Invariance of long-term visual priming to scale, reflection, translation, and hemisphere

József Fiser [a], Irving Biederman [b,*]

[a] *Department of Brain and Cognitive Sciences, Center for Visual Sciences, University of Rochester, Rochester, NY 14627, USA*
[b] *Department of Psychology and Neuroscience Program, Hedco Neurosciences Building, MC 2520, University of Southern California, Los Angeles, CA 90089-2530, USA*

Abstract

The representation of shape mediating visual object priming was investigated. In two blocks of trials, subjects named images of common objects presented for 185 ms that were bandpass filtered, either at high (10 cpd) or at low (2 cpd) center frequency with a 1.5 octave bandwidth, and positioned either 5° right or left of fixation. The second presentation of an image of a given object type could be filtered at the same or different band, be shown at the same or translated (and mirror reflected) position, and be the same exemplar as that in the first block or a same-name different-shaped exemplar (e.g. a different kind of chair). Second block reaction times (RTs) and error rates were markedly lower than they were on the first block, which, in the context of prior results, was indicative of strong priming. A change of exemplar in the second block resulted in a significant cost in RTs and error rates, indicating that a portion of the priming was visual and not just verbal or basic-level conceptual. However, a change in the spatial frequency (SF) content of the image had no effect on priming despite the dramatic difference it made in appearance of the objects. This invariance to SF changes was also preserved with centrally presented images in a second experiment. Priming was also invariant to a change in left–right position (and mirror orientation) of the image. The invariance over translation of such a large magnitude suggests that the locus of the representation mediating the priming is beyond an area that would be homologous to posterior TEO in the monkey. We conclude that this representation is insensitive to low level image variations (e.g. SF, precise position or orientation of features) that do not alter the basic part-structure of the object. Finally, recognition performance was unaffected by whether low or high bandpassed images were presented either in the left or right visual field, giving no support to the hypothesis of hemispheric differences in processing low and high spatial frequencies. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Object recognition; Spatial frequency; Shape representation; Ventral pathway; Hemisphere laterality

1. Introduction

Visual repetition priming has been a major tool for assessing the nature of the representation mediating object recognition by humans (Bartram, 1974; Intraub, 1981; Cooper, Biederman, & Hummel, 1992). Priming refers to the phenomenon that the identification (often assessed by naming) of a briefly presented picture of an object is faster and more accurate on its second presentation than control items not initially shown.[1] That the priming is visual and not just verbal or conceptual is evidenced by the greater facilitation in the naming of the identical image compared with one that has the same name (and belongs to the same basic-level class) but a different shape, e.g. an upright piano as a prime

* Corresponding author. Tel.: +1-213-7406094; fax: +1-213-7405687.
E-mail addresses: fiser@bcs.rochester.edu (J. Fiser), bieder@usc.edu (I. Biederman).

[1] This type of *repetition* priming in which there is perceptual facilitation is generally regarded as an implicit memory effect (Schacter & Tulving, 1994). There is also an explicit memory effect by which one might recall the stimulus or recognize it as one that has been presented previously. A number of operations and clinical conditions can distinguish these two types of memory representations (Biederman & Cooper, 1992; Schacter & Tulving, 1994).

for a grand piano.[2] Visual priming can be long term. In most studies it is measured over minutes between first and second presentations (e.g. Bartram; Cooper, Biederman, & Hummel) but has been documented over months (Cave, 1997). Several studies have shown that the magnitude of long-term visual priming is independent of change in the retinal size, position, mirror reflection, and orientation in depth (as long as the original parts can be easily resolved) of the primed object compared with that of the prime (Biederman & Cooper, 1991a, 1992; Biederman & Gerhardstein, 1993; Cooper, Biederman, & Hummel; Fiser & Biederman, 1995)[3].

The objective of the present study was to (a) extend the above results by investigating the specificity of the representation mediating long-term visual object priming in terms of spatial frequency (SF) content and (b) to assess the locus of this representation within the ventral pathway presumed to be mediating visual object recognition. To this end, we conducted an experiment in which subjects named briefly flashed (185 ms), band-passed images of common objects on two presentations, separated by several minutes. The image on the second presentation could differ from that on the first in SF content (2 vs. 10 cpd center frequency) and position (5° left or right of fixation) in the visual field.

---

[2] Consistent with these results are findings from a number of studies showing that reading the names of objects a few minutes before naming pictures or judging whether they are real objects or not has no effect on RTs or error rates; e.g. Kroll and Potter (1984), Biederman (1987). To distinguish explicit memory from implicit priming effects, subliminal priming tasks have often been employed. In these paradigms, conceptual priming of stimulus identification has been shown to be short-lived, not exceeding a few 100 ms (Greenwald, Draine, & Abrams, 1996) but visual priming of pictures extends at least over 10 min (Bar & Biederman, 1998). Biederman and Cooper (1991b) showed that there was no contribution to visual priming at the level of subordinate-level concepts (as opposed to the percept of the sub-ordinates parts) either, as would be expected if the concept of an upright piano, for example, would facilitate the naming of an image of an upright piano more than a grand piano. Their evidence was that a complex object image with half its parts did not prime its complementary image, i.e. an image with the other half of the parts, any more than it primed the same name, different shaped exemplar, i.e. an image of half the parts of an upright piano.

[3] Short-term (or 'hot trail') priming, assessed by same-different matching with brief (100 ms) ISIs, does show costs when the second stimulus is presented at a different position or size (Ellis & Allport, 1986; Magnussen & Greenlee, 1999). These costs are eliminated, however, with an intervening mask or a longer (1 s) ISI (Ellis & Allport, 1986). It is possible that the costs are a consequence of the subject being able to respond 'same' if a transient is not present between the two presentations (Biederman & Bar, 1999). The transient could be produced either by a change in shape, to which the subject should respond 'different,' or the changes in position or size of the same stimulus, to which the subject should respond 'same.' The greater uncertainty on trials with a transient present would serve to increase same RTs and lead to an apparent cost of position or size changes.

1.1. Inferring the locus in the ventral pathway of the representation mediating priming

There is widespread consensus that the visual areas most heavily involved in object recognition are located along the ventral pathway of the primate visual cortex and include areas V1, V2, V4, TEO and TE (Mishkin, Ungerleider, & Macko, 1983; Van Essen, 1985; Maunsell, 1995). There are several findings that can help to assess the locus of visual priming within this pathway. First, receptive field sizes in the foveal and parafoveal region in V1 are between 0.1 and 2° in visual angle and in V2 between 0.2 and 3° (Dow, Snyder, Vautin, & Bauer, 1981; Roe & Ts'o, 1995). It is only in V4 where receptive field sizes can exceed 7° in diameter (Van Essen). Second, the two visual hemifields are represented separately in the two hemispheres in V1 and V2 with an overlap along the vertical midline that the size of one RF. Third, although the RFs in areas V4 and TEO are much larger than in V1 and V2, in monkeys there exists a strong correlation between eccentricity and RF sizes with smaller RFs being closer to the fovea. As a result, even in V4 and TEO the visual representation in one hemisphere extends only slightly over the vertical midline, typically less than 2° for cells with RF centers within 10° of eccentricity (Boussaoud, Desimone, & Ungerleider, 1991). Finally, in the macaque, TE is the first area to have full field, feed forward connections. The homologue to TE is regarded by some authors to be the anterior portion of the fusiform gyrus, which has full field enervation from the contralateral hemisphere (Halgren et al., 1999).

Thus anything presented in the right visual field at more than 3° of eccentricity, and therefore initially processed in the left hemisphere, cannot directly (i.e. in a feed-forward manner) activate cells in the right hemisphere in V1 and V2 (and can do so only by feedback from more anterior areas). Similarly, such stimuli can excite cells in V4 or TEO[4] only indirectly by feedback and/or via callosal connections. If priming requires that a substantial proportion of the same cells or cell groups be activated during the two presentations of the stimulus, then pictures presented at widely different positions (> 8° apart) on the retina in different hemifields should not prime each other unless priming occurs beyond TEO, or there are very specific feedback or callosal connections involved. Therefore, assuming substantial overlap in the cells activated by first and second presentations for sizable visual priming, a priming study with laterally translated images could address the question as to the locus of the representation mediating priming in the ventral pathway.

---

[4] In referring to areas TEO and TE in humans, we are referring to human homologues of these areas (assuming that they exist).

Biederman and Cooper (1991a) tested the effect of translation and reflection on priming using line drawings with a maximum extent of 4°, positioned 2.4° left or right of the center of fixation. They found that a line drawing image could prime itself equally well regardless of whether in the second block it appeared in the exact same position and orientation or was presented equally far but on the other side of the fixation point and mirror reflected along its vertical midline. Biederman and Cooper concluded that the primed representation must be position and orientation independent.

We set out to replicate and extend Biederman and Cooper's (1991a) experiment with SF filtered, gray level images. There were two aspects of our design that allowed a more stringent test of the role of early ventral stages than that employed by Biederman and Cooper. First, the 2.4° of translation for an image that was 4° in maximum extent in their study might have been insufficient to completely exclude a partial contribution of even V2 as the anatomical locus of priming, if the receptive fields in V2, indeed, extend up to 2° in the opposite hemifield. We used 10° of separation between the centers of images with a maximum extent of 6° presented on the left or right side of fixation. Second, by testing priming across images which differed by such a large extent in their position, SF content, and orientation we could assess whether any of the low level attributes of an image (defined as simple combinations of those attributes) would provide even a modest contribution to visual priming.

Our study also allowed us to test a claim of hemispheric specialization for the utilization (rather than just detection) of high versus low spatial frequencies. Several authors have argued that information presented at low SF is more efficiently used in the right hemisphere, whereas the left hemisphere is better at high spatial frequencies (e.g. Sergent, 1982; Sergent, 1987; Jonsson & Hellige, 1986; Robertson & Lamb, 1991; Kitterle, Christman, & Conesa, 1993; for an overview see, Hübner, 1997). This would suggest that low bandpass images would be better identified in the left visual field, and high bandpass images in the right visual field.

2. Experiment 1: priming with laterally presented images

2.1. Subjects

Sixty-four native English speakers with normal or corrected-to-normal vision participated for credit points in their Introductory Psychology course at the University of Southern California. In both experiments in this investigation, the subjects were naive with respect to the goals of the experiment.

2.2. Stimuli

Sixty-four gray-level images of everyday objects from 32 categories were used in the experiment. There were two visually distinct images of each category, such as a high-heeled woman's shoe and a man's walking shoe for the category 'shoe.' The inclusion of these same-name-different-shaped exemplars allowed assessment of the extent to which the priming (if any) could be attributed to visual versus verbal/semantic factors. The maximum extent of each object was normalized to 6° in diameter. Two SF filtered versions of each image were created by the following method using a commercial image-processing package (KBVISION).

The images were Fourier transformed and bandpass filtered cutting off high frequencies above 16.4 cpd and low frequencies below 1 cpd in the Fourier-domain. This filtering left the Fourier coefficients within a wide ring of the Fourier domain intact, and erased all the coefficients outside. When this representation was transformed back to the space domain, the original image was obtained with very little degradation in quality. Two narrower rings were selected (by two bandpass filtering operations) within the wide ring of coefficients in the Fourier domain. Both rings were 1.5 octaves wide, and there was an octave wide gap between the two rings (Fig. 1). Ten and two cycles per degree were the center frequencies for the high frequency and low frequency bandpass filters, respectively.

There were three slightly incompatible measurements for selecting these filtering parameters, all related to our goal of having a reasonable level of identification performance with both low- and high-bandpassed images while maintaining the maximum possible distinctiveness in their SF. First, the center frequencies had to be preferentially symmetrical to the known peak value of human contrast sensitivity function approximately 3–5 cpd (Wilson, Levi, Maffei, Rovamo, & DeValois, 1990). Second, the two bands had to have sufficient and approximately equal amounts of information so that the bandpass images obtained after the inverse Fourier transformation would be subjectively equivalent in identifiability. This measurement required as wide rings as possible, and the center of the low frequency band to be at a relatively high absolute frequency so that the low bandpass images could be identified at all given our selected presentation duration of 185 ms. This was the longest possible presentation time given our displays that would still be sufficiently brief to prevent subjects from making a saccade before the image disappeared from the screen.

The third measurement was based on the known 1/f energy distribution of spectra of natural images (Field, 1987). According to this measurement, to ensure relatively equal amounts of energy in the high- and low-pass images, the boundaries of the spectra had to be
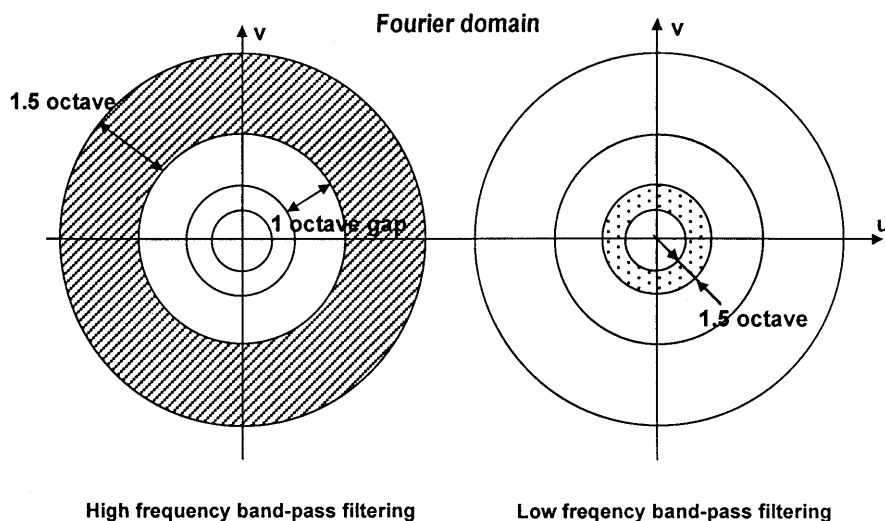
Fig. 1. The Fourier representations of the two SF filtered versions for each image. The hatched and shaded areas show the coefficients which were used in the inverse Fourier transformation to obtain the two bandpass filtered images. The one octave gap between the spectrum of the two images ensured minimal overlap between the simple cell population responding to both of the two images.

selected to be proportionally further away from each other for the high pass images than for the low pass ones.

In addition, we needed to have a large gap between the spectra of the high and low bandpass filtered images in order to prevent cells in the primary visual cortex from receiving input from both bandpass filtered versions of the image as much as possible. The average SF bandwidth in macaque V1 is 1–1.4 octaves with larger bandwidths at lower center frequencies and narrower bandwidths at higher center frequencies (DeValois & DeValois, 1988). We created a one-octave gap by removing all information from the image between 3 and 6 cpd.

The final set of parameters was almost the only compromise to satisfy at least partially all these requirements. Fig. 2 shows two examples of SF filtered images used in the experiment. The images were presented on a 16-in. Apple monitor (832×624 pixels resolution) from a 1 m viewing distance.

2.3. Procedure

The subjects pressed a mouse button to start each trial. A fixation cue would then be presented in the middle of the screen for 500 ms, followed by a 185 ms presentation of the object picture, and then by a mask for 500 ms (Fig. 3). The fixation cue was a bandpass filtered dot. The lower and higher limits of the band were the same as the lower and the higher limits of the one octave gap between the low and high filtered images, so the filtered cue dot had no orientation or SF bias. The images were presented 5° left or right of the fixation cue.

When an image was translated to the opposite hemifield, it was always mirror reflected as well. For example, if an airplane presented left of fixation was pointing toward the center it would also point toward the center when presented on the right of fixation. It is known that with increasing eccentricity the retinal and cortical sampling density of a given area decreases monotonically (Rovamo & Virsu, 1979; Wässle, Gruenert, Rohrenbeck, & Boycott, 1991). This means that when a large image that has more characteristic features on one side is positioned at the same eccentricity
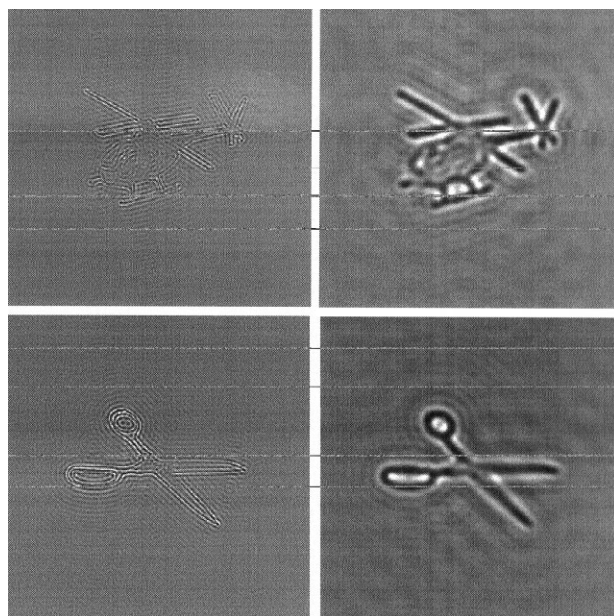

Fig. 2. Two pairs of high and low bandpass filtered images used in the experiment, shown at reduced sizes. The 'rings' around the objects are artificially amplified due to the size reduction of these images. In the original stimuli they were not apparent.
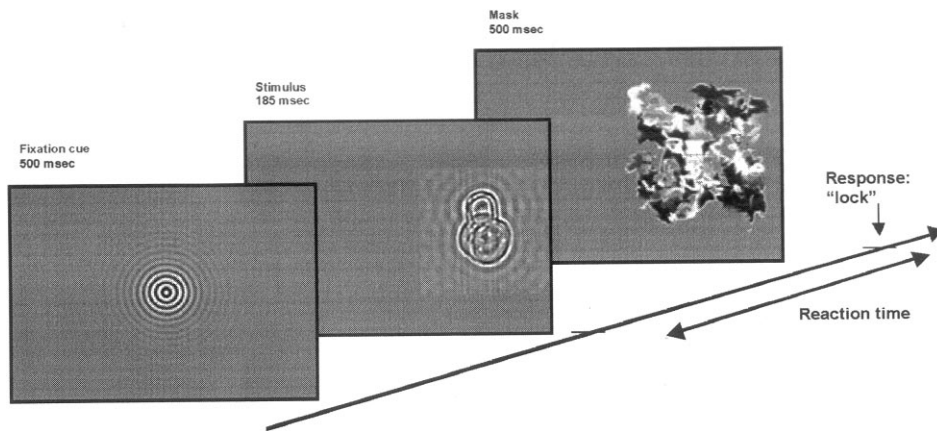
Fig. 3. Schematic representation of one trial of the experiment. The bandpass filtered, unbiased cue was followed by a bandpassed image positioned 5° left or right of center which, in turn, was followed by a mask in the same position.

left or right of fixation without mirror reflection, it will be easier or harder to identify, depending on whether the cue-rich side is closer to or farther from the midline (Cooper et al., 1992). Thus mirror reflection eliminated potential differences due to any uneven distribution of features on the two sides of the object. In addition, the reflection altered the local orientation of every non-horizontal or vertical feature (contour segment, corner, etc.) when the same image was presented on different sides.

Four rotated gray-scale masks were used in randomized order. The masks were created by superimposing small segments of different images and random patterns blurred to different degrees, so that the energy distribution of the masks were comparable to a natural image, and they had similar luminance structures to that of everyday object images.

The subject started each trial by pressing the mouse button. They were instructed to name the image as quickly as possible using basic-level names and to ignore any variation in SF or the level of blur or position. The category names were not given in advance, but subjects were told that the objects would be everyday objects. They were also told that the object images would be filtered in different ways, and they would look like incomplete imprints in sand rather than clear images. The naming RTs were measured by means of a voice key. Subjects were given feedback of the correct name and their RT after each trial. There were 12 familiarization trials prior to the experiment with images not used in the main experiment.

2.4. Design

Each subject named 32 pictures of objects on each of the two blocks. In each block half of the objects were high, the other half were low bandpass filtered. Half of the images were presented 5° left and the other half were presented 5° right of the center of fixation. Each subject saw only two images of each basic-level category in the entire experiment, one in the first and the other in the second block. For each subject, half of the objects in the second block were different exemplars of the objects shown in the first block, the other half were identical[5]. One eighth of the pictures were in one of the eight possible conditions (two exemplars × left or right position × high or low SF) for each subject. The sequences of images were balanced across subjects so that every object appeared equally often in the eight conditions. Approximately 7 min intervened between the first and the second presentation of an object.

Four analyses of variance (ANOVAs) were computed on the RTs and error rates of the first and second blocks. The fixed factors of the ANOVA of the first block were side (left vs. right) and SF (high vs. low). The fixed factors of the second block were exemplar (same vs. different), side (same-different), and SF (same vs. different). The random factor for all the analyses was subjects.

2.5. Results

Reaction times (RTs) and error rates on the first block and the second block are shown in Fig. 4. The analysis of variance (ANOVA) of the first block RTs revealed no main effect for either side, $F(1, 63) < 1.00$, ns., or for SF filtering, $F(1, 63) = 1.28$, ns., and there

---

[5] This means that there were no pictures of new object categories in the second block. Such images would allow assessment of the lexical/semantic factors to the overall improvement from block 1 to block 2. Our interest was in the visual portion of priming, for which a lower bounds estimate can be obtained by the difference in priming between the identical picture and a same name, different-shaped exemplar of the same basic-level object class. It is a lower bounds estimate in that different exemplars from the same basic-level class tend to be more similar than objects from different basic-level classes. Consequently, a portion of the block 1 to block 2 improvement for the different exemplars can be attributed to visual priming as well.

was no interaction between the factors, $F(1, 63) = 1.36$, ns. Similarly, there was no effect of side or SF on error rates, $F(1, 63) = 1.38$, ns., and $F(1, 63) = 1.59$, ns., respectively, and no interaction between them, $F(1, 63) = 0.67$, ns. In other words, we found no advantage in object naming of presenting high or low bandpassed images to a particular hemifield, nor did we find an advantage of high over low bandpassed versions of an image or vice versa when presenting them at an eccentricity of 5°. The absence of effects of side, SF, and their interaction was also evident on the second block

(Fig. 4 bottom). All the $Fs(1, 64)$ for these variables were < 1.00 for both RTs and error rates.

The bandpass filtering and lateralized presentation made the identification of the objects much more difficult than would be expected from centrally presented, unfiltered images, with first block accuracy of approximately 60% (chance basic-level naming would be less than 0.01%). There was a significant reduction in both RTs and error rates from the first block to the second one, $t(63) = 6.26$, $P < 0.001$, and error rates, $t(63) = 10.79$, $P < 0.001$ (Fig. 5). This analysis included
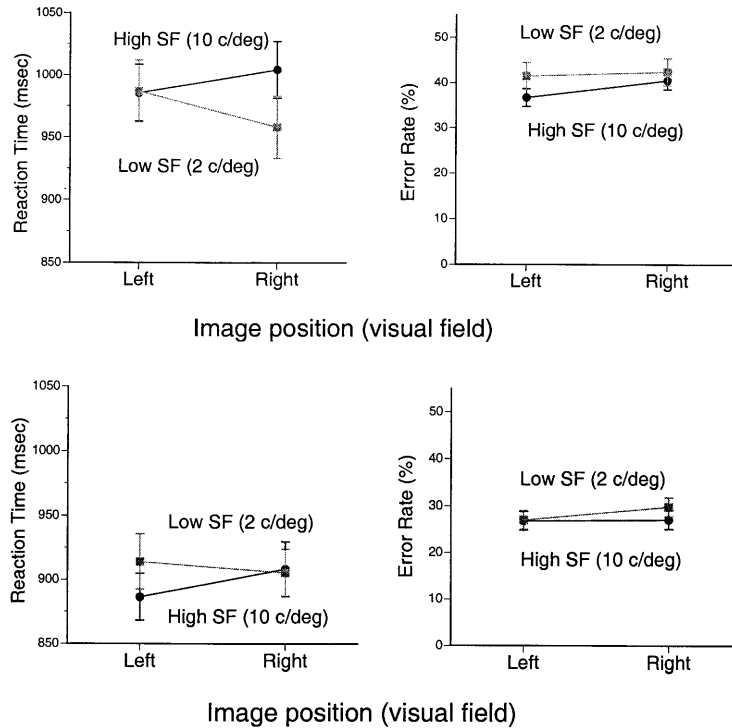

Fig. 4. First block (top) and second block (bottom) mean correct naming RTs (left panels) and error rates (right panels) as a function of Field of Presentation (left or right) and the SF of the image (high or low passed). Note that the *y*-axis in the left graphs starts at 850 ms, rather than at zero. Error bars in this figure and in all the following figures represent standard errors.
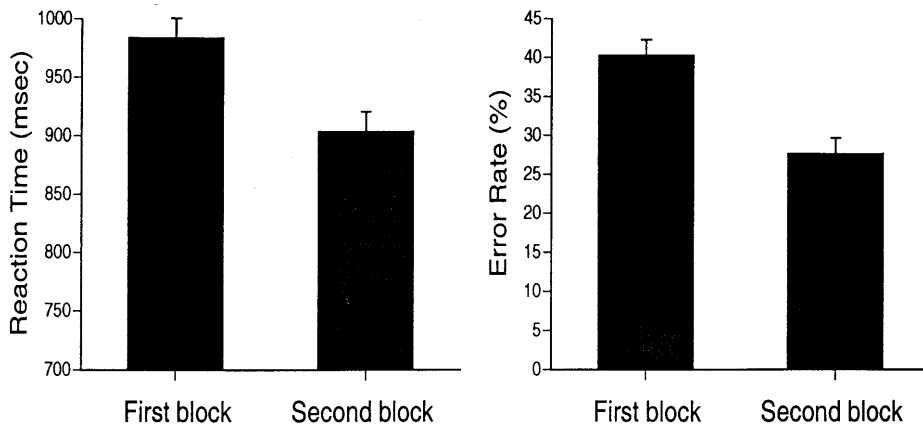

Fig. 5. Mean correct naming RTs and error rates for the first and second blocks in Experiment 1. The reduction in both measures was highly significant (both *P*s < 0.001).
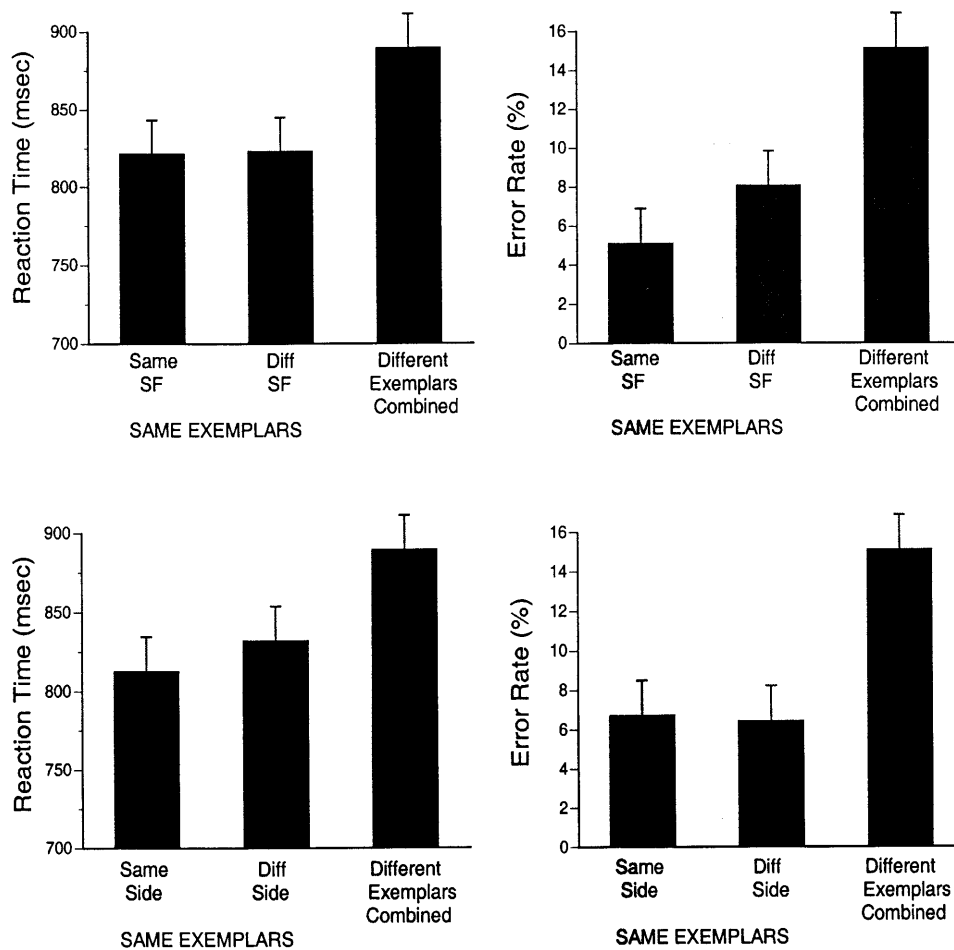
Fig. 6. Top, second block results of Experiment 1 collapsed across presentation position (left or right). Bottom, second block priming results collapsed across SF filtering (high or low passed). No main effect of either SF or size was found.

all trials in the second block irrespective of whether the image was identified correctly or not in the first block.

To assess the specificity of visual priming, the second block results were analyzed only for trials whose objects were successfully identified in the first block irrespective of their side, exemplar, or SF. Fig. 6 shows the results of these second block trials collapsed across left and right presentations (top), and the same data collapsed across different spatial frequencies (bottom). The significantly lower RTs and error rates for the same exemplars compared to the different exemplars in the $2 \times 2 \times 2$ (exemplar $\times$ side $\times$ SF) ANOVA indicated that there was visual priming in the second block, and not just semantic or verbal priming. $F(1, 63) = 12.71$, $P < 0.001$ for RTs, and $F(1, 63) = 22.54$, $P < 0.001$ for error rates.

No main effect of side change was found either in RTs, $F(1, 63) = 1.27$, ns., or in error rates, $F(1, 63) = 1.14$, ns. Similarly, changing SF content had no effect on RTs or error rates, both $Fs(1, 63) < 1.00$, ns. None of the two-way or three-way interactions were significant at the 0.05 level for either RTs or error rates. Thus

presenting the same image at opposite sides of the fixation point or in widely different SF ranges in the first and the second block had no effect on the magnitude of priming, whereas presenting a different exemplar of a given category in the second block led to a large and highly significant reduction of priming. This result — the lack of an effect from a change in SF content — would appear to pose a challenge to those theories that assign a central role to 'appearance-based' representations in object recognition (e.g. Poggio & Edelman, 1990), since the change in SF content resulted in a dramatic change in the appearance of the object.

The different exemplar trials were collapsed into a single bar in the graphs of Fig. 6 since different exemplars in the second block meant that the subjects did not see the actual image in any form in the first block (they saw instead a same name, different shaped object). Thus breaking down the different exemplar category into same/different SF and side in a graph showing aspects of visual priming had little meaning. Nevertheless, it is interesting to ask whether there was a hint of side or SF effect in semantic priming in our data

— even if just a non-significant one. This type of priming could cause the appearance of a chair — any chair — in the second block to benefit from the fact that a chair in the first block appeared on the left side, so 'chair' and 'left' somehow got associated. However this was not the case in our experiment. None of the F values for the two- and three-way interactions for RTs and error rates were significant, typically below or close to 1.00 (with the exception of a two-way interaction in RTs between side and SF $F(1, 63) = 3.87$, $P = 0.054$), and showed no similar tendencies for RTs and errors for any of the interactions.

The advantage of the same over the different exemplars in the block 2 analyses described above excluded those pictures that a participant missed on the first block. If all the block 2 data are included, independent of whether the picture was correctly named on block 1, then the RT advantage of same over different exemplars observed for the restricted data (67.2 ms) is reduced in the full data set to 36.9 ms but remains significant, $F(1, 63) = 6.15$, $P < 0.02$. However, the 8.5% advantage for same over different exemplars for errors is reduced to 1.9%, which is no longer significant, $F(1.24)$, $P = 0.26$. The absence of costs for changes in position and SF remain, with all Fs below 0.2 for errors and just slightly above 1.00 for RTs.

We have no explanation (other than bad luck) as to why inclusion of all the second block data, independent of whether that picture was accurately identified on the first block, reduced the advantage of same over different exemplars for the error rates to non-significance. If there was no priming activation from missed block 1 stimulus exposures, at worst we would have expected the exemplar effect to be reduced by about 40% (the proportion of trials that were in error) to approximately 6% rather than the observed 2%. We note that in Experiment 2, as well as other studies of priming (e.g. Biederman & Cooper, 1991b), there was no effect on the magnitude of the exemplar effect from excluding data from block 2 trials where the picture was not accurately named on block 1.

One possible reason why the exemplar effect was magnified with exclusion of trials on block 2 of pictures that were missed on block 1 could be that for many of the objects, the two exemplars were not equally easy to identify. On the first block, subjects would identify the easy versions more frequently than the hard ones. Since in the second block analysis only those objects were considered that were successfully identified in the first block, a greater proportion of the second block same-exemplar pictures would be easy compared with the different exemplar object pictures, as the latter would have a higher proportion of difficult pictures which had the easier exemplar on block 1. Thus the main effect of exemplar could simply have been the result of unequal baseline difficulty in identification rather than that of weaker priming.

In order to test this possibility, the data were reanalyzed using OBJECTS rather than SUBJECTS as the random variable. This was possible because our design completely balanced objects across subjects. As a result, in the second block we computed average errors and RTs for each exemplar of each object rather than for each subject. This means that each exemplar of each object weighted equally in shaping the second block results, because one mean RT and one mean error rate was calculated for each one, given there were no objects that were missed by all the subjects. However, some of the exemplars of some objects were missed by all the subjects in a given condition (27 for error rates and 34 for RTs out of the 256 conditions of block 2). For these objects we excluded from the analysis the data of both the missed (hard) exemplar and that of the other (easy) exemplar of the object. This way of excluding trials with objects which were not identified successfully on the first block could not introduce any bias based on exemplar selection in the analysis, but instead only weakened the power of the ANOVA test.

The results of the first block errors were exactly the same as in the analysis by subjects since summing the data by objects or subjects makes no difference. The mean RTs of the first block were 1041, 1050, 1050 and 1067 ms for left-high, left-low, right-high and high-low conditions, respectively. The small differences between these means and those in Fig. 4, were due to object exclusions. The ANOVA of the first block RTs revealed no significant main effect for either side or for SF, and there was no interaction between the factors, all $F(1, 52) < 1.00$, ns.

The second block error rates were 6.9, 13.9, 7.0 and 12.3% for the same exemplar left-high, left-low, right-high and high-low conditions, respectively, and was 19.6% for the different exemplars combined. The $2 \times 2 \times 2$ (exemplar $\times$ side $\times$ SF) ANOVA found a strong main effect of exemplar, $F(1, 38) = 6.66$, $P < 0.014$, but no effect for side, $F(1, 38) < 1.00$, ns., or for SF $F(1, 38) < 1.00$, ns. None of the two-way or three-way interactions were significant except for a two-way exemplar $\times$ SF interaction, $F(1, 38) = 4.74$, $P < 0.05$.

The second block RTs were 840, 875, 834 and 867 ms for the same exemplar left-high, left-low, right-high and high-low conditions, respectively, and was 923 ms for the different exemplars combined. The $2 \times 2 \times 2$ (exemplar $\times$ side $\times$ SF) ANOVA found an almost significant exemplar main effect, $F(1, 31) = 4.07$, $P = 0.052$, but no effect for side $F(1, 31) = 2.32$, ns., or for SF $F(1, 31) = 2.02$, ns. None of the two-way or three-way interactions were significant.

These results show that when the exemplar differences due to uneven performance of the subjects are completely balanced across all objects, the second block exemplar effect still remains just as in the original analysis. This indicates that the significant difference in

performance between same and different exemplars in the original analysis cannot be attributed to stimulus selection biases.

3. Experiment 2: foveal versus lateralized stimulus presentation

3.1. Introduction

Recognition of objects is strongly facilitated by presenting the stimuli at fixation as opposed to lateralized presentations, suggesting somewhat different processing of visual information within versus outside the fovea. Jüttner and Rentschler (Jüttner & Rentschler, 2000) recently demonstrated that a classification of variations from three Gabor prototypes that could be learned by human subjects at foveal presentations could not be learned if presented at an eccentricity of 2.5°. This striking interaction between cognition and visual eccentricity held true despite size scaling of the patterns so they were equally discriminable at 0° and 2.5° eccentricity. It is possible that whatever is responsible for the difficulty of learning classes in the Jüttner and Rentschler task is also serving to reduce the specificity of priming. Would invariance to a change in SF content still characterize visual priming if the stimulus presentations were at fixation?

3.2. Method

To address the above question, we ran a group of 64 subjects in Experiment 2 with foveal presentation of stimuli. Experiments 1 and 2 were identical except for two aspects. First, images in Experiment 2 were presented at fixation. Second, in half of the cases the images were not SF bandpassed but filtered according to their orientation. Thus in this experiment, instead of high or low SF bandpassed images on the left or right side of fixation, the subjects saw centrally positioned, high or low SF bandpassed images, or images bandpass filtered at one of two orientation bands in the 32 trials. For the present discussion, however, only the SF bandpassed images are relevant, which were exactly the same images that were used in Experiment 1, and which were presented under identical circumstances, except at fixation. The question we addressed in both experiments was whether performance with different SF images in the second block would be more similar to that with the same SF (i.e. identical) images, or closer to those with the different exemplar images.

3.3. Results

Fig. 7 shows the second block results of Experiment 2 with centrally presented images, in comparison with the second block results of Experiment 1 with lateralized presentations. As in Experiment 1, the second block priming analysis in Experiment 2 was based on only those trials for which the objects were correctly identified in the first block. However, due to the significantly lower error rates in Experiment 2 compared with Experiment 1 (see Fig. 8), all the results of Experiment 2, including the significant exemplar effect, remained the same when all trials were included, not only the ones with objects successfully identified in the first block.

Similar to Experiment 1, the same exemplars of the SF filtered images in Experiment 2 had a significant advantage in RTs, 52 ms, $F(1, 63) = 7.34$, $P < 0.01$, and in error rates, 6.2%, $F(1, 63) = 12.17$, $P < 0.001$, over the different exemplars, indicating significant visual priming. As with laterally positioned images, there was no main effect of SF filtering (same vs. different SFs) with centrally positioned images; $F(1, 63) < 1.00$, ns., for both RTs and errors. Finally, the ANOVA showed no significant interactions between exemplar and SF in Experiment 2 for either RTs, $F(1, 63) = 1.22$, $P > 0.27$, or for error rates, $F(1, 63) = 3.7$, $P > 0.058$. In summary, shifting the images 5° to the side barely increased the RTs and error rates in the second block, and it left the structure of the priming (complete cross-priming across spatial frequencies, and a highly reliable difference between same and different exemplars) intact.

To make the analysis between the two experiments complete, Fig. 8 presents the first block results of the two experiments. In the first block of Experiment 2 the difference in RTs between centrally positioned high and low bandpass filtered images (high = 987 ms vs. low = 1027 ms) was close to significant, $t(63) = 1.71$, $P > 0.09$, and the difference in error rates (high = 18.9% vs. low = 30.5%) was highly significant, $t(63) = 4.19$, $P < 0.001$, both in favor of the high pass images. This is in contrast with the results of Experiment 1 where we found no advantage of high bandpassed images over low bandpassed ones (Fig. 4).

In order to quantify this difference, the first block error results of Experiment 1 collapsed across left and right presentations and the results from the first block Experiment 2 were reanalyzed by a 2 × 2 ANOVA with fixed factors of presentation position (Lateral vs. Central) and SF Filtering (high vs. low, Fig. 8). Lateralized presentations led to significantly higher error rates, suggesting a general decrement in performance as stimulus presentation is moved more toward the periphery, $F(1, 63) = 79.3$, $P < 0.001$. In addition, the error rates for identifying high bandpassed images were affected significantly more by positioning the images laterally compared with those of low bandpassed images. This additional increase in error rates for the laterally presented high pass images led to a significant interaction between the presentation position and SF Filtering
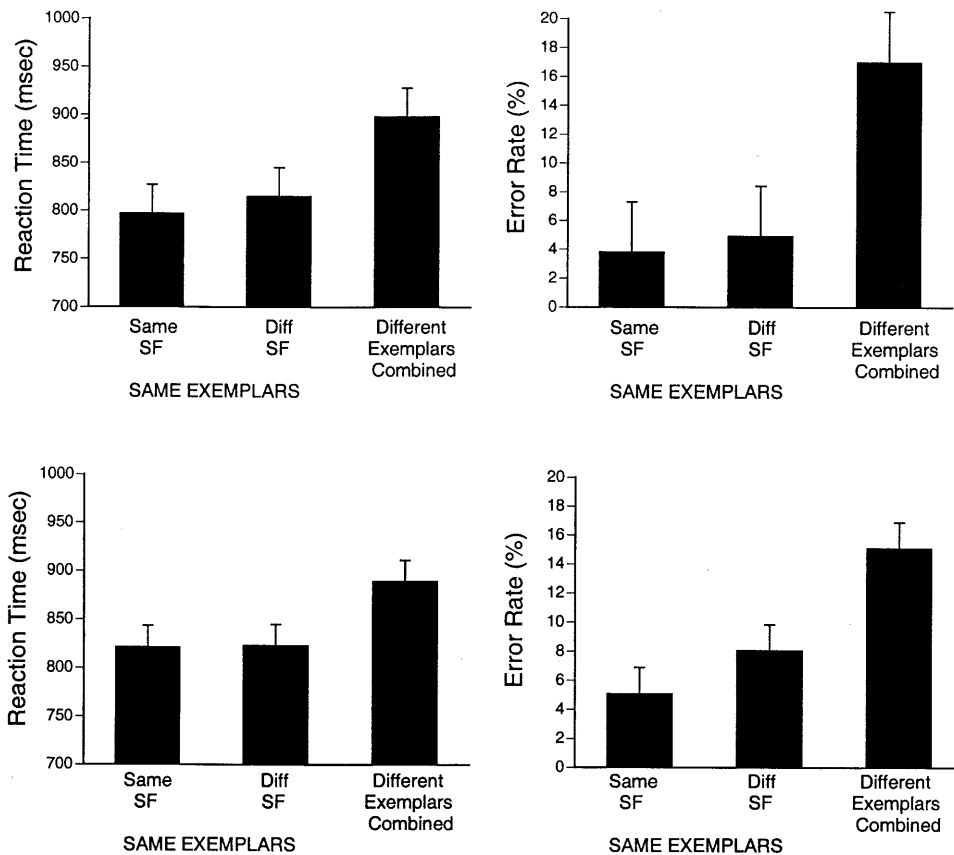
Fig. 7. Comparing centrally and laterally presented SF filtered images. Top, second block results from Experiment 2 study with images presented centrally. Bottom, second block results from Experiment 1 collapsed across presentation position (left or right).

variables in the ANOVA, $F(1, 63) = 4.76$, $P < 0.05$. A similar ANOVA of the RTs found no effects of presentation position, $F(1, 63) = 1.27$, ns., or SF filtering, $F(1, 63) < 1.00$, ns., indicating that a speed-for-accuracy tradeoff could not account for the increase in error rates between central and lateralized positions.

4. Discussion

The primary result of this study is that there is as much visual priming from an identical image of an object as there is from another image of the same object that is displaced by 10° to the corresponding position in the contralateral hemifield, reflected about its vertical midline, and has its SF shifted by more than two octaves. The magnitude of this priming is significantly greater than from an image that is in the same location and frequency spectrum as the original image, is oriented the same way, has the same name, but has a different shape; i.e. is a different exemplar.

This suggests that in contrast to semantic priming that applies to all the images in the experiment — identical, transformed (in position and/or frequency range), and the different-shaped exemplars — the orig-

inal and the transformed version of the original image activate essentially the same visual representation, as indicated by equal visual priming, whereas the different exemplar for that image activates a different visual representation.
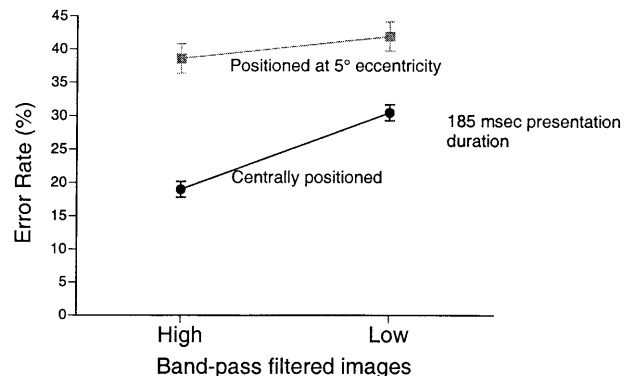

Fig. 8. The effect of SF and eccentricity on object naming error rates. The data are from the first block of Experiments 1 and 2 where the same bandpass filtered images were shown. Apart from an overall increase in error rates in laterally as compared with centrally presented images, there is an interaction signaling that presentations at larger eccentricity result in a greater increase in error rates for the high bandpassed images compared with the low bandpassed images.

Our comparison of subjects' performance with laterally and centrally presented images provides additional evidence that the primed representation is independent of low-level feature attributes. Despite large differences in initial identification performance the invariance of priming to SF changes was evidenced in both experiments. More specifically, the comparison of first block results (Fig. 8) shows a general decrement of performance with laterally presented images, which might be attributed partially to an eccentricity effect, and partially to the uncertainty of whether the picture would appear to the right or to the left of fixation. In addition to this effect, the comparison also showed selectively better performance with high bandpass images than with low bandpass images in the center compared with lateral presentations. This improvement is probably due to the increased cortical and retinal sampling and higher resolution of cells in the fovea compared with the parafoveal region. Thus, identification performance in the first block was profoundly affected by known differences in uncertainty and the sampling density of the input.

In contrast, the priming results in the two experiments (Fig. 7) were nearly identical. It did not matter whether on the first block there was a large difference in performance with low and high bandpassed images as in the case of Experiment 2 or no difference at all as in Experiment 1, the magnitudes of priming from low to high filtered images and vice versa were equal. The picture that emerges from these results is consistent with the earlier proposal of Biederman and Cooper (1991b) that visual object priming is completely mediated by an intermediate representation that can be activated by a variety of different specific image features.

There are two implications of the results of the present study. First, the manipulation introduced in this experiment affected the three arguably most important dimensions of the spatial filters presumed to mediate the initial representation of shape in the visual system, position, local SF, and local orientation. Virtually all proposed recognition schemes draw heavily on these attributes of an image in order to derive object descriptions. Our results suggest that the activated representation used in identification tasks, which is responsible for visual priming, is essentially independent of these low-level attributes of the input image. Consequently, the intermediate representation that gets primed must encode object information in a more abstract form, probably in terms of invariant descriptions of parts and spatial relations of parts as was suggested by Biederman (1987). Note that this conclusion does not exclude the possibility that other non-invariant representations exist in the brain and are used for different recognition or classification purposes, not to mention motor interactions. It does suggest that for

the task of entry-level recognition of objects with well-articulated shape the relevant representation is an intermediate one. A similar conclusion was reached by Fiser, Biederman, and Cooper (1996) based on the failures of a local spatial filter-based object recognition system to account for human experimental data on entry-level object recognition. The present study extends the implications of the Biederman and Cooper (1991b) study with complementary contour-deleted line drawings — that visual priming of objects is invariant to low-level features — to include the fundamental dimension of SF of gray-level images.

Supporting the above conclusion is single-unit evidence that it is in the anterior portions of the ventral pathway, viz., area TE, where cell tuning reflects invariance of object representations to variation of early 'appearance' features. Sáry, Vogels, and Orban (1993) reported that 28% of TE cells responding to a particular complex shape, retained their selectivity (preference over that shape to others) independent of whether the shapes were defined by differences in luminance, motion, or texture. Similarly, Kovács, Vogels, and Orban (1995) showed that TE cells retained their selectivity to (a) solid and line drawing versions of various shapes, and (b) partially occluded shapes that humans and monkeys could still identify as a version of an originally learned intact shape.

The second implication of the present study derives from the extent of the lateral shift employed in Experiment 1. Ten degrees of change in center position means at least 4° of translation for any feature of an object with an extent of 6°. This is not only larger than the shift that Biederman and Cooper (1991a) used in their studies, but it is also larger than the reported extent of any receptive field in primate visual areas V1 and V2. It is also larger than any suggested overlap between the hemifield representations in those two areas. Assuming receptive field sizes in the human visual cortex are comparable to those in monkeys, this finding excludes any parsimonious argument of visual priming occurring in V1 and V2. In theory, it is possible that information which is known to cross between hemispheres via the corpus callosum only starting with area V4 (Clarke & Miklóssy, 1990) is propagated back to earlier visual areas to induce priming there. In practice, this scenario is not plausible. In our experiment, such an explanation would require that information propagated back would affect cells which cover visual areas very distant from the original source of information, and which prefer ranges of SF and orientation very different from those preferred by the priming cell. It is hard to see how such a scheme would act selectively enough to distinguish between different exemplars with the same name.

The situation is less conclusive in areas V4 and TEO. In these areas there are cells whose RF size exceeds the lateral shift employed in our study. In theory, such cells could receive input during both left and right presentations of our images, and thus be responsible for the visual priming effect. However, due to the strong correlation between RF size and eccentricity of cells in those areas, neurons with large RFs tend to be further away from the fovea resulting in almost as little representation of the opposite hemifield as in the case of V1 and V2. The regression lines between RF size and eccentricity provided by Boussaoud et al. (1991) suggest that the average overlap across the midline in the foveal and parafoveal regions is around 1° in V4 and around 2° in TEO. This is in sharp contrast with cells in TE whose very large ($> 20°$) receptive fields almost always include the center of gaze and extend substantially to the opposite hemifield (Gross, Rocha-Miranda, & Bender, 1972). In addition, although there are callosal connections connecting the two hemispheres in V4 and TEO, the functional role of these connections may be largely suppressive to facilitate global color constancy figure-ground segregation in V4, rather than specific shape encoding (Desimone & Ungerleider, 1989). Based on this evidence, we suggest that the visual priming effect we found is mediated by cells that reside beyond an area homologous to area TEO in monkeys.

The suggestion that visual priming occurs beyond TEO is supported by recent PET and fMRI findings that priming affects cell activation in the anterior regions of the temporal cortex (Ungerleider, 1995; Martin, Wiggs, Ungerleider, & Haxby, 1996; Buckner et al., 1998). A picture that emerges from these results and the present findings is that the representation mediating visual priming in object recognition involves visual areas anterior to the human homologues of V4 and TEO but posterior to anterior TE.

For additional insight, it is instructive to compare the present results to those of Bar and Biederman (1998), who also studied visual priming of translated images. They employed a methodology that closely approximated our own with one important difference. The presentation time in their study averaged 47 ms compared with our presentation time of 185 ms. Even though they used line drawings rather than gray-level images, their exposure time was typically too short to identify images in the first block, given their extremely effective masks which were custom designed for each stimulus. If the subject could not identify the image, she/he had to pick one of four possible names after each trial. One of the names was the true object just presented, one was a different exemplar, one was a different object with a similar shape, and the last was an unrelated object with a different shape. Subjects performed poorly in the first block, they could name

only 13.5% of the images, and their forced-choice accuracy was at chance when their initial naming attempt was an error. However, the second block results showed significant subliminal priming (21% increase in correct answers) indicating that information during the presentations in the first block did activate some representations in the visual system, and this information helped improve performance in the second block (this increase was independent of whether subjects received a forced-choice test). Performance with same name, different shaped exemplars on the second block was equivalent to that of new objects. Most important, Bar and Biederman found that this subliminal priming effect was position specific: images shown at the same position in the second block enjoyed significantly stronger priming than images that were translated. In a follow-up study, Bar and Biederman (1999) showed that if the same amount of translation took the image across the vertical midline, subliminal priming was eliminated altogether.

A comparision of these results to ours suggests that if the image in the first block is identified successfully, priming becomes translation and hemisphere invariant. This might indicate that the dominant representations that are responsible for priming in the two cases (although both visual) are different and might reside at different levels in the visual system. Based on the effect of translation on priming, Bar and Biederman (1998, 1999) proposed that their subliminal priming affected intermediate visual areas homologous to TEO. The invariance to translation that we observed suggests that the lows of the representation mediating supraliminal priming lies downstream from area TEO, most likely in an area (or areas) homolgous to TE.

Our result of equal priming across SF bands is in agreement with Schyns and Oliva (1997) findings. Since the focus of their study was on the dynamical selection of SF scales during recognition, Schyns and Oliva used a rapid visual presentation paradigm, and thus their results could be regarded as short term (or 'hot trail') priming. They found that when recognizing SF filtered complex images, subjects select a scale that is best for recognition and their performance is better if the next image is presented at the selected scale. However, images presented in the non-attended scale could facilitate recognition of subsequent target images just as well as images presented at the attended scale (their Experiments 3 and 4). They proposed that this facilitation could not be based on semantics, because the subjects were not even aware of the object presented at the non-attended scale.

Finally, our results do not support the idea suggested by several researchers (Sergent, 1982; Jonsson & Hellige, 1986; Kitterle et al., 1993; Hughes, Nozawa, & Kitterle, 1996), that the two hemispheres differ in their

preferred SF ranges. According to this notion, the left hemisphere is more specialized in local processes and has an advantage in processing high SF information, whereas the right hemisphere is more responsible for global information and is at an advantage with low SF bandpass images. In our object recognition tasks, subjects performed equally well on either side of the fixation point with low and high bandpass images. Although there is ample evidence for hemispheric specialization in local-global processing (Bradshaw & Nettleton, 1981; Robertson & Lamb, 1991), we conjecture that these hemispheric differences might be due to attentional mechanisms, and they may not be related to processing low and high spatial frequencies per se (c.f. Hübner, 1997). Whatever the status of the SF bias across the hemispheres, it does not appear to be on a critical path for object recognition.

In conclusion, visual priming of gray level images of objects is insensitive to large variations in the fundamental dimensions specified by the early spatial filtering of the visual system — position in the visual field, local orientation, and SF — but is highly sensitive to shape features of the presented object. The shape specific intermediate representations likely reside beyond an area homologous to TEO in the ventral pathway.

References

Bar, M., & Biederman, I. (1998). Subliminal visual priming. *Psychological Science*, 9(6), 464–469.

Bar, M., & Biederman, I. (1999). Localizing the cortical region mediating visual awareness of object identity. *Proceedings of the National Academy of Sciences*, 96, 1790–1793.

Bartram, D. J. (1974). The role of visual and semantic codes in object naming. *Cognitive Psychology*, 6, 325–356.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review*, 94(2), 115–147.

Biederman, I., & Cooper, E. E. (1991a). Evidence for complete translational and reflectional invariance in visual object priming. *Perception*, 20, 585–593.

Biederman, I., & Cooper, E. E. (1991b). Priming contour-deleted images: evidence for intermediate representations in visual object recognition. *Cognitive Psychology*, 23, 393–419.

Biederman, I., & Cooper, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology-Human Perception and Performance*, 18(1), 121–133.

Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology-Human Perception and Performance*, 19(6), 1162–1182.

Biederman, I., & Bar, M. (1999). One-shot invariance in matching novel objects. *Vision Research*, 39, 2885–2899.

Boussaoud, D., Desimone, R., & Ungerleider, L. G. (1991). Visual topography of area TEO in the macaque. *Journal of Comparative Neurology*, 306(4), 554–575.

Bradshaw, J. L., & Nettleton, N. C. (1981). The nature of hemispheric specialization in man. *Behavioral and Brain Sciences*, 4, 51–63.

Buckner, R., Goodman, J., Burock, M., Rotte, M., Koutstaal, W., Schacter, D. L., Rosen, B. R., & Dale, A. (1998). Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. *Neuron*, 20, 285–296.

Cave, C. B. (1997). Very long-lasting priming in picture naming. *Psychological Science*, 8, 322–325.

Clarke, S., & Miklóssy, J. (1990). Occipital cortex in man: organization of callosal connections, related myleo- and cytoarchitecture, and putative boundaries of functional visual areas. *The Journal of Comparative Neurology*, 298, 188–214.

Cooper, E. E., Biederman, I., & Hummel, J. E. (1992). Metric invariance in object recognition: a review and further evidence. *Canadian Journal of Psychology*, 46, 191–214.

Desimone, R., & Ungerleider, L. G. (1989). Neural mechanisms of visual processing in monkeys. In F. Boller, & J. Grafman, *Handbook of neurophysiology* (pp. 267–299). Amsterdam: Elsevier.

DeValois, R. L., & DeValois, K. K. (1988). *Spatial vision*. New York, NY: Oxford Press.

Dow, B. M., Snyder, A. Z., Vautin, R. G., & Bauer, R. (1981). Magnification factor and receptive field size in foveal striate cortex of the monkey. *Experimental Brain Research*, 44, 213–228.

Ellis, R., & Allport, D. A. (1986). Multiple levels of representation for visual objects: a behavioural study. In A. G. Cohn, & J. R. Thomas, *Artificial intelligence and its applications* (pp. 245–257). New York: Wiley.

Field, D. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, *Optics and Image Science*, 4, 2379–2394.

Fiser, J., & Biederman, I. (1995). Size invariance in visual object priming of gray scale images. *Perception*, 24(7), 741–748.

Fiser, J., Biederman, I., & Cooper, E. E. (1996). To what extent can matching algorithms based on direct outputs of low level generic descriptors account for human object recognition. *Spatial Vision*, 10(3), 237–271.

Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science*, 273, 1699–1702.

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology*, 35, 96–111.

Halgren, E., Dale, A. M., Sereno, M. I., Tootell, R. B. H., Marinkovic, K., & Rosen, B. R. (1999). Location of human face-selective cortex with respect to retino-topic areas. *Human Brain Mapping*, 7, 29–37.

Hübner, R. (1997). The effect of spatial frequency on global precedence and hemispheric differences. *Perception and Psychophysics*, 59(2), 187–201.

Hughes, H. C., Nozawa, G., & Kitterle, F. (1996). Global precedence, spatial-frequency channels, and the statistics of natural images. *Journal of Cognitive Neuroscience*, 8(3), 197–230.

Intraub, H. (1981). Identification and naming of briefly glimpsed visual scenes. In D. F. Fisher, R. A. Monty, & J. W. Senders,

Eyemovements: cognition and visual perception (pp. 181–190). Hillsdale, NJ: Erlbaum.

Jonsson, J. E., & Hellige, J. B. (1986). Lateralized effect of blurring: a test of the visual spatial frequency model of cerebral asymmetry. *Neuropsychologia*, *24*, 351–362.

Jüttner, M., & Rentschler, I. (2000). Scale-invariant superiority of foveal vision in perceptual categorization. *European Journal of Neuroscience*, *12*, 353–359.

Kitterle, F. L., Christman, S., & Conesa, J. (1993). Hemispheric differences in the interference among components of compound gratings. *Perception and Psychophysics*, *54*, 785–793.

Kovács, G., Vogels, R., & Orban, G. A. (1995). Selectivity of macaque inferior temporal neurons for partially occluded shapes. *Journal of Neuroscience*, *15*(3), 1984–1997.

Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures, and concepts: a comparison of lexical, object, and reality decisions. *Journal of Verbal Learning and Verbal Behavior*, *23*, 39–66.

Magnussen, S., & Greenlee, M. W. (1999). The psychophysics of perceptual memory. *Psychological Research — Psychologische Forschung*, *62*, 81–92.

Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, *379*, 649–652.

Maunsell, J. H. R. (1995). The brain's visual world: representation of visual targets in cerebral cortex. *Science*, *270*, 764–768.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*, *6*, 414–417.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, *343*, 263–266.

Robertson, L. C., & Lamb, M. R. (1991). Neurophysiological contributions to theories of part/whole organization. *Cognitive Psychology*, *23*, 299–330.

Roe, A. W., & Ts'o, D. Y. (1995). Visual topography in primate V2: multiple representations across functional stripes. *Journal of Neuroscience*, *15*(5), 3689–3715.

Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, *37*, 495–510.

Sáry, G., Vogels, R., & Orban, G. A. (1993). Cue invariant shape selectivity of macaque inferior temporal neurons. *Science*, *260*, 995–997.

Schacter, D. L., & Tulving, E. (1994). *Memory systems*. Cambridge, MA: MIT Press.

Schyns, P. G., & Oliva, A. (1997). Flexible, diagnosticity-driven, rather than fixed, perceptually determined scale selection in scene and face recognition. *Perception*, *26*, 1027–1038.

Sergent, J. (1982). The cerebral balance of power: confrontation or cooperation. *Journal of Experimental Psychology Human Perception and Performance*, *8*, 253–272.

Ungerleider, L. G. (1995). Functional brain imaging studies of cortical mechanisms for memory. *Science*, *270*, 769–775.

Van Essen, D. C. (1985). Functional organization of the primate visual cortex. In A. Peters, & E. G. Jones, *Cerebral cortex* (pp. 259–329). New York: Plenum Press.

Wässle, H., Gruenert, U., Rohrenbeck, J., & Boycott, B. (1991). Retinal ganglion cell density and cortical magnification factor in the primate retina. *Vision Research*, *30*, 1897–1911.

Wilson, H. R., Levi, D., Maffei, L., Rovamo, J., & DeValois, R. (1990). The perception of form: retina to striate cortex. In L. Spillmann, & J. S. Werner, *Visual perception: the neurophysiological foundations* (pp. 231–272). San Diego, CA: Academic Press.