiotics and analgetics were delivered i.m. after the surgery for several days.

Experimental apparatus

During each session, the monkeys were seated in a primate chair with their head fixed facing a 21-inch monitor. Stimulus presentation and the behavioral task were under control of a PC running house-made software under MSDOS. Eye position was measured through the scleral search coil technique (Robinson, 1963). IT cortex (area TE) was accessed through the recording chamber allowing a dorsal approach. At the start of each recording session, one tungsten microelectrode (FHC, Bowdoinham, ME, USA) was lowered with a hydraulic Narishige microdrive containing a 19 mm grid with a spacing of 1 mm between holes (Crist Instruments) to control electrode position. Signals were amplified and filtered and action potentials from a single neuron were isolated on-line using a Plexon system (Plexon Inc., Dallas, TX, USA). The timing of the action potentials and the stimulus and behavioral events were

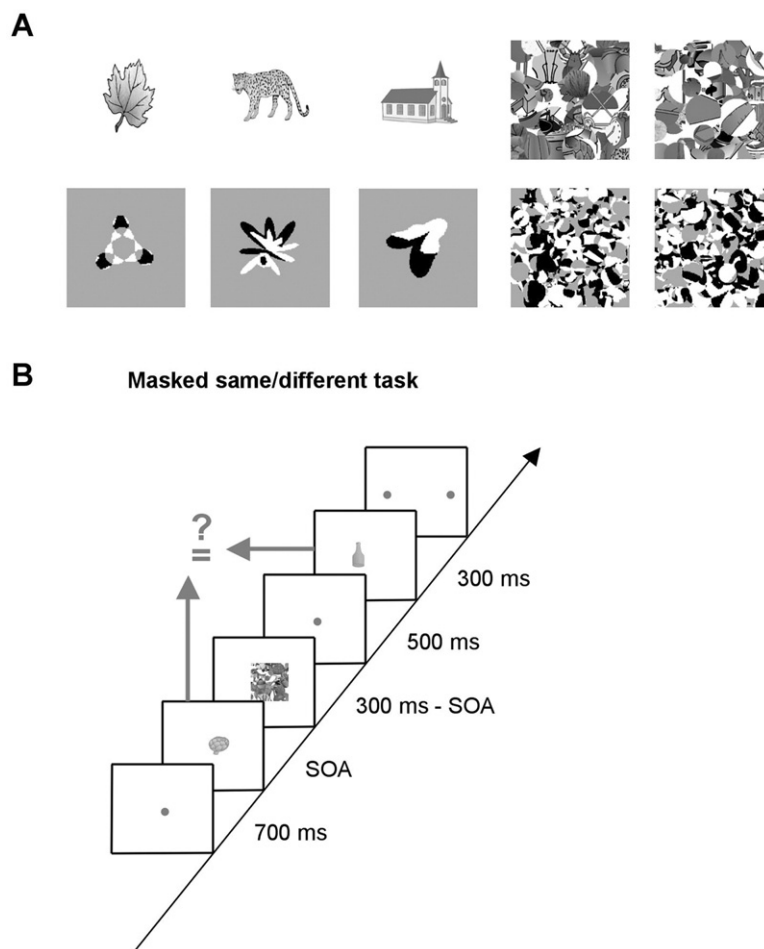


Fig. 1. Stimuli and training task. (A) Examples of the object images and masks. (B) The masked same/different task in which monkeys had to compare two object images.

Table 1. Chronological progression of experiments for each monkey

Monkey	Task			
	Same/different	Fixation	Masked same/different	Contrast same/different
1A	LS		LS TR BT1 TR1 BT2 REC TR2 REC TR3 REC	LS TR4 REC TR5 REC
1B	LS		LS TR BT1 TR1 BT2 REC	
2A	LS	REC	REC	
2B	LS	REC		

Each row shows the phases of each task in chronological order for each monkey. Abbreviations: BT1, period of behavioral testing before first training period; BT2, period of behavioral testing after first training period; LS, period of obtaining a learning set for task requirements; REC, recording period; TR, period of training with objects at short SOA.

stored by a PC with 1 ms resolution for later offline analysis. This PC also showed raster displays and spike histograms.

Stimuli

We used two large stimulus sets (see Fig. 1A for examples). Set 1 consisted of 120 grayscale images of objects (modified from Rossion and Pourtois, 2004), and set 2 of 50 black/white patterns (similar to those used by Op de Beeck and Vogels, 2000). Stimulus set 1 was used for all experiments, except for period TR3 (see Table 1) and the following recording phase in monkey 1A when stimuli from set 2 were used. Training effects were similar for the two stimulus sets. All objects in a set had the same average luminance, and the sets were randomly divided into groups of 10 exemplars. The masks consisted of many circular patches (with varying position and radius) from all stimuli in a set. We regularly changed these masks during training and recording to prevent that training effects would be specific to the individual mask images. The recording sessions involved 20 and 14 different mask images in monkeys 1A and 1B respectively (a mask image was used for 1.7 days on average), and 18 and 10 different mask images in monkeys 2A and 2B respectively (a mask image was used for 1.9 days on average). The maximum stimulus size was 5.4°, the masks measured 7° by 7°. The stimuli of set 1 were presented on a white background (luminance approximately 32 cd/m²), set 2 was presented on a gray background (luminance approximately 16 cd/m²). The frame rate of the monitor was 80 Hz, so the duration of objects and masks could be manipulated with intervals of 12.5 ms.

Behavioral tasks

The tasks were learned using standard operant conditioning techniques (positive reinforcement with water while monkeys were under restricted access to water; monkeys received dry food *ad libitum*). The experiment involved several behavioral tasks whose chronological order is shown for each monkey in Table 1. Each task involved fixation control throughout each trial, and data from aborted trials were not included.

Same/different task. All monkeys were trained in a successive same/different task with fixation control before the start of the actual backward masking experiments (period “LS” in Table 1). In this task, two images were presented successively during fixation (fixation window diameter approximately 2.5°) and subjects were trained to make a leftward or rightward saccade to one of two target spots presented after the offset of the second image. The correct saccade direction, rewarded with a small drop of apple juice, depended on whether the two stimuli matched or not.

Masked same/different task. In this task (Fig. 1B), the first stimulus period in a trial had a duration of 300 ms and was composed of an object image immediately followed by a mask

(object image duration indicated by the SOA). A second object image was presented for a duration of 300 ms without mask. The inter-stimulus interval (between mask offset and onset of the second object image) was 500 ms. The position of each object image varied with a maximum deviation of 2° from the fixation spot (only during training, position was fixed during recordings), and the position of the mask varied with a maximum deviation of 1.2° from the position of the first object (during training and recordings). Thus, monkeys could not perform the task by focusing on the presence of particular object features in particular locations of the visual field or on particular juxtapositions of objects and mask. Monkeys learned the specific requirements of the masked same/different task during several weeks with a variety of stimuli and masks (all different from those used in the actual experiment) at relatively long SOAs (at least 200 ms). This is period “LS” for monkeys 1A and 1B in Table 1. Before the start of the actual experiments, these monkeys were made familiar with intervening mask patterns at short SOAs using a set of simple shape silhouettes (16 and 20 daily sessions in monkeys 1A and 1B, respectively), this is period “TR” in Table 1. Thus, at the start of the actual experiments, the monkeys had learned the requirements of the masked same/different task: ignoring the mask pattern and attending to the preceding image.

Contrast same/different task. Part of the data from monkey 1A were obtained in the contrast same/different task (Fig. 10A) in which the masked objects were irrelevant. Training and recordings in this task took place after the data collection with the masked same/different task (see Table 1). Before as well as after the masked object, an enlarged fixation spot with a variable luminance was shown for 100 ms. The monkey had to compare the luminance of the two enlarged spots. The monkey was trained for several months in this task (period “LS” in Table 1) with a variety of timing, size, and contrast parameters, and with a variety of intervening stimuli (that were never used during recordings). We used this task instead of a passive fixation task in which the monkey might be doing the masked same/different task covertly. To avoid that the contrast same/different task would be too easy to serve as a good attentional control, we reduced the size and contrast change of the relevant enlarged fixation spot at the end of period “LS” to make the task relatively difficult for the monkey during training and recording periods (average performance during recordings was 81% correct).

Passive fixation task. Most data of monkeys 2A and 2B were collected during passive fixation. Only one stimulus period of 300 ms (object plus mask) was included in each trial, and continuous fixation during the trial was rewarded immediately after the stimulus period. Although these two monkeys had been trained on a (unmasked) same/different task before, they had never received any training with objects at short SOAs before the recordings under passive fixation.

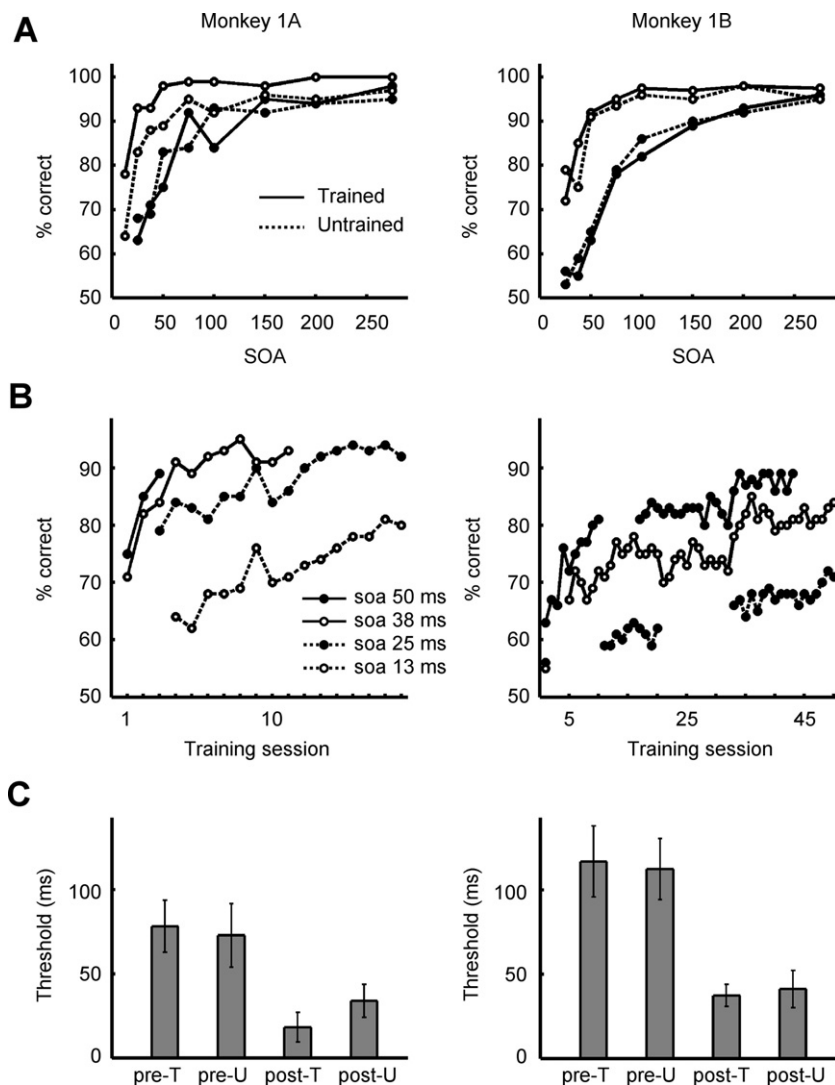


Fig. 2. Behavioral performance of monkeys before and after the training in the recognition of backward-masked objects. (A) Psychometric curve before training (filled symbols) and after training (open symbols) for two groups of 10 objects (trained and untrained). (B) Performance as a function of training session and SOA. (C) Eighty-five percent thresholds pre- and post-training for two groups of 10 objects, of which one was trained in-between (pre-T and post-T) and one not (pre-U and post-U). Error bars indicate 95% confidence intervals.

Behavioral training

In monkeys 1A and 1B (referred to as the “trained monkeys”), a period of extensive behavioral training with a group of 10 objects (period TR1 in Table 1) preceded the first recording session. We tested the effect of shortening the SOA on the performance of monkeys 1A and 1B before and toward the end of this training period (periods BT1 and BT2 in Table 1). During these behavioral tests, we tested the performance of each monkey for different SOAs (range: 25–275 ms) with two groups of objects from set 1 (the actual groups differed between monkeys). At all times (during this test and also during later training phases and recordings), the two objects in a trial belonged to the same object group and trials of different object groups were interleaved randomly. Before training, the two object groups were associated with very similar performance (see Fig. 2A). To have a conservative test of training effects, we opted for the one associated with slightly worse performance in BT1 as the trained group that was subsequently used during tens of training sessions (period TR1 in Table 1). During training, the presented SOAs were adapted in each session so

that at least one SOA was long enough to produce a performance around or above 85% correct (the one or two other SOAs were shorter). The effect of training and its stimulus specificity was determined by testing performance for the same two stimulus groups after this extensive training phase (period BT2 in Table 1), using the same SOAs as in the pre-training test (with additional testing at SOA 13 ms for monkey 1A). The group of stimuli presented during period TR1 is referred to as the “trained group,” the other stimuli as the “untrained group.” The data included an average of 120 trials per data point in Fig. 2A, that is more than 8000 trials in total. We determined thresholds corresponding to a performance level of 85% correct and their 95% confidence intervals (taking into account the number of trials) by fitting a cumulative Gaussian function to the percent correct responses at the different SOAs using the methodology of Wichmann and Hill (2001).

The experiment in monkey 1A involved several training periods interleaved with recording periods. The recordings were interrupted twice to train this monkey extensively on two new groups

of 10 objects (periods TR2 and TR3 in Table 1, these periods involved at least as much training trials as period TR1), and each recording period involved the most recently trained group of stimuli as the “trained group.”

Two behavioral training periods in monkey 1A were done in the contrast same/different task (periods TR4 and TR5 in Table 1), again with one group of 10 objects in each training period. These 10 objects were labeled as the “trained objects” for the following recording period. The objects were shown for 300 ms without a mask in half of the trials during the training periods, in the other trials the SOA was mostly 13 ms. The number of the trials with short SOA before the start of recordings in the contrast same/different task was at least as large as during the training periods with the masked same/different task (TR1-3). The timing parameters during recordings with the contrast same/different task were those of Fig. 10A, while during most of the training the fixation interval after the masked stimulus was 100 ms instead of 450 ms.

Recordings

At the start of each recording session, a guiding tube containing a micro-electrode was inserted in the brain, and the electrode was slowly lowered to IT cortex. We estimated the anatomical position of each recording site from the position of the guiding tube in the grid in the microdrive, the transitions of gray and white matter encountered when lowering the electrode, the depth at which neurons were encountered relative to the depth of bone, and a superposition of anatomical MRI images and CT scans (with guiding tubes positioned in the grid). The estimated location of recorded neurons ranged between 12 and 21 mm anterior to the external auditory meatus.

We recorded the responses of single IT neurons in each monkey while they were performing the masked same/different task (monkeys 1A, 1B, and 2A), the contrast same/different task (monkey 1A), and the fixation task (monkeys 2A and 2B). In the masked same/different task, the recordings in monkeys 1A and 1B followed an extensive training period with a group of 10 trained objects. Each monkey performed a few hundred trials of the same/different task using the trained objects as stimuli at the start of each recording session to give the monkey even more training with these objects during the recordings. Stimulus-responsive neurons were searched with two groups of 10 objects, one group being the trained objects. Objects were presented at long SOAs (275 ms) during this search test, and cells that appeared responsive upon visual inspection of the histograms were tested further at this SOA for at least six masked presentations of each stimulus (the full set of stimuli used during the search procedure). Further testing was done with all objects presented at two shorter SOAs. The longest SOA was always 150 ms, and the choice of the shortest SOA depended on the performance of the monkey at the end of the behavioral training (13 ms for monkey 1A and 25 ms for monkey 1B; a different SOA was chosen because in monkey 1A performance was almost 100% at SOA 25 ms and in monkey 1B performance was near chance level at SOA 13 ms).

All our tests included a high number of different stimuli because we did not select a few stimuli based on neuronal responses (e.g. the stimuli that elicited the strongest and weakest responses). With only a few stimuli, monkeys might switch strategies and attend to a simple feature that is sufficient to discriminate the few stimuli. Our procedure avoids that the monkeys could switch strategies, but it was not feasible to combine this high number of stimuli with a high number of SOAs, or to get a very high number of trials per stimulus/SOA combination. This precludes the acquisition of detailed neurometric curves (responses as a function of SOA), or trial-by-trial analyses of correct versus incorrect trials.

We compared many different groups of 10 objects to minimize a priori differences between stimuli as a confounding factor. Furthermore, each group of untrained objects was used for less than

three daily sessions on average to avoid that the objects would become trained. Even on the first day of recording with a new group, the untrained objects were already presented to a monkey tens of times before spikes were collected in a systematic test because the search and isolation of single cells required some time. Thus, no recordings involved the first few presentations of a new object, and as such our untrained and relatively unfamiliar objects were not “novel” in an absolute sense. The recordings in monkey 1A involved a comparison of 3×10 trained objects with 13×10 untrained objects (three groups of trained objects because the recordings were interrupted twice to train this monkey intensively on two new groups of objects, see above). The recordings in monkey 1B involved a comparison of 1×10 trained objects with 8×10 untrained objects.

In the contrast same/different task (monkey 1A), neurons were searched for with two groups of 10 objects, one group being the objects that were used during the previous training period (recordings involved the comparison of 2×10 trained with 10×10 untrained objects). Objects were presented at long SOAs (275 ms) during the search test, and further testing was done with all stimuli presented at two different SOAs (13 ms and 150 ms). The difference in exposure/familiarity between trained and untrained objects was greater in the contrast same/different task compared with the masked same/different task (at least as many presentations of the trained objects during training, and less presentations of the untrained objects during recordings because we omitted the test with the longer SOA 275 ms; this SOA was only used during search).

In the passive fixation task, spikes were recorded until 200 ms after stimulus offset. Neurons were searched with one group of 10 objects presented at long SOAs (275 ms), and responsive cells were tested further with these objects presented at two shorter SOAs (25 ms and 150 ms). In monkey 2B, cells that appeared to be unresponsive to this one group of 10 objects were screened a second time with another group of 10 objects.

Analyses of neural responses

We focused on the responses to the masked stimuli in the test that interleaved a short and longer SOA. For each neuron we had at least six trials for each stimulus at each SOA (average of 13 trials per stimulus/SOA combination). All non-aborted trials were included in the analyses. In the active task conditions, monkeys had to respond for the trial to be counted as non-aborted. We ranked the 10 stimuli from each group according to net response strength in the condition with SOA 150 ms. Net response was calculated by subtracting the baseline activity in the 200 ms before stimulus onset from activity 60–260 ms after stimulus onset. This ranking was used to select the best and the worst stimulus for each cell. The same ranking was applied to the shorter SOA. This analysis assumes that neurons have similar preferences at different SOAs. This was certainly true at the population level, and analyses of the data of shorter SOAs using a ranking based on the data at these short SOAs revealed similar results.

Analyses of object-specific training effects were performed on data that were corrected for the response latency of each individual cell because previous work showed that the timing of stimulus selectivity during backward masking is locked to the response onset of individual neurons (Kovacs et al., 1995). This latency was determined by a Poisson spike train analysis (Hanes et al., 1995) of the responses to the best stimulus at 150 ms SOA in combination with a visual inspection of the histograms to eliminate false alarms. One measure of response latency was obtained for each neuron (so the same correction was done for trained and untrained objects). In the recorded population of neurons, the computed latencies ranged from 60 ms to 180 ms (mean: 99 ms; standard deviation: 23). It is important to note that the training-related increase in object selectivity in monkey 1A as reported in the Results section was not an artifact of a better latency correc-

tion for trained objects because (i) similar effects were noted without any latency correction, and (ii) we would expect the opposite, that is, worse latency correction for trained objects, since for most neurons the object associated with strongest responses (the condition used for determining a neuron's latency) was an untrained object (see Results).

Analyses of mask responses (object-independent training effects) were done on data aligned at stimulus onset. While Kovacs et al. (1995) showed that the timing of stimulus selectivity is more precise in data that are aligned at response onset, preliminary analyses of our data suggested that the opposite is true for the timing of mask responses. For example, in population histograms as those shown in Figs. 7–8, the onset time of the mask responses in untrained monkeys was most clear with data aligned at stimulus onset.

Population time histograms (e.g. Fig. 4) are based on normalized responses. The absolute response (i.e. baseline firing rate not subtracted) of each cell to the best stimulus summed in the interval from 0 to 200 ms after response onset was normalized to 20 (average height of a 10 ms bin: 1.0). The same normalization factor was used for the two stimulus groups, so differences in response strength between trained and untrained objects would be preserved. In addition, we constructed difference histograms for each stimulus group by subtracting the histogram for the worst stimulus from the histogram for the best stimulus. Effects of training on these best-worst histograms and their dependence on "time from response onset" (Fig. 4B) were assessed by a repeated-measures design with two within-neuron factors: trained vs. untrained, and time (20 bins in the interval 0–200 ms after response onset).

We mention two selectivity indices in the Results section. The d' index is a variability-corrected measure of selectivity between two stimuli, and is computed as:

$$d' = (R(\text{best}) - R(\text{worst})) / \sqrt{((\text{var}(R(\text{best})) + \text{var}(R(\text{worst}))) / 2)}$$

with $R(\text{best})$ and $R(\text{worst})$ the responses to the best and worst stimulus, and $\text{var}(R)$ the variance of these responses across trials.

The second selectivity index is the "depth of selectivity" (DOS; see Rainer and Miller, 2000):

$$\text{DOS} = (N - \text{sum}(R) / \text{max}(R)) / (N - 1)$$

with N the number of stimuli, $\text{sum}(R)$ the sum of the N responses, and $\text{max}(R)$ the maximum response out of N .

RESULTS

The neural correlates of object-specific learning

The initial aim of our experiments was to investigate the neural correlates of object-specific training effects. Two monkeys (1A and 1B) were trained with a particular set of stimuli, and we compared behavioral performance and neuronal responses after training for this trained set of stimuli and untrained sets of stimuli.

The monkeys performed a masked same/different task in which they compared two objects of which the first was masked (Fig. 1). The effect of SOA was measured for two groups of 10 objects in each monkey. Before training, each monkey performed above 90% correct at long SOAs (Fig. 2A). The performance of monkey 1A was impaired for SOAs below 100 ms, although he was still able to do the task at around 65% correct at SOA 25 ms. Monkey 1B showed a more pronounced effect of masking and needed an SOA of 50 ms to perform around 65% correct.

The monkeys were subsequently trained with one of the two groups of 10 objects. Before training, there was no significant difference between the two groups in performance level (Fig. 2C), and in each monkey the to-be-trained group was chosen to be the one with slightly worse performance pre-training. Monkey 1A was trained for 17 daily sessions (about 5000 masked presentations of each object) and monkey 1B for 52 sessions (more than 15,000 masked presentations of each object). During training, the performance of each monkey improved gradually (Fig. 2B). The performance for these trained and highly familiar objects was compared after this training period with the performance for the group of untrained objects. The comparison of the 85% performance thresholds before and after training for the two groups of objects (Fig. 2C) revealed that only half of the training-related improvement for trained objects in monkey 1A generalized to the untrained objects. In this monkey, the 85% threshold post-training was significantly lower for trained than for untrained objects ($P < 0.01$; bootstrapping test as described in Wichmann and Hill, 2001). Thus, as observed before in human subjects using a similar paradigm (Grill-Spector et al., 2000), the training in this animal resulted in a training effect that was partially object-specific. Monkey 1B showed also a large effect of training, but he did not reach the performance level of monkey 1A (despite the longer training). Furthermore, the training effects in this monkey transferred almost completely to untrained objects, with no significant difference in behavioral threshold between trained and untrained objects after training (bootstrapping test; $P > 0.05$). Thus, one monkey showed a difference in performance between trained and untrained objects, the other monkey did not.

We investigated the neural correlates of these behavioral effects of training with single-unit recordings in IT cortex (monkey 1A: 46 neurons; monkey 1B: 45 neurons). Each neuron was tested with two groups of 10 objects, one trained group and one untrained. The behavioral performance during the recordings was similar to that in Fig. 2. The performance at SOA 275 ms was above 98% for each object group in each monkey. Monkey 1A performed at a level of 84% and 69% correct at SOA 13 ms for trained and untrained objects respectively (difference significant according to a bootstrapping test; $P < 0.001$), while monkey 1B performed at a level of 68% and 73% correct at SOA 25 ms for trained and untrained objects, respectively. This slightly albeit significantly ($P < 0.05$) worse performance for trained stimuli in monkey 1B might be related to the trained objects being slightly more difficult prior to training.

Fig. 3 shows the responses of one IT neuron to objects at a short and longer SOA. This neuron responded selectively at each SOA, preferring the same object at each SOA, but the duration and strength of the object-selective response were much smaller at the short SOA. Kovacs et al. (1995) showed before that the exact timing of the interruption of the transient stimulus response by the mask depends on a neuron's response latency. So, we report the results of analyses after correcting for the response latency of each individual neuron (see Experimental Proce-

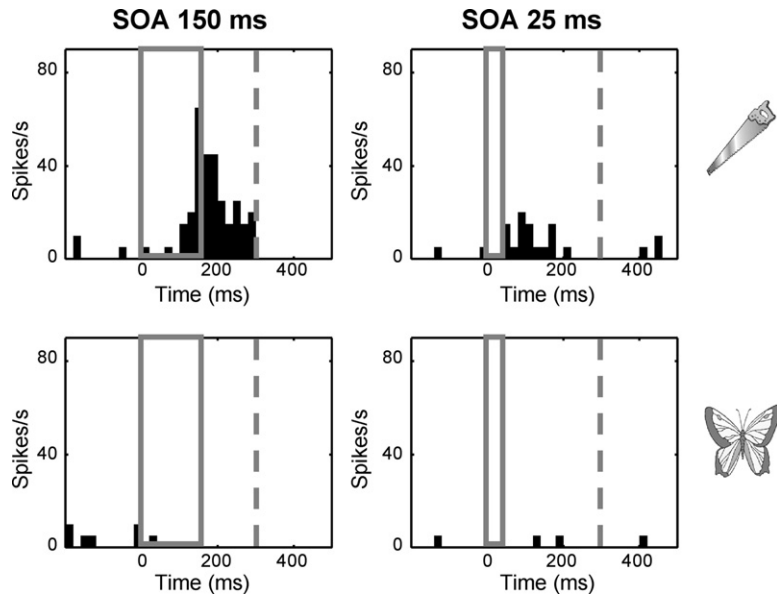


Fig. 3. Peri-stimulus time histograms (bin width 20 ms) for a single IT neuron from monkey 1B at two SOAs for the “best” object (top row) and the “worst” object (bottom row) out of 10 objects. The gray box represents the time interval during which the object was presented. The masking pattern was presented immediately after this interval and its offset time is shown by the gray striped line.

dures). **Fig. 4** shows the population time histograms based on normalized responses for the best and worst stimulus conditions and best minus worst for trained and untrained objects, all aligned at response onset. Shortening the SOA had the same qualitative effect in each monkey for both trained and untrained objects: the amount of object selectivity was reduced as was the time interval during which clear object-selective responses were present.

To what extent does this neural effect of SOA differ between trained and untrained stimuli? We investigated the difference between best and worst stimuli as shown in **Fig. 4B** as a function of training (trained stimuli versus untrained stimuli) and time (repeated-measures analysis of variance (ANOVA); see Experimental Procedures). A main effect of training would indicate a difference in selectivity for trained compared with untrained objects, and an inter-

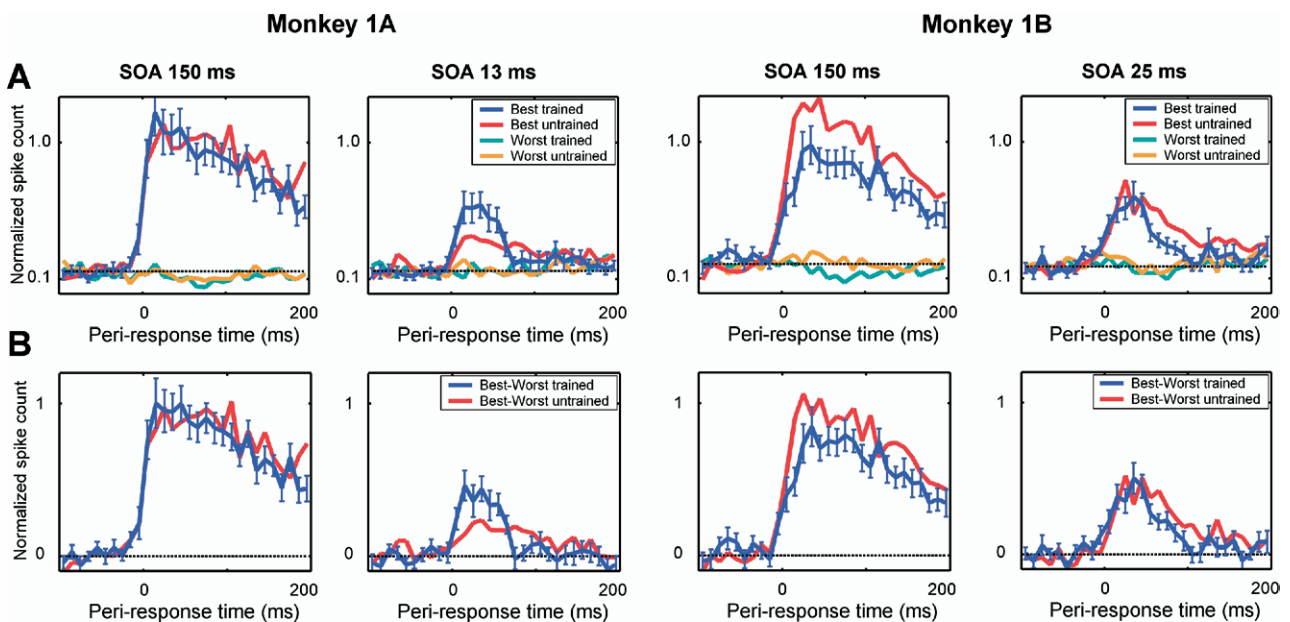


Fig. 4. Normalized peri-response time population histograms for monkeys 1A and 1B in the masked same/different task. (A) Histograms for the best and the worst stimulus in each group of objects (trained and untrained). The horizontal line represents the average activity from 100 to 20 ms before response onset. (B) Histograms for the difference between best and worst stimulus. The horizontal line represents zero difference. Bin width is 10 ms. Whiskers represent the standard error of the mean across neurons.

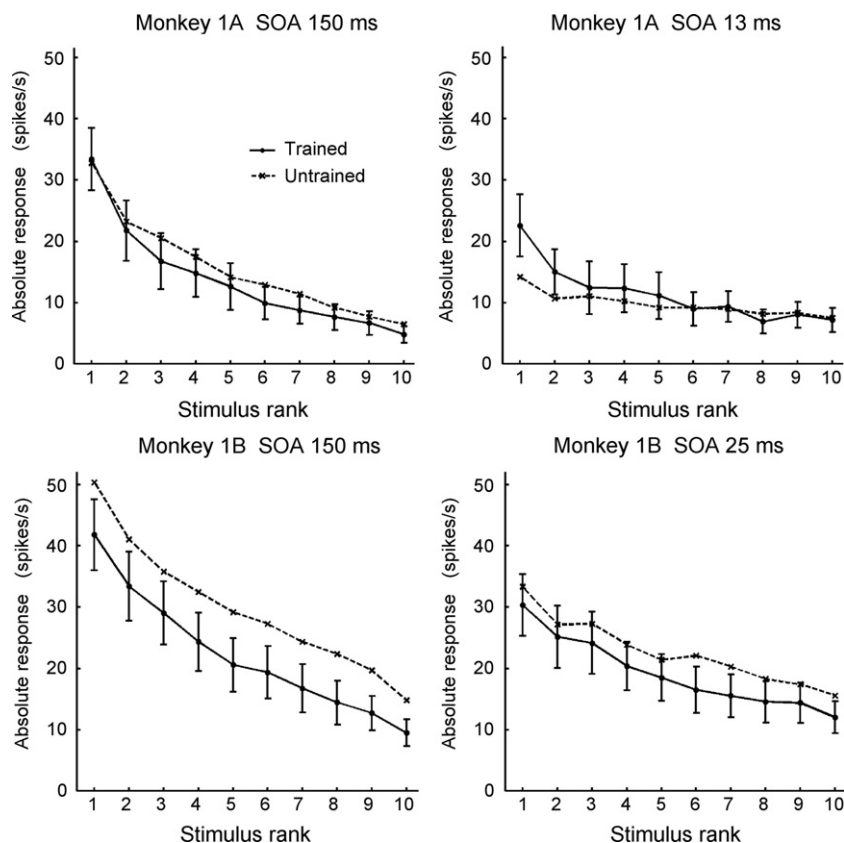


Fig. 5. The average response during the masked same/different task as a function of stimulus rank for monkeys 1A and 1B in different SOAs. Average response was calculated in the interval 0–150 ms and 0–50 ms after response onset for SOA 150 ms and SOA 13/25 ms, respectively. Whiskers represent standard error of the mean across neurons.

action between training and time would indicate that such differences depend on the time from response onset. At long SOA, this analysis showed no significant main effect of training ($P > 0.5$), and no interaction between training and time for monkey 1A ($P > 0.4$). A marginally significant main effect of training (stronger selectivity for untrained objects) was found for monkey 1B at long SOA ($P = 0.04$), without any interaction ($P > 0.3$). At short SOAs, we found a significant interaction between training and time for monkey 1A ($P < 0.005$), mostly due to a significantly higher selectivity for familiar objects in the interval 0–50 ms after response onset ($P < 0.02$; normalized difference in response between best and worst was 0.37, standard error of the mean (S.E.M.) 0.06 for trained objects and 0.17, S.E.M. 0.05 for untrained objects). In monkey 1B, there was a trend toward stronger selectivity for untrained objects ($P = 0.06$), without an interaction between training and time ($P > 0.3$). Thus, monkey 1A showed a strong and transient training-related increase in object selectivity at the shortest SOA. In contrast, for monkey 1B object selectivity tended to be slightly worse for trained objects than for untrained objects.

The transient training-related increase in selectivity at short SOA for the best versus the worst stimulus was also found in monkey 1A with d' , a measure of selectivity that takes into account response variance (for the interval 0–50 ms after response onset: for trained $d' = 1.01$,

S.E.M. = 0.17; for untrained $d' = 0.46$, S.E.M. = 0.11; difference between trained and untrained $P < 0.02$). Furthermore, the difference was not caused by a difference in selectivity between correct and incorrect trials. We found similar results when the analysis was restricted to correct trials, which comprise 84% and 69% of the trials with trained and untrained objects respectively (normalized difference in response between best and worst was 0.37, S.E.M. 0.06 for trained objects and 0.19, S.E.M. 0.05 for untrained objects; difference between trained and untrained significant at $P < 0.02$).

All previous analyses focused on selectivity, and for that reason we showed normalized responses for the best and worst stimuli. In order to provide a broader picture of our results, Fig. 5 shows absolute, non-normalized responses for all stimuli in the initial, stimulus-selective part of the responses (0–150 ms after response onset for SOA 150 ms and 0–50 ms for the short SOA). Plotting the data in this new format confirms the difference between monkeys in the effect of training on selectivity at short SOA. In particular, a repeated-measures analysis with two factors (rank and stimulus group: trained or untrained) revealed that the interaction between rank and stimulus group was significant across neurons in monkey 1A ($P < 0.005$), but not in monkey 1B ($P > 0.5$).

In addition, the non-normalized absolute responses reveal that the absolute responses at the long SOA tended

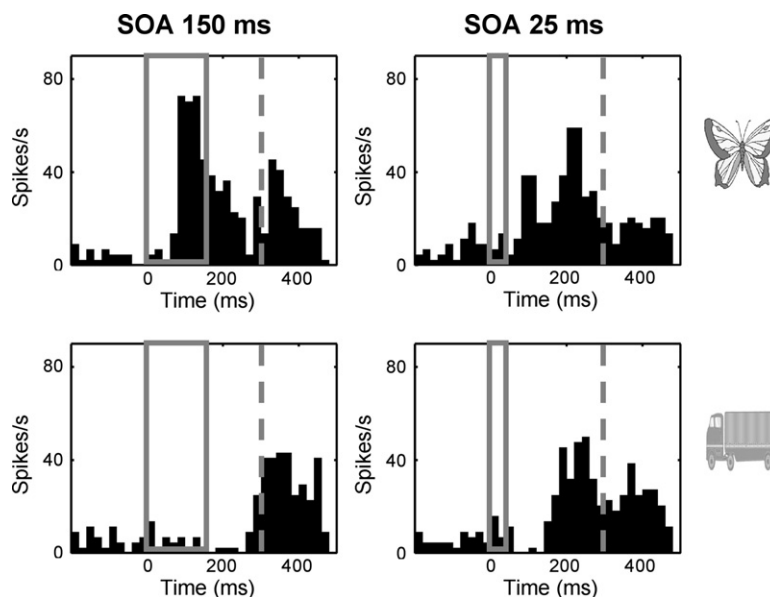


Fig. 6. Peri-stimulus time histograms (bin width 20 ms) for a single IT neuron from monkey 2A at two SOAs for the “best” object (top row) and the “worst” object (bottom row) out of 10 objects. The gray box represents the time interval during which the object was presented. The masking pattern was presented immediately after this interval and its offset time is shown by the gray striped line.

to be smaller for trained than for untrained objects. For SOA 150 ms, a repeated-measures analysis with two factors (rank and stimulus group: trained or untrained) revealed a main effect of stimulus group in monkey 1B ($P < 0.001$). This effect did not reach significance in monkey 1A ($P = 0.10$). The object associated with the strongest response was an untrained object for 25 of 46 neurons (54%) in monkey 1A, and for 30 of 45 neurons (67%) in monkey 1B. At a short SOA, the reduction in mean responses for trained compared with untrained objects was still found for monkey 1B ($P < 0.001$), but overall responsiveness was the same for the two object groups in monkey 1A ($P > 0.5$).

The neural correlates of object-independent learning

Our initial experiment was set up to investigate object-specific training effects, but the behavioral improvement due to training turned out to transfer to untrained objects for the most part (Fig. 2). We wondered what neural changes might be associated with the large object-independent part of the behavioral improvement due to training. One striking aspect of the data from monkeys 1A and 1B was the absence of clear responses to the mask patterns. Kovacs et al. (1995) reported that object selectivity was interrupted by a mask pattern. In addition, however, many IT neurons were still responsive after this interruption because they responded to the mask patterns. In our present data, the surprising absence of mask responses in monkeys 1A and 1B can be seen in the population histograms for short SOAs in Fig. 4A. After a visual response of less than 100 ms, the responses fell back to a level close to the baseline level of activity of before any visual response appeared. The mask pattern that was on the screen for more than 250 ms did not elicit a clear population response.

There are many differences between our experimental methods with monkeys 1A and 1B and the study by Kovacs et al. (1995) who worked with different stimuli and masks (line patterns), and with passively fixating monkeys that received no training in backward masking. Thus, we decided to record in conditions more comparable to this previous work but with the same stimuli as used for our trained monkeys. We recorded the responses of IT neurons to masked objects in two untrained, passively fixating monkeys (monkey 2A: 63 cells; 2B: 53 cells). We compared these data to our data obtained in the two extensively trained monkeys while these monkeys were performing the masked same/different task with one set of trained objects and one set of untrained objects (monkey 1A: 46 cells; 1B: 45 cells).

The object and mask responses of a single neuron in the fixating animals are shown in Fig. 6. The data of this single neuron showed strong responses to the presentation of an image of a butterfly at a longer SOA (150 ms), but not when an image of a truck was presented. However, there was a response in both conditions to the mask pattern (visible as a second bump in the time histogram of the butterfly condition). At a short SOA, there was a short transient response to the butterfly, with no initial response to the truck. However, as expected the shortening of the SOA shifted the mask response to an earlier time point. This is again visible as a second bump in the butterfly condition. The shorter latency relative to stimulus onset of the response in the truck condition with a shorter SOA clearly shows that this response is a mask response: it is linked to mask onset but not stimulus onset.

Fig. 7 shows the peri-stimulus population histograms at short and long SOA for the best and worst stimulus conditions in each monkey. There was a very striking difference between trained and untrained monkeys at the point

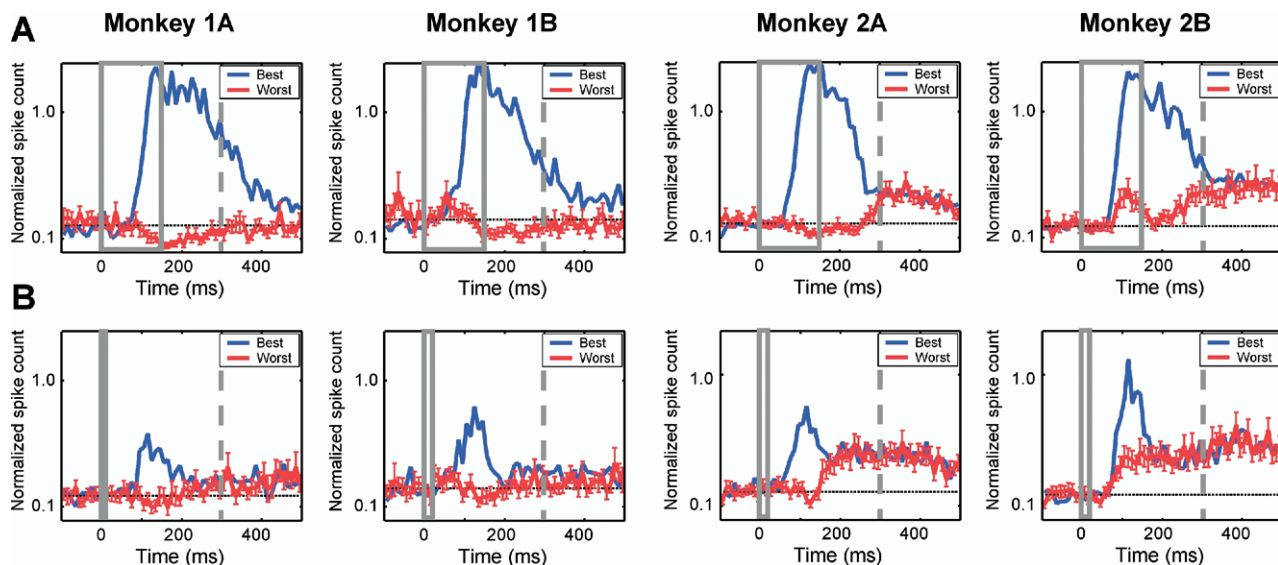


Fig. 7. Normalized peri-stimulus time population histograms for each monkey at long SOA (A) and short SOA (B). Responses to the best out of 10 stimuli at SOA 150 ms are shown in blue, responses to the worst stimulus in red. Histograms are normalized to the best response at SOA 150 ms (see Experimental Procedures). The horizontal line represents the average activity from 100 to 20 ms before stimulus onset. Bin width is 10 ms. Whiskers represent the standard error of the mean. The gray box represents the time interval during which the object was presented. The masking pattern was presented immediately after this interval and its offset time is shown by the gray striped line.

in time at which most object selectivity has disappeared. The average response, in particular for the worst stimuli, returned to a level near the baseline in trained monkeys, while it remained relatively high for untrained monkeys. This indicates that the population of neurons in monkeys 2A and 2B was responsive to the masking patterns, similar to the results described by Kovacs et al. (1995).

We computed the net mask response of each neuron from the worst object condition at short SOA by subtracting the baseline activity in the 200 ms before stimulus onset from the response in the interval 270–370 ms after stimulus onset (similar results were found using the later interval 370–470 ms after stimulus onset). We decided to take this relatively late interval (the mask onset is more than 240 ms before the start of this interval) to avoid any stimulus-related responses. The baseline activity did not differ systematically between trained monkeys (spikes/s, mean \pm S.E.M.; 1A: 5.59 ± 0.84 ; 1B: 7.87 ± 1.13) and untrained monkeys (spikes/s, mean \pm S.E.M.; 2A: 6.77 ± 0.82 ; 2B: 5.56 ± 0.61). In monkeys 1A and 1B, only 15 of 91 neurons (16%) showed a significantly positive net mask response ($P < 0.05$; tested per neuron with a *t*-test across all trials with the worst object condition at short SOA). In monkeys 2A and 2B, 62 of 116 neurons (53%) showed significant mask responses. The mean net mask response was 1.3 ± 0.5 , 2.1 ± 1.1 , 9.5 ± 2.1 , and 8.7 ± 1.2 spikes/s in monkey 1A, 1B, 2A, and 2B, respectively. The difference in the net mask response between neurons from trained and untrained monkeys was strongly significant (independent *t*-test, $P < 0.001$). The difference in the mask response between individual trained and untrained animals was significant ($P < 0.01$) for each of the four possible pairings of the two trained and two untrained animals. Thus, unlike the object-specific effect (see above), the difference in mask

response was present in each animal. Interestingly, mask responses in the trained monkeys were as small when the mask followed a trained and highly familiar object (mean mask response: 1.7, S.E.M. = .6) as when the mask followed an untrained object (mean mask response: 1.7, S.E.M. = .6). These results suggest that training resulted in strongly reduced responses to mask patterns in an object-independent way.

This strong difference in mask responses was not related to a difference between trained and untrained monkeys in the amount of object selectivity at long SOA. We quantified object selectivity by the DOS index (DOS; see Experimental Procedures), and there was no consistent difference between trained and untrained monkeys in selectivity (monkey 1A: DOS = .66; monkey 1B: DOS = .57; monkey 2A: DOS = .61; monkey 2B: DOS = .49). Furthermore, an ANCOVA (analysis of covariance) with training as a between-neuron factor and DOS as a covariate revealed a significant effect of training ($P < 0.001$). Thus, the effect of training was still found when object selectivity was controlled for statistically.

Trained monkeys performed a same/different task during recordings, whereas the untrained monkeys were passively fixating. Was the difference in mask responses between these two groups of animals due to the difference in task? To answer this question, we recorded from 27 neurons in monkey 2A while this monkey was performing the masked same/different task. This monkey (that performed above 90% correct in a same/different task without masks) received a few days of training with masked objects at a long SOA (275 ms; average performance 86% correct) to make sure that he followed the general task requirements (that is, to ignore the short mask). He was not trained at SOAs below 200 ms, explaining why his performance was

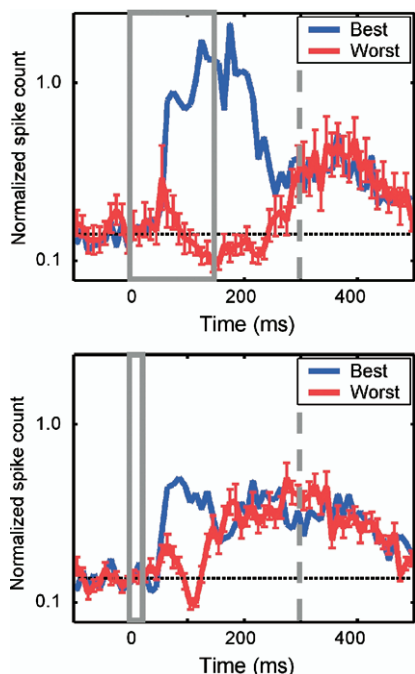


Fig. 8. Normalized peri-stimulus time population histograms for monkey 2A in the masked same/different task at SOA 150 ms (top panel) and SOA 25 ms (bottom panel). Conventions as in Fig. 7.

low during the recordings for intermediate and short SOAs (SOA 150 ms: 74%; SOA 25 ms: 54%). The data are shown in Fig. 8. The mean mask response in the recorded population of 27 neurons was 11.4 spikes/s (S.E.M.=2.3), which is significantly stronger than the mask response in each of the trained monkeys (independent *t*-test, $P < 0.001$), but similar to those observed during the passive fixation task. Twelve out of 27 neurons (44%) showed significant mask responses ($P < 0.05$). Thus, we found robust mask responses in a monkey performing the masked same/different task without any extensive training with masked objects (there was no difference in these results between correct and incorrect trials). The latter results also demonstrate that the presence of strong mask responses in the fixating, untrained animals is not due to any of the differences between the fixation task and same/different task, such as the fact that monkeys were rewarded for fixation in the passive fixation task at stimulus offset (whereas the reward in the same/different task was only delivered after the second stimulus).

An interesting aspect of our design is that the recordings in monkey 2A with the masked same/different task were performed after all recording sessions (20 sessions) in the passive fixation task. Thus, the finding of strong mask responses in the active task condition does also show that mask responses do not decline over time when many mask images have been seen already on previous days. Further analyses also suggest that mask responses were unaffected by the relative familiarity of specific mask images. In these analyses, we divided all data in monkeys 2A and 2B according to the relative familiarity of the mask image (first day that a mask was seen versus later days).

The average response to first-day masks was 7.6 spikes/s (S.E.M.=1.0), which is not different from the response of 9.7 spikes/s (S.E.M.=1.5) to more familiar masks (independent *t*-test, $P = 0.23$).

Since object-independent effects of training were investigated by comparing different monkeys, we have to rule out any systematic difference in the recording position in IT cortex between trained and untrained monkeys. As shown in the reconstruction of our recording positions in the different monkeys in Fig. 9, the recordings in our trained monkeys were in a range of positions covered by the recording positions in untrained monkeys. Furthermore, the relatively widespread recordings in the untrained monkeys suggested that strong mask responses are found across the whole IT cortex with little relation between the strength of mask responses and anterior/posterior, dorsal/ventral, and medial/lateral coordinates. There was a substantial range on each of these dimensions in monkey 2B. This monkey showed strong mask responses in the 11 most anterior neurons (9.2 spikes/s, S.E.M.=2.3) as well as in the eight most posterior neurons (9.5 spikes/s, S.E.M.=4.7; difference in anteriority 6 mm); in the 22 neurons located in the lower bank of the superior temporal sulcus (6.1 spikes/s, S.E.M.=1.7) as well as in the 26 neurons in ventral IT cortex (12.4 spikes/s, S.E.M.=1.6); and in ventral IT cortex in the 15 most lateral neurons (9.5 spikes/s, S.E.M. 2.0) as well as in the eight most medial neurons near or in the anterior middle temporal sulcus (13.7, S.E.M. 2.5; average difference in laterality 3 mm). There are some differences between these sub-populations, but most importantly each of these sub-populations showed significantly stronger mask responses ($P < 0.01$) than found in the trained monkeys. This illustrates the robustness of the difference between monkeys: it is seen consistently with smaller sub-populations of neurons, independently of the anatomical position of these neurons. Thus, we can exclude a different recording position as causing the difference in mask responses between trained and untrained monkeys.

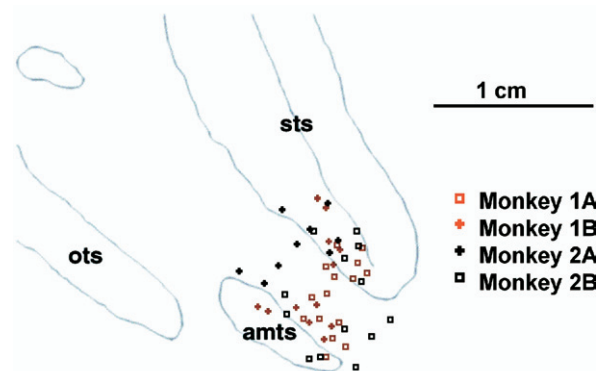


Fig. 9. Reconstruction of recording positions in IT cortex in the four monkeys. Abbreviations: amts, anterior middle temporal sulcus; ots, occipito-temporal sulcus; sts, superior-temporal sulcus.

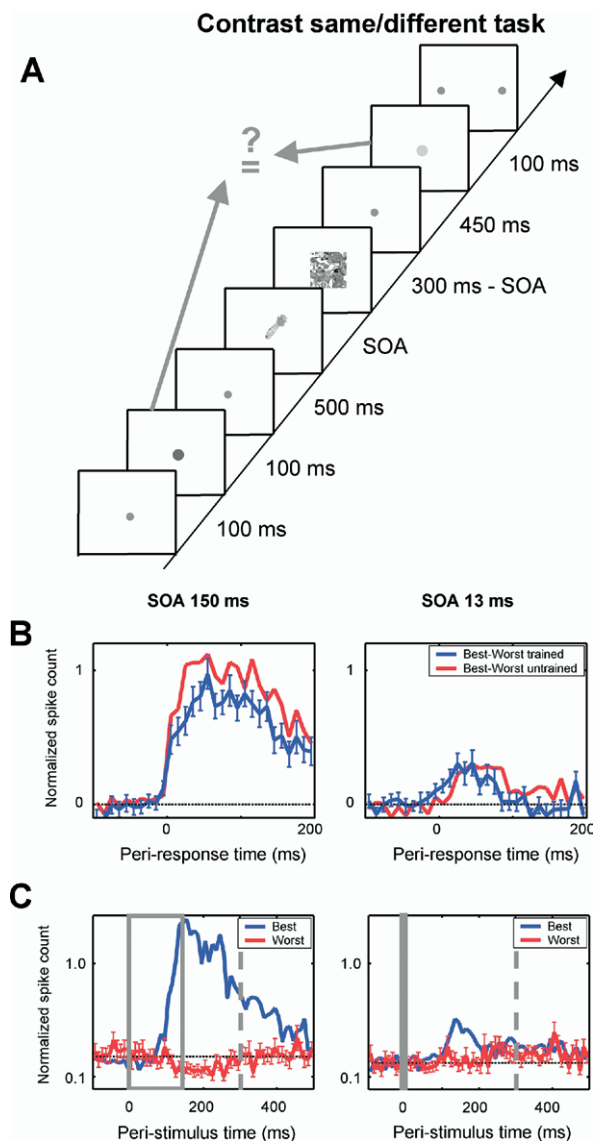


Fig. 10. Contrast same/different task. (A) The task procedure. The monkey had to compare the contrast of the two enlarged fixation spots. (B) Normalized peri-response time population histograms for the difference in response between best and worst stimulus in each group of objects (trained and untrained). The horizontal line represents zero difference. (C) Normalized peri-stimulus time population histograms that reveal the absence of mask responses. Conventions as in Fig. 7.

Neural responses to irrelevant masked objects

The reduction of mask responses could be the result of an active, top-down inhibition of processing of the mask in order to avoid interference with the task-relevant object. Such top-down inhibition would not be necessary when the object is no longer relevant for the task, so in that context the reduction of mask responses might disappear.

We tested this possibility in 49 IT neurons in monkey 1A in the contrast same/different task (Fig. 10). In this task, the masked objects were irrelevant for task performance. This monkey performed this new task in the training period before the recordings as well as during the recordings.

Thus, the “trained objects” were not only irrelevant during recordings but also in the training period (highly familiar, but never relevant). Note that in this task object and mask have the same behavioral status: both are irrelevant to the task but can interfere. If the response reduction for the mask in the masked same/different task was caused by active, top-down inhibition then this response reduction might disappear (if the inhibition is contingent upon the relevance of the object), or a similar response reduction might occur for the objects (since they are as irrelevant as the masks).

There was a similar absence of mask responses in the contrast same/different task (mean mask response: 2.8 spikes/s, S.E.M.=0.8) as in the masked same/different task (mean mask response: 1.2 spikes/s, S.E.M.=0.8), and these mask responses were significantly smaller than the mask responses in the untrained monkeys ($P<0.01$). Only 5 of 49 neurons (10%) showed significant mask responses ($P<0.05$). Note that the same neurons did respond to the objects, thus the near absence of mask responses in both the contrast same/different task and the masked same/different task is also clear when responses are normalized relative to the object responses (see Fig. 7 and Fig. 10C). In sum, near absence of mask responses was also found in a previously mask-trained monkey that was afterward trained for months in a task that required not only the monkey to ignore the masks, but also the preceding object image.

Concerning object-specific training effects, the data obtained in the contrast same/different task showed the same effects of training on object responses as found in the *other* trained monkey, monkey 1B, in the other task, the masked same/different task. The best-worst histograms for the short SOA revealed no interaction between training and time ($P>0.2$), as found for monkey 1B, but in contrast to the data of monkey 1A in the masked same/different task. This was confirmed in a combined ANOVA on the data of monkey 1A from the two tasks at short SOA: there was a significant three-way interaction between task, training, and time ($P<0.05$). Thus, the object-specific transient increase in object selectivity was not found in the contrast same/different task.

Relation between mask responses and “sustained” object selectivity

The reduction of the mask response of a particular neuron might increase the neuron’s potential to relay object information without mask interference. This effect is very clear at SOA 150 ms (Fig. 7A). During the late time interval that was used to quantify mask responses (270–470 ms after onset of the object image), neurons from the trained monkeys showed a strong “sustained” selectivity in the comparison of best and worst objects: the difference between best and worst condition was 9.8 spikes/s (S.E.M.=1.7) and 8.2 spikes/s (S.E.M.=1.4) in monkeys 1A and 1B respectively (strongly significant selectivity in each monkey, $P<0.0001$). In the untrained, passively fixating monkeys, there was hardly any object selectivity in the same time interval: the difference between best and worst con-

dition was 0.91 spikes/s (S.E.M.=0.9; $P=0.33$) and 2.5 spikes/s (S.E.M.=1.1; $P=0.02$) in monkeys 2A and 2B respectively. However, the relation between the reduction in mask responses and sustained object selectivity was not found at short SOA (Fig. 7B). Despite the near absence of mask responses in trained monkeys, these monkeys showed at most a weak sustained object selectivity at very short SOA: the difference between best and worst condition was 0.40 spikes/s (S.E.M.=0.7; $P>0.3$) and 1.8 spikes/s (S.E.M.=0.8; $P<0.05$) in monkeys 1A and 1B respectively. In sum, training is associated with more sustained object selectivity at longer SOAs, but this sustained selectivity is not consistently found at very short SOAs.

Furthermore, if the strength of mask responses and sustained object selectivity would be related, then we would expect a correlation between the mask response of a neuron and its sustained object selectivity. We tested for such a correlation in the untrained monkeys, in which the recorded population of neurons showed a high diversity in mask responses (no effects were seen for trained monkeys either). After subdividing this population in half according to the strength of the mask response, we found no significant difference in sustained object selectivity at SOA 150 ms (independent t -test, $P>0.3$) between neurons with a strong mask response (mean mask response: 17.4 spikes/s) and neurons with a weak mask response (mean mask response: 0.9 spikes/s). Thus, our data provide no evidence that sustained object selectivity and the strength of mask responses are correlated properties of IT neurons.

DISCUSSION

We found that training in recognizing masked objects improved monkeys' recognition performance. A substantial part of this training effect transferred to untrained objects. The small object-specific component of training-related improvements, found in only one monkey, was associated with a transient, object-specific increase in neuronal selectivity. A comparison of trained and untrained monkeys suggested that the strong object-independent behavioral effects were associated with a reduction of the response to the irrelevant masking patterns. These object-independent neuronal effects were still found when a monkey, previously trained to ignore the mask but not the objects, was retrained in a task in which both objects and masks were irrelevant.

We initially designed our study to investigate object-specific training effects, based on previous human studies that reported that the effect of training in backward masking was mostly object-specific (Furmanski and Engel, 2000; Grill-Spector et al., 2000). As far as we found evidence for object-specific training effects in one of two monkeys, it is tempting to conclude that these behavioral effects in the two species reflect similar neural changes. Grill-Spector et al. (2000) found a training-related increase in functional activation in human object-selective cortex to backward-masked objects presented at a short SOA. However, the object-specific effect that we found at the neural level involved a change in object selectivity without a major

effect on response strength at short SOA, and in addition this effect was very transient. It is doubtful that such a small and transient effect at the single-neuron level would give rise to a major change in the BOLD response as measured by Grill-Spector et al. (2000).

Most importantly however, most of the behavioral effect of training in monkeys transferred to untrained objects. Each monkey showed more transfer than any of the subjects in previous human studies with similar training paradigms (Furmanski and Engel, 2000; Grill-Spector et al., 2000). The discrepancy at the behavioral level makes it even more difficult to compare our results directly with those of Grill-Spector et al. (2000). With our current knowledge, many factors might be responsible for the difference, including the design (e.g. different stimuli and task), the species (monkeys versus human subjects), and the methods (functional imaging versus extra-cellular recording of the spiking output of a selected sub-population of neurons).

An additional difficulty is that the amount of transfer was very different across monkeys. We do not know the cause of these inter-individual differences in our limited sample size (two trained monkeys). Human perceptual learning studies have also reported large inter-individual differences in performance level and in the specificity of learning effects. These differences might be caused by many factors, even unrelated to perceptual processes (Ahissar and Hochstein, 1996). In our study, monkey 1B showed much larger effects of intermediate levels of backward masking than monkey 1A (see Fig. 2A). There might be a causal relationship between this difference in performance and the fact that training effects in monkey 1B were less specific than in monkey 1A. A similar relationship between overall behavioral performance and specificity of training effects was suggested in another study of perceptual learning in macaque monkeys (Yang and Maunsell, 2004).

We identified a neural change that might underlie the large part of the training effects that transferred to untrained objects: a strong reduction of the responses to mask patterns in the recorded population of neurons (selected to be responsive to the object images). One possible interpretation is that this effect would reflect an active, attention-based suppression of the irrelevant masking stimulus while executing the task. However, two task manipulations suggest that the almost complete disappearance of mask responses seemed to require a long-term training in recognizing an object presented in close temporal proximity to the masking pattern that needed to be ignored. First, this training had to be long-term because we did not see a reduction of mask responses when a monkey (monkey 2A) performed the masked same/different task without prior training in recognizing objects at very short SOAs (such effects were also absent at intermediate SOAs for which the monkey showed above-chance performance). As such, these training effects have a different time course than the effects reported in a previous study in ventral IT cortex (Jagadeesh et al., 2001), in which it was shown that one-session learning can suppress the re-

sponses to irrelevant, non-rewarded distractor stimuli in a discrimination task. At least, training has to be more intensive than what we did in that monkey (few days of training at intermediate SOAs). Second, we found that once a monkey was trained extensively in the “relevant object plus irrelevant mask” task, the reduction in mask responses was also present in the contrast same/different task, an attentionally demanding task for which neither the object nor the mask was relevant (as in the fixation task). Thus, the reduction in mask responses not only requires long-term training, but once established, it remains present after training in another task with irrelevant objects. At least, training in this new task would have to be longer than a few months in order to find strong mask responses. Each of these task manipulations was done in only one monkey (a different one for each manipulation), but each of them supports the conclusion that the mask responses are not modulated easily by changing task demands.

Nevertheless, our data suggest that attentional processes might have an important role in the emergence of the reduction of mask responses. If not, then the same effect would occur as a consequence of passive exposure to the mask stimuli without any need to ignore or filter out these stimuli. It has been shown that passive exposure to stimuli leads to response reduction (Desimone, 1996; Ringo, 1996). Even though we regularly changed the mask stimuli, this reduction might generalize to the whole class of mask stimuli once many individual exemplars have been presented. However, if passive exposure were important, then it would also affect mask responses in untrained monkeys when they have passively viewed many mask stimuli. This is not what we found: The control experiment in monkey 2A in the masked same/different task was performed after all of the recordings in passive viewing. Nevertheless, there was no tendency toward weaker mask responses in this control experiment. Thus, even though we need further experiments to determine the exact role of attentional and task-related processes for the emergence of the reduction of mask responses, the present data suggest that this reduction does not occur as a consequence of passive exposure. In addition, there are other unanswered questions about the reduction in mask responses. For example, we did not investigate the time course of this effect as a function of the length of the training interval, and we do not know how the reduction in mask responses generalizes to different sorts of mask stimuli, that is, how mask-specific the effect is.

Together, the present findings reveal how the visual system learns to optimize its processing of very transient and masked visual input, far beyond the general ability in fast object recognition and categorization in naive, untrained primates (Kovacs et al., 1995; Rieger et al., 2005; Bacon-Macé et al., 2005). A remaining question is to what extent the effects are specific to the backward masking paradigm, that is, would the same effects be found with another perceptual learning paradigm? At a general level, we always expect visual training to result in a more optimized processing of the relevant aspects of the visual input. In our results with the backward masking paradigm,

this is reflected by an object-specific increase of selectivity for the relevant objects in one monkey, together with an overall reduction of mask responses. Each of these training effects makes the response to the object plus mask combination more driven by the relevant object and less by the irrelevant mask. However, the sort of learning-induced neural changes might depend on how visual processing is challenged during the training period. The specific challenge posed by the backward masking paradigm is that objects are only presented for a very short time, and after that (during the mask presentation), no useful information is left on the screen, but instead the mask can interfere with stimulus processing. The visual system can increase the time window available for processing the relevant object by shutting down the discrepant new input (the mask) and re-iterating the signal from the presented objects. Recent psychophysical and neurophysiological work (Di Lollo et al., 2000; Enns and Di Lollo, 2002; Lamme et al., 2002; Rieger et al., 2005) has suggested that it is exactly this kind of re-entrant processing that is disrupted by masking, and our results reveal how visual processing can be tuned through an extensive training period in order to decrease the amount of interference of the mask on the processing of the object. Other perceptual learning paradigms, such as adding noise to object images (Rainer et al., 2004) pose other challenges to the visual system, and as such the way in which the processing of relevant information is optimized is expected to be different.

CONCLUSION

In conclusion, our results suggest that training in backward masking is associated with a marked decrease in mask responses, in addition to a transient object-specific increase in object selectivity at short SOA in one monkey. These new results reveal how the visual system deals with the specific constraints of a perceptual learning paradigm, such as backward masking, in order to enhance the processing of degraded but relevant objects and to suppress the processing of salient but irrelevant stimuli.

Acknowledgment—We thank M. De Paep, P. Kayenbergh, G. Meulemans, and G. Vanparrys for technical assistance, and C. Baker and D. Cox for their helpful comments on a previous version of the manuscript. This work was funded by GSKE and IUAP 5/10 (R.V.), IDO/02/004 (R.V. and J.W.), Human Frontier Science Program (RGP 18/2004 (R.V.)). H.O. was supported by the Human Frontier Science Program and the Fund for Scientific Research Flanders.

REFERENCES

- Ahissar M, Hochstein S (1996) Learning pop-out detection: Specificities to stimulus characteristics, II: Retinotopic organization. *Vision Res* 36:3487–3500.
- Bacon-Macé N, Macé MJ, Fabre-Thorpe M, Thorpe SJ (2005) The time course of visual processing: Backward masking and natural scene categorisation. *Vision Res* 45:1459–1469.
- Baker CI, Behrmann M, Olson CR (2002) Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nat Neurosci* 5:1210–1216.

- Desimone R (1996) Neural mechanisms for visual memory and their role in attention. *Proc Natl Acad Sci U S A* 26:13494–13499.
- Di Lollo V, Enns JT, Rensink RA (2000) Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *J Exp Psychol Gen* 129:481–507.
- Enns JT, Di Lollo V (2002) What's new in visual masking? *Trends Cogn Sci* 4:345–352.
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2006) Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cereb Cortex* 16:1631–1644.
- Furmanski C, Engel S (2000) Perceptual learning in object recognition: Object specificity and size invariance. *Vision Res* 40:473–484.
- Gilbert CD, Sigman M, Crist RE (2001) The neural basis of perceptual learning. *Neuron* 31:681–697.
- Gold J, Bennett PJ, Sekuler AB (1999) Signal but not noise changes with perceptual learning. *Nature* 402:176–178.
- Goldstone RL (1998) Perceptual learning. *Annu Rev Psychol* 49: 585–612.
- Grill-Spector K, Kushnir T, Hendler T, Malach R (2000) The dynamics of object-selective activation correlate with recognition performance in humans. *Nat Neurosci* 8:837–843.
- Hall G (1991) *Perceptual and associative learning*. Oxford, England: Clarendon Press.
- Hanes DP, Thompson KG, Schall JD (1995) Relationship of presaccadic activity in frontal eye field and supplementary eye field to saccade initiation in macaque: Poisson spike train analysis. *Exp Brain Res* 103:85–96.
- Jagadeesh B, Chelazzi L, Mishkin M, Desimone R (2001) Learning increases stimulus salience in anterior inferior temporal cortex of the macaque. *J Neurophysiol* 86:290–303.
- Kobatake E, Wang G, Tanaka K (1998) Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *J Neurophysiol* 80:324–330.
- Kovacs G, Vogels R, Orban G (1995) Cortical correlate of pattern backward masking. *Proc Natl Acad Sci U S A* 92:5587–5591.
- Lamme VA, Zipser K, Spekreijse H (2002) Masking interrupts figure-ground signals in V1. *J Cogn Neurosci* 14:1044–1053.
- Li L, Miller EK, Desimone R (1993) The representation of stimulus-familiarity in anterior inferior temporal cortex. *J Neurophysiol* 69:1918–1929.
- Logothetis NK, Sheinberg DL (1996) Visual object recognition. *Annu Rev Neurosci* 19:577–621.
- Op de Beeck H, Vogels R (2000) Spatial sensitivity of macaque inferior temporal neurons. *J Comp Neurol* 426:505–518.
- Op de Beeck H, Wagemans J, Vogels R (2001) Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nat Neurosci* 4:1244–1252.
- Rainer G, Lee H, Logothetis NK (2004) The effect of learning on the function of monkey extrastriate visual cortex. *PLoS Biol* 2:E44.
- Rainer G, Miller EK (2000) Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* 27:179–189.
- Rieger JW, Braun C, Bühlhoff HH, Gegenfurtner KR (2005) The dynamics of visual pattern masking in natural scene processing: A magnetoencephalography study. *J Vision* 5:275–286.
- Ringo JL (1996) Stimulus specific adaptation in inferior temporal and medial temporal cortex of the monkey. *Behav Brain Res* 76:191–197.
- Robinson DA (1963) A method of measuring eye movements using a scleral search coil in a magnetic field. *IEEE Trans Biomed Eng* 101:131–145.
- Rolls ET, Tovee MJ (1994) Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proc R Soc Lond B Biol Sci* 257:9–15.
- Rolls ET, Tovee MJ, Panzeri S (1999) The neurophysiology of backward visual masking: information analysis. *J Cogn Neurosci* 11:300–311.
- Rossion B, Pourtois G (2004) Revisiting Snodgrass and Vanderwart's object databank: the role of surface detail in basic level object recognition. *Perception* 33:217–236.
- Sakai K, Miyashita Y (1994) Neuronal tuning to learned complex forms in vision. *Neuroreport* 5:829–832.
- Sigala N, Logothetis NK (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415:318–320.
- Tanaka K (1996) Inferotemporal cortex and object vision. *Annu Rev Neurosci* 19:109–139.
- Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Nature* 6:520–522.
- Vogels R (1999) Categorization of complex visual images by rhesus monkeys. Part 2: single-cell study. *Eur J Neurosci* 11:1239–1255.
- Wichmann FA, Hill NJ (2001) The psychometric function I: fitting, sampling and goodness-of-fit. *Percept Psychophys* 63:1293–1313.
- Yang T, Maunsell JHR (2004) The effect of perceptual learning on neuronal responses in monkey visual area V4. *J Neurosci* 24:1617–1626.

(Accepted 18 December 2006)
(Available online 9 February 2007)