

Effects of spatial frequency bands on perceptual decision: It is not the stimuli but the comparison

Pia Rotshtein

School of Psychology, University of Birmingham,
Birmingham, UK



Andrew Schofield

School of Psychology, University of Birmingham,
Birmingham, UK



María J. Funes

School of Psychology, University of Birmingham,
Birmingham, UK, &
Facultad de Psicología, Universidad de Granada,
Granada, Spain



Glyn W. Humphreys

School of Psychology, University of Birmingham,
Birmingham, UK



Observers performed three between- and two within-category perceptual decisions with hybrid stimuli comprising low and high spatial frequency (SF) images. We manipulated (a) attention to, and (b) congruency of information in the two SF bands. Processing difficulty of the different SF bands varied across different categorization tasks: house–flower, face–house, and valence decisions were easier when based on high SF bands, while flower–face and gender categorizations were easier when based on low SF bands. Larger interference also arose from response relevant distracters that were presented in the “preferred” SF range of the task. Low SF effects were facilitated by short exposure durations. The results demonstrate that decisions are affected by an interaction of task and SF range and that the information from the non-attended SF range interfered at the decision level. A further analysis revealed that overall differences in the statistics of image features, in particular differences of orientation information between two categories, were associated with decision difficulty. We concluded that the advantage of using information from one SF range over another depends on the specific task requirements that built on the differences of the statistical properties between the compared categories.

Keywords: face perception, gender, ordinate categorization, subordinate categorization, hybrid stimuli, attention and congruency manipulations

Citation: Rotshtein, P., Schofield, A., Funes, M. J., & Humphreys, G. W. (2010). Effects of spatial frequency bands on perceptual decision: It is not the stimuli but the comparison. *Journal of Vision*, 10(10):25, 1–20, <http://www.journalofvision.org/content/10/10/25>, doi:10.1167/10.10.25.

Introduction

Looking for a friend in a crowded train station and in a forest is likely to involve different strategies. Efficient strategies for object recognition and search need to depend on the visual properties of the background environment, the target stimulus, and the potential distracters (other people as opposed to trees, respectively). A recent hypothesis proposes that the diagnostic value of various visual features for a given task will determine their usage in that task (Morrison & Schyns, 2001; Oliva & Schyns, 1997, 2000; Ruiz-Soler & Beltran, 2006; Schyns, 1998; Schyns & Gosselin, 2003; Schyns & Oliva, 1997, 1999). To test the diagnostic hypothesis, we revisited the differential use of high and low spatial frequency (SF) components in different visual categorization tasks. We systematically investigated various factors that may lead

to an advantage of using one SF range over another. Our specific interest was to test the degree of flexibility in the visual system and the potential factors that affect this flexibility, as implemented through the processing of different SF components.

Early processing stages in the visual system are known to dissociate in terms of the SF range of the information they preferentially extract. In particular, magno-cellular and parvo-cellular visual pathways have different SF preferences, with the former reaching the cortex faster and being more sensitive to low SF components, while parvo-cellular pathways are more sensitive to high SF bands (e.g., Bullier, 2001; Lamme, 2001; Livingstone & Hubel, 1988). Neurophysiology studies show that these pathways project to distinct cortical regions (Shipp, 2001; Shipp & Zeki, 1995) and neuroimaging studies confirm that high and low SF components of face stimuli are processed in dissociated brain regions in posterior occipital

cortices (Eger, Schyns, & Kleinschmidt, 2004; Rotshtein, Vuilleumier, Winston, Driver, & Dolan, 2007; Vuilleumier, Armony, Driver, & Dolan, 2003)—a separation that is also observed in associative visual regions (Gauthier, Curby, Skudlarski, & Epstein, 2005; Peyrin et al., 2005; Rotshtein et al., 2007; though see Eger et al., 2004).

The neurophysiological evidence for dissociable SF processing routes has triggered considerable research into the roles of high and low SF components in visual recognition. Traditionally, it has been argued that there is a fixed order of coarse-to-fine integration, where low SFs (due to their faster arrival at the cortex) generate an initial coarse representation of an image that is used to guide the processing of the more detailed information, conveyed by high SF components (Bar, 2003; Blackmore & Campbell, 1969; Bullier, Hupe, James, & Girard, 2001; Lamme, 2001; Marr, 1982; McCarthy, Puce, Belger, & Allison, 1999; Oliva & Schyns, 1997; Parker & Costen, 1999; Schyns & Oliva, 1994). However, there is also increasing experimental evidence suggesting that the integration and usage of different ranges of SFs is flexible (Collin, 2006; Goffaux, Jemel, Jacques, Rossion, & Schyns, 2003; Morrison & Schyns, 2001; Oliva & Schyns, 1997; Peyrin et al., 2005; Ruiz-Soler & Beltran, 2006; Schyns & Oliva, 1994, 1997, 1999). In an elegant set of studies involving hybrid stimuli, in which contrasting information was conveyed by low and high SF channels, Oliva and Schyns (1997; Schyns & Oliva, 1994, 1999) demonstrate that information from low or high SFs can be equally influential in scene and face categorization. Additional findings suggest that the usage of information from different SF channels can be altered by previous experience. Previous experience was manipulated by altering the perceptual set (i.e. the tendency to use one SF over another; Schyns & Oliva, 1999) and by sensitization procedure to different SFs (i.e. pre-exposure to limited SF band; Oliva & Schyns, 1997; Ozgen, Payne, Sowden, & Schyns, 2006; Schyns & Oliva, 1999). Computational modeling also shows that algorithms that use SF components in a flexible manner can outperform fixed coarse-to-fine algorithms in recognizing objects (Mermillod, Guyader, & Chauvin, 2005).

This has led to the development of the diagnostic hypothesis, which proposes that observers use the most diagnostic SF band for any given task (Morrison & Schyns, 2001; Oliva & Schyns, 1997; Ruiz-Soler & Beltran, 2006; Schyns, 1998; Schyns & Gosselin, 2003; Schyns & Oliva, 1997, 1999). For example, identifying faces is assumed to rely on middle range frequencies (Näsänen, 1999; Ojanpää & Näsänen, 2003) with larger contribution from low than high frequency ranges (Harel & Bentin, 2009; Schyns & Oliva, 1999). Judging if a face has an expression or not is based on low SF information while judging if a face expresses a happy or an angry emotion relies on components in the high SF range, though this latter bias can be reversed by recent previous experience (Schyns & Oliva, 1999).

An open question is what makes an SF band diagnostic for a particular task, if at all. The “task-level-based” hypothesis suggests that low SF information is used for super-ordinate and ordinate categorizations and high SF information is used for subordinate or within-category discriminations (Collin, 2006; Collin & McMullen, 2005; Morrison & Schyns, 2001; Schyns, 1998). Though others argue that some categorization at the super-ordinate level rely on information from the middle and high SF ranges, as in the case of man-made vs. natural scene categorization (Oliva & Torralba, 2001). Alternatively, the “stimulus-based” hypothesis proposes that the SFs used in perceptual tasks depend on the stimulus. For example, tasks that involve discrimination between faces will rely on low SF components whereas discriminations between objects will rely on high SFs (Goffaux, Gauthier, & Rossion, 2003; Harel & Bentin, 2009). A stimulus-noise diagnostic hypothesis argues that the dependency of particular stimuli on particular SFs relates to how much each SF range in the stimulus diverges from the distribution of the same SFs in noise (Sowden & Schyns, 2006). Finally, the fully flexible usage hypothesis implies that the use of a specific range of SFs is not by default biased toward the usage of one SF range over another in any given task (Oliva & Schyns, 1997; Schyns & Oliva, 1994). Consistent with this later hypothesis, it has been shown that observers can be biased, by pre-exposure (Ozgen et al., 2006; Ozgen, Sowden, Schyns, & Daoutis, 2005; Sowden, Ozgen, Schyns, & Daoutis, 2003) or by perceptual set manipulations (Schyns & Oliva, 1999) to use one SF over another in variety of visual categorization tasks. Despite the wealth of research into the use of SF ranges in visual recognition it is unclear how much this flexibility can be controlled by the observer and how much it is affected by the information in the stimuli.

The aims of the studies reported here was to address the following questions: (i) Can observers flexibly report information from one SF range and ignore information in another SF range, based on task instructions? (ii) Is performing visual categorization decisions based on information from high and low SF ranges equally easy? (iii) Would information from the non-attended SF range interfere with categorization, and if so, at what level would the interference arise? Finally, (iv) what determines any task-related SF bias in different categorization tasks—to what extent is this built on low-level perceptual differences between stimuli?

To address the above questions, we presented participants with hybrid stimuli constructed from overlaid low and high SF filtered images. We used hybrid images as our stimuli to be able to test interference from one SF range upon another and also to ensure that both SF ranges were presented conjointly—avoiding confounding non-specific visual cues with the SF manipulation (Morrison & Schyns, 2001; Oliva & Schyns, 1997; Schyns & Oliva, 1994, 1997, 1999). We manipulated four factors: (1) the type of perceptual categorization required. We used five

categorization tasks: three at the ordinate level between-category discriminations: *face–flower*, *house–flower*, and *face–house* and two at the subordinate level, within-category discriminations of faces: positive–negative *valence* of facial expression and female–male *gender* of faces. (2) The SF range that should be categorized. Observers categorized the high or low SF components within the hybrid stimuli, and hence were instructed to attend and respond only to one range while ignoring the other. (3) The relations between the high and low SF information in the hybrid stimuli. The two SF ranges could convey congruent (Cong), task-relevant incongruent (TR-incong), task-irrelevant incongruent (TIR-incong), or neutral (baseline) information. For the baseline hybrids, high SF filtered images were overlaid with low SF noise and vice versa. Finally, we also manipulated (4) the exposure duration of the hybrid stimuli. We used two presentation durations: 30 and 200 ms. The stimulus exposures were determined from previous research demonstrating that, with this short duration, categorization of low SF stimuli is better than that of high SF stimuli, while this pattern is reverse for longer durations, i.e., 150 ms (Schyns & Oliva, 1994).

It is important to note that, in contrast to previous studies that used hybrid stimuli to investigate the flexible use of high and low SF components (e.g., Morrison & Schyns, 2001; Oliva & Schyns, 1997; Schyns & Oliva, 1994, 1997, 1999), here observers were made aware of the hybrid nature of the stimulus. The processing of information from high and low SF ranges was manipulated using explicit instructions. Biases in processing high vs. low SFs were measured by the ability (accuracy and response time) to make categorization decisions based on one or the other SF ranges while ignoring the information from the other range. This enabled us to compare the processing of high and low SF components as a function of the various tasks. The design also enabled us to test the effect of the non-attended SF. Here the non-attended range was defined according to the experimental manipulation, as the SF range that observers were instructed to ignore. Finally, we used only perceptual categorization tasks, while varying the categories that had to be discriminated. This was done to insure that any SF biases that arise are due to differences between the compared categories and not due to variability in the cognitive requirements of the task (e.g., categorization vs. detection).

The various hypotheses reviewed above lead to different predictions. The task-level-based flexible usage hypothesis predicts that an advantage for low SF targets would be observed for ordinate categorization, while subordinate categorization would be more efficient when stimuli are conveyed by high SFs. The stimulus-based flexible hypothesis predicts that the processing of faces would be more efficient for low than high SF targets and this would be independent of the compared category. The stimulus-diagnostic hypothesis (Sowden & Schyns, 2006) predicts that, for any given stimulus category, the same SF range

would always be more diagnostic, as diagnosticity is determined relative to the SF components of the noise. Finally, the fully flexible hypothesis predicts that visual categorization of low and high SF ranges will be equally good and there would be no interference from the non-attended SF range.

Study 1: Between-category discrimination

In Study 1, we used identical sets of stimuli: faces, houses, and flowers, in two different ordinate level categorization tasks. Observers were asked to decide whether they perceived: (i) *houses or flowers*; or (ii) *faces or flowers*. These two tasks were performed separately on the high and low SF components of the hybrid images (e.g., “does the high SF depict a house or a flower?”). Four types of hybrid were used: (i) congruent, (ii) task-relevant incongruent, (iii) task-irrelevant incongruent, and (iv) baseline hybrids. In the *house–flower* task, the task-irrelevant incongruent stimuli were faces and in the *face–flower* task they were houses. Note that, by using the same set of stimuli in both tasks, though with a different relation to the specific task instructions, we were able to test whether a processing advantage for particular SF components is determined by the hybrid stimuli themselves or by the task requirements.

Methods

Participants

Fifteen volunteers participated in this study (8 females, mean age: 28 years \pm 7 std). All had normal or corrected-to-normal vision. The study was approved by the local ethics committee. Participants were paid £5 for their time.

Stimuli

The stimulus set constituted 64 front view faces (32 females) with neutral expressions (from the Karolinska Directed Emotional Faces set: Lundqvist & Litton, 1998), 57 close-up of buildings and houses, and 66 flowers and plants pictures. Faces were cropped to extreme close-up and contained mostly “inner” facial features with minimal hair features. All photos were achromatic and resized to 256 \times 256 pixels using Photoshop 8.0. The “intensities” of the pixels making up each stimulus were normalized (using MatLab7.0) by subtracting the mean pixel value and dividing by the standard deviation of the pixel intensities. Image intensities were then rescaled to give a mean gray level of 128 and a range of 88–168. The stimuli were transformed into the frequency domain using a fast Fourier Transform algorithm and were then filtered using

Butterworth filters (Rotshtein et al., 2007; Winston, Vuilleumier, & Dolan, 2003). The filters were set to filter either high frequencies (SF > 24 cycle/image; viewed as SF > 3.4 cycle/degree) or low frequencies (SF < 8 cycle/image; viewed as SF < 1.14 cycle/degree). Note that with a Butterworth filter, the cutoff frequency corresponds to 50% of the magnitude of the filter. Therefore, to minimize overlap between frequency channels, the distance between the cutoff frequencies (i.e., bandwidth) was 1.5 octaves. These cutoffs were chosen to fit previous psychophysical findings suggesting that magno-cellular visual pathways are preferentially sensitive to SF below 1.5 cycles per degree, while parvo-cellular pathways are sensitive to SFs above this value (Skottun, 2000). In addition, for all the filtered stimuli, phase-scrambled images were generated, resulting in low and high SF noise images that were used for the baseline hybrid noise stimuli (see below).

Four types of hybrid were generated for each of the tasks: (1) congruent (low and high SF images were of the same category); (2) task-relevant incongruent (TR-incong, the high and low SFs were of opposing response relevant categories); (3) task-irrelevant incongruent (TIR-incong, overlaying the high SF image with a low SF image from a task-irrelevant category, or vice versa); and (4) baseline hybrid noise (BL-noise), overlaying the high SF image with a low SF phase-scrambled noise image from the opposing category, or vice versa. In order to preserve the

natural power spectrum distribution of the different SF ranges, the hybrids were generated by adding the two filtered images and no additional image manipulation was applied.

In the *house–flower task*, congruent hybrids contained low and high SF images of two different houses or two different flowers; in the TR-incong hybrids, the low and high SF stimuli depicted a house and a flower or vice versa; in the TIR-incong hybrids, the attended SF range depicted a house or a flower while the non-attended range depicted a face; finally, in the baseline hybrid, a low/high SF house was overlaid with a phase-scrambled high/low SF flower, respectively, and vice versa (Figure 1A).

In the *face–flower task*, congruent hybrids contained low and high SF images of two different faces of the same gender or two different flowers; the TR-incong hybrids had low and high SF images depicting faces and flowers and vice versa; the TIR-incong hybrids depicted face or flower in the attended SF range and a house in the non-attended range; finally, in the baseline hybrid, a low/high SF face was overlaid with a phase-scrambled high/low SF flower, respectively, and vice versa (Figure 1B).

Note that similar hybrids were used in the two different tasks: flower and face congruent hybrids appeared in both tasks, while face–flower, face–house, and house–flower hybrids appeared in both tasks either as the TR- or TIR-incong hybrids.

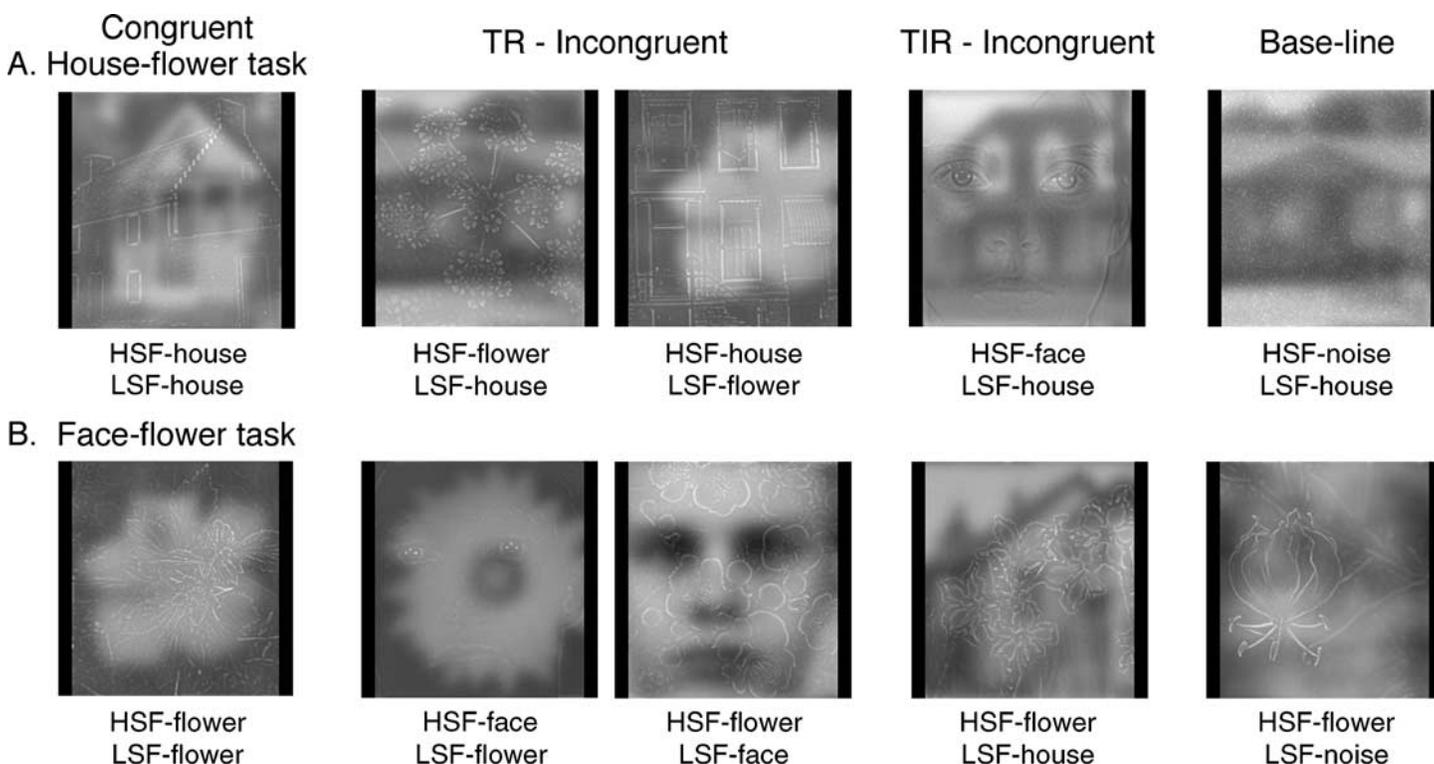


Figure 1. Study 1—Stimuli. (A) Examples of stimuli used in the *house–flower* task. (B) Example of stimuli used in the *face–flower* task. First column, examples of congruent hybrids; second and third columns, examples of task-relevant (TR) incongruent hybrids; fourth column, example of task-irrelevant (TIR) incongruent hybrids; and fifth column, examples of baseline hybrids; HSF, LSF: high and low spatial frequencies, respectively.

All the hybrids were generated by randomly pairing the stimuli. There were 5 different sets of pairings for each stimulus. The wide variety of stimuli used ensured that participants could not develop a stimulus-specific strategy to perform either of the tasks. To facilitate selective attention within the hybrids, especially critical for the face hybrids (see below) one of the filtered images was shifted randomly to the left or the right by 15 pixels (~1.8 degrees; Rotshtein et al., 2007). Black bars were overlaid on either side of the hybrids to ensure symmetry in amount of information on the left and right of the images (Figure 1).

Design and procedure

A within-subject fully factorial design was used with the following factors: task (*face–flower*, *flower–house*), attention (low, high SF), exposure duration (200 ms, 30 ms), and hybrid types (congruent, TR-incong, TIR-incong, baseline).

Thirty-six hybrids were presented in each of the conditions, half for each response category. The three first factors (task, attention, exposure duration) were manipulated across different blocks while hybrid types were presented in random order within each block. The block design had hierarchical structure and the order of blocks at each level was randomized. The block design was used to minimize interference due to task switching. Each task was run separately and was divided into two attention blocks. In each, participants were cued to attend either to the low or the high SF images, described, respectively, as the image that looks blurred or the image that looks like a line drawing. The task was to categorize the attended SF image while ignoring the other SF image in the stimulus. Note that, within a block, half the trials (i.e., the TIR-incong and baseline hybrids) conveyed task-relevant information only in the attended SF range; this adds a sensitization manipulation that is assumed to facilitate processing of the attended SF range (Ozgen et al., 2006, 2005) and potentially creates a biased perceptual set toward the attended SF in each block (Schyns & Oliva, 1999).

Each attention block started with 10 practice trials (depicting randomly the four hybrid types). The practice trials were presented until response and feedback was provided for each response. This was done to ensure that participants understood the task and that adequately attend to the SF that was relevant for the response. The hybrid nature of the stimuli was explicitly explained to the participants. Thus, any interference from the non-attended SF range would indicate a failure to suppress bottom-up effects from irrelevant information.

Each SF attention block was further divided to the two exposure duration blocks (200 ms, 30 ms). The first 10 hybrids in each duration block were excluded from the analysis to minimize task switching effects. At the end of each duration block, participants received feedback for

their overall accuracy and reaction times (RTs) and had a short break. Feedback was given to keep the participants motivated and alert.

A trial started with the presentation of a circle at fixation for 500 ms, then a hybrid stimulus appeared for 200 ms (or 30 ms), followed by a textual cue at fixation to remind participants of the attended SF range (“blurred”, “line drawing”). A chin rest was used to ensure that participants were at a constant distance from the screen (65 cm) across all trials and conditions. The viewing angle was 3.5°. Responses were carried out using the right index and middle fingers and the “1” and “2” buttons of the keyboard, randomly assigned for each category across participants. Stimulus presentation and data collection were realized using the Matlab-based toolbox: Cogent 1.25 and Cogent Graphic 1.24 (Wellcome Department of Imaging Neuroscience, UCL).

Data analysis was performed using Matlab7.0 and SPSS15.0. Averages of correct responses and of their median response times (RTs) per condition are reported in Table 1. Based on the LATER model (Reddi, Asrress, & Carpenter, 2003), it is assumed that variability in accuracy and RT responses arise from the same decision mechanism. Therefore, to avoid “trade-off” effects and to make the reported results more concise, accuracy and RT measures were combined to obtain the psychological-efficiency measure on a per condition basis: the median RTs were divided by the proportion of accurate response

	Congruent	TR-incong	TIR-incong	Baseline
<i>Task: Face–flowers 200 ms</i>				
Attend	572 (23)	612 (35)	624 (33)	563 (22)
LSF	<i>0.95 (0.01)</i>	<i>0.83 (0.04)</i>	<i>0.88 (0.03)</i>	<i>0.97 (0.01)</i>
Attend	584 (26)	641 (28)	578 (26)	579 (30)
HSF	<i>0.94 (0.01)</i>	<i>0.87 (0.03)</i>	<i>0.92 (0.02)</i>	<i>0.95 (0.01)</i>
<i>Task: Face–flowers 30 ms</i>				
Attend	508 (16)	514 (16)	548 (21)	516 (13)
LSF	<i>0.96 (0.01)</i>	<i>0.91 (0.02)</i>	<i>0.92 (0.02)</i>	<i>0.93 (0.02)</i>
Attend	672 (28)	806 (53)	720 (35)	650 (28)
HSF	<i>0.78 (0.03)</i>	<i>0.46 (0.06)</i>	<i>0.81 (0.03)</i>	<i>0.64 (0.03)</i>
<i>Task: Flowers–house 200 ms</i>				
Attend	726 (51)	800 (46)	708 (35)	668 (30)
LSF	<i>0.84 (0.02)</i>	<i>0.60 (0.05)</i>	<i>0.88 (0.02)</i>	<i>0.83 (0.02)</i>
Attend	546 (21)	564 (23)	601 (26)	553 (23)
HSF	<i>0.95 (0.01)</i>	<i>0.93 (0.02)</i>	<i>0.93 (0.02)</i>	<i>0.94 (0.01)</i>
<i>Task: Flowers–house 30 ms</i>				
Attend	696 (45)	771 (62)	690 (34)	672 (40)
LSF	<i>0.81 (0.02)</i>	<i>0.47 (0.04)</i>	<i>0.71 (0.03)</i>	<i>0.69 (0.04)</i>
Attend	572 (16)	576 (14)	618 (17)	562 (14)
HSF	<i>0.94 (0.02)</i>	<i>0.81 (0.03)</i>	<i>0.84 (0.02)</i>	<i>0.89 (0.02)</i>

Table 1. Response times (ms) of correct responses (in bold) and proportion of accurate responses (in italics): mean (SEM).

for each condition and each participant (Townsend & Ashby, 1983). Statistical analyses were performed on these combined scores, though similar patterns of results were obtained for the separate accuracy and RT measures (see Table 1). The analysis focused on the effect of the task on SF processing. We measured the psychological efficiency of responding to the attended SF and the amount of interference and benefit from the non-attended SF image. Interference was measured in the incongruent hybrid conditions (TR- and TIR-incong) by subtracting the corresponding responses to baseline hybrids (i.e., the same task, attention, and duration conditions). To facilitate comparisons across conditions, the measures were divided by the sum of the responses (incong – BL)/(incong + BL). Any benefit for congruent hybrids was measured by subtracting performance with these stimuli from that with baseline hybrids; again this was divided by the sum of the responses (BL – cong)/(BL + cong). The reliability of all the effects was determined using a repeated measures ANOVA; all reported results were Greenhouse–Geisser corrected to account for non-sphericity in the data and Bonferroni correction was applied to the simple effect tests.

Results and discussion

Accuracy in the *face–flower* task when attending to the low SF images was on average 91% (with 95% confidence interval (CI): 87.3–96.6%) and 79.6% when attending to the high SF stimuli (CI: 73–86%). In the *flower–house* task, accuracy for low SF images was 72% (CI: 64–81%) and 90% for high SF images (CI: 88–92%). Overall performance was well above chance, showing that observers were able to categorize *face–flower* and *flower–house* stimuli from either low or high SF components. Table 1 presents accuracy and RT responses for each condition. The follow-up analyses were carried out on the psychological-efficiency scores.

Overall performances

A repeated measure four-way ANOVA was computed on the psychological-efficiency measurements with the following factors: Task (*face–flower*, *flower–house*), attention (high, low SF), exposure duration (200 ms, 30 ms), and hybrid type (congruent, TR-incong, TIR-incong, baseline). Overall, participants were better in categorizing *face–flower* stimuli than *flower–house* stimuli ($F(1, 14) = 8.745$, $\eta_p^2 = 0.384$, $p < 0.05$). More importantly, we observed a reliable interaction between task and the attended SF ($F(1, 14) = 75.5$, $\eta_p^2 = 0.84$, $p < 0.001$). This interaction was observed even when the TR-incong hybrids were excluded from the analysis ($F(1, 14) = 78$, $\eta_p^2 = 0.84$, $p < 0.001$), confirming that differences in

responses were not driven by the extent of interference from the unattended SF in the TR-incong hybrids (see below). To unravel the sources of the interaction, we analyzed performances in each task separately.

In the *flower–house* task, performance was better when attention was directed to the high SF images than to the low SF images ($F(1, 14) = 37.9$, $\eta_p^2 = 0.73$, $p < 0.001$). In contrast during the *flower–face* task, performance was better when participants attended the low SF components in the image (versus high; $F(1, 14) = 48.7$, $\eta_p^2 = 0.77$, $p < 0.001$; Figure 2A). Exposure duration affected performance in the *face–flower* but not in the *house–flower* task (three-way interaction of task, attention, exposure duration: $F(1, 14) = 52$, $\eta_p^2 = 0.79$, $p < 0.001$). In the *face–flower* task, there was a reliable advantage for categorizing the low SF stimuli under short exposure durations (30 ms, $F(1, 14) = 54$, $\eta_p^2 = 0.794$, $p < 0.001$) but not when the hybrids were exposed for 200 ms ($F(1, 14) = 0.375$, $\eta_p^2 = 0.026$). These data suggest that *flower–house* categorization was easier to perform when based on information from high SF channels while *flower–face* categorization was easier to perform when based on low SF information; this latter effect was most apparent with short exposures of the stimuli.

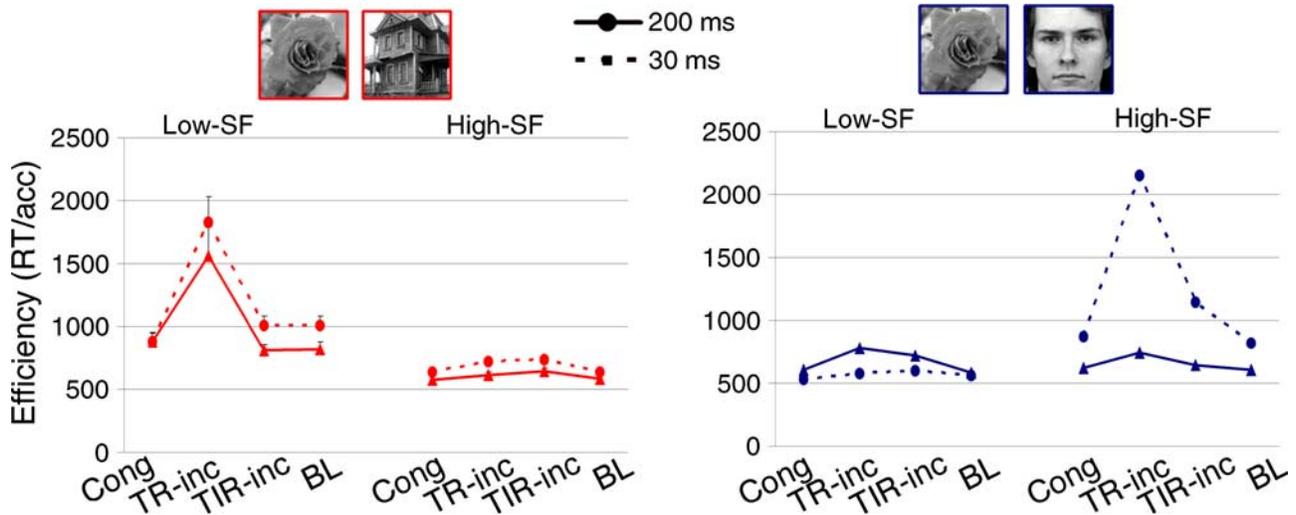
Effects of the unattended components

The next analysis tested the effects of the non-attended SF range on performance (i.e., effects of the SF range that was irrelevant to the explicit attention instructions and that observers were explicitly advised to ignore). The hybrid type affected responses in both tasks ($F(1.06, 14.9) = 29.7$, $\eta_p^2 = 0.68$, $p < 0.001$; see Figure 2A). TR-incong hybrids were more difficult to categorize (mean = 1119.7; CI: 948.3–1291.1) than TIR-incong hybrids (mean = 787, CI: 725.1–849.4) and baseline hybrids (mean = 702.5, CI: 650.6–754.3); the congruent hybrids were the easiest (mean = 700, CI: 646.8–753.7).

Across the conditions, there was no reliable evidence for a benefit when the high and low SF ranges provided congruent information ($F(1, 14) = 0.018$, $\eta_p^2 = 0.001$). The data did suggest that observers benefited from non-attended low SF information when categorizing high SF images under short exposures (*face–flower*: $t(14) = 2.4$, $\eta^2 = 0.3$, $p < 0.05$; *house–flower*: $t(14) = 2.2$, $\eta^2 = 0.26$, $p < 0.05$, though these effects did not survive Bonferroni correction). This pattern of result shows that, on the whole, any benefit from the non-attended SF was marginal in this experiment (occurring only in two out of eight conditions).

Interference from the unattended SF range was estimated for each incongruent hybrid (TR- and TIR-incong) relative to the baseline condition (see Methods section and Figure 2B). A repeated measures ANOVA was computed with the following factors: task (*face–house*, *face–flower*), attended SF (high, low), type of incongruence (TR, TIR), and exposure duration. There was a significant main effect

A. Task by attention by duration interaction



B. Interference effect: task by attention interaction

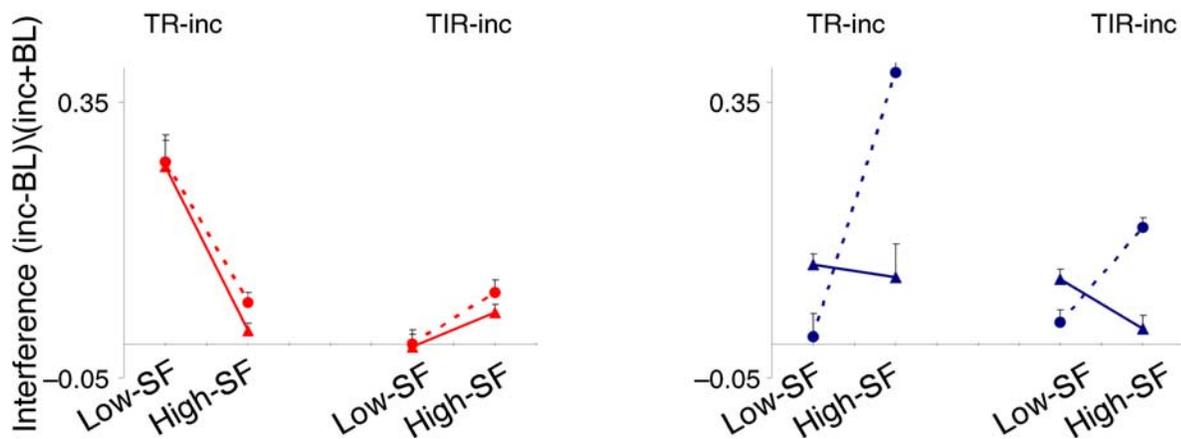


Figure 2. Study 1—Results. (A) Averaged responses across observers presented for each attention condition at each exposure duration (line = 200 ms; dash line = 30 ms) for each hybrid type, left plot for the *house-flowers* task, right for the *face-flowers* task. The y-axis reflects psychological-efficiency (RT/proportion of accuracy) measure of performances. (B) Interference effects presented separately for each task, SF attention, duration, and type of interference conditions. Low-SF and high-SF note the two spatial frequency attention conditions; Cong, congruent; TR-inc, task-relevant incongruent; TIR-inc, task-irrelevant incongruent; BL, baseline.

of interference across all conditions ($F(1, 14) = 122.6$, $\eta_p^2 = 0.90$, $p < 0.001$). However, this effect depended on the incongruence type ($F(1, 14) = 43.7$, $\eta_p^2 = 0.75$, $p < 0.001$), with greater interference when the unattended SF range depicted task-relevant information compared to when it depicted task-irrelevant information. A further significant four-way interaction suggested that the extent of interference varied depending on the categorization task, the attended SF, the exposure duration, and the type of distracter ($F(1, 14) = 5.29$, $\eta_p^2 = 0.27$, $p < 0.05$). To unravel the sources of this interaction, we performed separate analyses as a function of the type of incongruence.

Interference from task-relevant information was modulated by the exposure duration, categorization task, and the SF attention factors (three-way interactions: $F(1, 14) =$

18.8 , $\eta_p^2 = 0.57$, $p < 0.001$). For both exposure durations, task-relevant high SF stimuli interfered with *house-flower* categorization, while low SF stimuli interfered with *face-flower* categorization (2-way interaction: 200 ms, $F(1, 14) = 16.2$, $\eta_p^2 = 0.53$, $p < 0.001$; 30 ms, $F(1, 14) = 57.9$, $\eta_p^2 = 0.8$, $p < 0.001$). Paired t -tests showed that this pattern was reliable following 30-ms exposure duration (interference from low vs. high SF: *house-flower*, $t(14) > -6.1$, $\eta^2 > 0.72$, $p < 0.001$; *face-flower*, $t(14) > 6.9$, $\eta^2 > 0.77$, $p < 0.001$). At the longer exposure duration (200 ms), there was reliable interference from high SF stimuli in the *house-flower* task, but no reliable interference from the *face-flower* task (interference from low SF distracters on high SF targets: *house-flower*, $t(14) > -5.6$, $\eta^2 > 0.69$, $p < 0.01$; *face-flower*, $t(14) > -0.61$, $\eta^2 > 0.026$).

These results show that task-relevant information in the unattended SF range interfered when in a given task, its processing lead to faster and more accurate categorization decisions.

Interference from task-irrelevant information (albeit small, see Figure 2B) was also modulated by the task, the attended SF range, and the exposure duration ($F(1, 14) = 19.79$, $\eta_p^2 = 0.58$, $p < 0.001$). However, the pattern of interference from irrelevant distracters was different than that observed for relevant distracters. For *house–flower* categorization, the pattern of interference went in the opposite directions for task-relevant compared with task-irrelevant distracters (two-way interaction: $F(1, 14) = 65.3$, $\eta_p^2 = 0.82$, $p < 0.001$). High SF relevant distracters caused relatively more interference than low SF distracters (see above and Figure 2B), while interference from low SF irrelevant distracters was higher than that from high SF distracters (interference from TIR low vs. high SF: $t(14) > 2.67$, $\eta^2 = 0.33$, $p < 0.05$). Two potential explanations can account for this result. One is that, in the *house–flower* task, faces were the irrelevant distracters and it has been hypothesized that low SF components are particularly advantageous when processing faces (see further discussion below). A second explanation is that interference from low SF faces arose due to carry-over effects between tasks. Recall that we used a within-subjects design and hence irrelevant information in one task was relevant in the other. Thus, the irrelevant faces here were relevant in the *face–flower* task. To avoid such potential carry-over effects across tasks, in the next study (Study 2, see below), the task-irrelevant information was restricted to one category and was never relevant for other tasks.

The interference pattern from relevant and irrelevant distracters was similar for the *face–flower* task, though the effects were more pronounced with relevant incongruent hybrids (three-way interaction: $F(1, 14) = 11.97$, $\eta_p^2 = 0.46$, $p < 0.01$). For both incongruent conditions, there was larger interference from low than high SF stimuli with brief exposures (interference from TIR low vs. high SF: $t(14) = 7.1$, $\eta^2 = 0.77$, $p < 0.001$), while this effect was attenuated and even reversed for long exposure durations (interference from TIR low vs. high SF: $t(14) > -2.88$, $\eta^2 = 0.37$, $p < 0.05$). The reverse effect may relate to potential carry-over effects between tasks (see above).

In sum, Study 1 showed that the SF components used optimally for between-category perceptual discriminations varied according to the categorization task. In the *flower–house* task, high SF targets were easier to classify than low SF targets, and task-relevant high SF distracters interfered more with performance. Conversely, in the *flower–face* task responses were easier for low SF targets and task-relevant low SF distracters generated more interference when not attended, especially at short exposure durations. These results fit with previous studies showing that discrimination between images of man-made objects (here houses) and natural stimuli (here flowers) is mostly carried by high–medium SF ranges (Oliva & Torralba, 2001).

With respect to the questions we posed in the Introduction section, the results showed that (i) observers can attend flexibly to high and low SF stimuli when directed by explicit instructions. (ii) The ease of categorizing low and high SF stimuli varied depending on the categorization task. This was observed despite: the explicit attention instructions, the specific SF sensitization, and perceptual set biases arising within the blocks and despite the use of identical hybrids across the two tasks. (iii) The information in the unattended SF interfered with the task. Larger interference was observed when the unattended SF was the “preferred” range for the task and when it depicted a category that was relevant to the task. This suggests that much of the interference effect arose at the decision level. These findings support the flexible usage hypothesis—though, in contrast with the fully flexible usage account (Oliva & Schyns, 1997; Schyns & Oliva, 1994, 1999), we found that the ease of categorizing high and low SFs is not equal and that relevant information in the unattended SF range interfered with responses. This result arose even though observers could reliably classify stimuli in the low SF range (indicating that each range contained enough information to achieve reliable categorization).

The task-level-based hypothesis (Collin, 2006; Collin & McMullen, 2005; Morrison & Schyns, 2001; Schyns, 1998) predicted that the level of categorization should determine the preferential use of high and low SFs. However, here both categorization tasks were at the ordinate level though each task was associated with a different SF range. On the other hand, the stimulus-based hypothesis predicts that different SF ranges support different categories of stimulus. Specifically, it is argued that low SF is the “preferred” channel to process faces, probably because of their canonical configuration (Goffaux, Gauthier et al., 2003; Harel & Bentin, 2009). The results of Study 1 provide some support for that hypothesis, since categorizing *faces* vs. *flowers* was easier for low than high SF stimuli. Furthermore, in the case of this specific study, face rather than house and flower stimuli had a homogeneous composition (e.g., the face was always presented in a front view so the spatial locations of the eyes and mouth were predictable). Thus, it could be that when the stimulus configuration is predicted, the processing of low SFs is advantageous.

Study 2 aimed to test the stimulus-based hypothesis and the specific role of low SF in processing faces. To do this, we examined perceptual decisions to low and high SF components in three different tasks all involving face categorization, one at the ordinate level (categorizing *face* vs. *house*) and two at subordinate levels (categorizing the *valence* of facial expressions and the *gender* of faces). If the stimulus category and predictability of stimulus configuration are critical for determining which SFs generate higher psychological efficiency, then the same (low SF) components should be used optimally across all three face tasks.

Study 2: Categorizing faces

Methods

Unless otherwise mentioned, the methods was the same as in Study 1.

Participants

Eighteen volunteers participated (11 females, mean age: 27 years \pm 7 std). All had normal or corrected-to-normal vision. The study was approved by the local ethics committee.

Stimuli

The set of 57 houses, 66 flowers, and 64 neutral (32 female) pictures was the same as in Study 1. In addition, we included 70 faces (35 females) with negative and 70 positive (i.e., angry and happy, respectively) expressions (from the Karolinska Directed Emotional Faces set, see Lundqvist & Litton, 1998). Image processing and the generation of hybrids followed identical procedures to Study 1. Four hybrid stimuli were generated: congruent, TR-incong, TIR-incong, and baseline. In all three tasks, the task-irrelevant objects depicted in the TIR-incong hybrids were flowers (Figure 3).

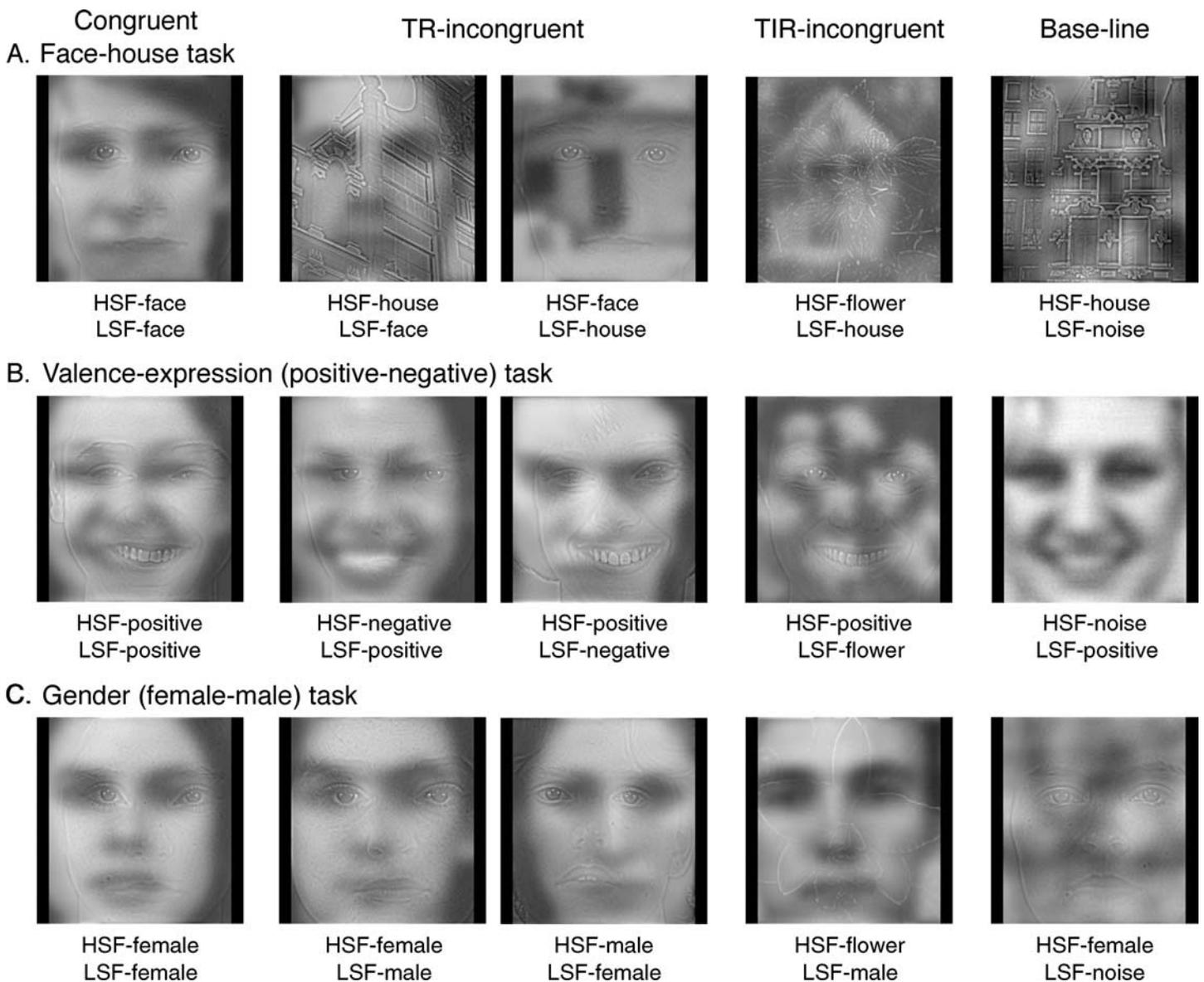


Figure 3. Study 2—Stimuli. (A) Examples of stimuli used in the *face–house* task. (B) Examples of stimuli used in the *valence–expression* task. (C) Examples of stimuli used in the *gender* task. First column, examples of congruent hybrids; second and third columns, examples of task-relevant (TR) incongruent hybrids; fourth column, example of task-irrelevant (TIR) incongruent hybrids; and fifth column, examples of baseline hybrids. HSF, LSF: low and high spatial frequencies, respectively.

Design and procedure

The design and procedure were similar to those of Study 1. Participants always started with the *face–house* categorization task and then the order of the *valence* and *gender* tasks were counterbalanced across observers. This order of presentation ensured that participants had similar levels of familiarity with faces and houses when they performed the *face–house* categorization task (as faces but not houses are used in the two other tasks). A within-subject factorial design was used with the following four factors: Task (*face–house*, *valence*, *gender*), SF attention (low, high), hybrid type (congruent, TR-incong, TIR-incong, baseline), and exposure duration (200 ms, 30 ms).

Results and discussion

Accuracy in the *face–house* task was 88.3% (CI: 68.3–100%); it was 77.6% in the *expression–valence* task (CI: 59.9–95.6%) and 69% in the *gender* task (CI: 52.2–87.6%). Thus, overall performance was well above chance in all cases, indicating that both high and low SF ranges contained sufficient information to perform all three categorization tasks. Table 2 presents averaged median RTs and proportions of correct responses for each condition.

Overall performance

A repeated measure ANOVA was computed with the following four factors: categorization task, attended SF, exposure duration, and hybrid type, using the psychological-efficiency measures as the dependent variable (see Methods section of Study 1 for details). The three tasks varied in their difficulty ($F(1.3, 21.7) = 7.923$, $\eta_p^2 = 0.3$, $p < 0.01$). Categorization was easiest for the *face–house* task and most difficult for the *face–gender* task. The exposure duration had a differential effect on performance according to the SF that participants were instructed to attend to (SF attention-by-duration interaction, $F(1, 17) = 4.485$, $\eta_p^2 = 0.21$, $p < 0.05$). Observers were better in categorizing low compared with high SF targets when the exposure duration was short (30 ms), while the opposite pattern emerged at the longer exposure durations (200 ms, Figure 4B, see below for more details). Critically, there was a reliable interaction between task and the attended SF ($F(1.5, 25.9) = 13.4$, $\eta_p^2 = 0.44$, $p < 0.001$). This interaction was significant even when the TR-incong hybrids were excluded from the analysis ($F(1.3, 22.5) = 17.1$, $\eta_p^2 = 0.5$, $p < 0.001$).

When participants attended to high SF components, the *face–house* categorization task was easier relative to the *expression–valence* task, which in turn was easier than the *gender* task (Figure 4A). In contrast, when participants attended to low SF targets the opposite pattern emerged: it was easier to categorize the *face–gender* than the

	Congruent	TR-incong	TIR-incong	BL + noise
<i>Task: Face–house 200 ms</i>				
Attend	672 (33)	765 (39)	648 (36)	619 (22)
LSF	0.87 (0.02)	0.65 (0.05)	0.94 (0.01)	0.94 (0.01)
Attend	509 (27)	526 (28)	525 (29)	516 (23)
HSF	0.94 (0.01)	0.91 (0.02)	0.95 (0.01)	0.94 (0.01)
<i>Task: Face–house 30 ms</i>				
Attend	651 (37)	737 (50)	677 (36)	644 (35)
LSF	0.91 (0.02)	0.63 (0.06)	0.88 (0.02)	0.90 (0.02)
Attend	542 (22)	560 (26)	553 (22)	547 (27)
HSF	0.94 (0.02)	0.87 (0.02)	0.89 (0.02)	0.93 (0.01)
<i>Task: Valence 200 ms</i>				
Attend	643 (31)	661 (36)	623 (29)	611 (27)
LSF	0.79 (0.03)	0.55 (0.04)	0.81 (0.02)	0.84 (0.03)
Attend	600 (27)	627 (28)	615 (31)	597 (23)
HSF	0.89 (0.01)	0.83 (0.02)	0.87 (0.02)	0.89 (0.01)
<i>Task: Valence 30 ms</i>				
Attend	587 (22)	600 (37)	589 (34)	596 (29)
LSF	0.79 (0.03)	0.60 (0.04)	0.80 (0.03)	0.79 (0.03)
Attend	603 (20)	602 (28)	619 (25)	586 (22)
HSF	0.82 (0.02)	0.60 (0.05)	0.73 (0.02)	0.80 (0.03)
<i>Task: Gender 200 ms</i>				
Attend	634 (26)	670 (37)	636 (21)	620 (21)
LSF	0.77 (0.03)	0.62 (0.03)	0.78 (0.03)	0.79 (0.03)
Attend	625 (30)	676 (43)	660 (32)	633 (28)
HSF	0.82 (0.02)	0.63 (0.04)	0.70 (0.03)	0.74 (0.02)
<i>Task: Gender 30 ms</i>				
Attend	610 (28)	617 (30)	617 (22)	598 (20)
LSF	0.77 (0.02)	0.75 (0.03)	0.77 (0.02)	0.75 (0.03)
Attend	636 (32)	677 (44)	660 (40)	652 (39)
HSF	0.68 (0.04)	0.42 (0.03)	0.57 (0.02)	0.61 (0.03)

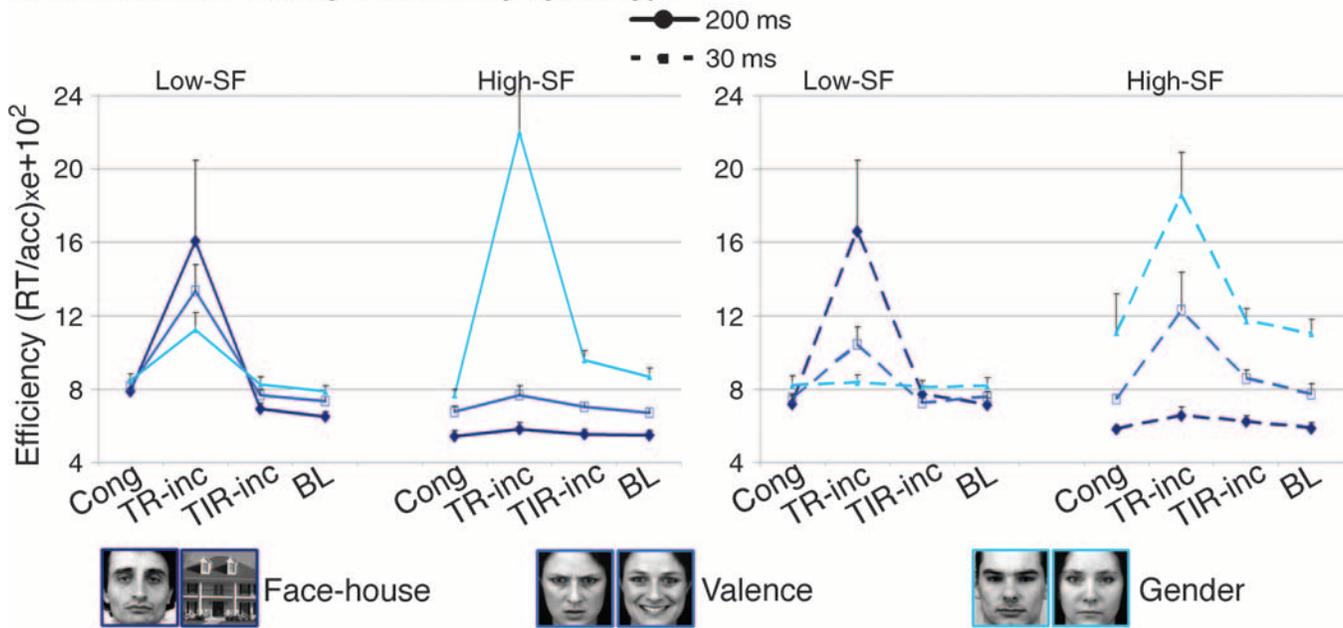
Table 2. Response times (ms) of correct responses (in bold) and proportion of accurate responses (in italics): mean (SEM).

expression–valence and this was in turn easier than the *face–house* discrimination task (Figure 4A). A follow-up analysis, which considered each task separately, clearly showed that each categorization task elicited a different pattern of responses for the high and low SF components.

Face–house and *expression–valence* discriminations were easier when participants attended to high (vs. low) SF targets (*face–house*: $F(1, 17) = 9.5$, $\eta_p^2 = 0.36$, $p < 0.01$ and *expression–valence*: $F(1, 17) = 4.6$, $\eta_p^2 = 0.21$, $p < 0.05$); while gender categorization was easier for low than for high SF targets ($F(1, 17) = 11.5$, $\eta_p^2 = 0.4$, $p < 0.001$).

Exposure duration interacted with performance on the attended SF for the *gender* and *valence* tasks but not for the *face–house* categorization task. Categorizing the *gender* of low SF faces were reliably easier than

A. Interaction of task by attention by hybrid type



B. interference effect:task by attention interaction

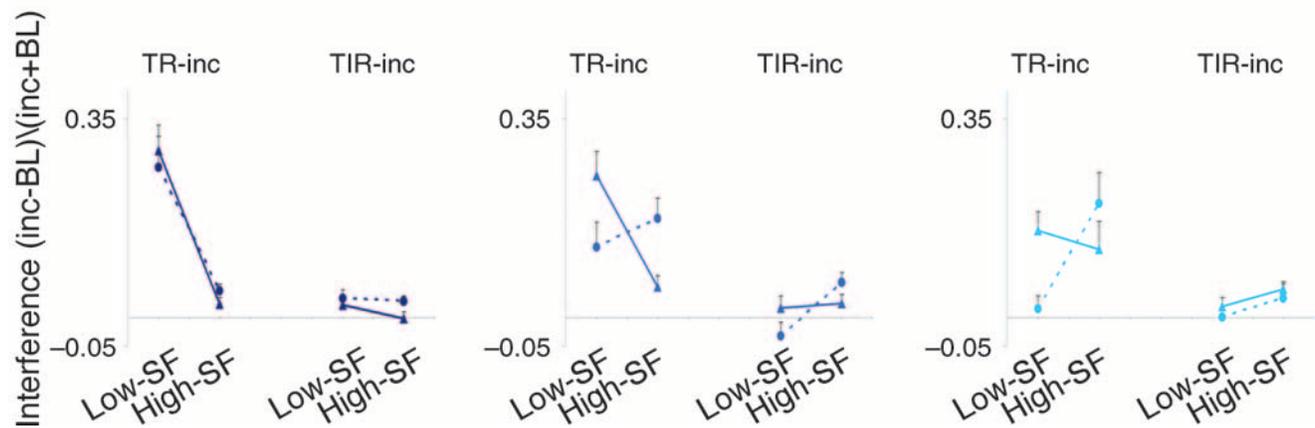


Figure 4. Study 2—Results. (A) Averaged responses across observers presented for each task, SF attention, duration, and hybrid type conditions; dark blue—*house–flower* task, light blue—*valence–expression* task, and cyan—*gender* task. Y-axis represents the efficiency scores (RT/proportion of accurate responses); right plot depicts the responses following 200 ms (line) and left plot for the 30-ms (dashed line) exposure durations. (B) Interference effects presented separately for each task, SF attention, exposure duration, and distracter conditions. Low-SF and high-SF denote the two spatial frequency attention conditions; Cong, congruent; TR-inc, task-relevant incongruent; TIR-inc, task-irrelevant incongruent; BL, baseline.

categorizing high SF faces with short (30 ms) exposures (low vs. high SFs: $F(1, 17) = 17.4$, $\eta_p^2 = 0.5$, $p < 0.001$) but not with longer (200 ms) exposures (low vs. high SF: $F(1, 17) = 0.93$, $\eta_p^2 = 0.05$). Categorizing *valence* under long (200 ms) exposures increased the advantage for high over low SF targets (high vs. low SFs: $F(1, 17) = 21.5$, $\eta_p^2 = 0.56$, $p < 0.001$), while this difference was eliminated with short (30 ms) exposures (high vs. low SF: $F(1, 17) = 0.7$, $\eta_p^2 = 0.04$). There was no effect of exposure duration on *face–house* categorization ($F(1, 17) = 0.07$, $\eta_p^2 = 0.004$).

Effects of the unattended components

Hybrid type affected the psychological-efficiency measure ($F(1.04, 17.6) = 13.6$, $\eta_p^2 = 0.44$, $p < 0.001$, see Figure 4A). TR-incong hybrids were the most difficult to categorize compared with all other types of hybrids, which differed only marginally. Similar to Study 1, the benefit from congruent unattended SF stimuli was not reliable ($F(1, 17) = 0.2$, $\eta_p^2 = 0.01$). There was only one condition in which unattended congruent stimuli facilitated responses. This occurred when observers made a

gender decision on high SF faces with long exposure duration ($t(17) = 3.9$, $\eta^2 = 0.47$, $p < 0.001$). In fact, when categorizing low SF images with long exposures, the categorization of congruent hybrids were more difficult than that of the baseline hybrids, for all three tasks (all $t(17) > -2.3$, $\eta^2 > 0.24$, $p < 0.05$). In comparison, the interference effect from congruent high SF images given long exposures only generated unreliable trends in Study 1 (*house–flower*, $t(14) = -1.2$, $\eta^2 = 0.09$; *face–flower*, $t(14) = -1.1$, $\eta^2 = 0.08$). We therefore conclude that this finding was not replicable under all categorization contexts and it will not be discussed further.

As in Study 1, there was a reliable interference effect on incongruent trials ($F(1, 17) = 161$, $\eta_p^2 = 0.9$, $p < 0.001$), and again the interference effect from task-relevant distracters was much larger than that from task-irrelevant distracters ($F(1, 17) = 146.4$, $\eta_p^2 = 0.89$, $p < 0.001$, Figure 4B). Similar to Study 1, the extent of interference was modulated by the categorization task, the attended SF, the exposure duration, and the relevance of distracting information ($F(1.7, 29.7) = 3.3$, $\eta_p^2 = 0.16$, $p = 0.05$). This interaction was decomposed by performing separate analyses for each task and each incongruence type, to better understand the source of it.

In the *face–house* task, high SF relevant distracters interfered more than low SF distracters ($F(1, 17) = 23.5$, $\eta_p^2 = 0.58$, $p < 0.001$), an effect that was similar for both exposure durations. In the *expression–valence* task, exposure duration affected the interference pattern ($F(1, 17) = 43.7$, $\eta_p^2 = 0.72$, $p < 0.001$), with larger interference from high SF relevant distracters only under long exposures ($t(17) = -4$, $\eta^2 = 0.48$, $p < 0.001$) but not with brief exposures ($t(17) = 0.8$, $\eta^2 = 0.03$). A complimentary interaction was observed in the gender task ($F(1, 17) = 10.9$, $\eta_p^2 = 0.87$, $p < 0.01$), with greater interference from low SF relevant distracters given brief exposures ($t(17) = 3.4$, $\eta^2 = 0.4$, $p < 0.01$), and no interference when hybrids were presented for longer durations ($t(17) = -0.57$, $\eta^2 = 0.02$).

Interference from task-irrelevant distracters in the *face–house* task was not modulated by any of the conditions and it was reliable only when the hybrids were presented for short exposures (interference from the unattended high SFs, $t(17) = 2.4$, $\eta^2 = 0.25$, $p < 0.05$; and from the unattended low SFs, $t(17) = 3.2$, $\eta^2 = 3.7$, $p < 0.05$). Similarly, interference effects in the *gender* task were not modulated by the conditions and were significant only when categorizing high SF images given long exposures ($t(17) = 3.9$, $\eta^2 = 0.47$, $p < 0.01$). In the *valence–expression* task, interference from irrelevant distracters was affected by the conditions (two-way interaction: $F(1, 17) = 8$, $\eta_p^2 = 0.32$, $p < 0.05$), but none of the simple comparisons was significant. Note that the reliability of the above simple effects did not survive Bonferroni corrections. This further confirms that interferences from response-irrelevant distracters were marginal.

To summarize, categorization of *faces vs. houses* was easier for high SF stimuli. In this task, the high SF distracters interfered only when they depicted the response relevant category. Similarly, categorizing *face–valence* was easier for high SF targets and high SF distracters caused more interference when the unattended information was relevant to the task. The advantage for high SF stimuli in the *valence* task was manifested at the long exposure duration. In contrast, categorizing *face–gender* was better for low SF targets and low SF distracters caused greater interference, especially at the short exposure duration and when the distracters depicted the task-relevant category.

The results of the subordinate face categorization tasks are in agreement with previous research that tested the use of high and low SFs in face perception. In particular, our data indicate that categorizing facial expressions is associated with the processing of high SFs (Schyns & Oliva, 1999), while categorizing face–gender is associated with low SFs (Goffaux, Jemel et al., 2003; though see Schyns & Oliva, 1999).

The main aim of Study 2 was to test the stimulus-based approach in relation to the flexible use account of the processing of high and low SF images. Specifically, we wanted to test whether faces are predominantly categorized using information from the low SF range. Our data categorically showed that no such relation exists between faces and low SF components. Furthermore, we showed that the advantage for high over low SF stimuli in object categorization, including faces, varies with the task, i.e., the categories that are compared.

One interpretation of the results of Studies 1 and 2 suggest that the effect of task arose from the saliency of the attended SF components relative to the non-attended stimuli (i.e., the level of distraction from the background; Navalpakkam & Itti, 2007). Alternatively, perceptual decisions may be based on separate comparisons between the categories at any given SF range. According to this later hypothesis, the high and low SF stimuli in the hybrid are compared with template representations of categories defined in terms of the visual properties required to categorize that target (perhaps based on specific features of high and low SF components; Duncan & Humphreys, 1989). If there is independent access to the templates for each SF, then performance should not depend on the distracting information present in the non-attended SF range. The results from Study 1 already provided some support for this latter hypothesis by demonstrating that, with the same hybrid combinations, the efficiency of discriminating low and high SF targets changes dramatically (see Study 1). The following analysis aimed to provide a more direct test for this hypothesis.

We next directly tested which of the above alternative explanations fit the data better (e.g., saliency of target against background, or the use of prior template knowledge). To that end, we compared hybrids that were

identical across tasks and required an identical response—e.g., recognizing faces in the congruent hybrids when the opposing target categories were flowers or houses. If the ease of categorizing a particular SF is driven by the target–background relation, then we would predict that responses to identical hybrids would be independent of the task. However, if the categorization task affects the decision criteria (templates) set up for each SF range, then responses to identical hybrids would vary depending on the categories that are being compared.

Comparisons across tasks (Studies 1 and 2)—Are psychological-efficiency effects driven by the stimuli or the task?

We assessed the effects of the categorization task and the attended SF on categorizing faces when embedded in the congruent hybrids (Figure 5A). A mixed ANOVA was used with task as a between-subjects factor and the attended SF as a repeated measures factor. Responses to face targets were affected by the task context (task by attention interaction: $F(1, 31) = 49.7$, $\eta_p^2 = 0.61$, $p < 0.001$). Observers found it easier to categorize the low compared with the high SF faces in the context of the *face–flower* task ($t(26.1) = 4.6$, $\eta^2 = 0.45$, $p = 0.001$) but easier to categorize the high than low SF faces in the context of the *face–house* task ($t(24.5) = -5.1$, $\eta^2 = 0.5$, $p < 0.001$).

The same pattern was observed for flower categorization. A repeated measures ANOVA tested effects of task and SF attention on categorizing flowers that were embedded in the congruent hybrids (Figure 5B). When congruent flower hybrids were presented, the decision that the attended SF range depicted flowers was affected by the task ($F(1, 14) = 38.2$, $\eta_p^2 = 0.73$, $p < 0.001$). Low SF flowers were easier to categorize in the *flower–face* task compared with the *flower–house* task ($t(14) = 5.2$, $\eta^2 = 0.66$, $p < 0.001$); in contrast, categorizing high SF flowers was easier in the *flower–house* compared with the *flower–face* task ($t(14) = 2.6$, $\eta^2 = 0.3$, $p < 0.05$).

Finally, the effects of task and the attended SF on house categorization were tested using a mixed ANOVA with task as a between-subjects factor and SF attended as a within-subjects factor. Here, high SFs were preferred for both tasks. There was no interaction between SF attention by task ($F(1, 31) = 1.1$, $\eta_p^2 = 0.03$), and only a main effect of the SF attended ($F(1, 31) = 44.5$, $\eta_p^2 = 0.59$, $p < 0.001$; Figure 5C).

The data demonstrated that the relations between the background and the target had a minimal effect on the difficulty to categorize targets. The more important factor is the two compared categories in any given task. Therefore, we propose that decisions are achieved by independent comparisons of template representations at the different SF ranges. This finding accords with the observations reported above in which interferences arose mostly from task-relevant incongruent distracters when the “preferred” SF range was to be ignored (e.g., larger interference from high SF components when categorizing low SF images of *faces and houses*). Taken together, the results suggest that the comparisons of stimuli to the discriminative templates were carried in parallel for the high and low SFs, with responses being affected by the ease of making a categorical decision based on one SF range or the other.

To summarize the arguments so far, our data do not support the task-level hypothesis because the different SFs were used differentially for ordinate level and subordinate categorizations. The data also do not support the stimulus-based hypothesis since categorization tasks for faces varied in their biases to high versus low SF. The advantage of SF range over another was also not associated with the overall ease of performing the task. The two most difficult tasks (categorizing gender and expressions) were performed most easily with low SF and high SF targets, respectively. This latter finding also rules out the possibility that high SF information is used only when information conveyed by low SF components provides insufficient details for the categorization task—a prediction implied by the coarse-to-fine hypothesis.

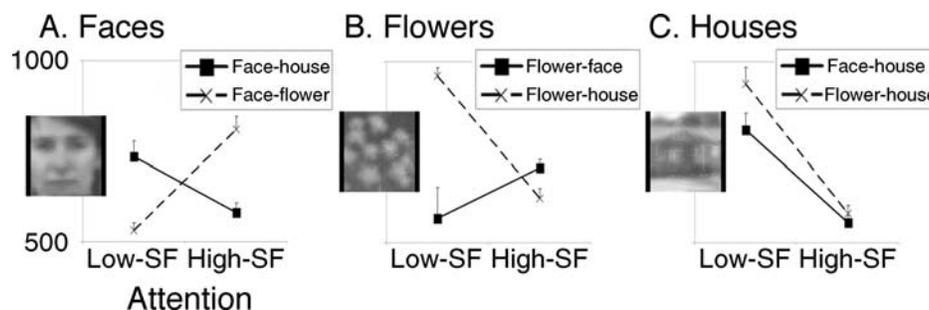


Figure 5. Comparison across tasks—Stimuli versus task. Averaged responses across observers for a given target category from identical congruent hybrids (i.e., identical distracting background information). Y-axis, psychological-efficiency scores; x-axis, attention conditions; different lines for the two different task contexts. (A) Responses to face targets in the face–face congruent hybrids. (B) Responses to flower targets in the flower–flower congruent hybrids. (C) Responses to house targets embedded in house–house congruent hybrids. Error bars are standard error of the means.

Finally, we demonstrated that the critical SF was not related to the background information per se suggesting that perceptual decisions were made separately for high and low SF images, consistent with an independent template account.

This raises the possibility that the information used to specify the templates for the categorization tasks may lie in the differences in the statistical properties between the stimulus categories. As a first attempt to identify the stimulus properties critical to the current tasks, a third study examined the statistics of the high and low images for each stimulus category and associated these with the observed behavior patterns.

Study 3—An analysis of the stimuli properties

The rationale of this analysis was inspired by theoretical models on perceptual decisions and search tasks, which suggest that perceptual decisions are based on computing the physical differences between sensory inputs (Heekeren, Marrett, Bandettini, & Ungerleider, 2004; Heekeren, Marrett, Ruff, Bandettini, & Ungerleider, 2006; Heekeren, Marrett, & Ungerleider, 2008; Navalpakkam & Itti, 2007; Romo & Salinas, 2003), with the efficiency of the decision affected by the size of the differences between sensory inputs (Romo & Salinas, 2003). Furthermore, computational models for object recognition that use unsupervised learning suggest that hidden layers in the models encode the statistical differences that distinguish between response categories (e.g., Fidler, Berginc, & Leonardis, 2006; Leibe, Leonardis, & Schiele, 2009). Here we assumed that the relatively large number of stimulus exemplars used per category will provide a reliable estimation of the statistical differences between the stimuli, as encoded in the mental templates. We focused on three low-level stimulus properties that are known to be encoded in the visual cortex: the overall energy in the image, pixel intensities, and orientation information (the axes along which luminance changes).

Methods

Statistical analysis of the stimuli was carried out separately on the low and high SF filtered images. For each stimulus, we first computed an overall energy level using the power spectral density (PSD) based on the spatial frequency representation of the image. The overall power was estimated as the mean of the power across all frequencies and all orientations for any given filtered image. Orientation maps were computed using Kovési's algorithm that uses phase coherence within Fourier space

to compute the orientation at each pixel (<http://www.csse.uwa.edu.au/~pk/>). Two measures were derived for the orientation property: orientation maps and overall orientation information (e.g., amount of vertical lines). The later was computed by counting the number of pixels with a given orientation (0–90 degrees) across each image. We note that previous reports suggest that orientation information independent of spatial location may be sufficient for discriminating between scenes (Oliva & Torralba, 2001).

Next we quantified the reliability of differences between every two categories (e.g., high SF: face vs. house; low SF: face vs. house; high SF: female vs. male; low SF female vs. male). Reliability was tested using independent two-sample *t*-tests, assuming the stimulus samples have unequal variance, separately for each stimulus property: energy level, orientation, and pixel intensity. To facilitate comparisons across studies, the *t*-values were transformed to reflect the effect size (η^2) for each comparison. Differences in orientations were computed by calculating differences in the number of pixels having a specific orientation range between any two categories (0–90 degrees, binned into 10-degree ranges: 0–10, 11–20, etc.). A mean of the η^2 values was computed across all orientations for each comparison. Differences in pixel intensity and pixel orientation were computed using SPM5 (<http://www.fil.ion.ucl.ac.uk/~spm>). The mean value of *F* across all images was used to compute an overall effect size.

Finally, we used the generalized linear model within SPSS16.0 to test whether differences in the statistics of the features predicted behavioral performance. In the model, we used the mean psychological-efficiency measure for each SF attention and task ($N = 10$) as the dependent variable and, as predictors, the measures of differences in stimulus statistics (i.e., the η^2 values): overall energy, pixel intensity, and orientation information. We used a linear function link and likelihood ratio chi square statistics. To identify the most likely model, we used a forward inclusion approach. The generalized linear model was used as the variables were not normally distributed.

Results and discussion

Differences in image statistics between the categories are presented in Figure 6. Comparisons across pixel intensity (Figures 6A–6D) revealed that low SF images were by far more diagnostic than high SF images for all the pair-wise comparisons (η^2 range low SF: 0.14–0.34; η^2 range high SF: 0.005–0.07). If participants had used this measure alone, then categorizing low SF images would always be better than categorizing high SF images. Intriguingly, this was not the case in the present data.

We next tested for differences in the overall orientations that characterized each category. As can be seen in Figures 6B and 6E, the largest differences are found for vertical and horizontal orientations. For example, high SF

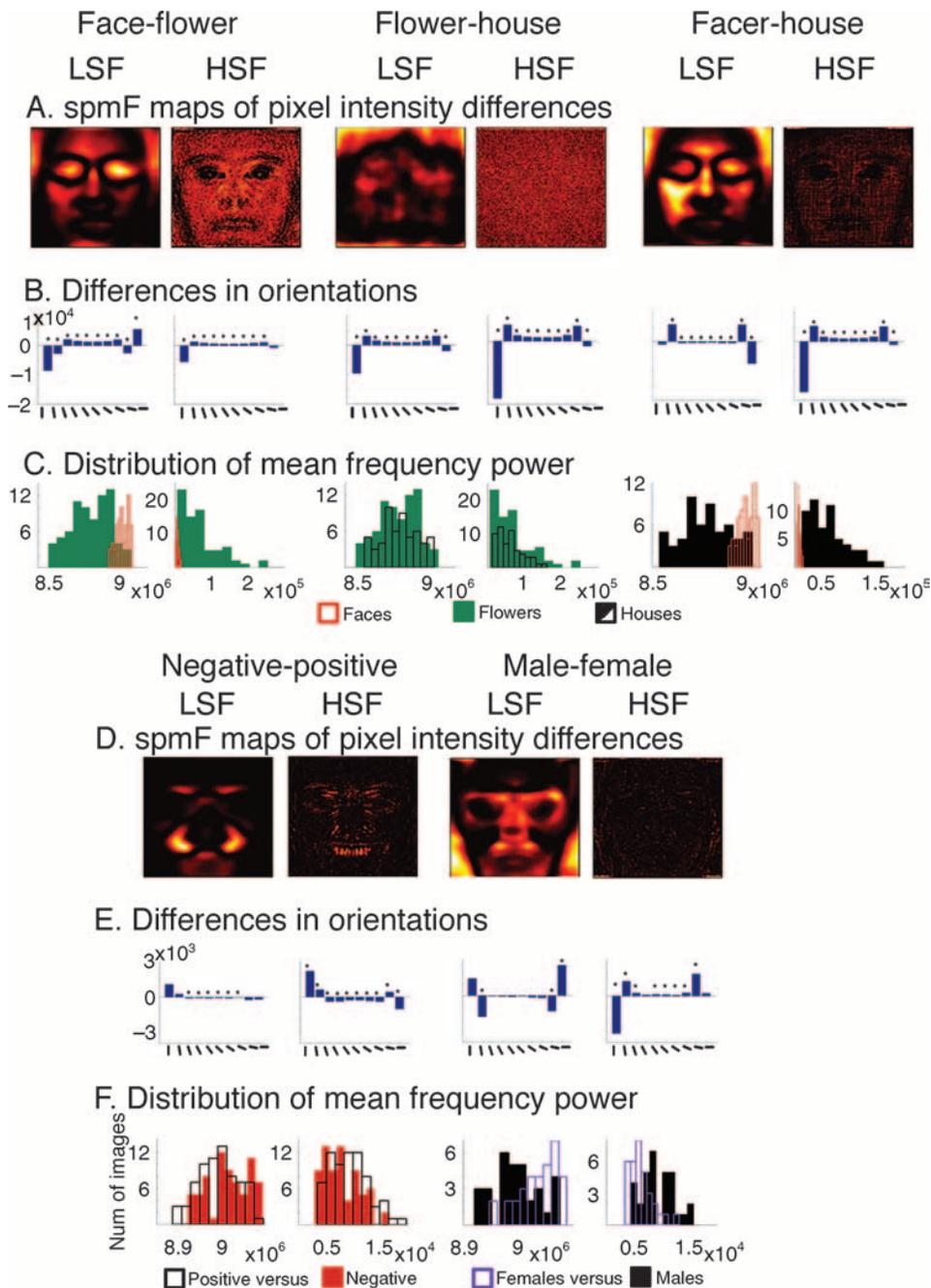


Figure 6. Comparisons of stimulus statistics. (A, D) SPM F-maps presenting pixel-wise comparisons between filtered images for the different stimulus categories. For example, in (A), leftmost column presents the difference between low SF faces and low SF flowers; in (D), leftmost column shows the difference between low SF positive and negative expressions. The maps are thresholded at $p < 0.05$. (B, E) Bars representing the differences in number of pixels (y-axes) for each orientation (depicted in the x-axes) between filtered images for the different categories. For example, in (B), leftmost column presents differences between low SF faces and low SF flowers. * indicates a significant difference at $p < 0.05$. (C, F) Histograms presenting the distribution of mean frequency power, i.e., the energy level of the SF filtered stimulus. Y-axes, number of images; x-axes, overall frequency power.

houses have a larger number of vertical elements (pixels) than faces and flowers. Similarly, larger numbers of vertical elements are common in high SF angry faces compared with happy faces and in high SF female compared with male faces. To quantify the diagnostic value of this property, we computed a mean effect size

across all 10 orientations. Overall orientation differences between faces and flowers were more reliable for the low than the high SF images ($\eta^2 = 0.62, 0.55$, respectively). In contrast, more reliable differences in orientation were observed between high SF houses and faces ($\eta^2 = 0.63, 0.34$, respectively) and between high SF houses and

flowers ($\eta^2 = 0.59, 0.37$, respectively). Not surprisingly, the average differences in the orientation elements were much weaker and less reliable between subface categories. The average reliability of differences in orientation across all elements between angry and happy faces was $\eta^2 = 0.02$ for high SF faces and $\eta^2 = 0.13$ for low SF faces. For male and female faces, the values were: high SF $\eta^2 = 0.14$ and low SF faces $\eta^2 = 0.09$. We also computed SPM statistics directly on the orientation maps. The SPM approach compares the orientation information at each location (pixel by pixel) as opposed to the comparison of the occurrences of particular orientations overall. However, affects of the orientation differences pixel by pixel was unreliable for all comparisons ($\eta^2 < 0.01$).

The final measure was the overall energy in the image. The distributions of this parameter across the high and low SF images for a given category are presented in Figures 6C and 6F. Consistent with common knowledge (Field & Brady, 1997), in our sample stimuli the low SF images had much higher energy levels than high SF images. However, our main interest here was whether the amount of energy can be a diagnostic feature to differentiate between two categories. In the low SF filtered images, faces had the highest energy ($9.02e + 006$), then flowers ($8.798e + 006$), then houses ($8.795e + 006$). This low SF face advantage was reliable when compared with flowers ($t(91) = 13$, $\eta^2 = 0.615$, $p < 0.01$) and houses ($t(67) = 11.98$, $\eta^2 > 0.68$, $p < 0.01$). Low SF flowers did not differ from low SF houses ($t(126) = 0.36$, $\eta^2 = 0.001$). In the high SF images, the highest energy was observed for flowers ($5.37e + 004$) then houses ($5.26e + 004$) and then faces ($0.722e + 004$). The overall difference in energy between high SF flowers and houses was not significant ($t(128) = 0.87$, $\eta^2 = 0.006$), but the energy of high SF faces differed significantly from flowers ($t(75.5) = 9.6$, $\eta^2 = 0.55$, $p < 0.01$) and from houses ($t(56.7) = 11.3$, $\eta^2 = 0.693$, $p < 0.01$). This suggests that, in the case of the current experiment, the energy level of high and low SFs provided sufficient diagnostic information to distinguish between faces and flowers, and faces and houses, but not between houses and flowers.

We next compared the energy level of the stimuli used in the different face categorization tasks. Low SF female faces ($9.04e + 006$) had reliably higher energies than low SF male faces ($8.99e + 006$; $t(61) = 3.9$, $\eta^2 = 0.2$, $p < 0.01$), while high SF male faces ($8.7e + 003$) had higher energy levels than high SF female faces ($5.7e + 003$; $t(56.4) = 4.5$, $\eta^2 = 0.26$, $p < 0.01$). Low SF negative expressions ($9.01e + 006$) had reliably more energy than low SF positive expressions ($8.99e + 006$; $t(137.8) = 2.8$, $\eta^2 = 0.05$, $p < 0.05$), while high SF positive expressions ($9.45e + 003$) had reliably higher energy than low SF negative expressions ($7.47e + 003$; $t(137.9) = 2.5$, $\eta^2 = 0.04$, $p < 0.05$). Interestingly, despite the high similarity between the subface categories, energy levels reliably differentiate between the face categories for both the low and high SF stimuli, though the effect size for differences

between high SF images were notably higher compared to low SF.

To test whether any of these diagnostic measures predicted the behavioral performances, we computed a generalized linear model. The independent variable was the psychological-efficiency measure (RT/Acc) in each categorization task and SF attention condition: 5 tasks * 2 SF attention. We built the model gradually, each time adding an additional predictor and assessing whether it improves the predictions. Note that, given the low number of parameters ($n = 10$), the sensitivity of the model was limited. Including a predictor that reflected differences in overall orientation information significantly improved the model fit when compared to a model based only on the intercept (likelihood ratio $\chi^2 = 4.703$, $p < 0.05$). In addition, the orientation regressor in that model significantly contributed to the predictions of the behavioral performance (e.g., its β was significantly different than zero; Wald $\chi^2 = 6$, $p < 0.05$). Adding any other predictor (e.g., energy, intensity, orientation-by-pixel and interaction between factors) did not improve the model. Furthermore, none of the other predictors on their own explained the data better than the intercept ($p > 0.1$). From this, we conclude that overall orientation information was the most likely diagnostic feature used by observers in the current categorization tasks.

General discussion

The current studies demonstrate that a measure of psychological efficiency for perceptual decisions depended on the interaction between the SF information provided and the specific comparison at hand. We further showed that the processing of the non-attended SF range interfered with processing of the attended SF, particularly when the non-attended SF range was the “preferred” range in that task and when it provided task-relevant information. Differences in categorizing low and high SF images were observed across tasks even when we compared responses to identical hybrids. Exposure duration accentuated the preferential use of one SF over another; for example, the advantage for low SF stimuli in gender categorization was stronger for briefly presented hybrids. Surprisingly, interference from task-irrelevant objects was negligible as was the benefit from congruent information in both SF ranges. Finally, exploratory analyses of the statistical properties of the stimuli suggested that the relative advantage of using one particular SF component for a given task was associated with differences in overall orientation information. We conclude that the diagnostic utility of contrasting SFs depends on the relative differences of the low-level visual properties between the target categories.

This study was designed to answer four questions. We first asked whether observers can flexibly report

information from different SFs based on explicit instructions. We showed that observers can direct attention to either low or the high SF components independently and that each set of components provided sufficient information to complete all the current categorization tasks. More interestingly, in response to our second question, we demonstrated that the ease of categorization responses of low or high SF components depended on the compared categories. This suggests that the use of SF is not fully flexible and is susceptible to specific biases toward one SF over another. This result partially accords with a previous study (Schyns & Oliva, 1999) that used identical face hybrids and showed that SF biases changes with the task demand. Here we replicate these findings and extended them to other stimuli and categorization tasks, using different bias measures. However, in contrast to previous findings (Schyns & Oliva, 1999) we demonstrate that SF biases for different comparisons persist beyond explicit (attention instructions) and implicit (e.g., sensitizations, perceptual sets) manipulations of processing.

The data showed that visual processing is highly adaptive and various properties of stimuli can be used in order to optimize task performance. Therefore, any studies of visual recognition must pay careful attention to the task given to the observers, as a small change in the task (i.e., categorizing faces from flowers or faces from houses) give rise to opposite pattern of results, suggesting that these tasks were mediated by different processes. We have demonstrated this flexibility here with respect to the use of SF range, though we believe a similar flexibility could hold for other properties of stimuli. In other words, the visual system cannot be strictly bottom-up and hardwired.

Our third aim was to test the effects of the non-attended SF ranges. We tested two types of interference: one from a task-relevant category and another from a task-irrelevant category. The data showed that the non-attended SF range interfered mostly when it provided information that conflicted with the response to the target, and there was negligible interference when the information in the non-attended SF was irrelevant to the task (Figures 2C and 4B). Furthermore, greater interference was observed from the “preferred” SF range for a given task.

The greater interference from task-relevant relative to irrelevant distracters (an effect of response conflict) implies that the high and low SF images were analyzed separately and in parallel up to the level of response selection and that each SF component was categorized independently. In support of this idea, we note that there were no facilitation effects when the low and high SF components provided congruent information. It further suggests that a given task set (the comparison at hand) “activated” priors indicating the likely differences between the target categories (e.g., houses tend to have more vertical lines than faces), which biased the sensory processing. The pattern of interference suggests that the priors are based on overall stimulus statistics and are not limited to specific SF range. Hence, task-relevant

conflicting information conveyed by different SF ranges was not suppressed at the level of sensory-perceptual processing and the information from the two SF ranges competed at the decision level. Decisions based on the SF range that are characterized by larger statistical differences were easier and faster, while this task-informative SF components need to be suppressed if these components characterize the non-attended stimuli.

Exposure duration affected performance mostly in the relatively difficult tasks. In the *face–flower* and *gender* tasks, the advantage for low over high SF targets was increased with brief exposures (30 ms) compared to longer exposures (200 ms); in contrast, in the *valence–expression* task the advantage for high over low SF targets was attenuated for brief exposures (30 ms). These results concur with previous findings demonstrating that 30-ms exposures bias perception toward the low SF ranges (Schyns & Oliva, 1994). This is consistent with the observation that low SF components are processed more rapidly than high SF components (Bullier, 2001; Lamme, 2001; Livingstone & Hubel, 1988). However, in the *face–house* and *flower–house* tasks, where responses were more efficient with high (rather than low) SF components, exposure durations did not affect the overall pattern of performance (Figures 2 and 4). These results show that despite low SF information being processed more rapidly, recognition followed the more diagnostic SF range for the task. Furthermore, these findings hint that the rapid projection of low SF information to the brain may have a different functional role than creating coarse representation of the visual scene. We postulate for example that the rapid processing of low SFs may be important for functions associated with dorsal stream functions: navigating in space, perceiving motion, and motor planning, but this is less important for the ventral stream function of recognizing objects (Ungerleider & Haxby, 1994). This latter idea is supported by the observations that majority of magno-cellular pathways terminate at regions associated with the dorsal stream processes (Shipp, 2001; Shipp & Zeki, 1995).

Our fourth question asked: what are the potential sources for the variability in SF biases between the different categorization tasks. We tested whether variability depends on the information in the stimuli or the task at hand. In other words, is an SF bias determined by the saliency of the target relative to the information in the background (Navalpakkam & Itti, 2007), or is it based on comparing input stimulus to mental templates (Duncan & Humphreys, 1989)? Our design allowed us to address this question by comparing performance across tasks where the relations between the attended SF (the target), the non-attended SF (the distractor), and the response were fixed but only the task context changed. We found that, for identical hybrids, the privileged SF range depended on the categorization task. For example, categorizing low SF faces from congruent face hybrids was more efficient in the *face–flower* than in the *face–house* task, while the

opposite was true for categorizing high SF faces (Figure 5). Similar effects were observed for flowers. These results demonstrate that the diagnostic value of one SF range over another was not affected by the properties of the background distractor and that perceptual decisions were based on separate comparisons between categories at each SF range. Consequently, in Study 3, we explored the types of information in the high and low SF images that corresponded with the variations in performance efficiency for the tasks.

The analysis of image statistics assumed that psychological efficiency is linked to differences in low-level visual information (Heekeren et al., 2008), with larger visual differences between categories being associated with easier categorization decisions. We observed that the pixel intensities of the low SF images provide the largest and most reliable differences between each pair of categories (Figure 6). However, despite the diagnostic value of this property it was not associated with participants' performances. This suggests that stimulus statistics alone are not sufficient predictors of perceptions, and performance does not necessarily follow the most efficient computation pattern. Instead, our analysis revealed that performance was best predicted by orientation information in the image. For example, the number of vertical lines was reliable for dissociating between houses and faces, and houses and flowers, and it also discriminated well between facial expressions and gender. Interestingly, the pixel-by-pixel comparison of the differences in orientation information measured using SPM did not yield reliable results. In contrast, the summation of the numbers of pixels in an image with a specific orientation (e.g., number of pixel that presented vertical lines) generated robust differences between two image categories that were associated with performance difficulty. This shows that diagnostic orientation information in the images was not based on computations at a specific location in the visual field; rather it was more related to summation across an image of the overall responses to a given orientation that were "activated" by a given stimulus. These findings fit with results reported in previous computational studies (Oliva & Torralba, 2001) where reliable categorization of different scenes relies on summation of orientation information across the visual field.

A final point to consider is whether the data could reflect a factor such as the perceptual load of the task (cf., Lavie, 1995). Possibly effects of certain properties only becomes apparent at particular levels of task load, and perhaps variations in task load across the categorization tasks contributed to performance. However, predictions derived from the perceptual load account do not match the data. Generally, there is less processing of distracters under high load conditions (Lavie, 1995). This predicts less interference from the "unattended" stimulus as the difficulty of discriminating the target increases. Contrary to this, we found a larger interference effect when target discrimination was more difficult. For example, in the

face–house task there were larger interferences from unattended high SF, when categorizing the more difficult range of low SF. This rules out the possibility that perceptual load was a confounding factor in our design.

In conclusion, we established that the utility of low and high SF components in images is determined by the task context, i.e., the categories that have to be compared. We showed that the SF range used for performance was not determined by the level of comparison (basic vs. subordinate level), by the category of the stimulus (e.g., faces, flowers), or by the background per se. In addition, the "preferred" SF range for the task generated most interference when it fell in the non-attended stimulus. Over and above this, biases on performance based on particular SFs appeared to be linked to the size of differences in low-level visual properties between stimulus categories (most notably of orientation information). We propose that (1) diagnostic features for visual categorization are determined by a comparison between targets and stimulus templates (or a difference template) set up according to the discrimination required of the observer, rather than are fixed to absolute values of the stimuli, and (2) this effect may impact on unattended as well as attended stimuli. Our results imply that visual recognition is a flexible process that aims to optimize performance. Hence, understanding the mechanisms that underlie object recognition must take into account the context in which the processes are tested, as different contexts may yield contrasting results.

Acknowledgments

This work was supported by grants from the MRC and a Leverhulme fellowship to P. R. We thank C. Hutton, S. Dakin, and C. Collin for advice on preparing the stimuli and J. J. Geng for fruitful discussions.

Commercial relationships: none.

Corresponding author: Pia Rotshtein.

Email: p.rotshtein@bham.ac.uk.

Address: School of Psychology, University of Birmingham, Birmingham B15 2TT, UK.

References

- Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, *15*, 600–609.
- Blackmore, C., & Campbell, F. W. (1969). On the existence of neurons in the human visual system selectivity sensitive to orientation and size of retinal images. *The Journal of Physiology*, *203*, 237–260.
- Bullier, J. (2001). Integrated model of visual processing. *Brain Research and Brain Research Reviews*, *36*, 96–107.

- Bullier, J., Hupe, J. M., James, A. C., & Girard, P. (2001). The role of feedback connections in shaping the responses of visual cortical neurons. *Program Brain Research*, *134*, 193–204.
- Collin, C. A. (2006). Spatial-frequency thresholds for object categorisation at basic and subordinate levels. *Perception*, *35*, 41–52.
- Collin, C. A., & McMullen, P. A. (2005). Subordinate-level categorization relies on high spatial frequencies to a greater degree than basic-level categorization. *Perception & Psychophysics*, *67*, 354–364.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.
- Eger, E., Schyns, P. G., & Kleinschmidt, A. (2004). Scale invariant adaptation in fusiform face-responsive regions. *Neuroimage*, *22*, 232–242.
- Fidler, S., Berginc, G., & Leonardis, A. (2006). Hierarchical statistical learning of generic parts of object structure. *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *1*, 182–189.
- Field, D. J., & Brady, N. (1997). Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes. *Vision Research*, *37*, 3367–3383.
- Gauthier, I., Curby, K. M., Skudlarski, P., & Epstein, R. A. (2005). Individual differences in FFA activity suggest independent processing at different spatial scales. *Cognitive Affective Behavioral Neuroscience*, *5*, 222–234.
- Goffaux, V., Gauthier, I., & Rossion, B. (2003). Spatial scale contribution to early visual differences between face and object processing. *Brain Research and Cognitive Brain Research*, *16*, 416–424.
- Goffaux, V., Jemel, B., Jacques, C., Rossion, B., & Schyns, P. (2003). ERP evidence for task modulations on face perceptual processing at different spatial scales. *Cognitive Science*, *27*, 313–325.
- Harel, A., & Bentin, S. (2009). Stimulus type, level of categorization and spatial frequencies utilization: Implications for perceptual categorization hierarchies. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1264–1273.
- Heekeren, H. R., Marrett, S., Bandettini, P. A., & Ungerleider, L. G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature*, *431*, 859–862.
- Heekeren, H. R., Marrett, S., Ruff, D. A., Bandettini, P. A., & Ungerleider, L. G. (2006). Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 10023–10028.
- Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience*, *9*, 467–479.
- Lamme, V. A. F. (2001). Blindsight: The role of feedforward and feedback cortical connections. *Acta Psychologica*, *107*, 209–228.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 451–468.
- Leibe, B., Leonardis, A., & Schiele, B. (2009). Robust object detection with interleaved categorization and segmentation. *International Journal of Computer Vision*, *77*, 259–289.
- Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, *240*, 740–749.
- Lundqvist, D., & Litton, J. E. (1998). The Averaged Karolinska Directed Emotional Faces—AKDEF [AKDEF CDROM]. Stockholm, Sweden: Psychology, Karolinska Institute.
- Marr, D. (1982). *Vision*. San Francisco: Freeman Publisher.
- McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex*, *9*, 431–444.
- Mermillod, M., Guyader, N., & Chauvin, A. (2005). The coarse-to-fine hypothesis revisited: Evidence from neuro-computational modeling. *Brain Cognitive*, *57*, 151–157.
- Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin & Review*, *8*, 454–469.
- Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research*, *39*, 3824–3833.
- Navalpakkam, V., & Itti, L. (2007). Search goal tunes visual features optimally. *Neuron*, *53*, 605–617.
- Ojanpää, H., & Näsänen, R. (2003). Utilisation of spatial frequency information in face search. *Vision Research*, *43*, 2505–2515.
- Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, *34*, 72–107.
- Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology*, *41*, 176–210.
- Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial

- envelope. *International Journal of Computer Vision*, 42, 145–175.
- Ozgen, E., Payne, H. E., Sowden, P. T., & Schyns, P. G. (2006). Retinotopic sensitisation to spatial scale: Evidence for flexible spatial frequency processing in scene perception. *Vision Research*, 46, 1108–1119.
- Ozgen, E., Sowden, P. T., Schyns, P. G., & Daoutis, C. (2005). Top-down attentional modulation of spatial frequency processing in scene perception. *Visual Cognition*, 12, 925–937.
- Parker, D. M., & Costen, N. P. (1999). One extreme or the other or perhaps the golden mean? Issues of spatial resolution in face processing. *Current Psychology*, 18, 118–127.
- Peyrin, C., Schwartz, S., Seghier, M., Michel, C., Landis, T., & Vuilleumier, P. (2005). Hemispheric specialization of human inferior temporal cortex during coarse-to-fine and fine-to-coarse analysis of natural visual scenes. *Neuroimage*, 28, 464–473.
- Reddi, B. A., Asrress, K. N., & Carpenter, R. H. (2003). Accuracy, information, and response time in a saccadic decision task. *Journal of Neurophysiology*, 90, 3538–3546.
- Romo, R., & Salinas, E. (2003). Flutter discrimination: Neural codes, perception, memory and decision making. *Nature Reviews Neuroscience*, 4, 203–218.
- Rotshtein, P., Vuilleumier, P., Winston, J., Driver, J., & Dolan, R. (2007). Distinct and convergent visual processing of high and low spatial frequency information in faces. *Cerebral Cortex*, 17, 2713–2724.
- Ruiz-Soler, M., & Beltran, F. S. (2006). Face perception: An integrative review of the role of spatial frequencies. *Psychology Research*, 70, 273–292.
- Schyns, P. G. (1998). Diagnostic recognition: Task constraints, object information, and their interactions. 147–179.
- Schyns, P. G., & Gosselin, F. (2003). Diagnostic use of scale information for componential and holistic recognition. In M. A. Peterson & G. Rhodes (Eds.), *Perception of faces, objects, and scenes: Analytic and holistic processes* (pp. 120–148). New York: Oxford University Press.
- Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, 5, 195–200.
- 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.
- Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, 69, 243–265.
- Shipp, S. (2001). Corticopulvinar connections of areas V5, V4, and V3 in the macaque monkeys: A dual model of retinal and cortical topographic. *Journal of Comparative Neurology*, 439, 469–490.
- Shipp, S., & Zeki, S. (1995). Segregation and convergence of specialised pathways in macaque monkey visual cortex. *Journal of Anatomy*, 187, 547–562.
- Skottun, B. C. (2000). The magnocellular deficit theory of dyslexia: The evidence from contrast sensitivity. *Vision Research*, 40, 111–127.
- Sowden, P. T., Ozgen, E., Schyns, P. G., & Daoutis, C. (2003). Expectancy effects on spatial frequency processing. *Vision Research*, 43, 2759–2772.
- Sowden, P. T., & Schyns, P. G. (2006). Channel surfing in the visual brain. *Trends in Cognitive Science*, 10, 538–545.
- Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modelling of elementary psychological processes*. Cambridge, UK: Cambridge University Press.
- Ungerleider, L. G., & Haxby, J. V. (1994). “What” and “where” in the human brain. *Current Opinion in Neurobiology*, 4, 157–165.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, 6, 624–631.
- Winston, J. S., Vuilleumier, P., & Dolan, R. J. (2003). Effects of low-spatial frequency components of fearful faces on fusiform cortex activity. *Current Biology*, 13, 1824–1829.