Spatial frequencies in short-term memory for faces: A test of three frequency-dependent hypotheses

MICHAEL J. WENGER
University of Notre Dame, Notre Dame, Indiana

and

JAMES T. TOWNSEND
Indiana University, Bloomington, Indiana

Spatial frequency manipulations have been shown to have significant utility in ascertaining the various types of information that might be important in object identification and recognition tasks. This utility suggests that, given some mapping between ranges of spatial frequencies and different types of psychological information, it should be possible to examine the roles of these different types of psychological information by way of spatial frequency manipulations. One potential problem, however, is that there is no well-specified, unambiguous mapping between the distinctions in the frequency domain and the distinctions in the informational domain. Three experiments provide tests of three general hypotheses regarding the ways in which different spatial frequencies might map to different information in facial perception and memory tasks: (1) the low-frequency dominance hypothesis, which proposes that low-frequency information should be superior (to high-frequency information) as a cue to perception and memory; (2) the distinct informational roles hypothesis, which holds that high spatial frequencies should carry featural information while low spatial frequencies should carry information about the configuration of those features; and (3) the task-dependent information hypothesis, which suggests that high-frequency information should be best suited to discrimination tasks while low-frequency information should be best suited for recognition tasks. Results generally contradict the first two of these hypotheses, while providing support for the third. Implications with regard to the various issues related to the mapping between spatial frequencies and the informational content of faces, as well as the need to consider important interactions among perceptual and memory processes, are discussed.

Nixon's nose and Hitchcock's profile are potent icons, and their effectiveness as cues in recognition and identification seems to suggest that specific features like the nose, as well as overall shape characteristics, are important in facial perception and memory (see, e.g., Bruce, 1988; Bruce & Young, 1986). Yet it is notoriously difficult to assay the relative importance, much less the informational coding, of such aspects of stimulus information in facial

cognition or, indeed, in object cognition in general. However, over the past 50 years, beginning in the late 1940s (Schade, 1956; Selwyn, 1948) and later with the seminal work of Campbell and Robson (1968), spatial frequency analysis has seen widespread application as a means of segregating components of stimulus information. Notable examples of the utility of spatial frequency analyses can be found in the numerous studies of object identification and recognition (e.g., Costen, Parker, & Craw, 1996; Ginsburg, 1986; Harmon & Julesz, 1973; Schyns & Oliva, 1994; Uttal, 1988; Uttal, Baruch, & Allen, 1995a, 1995b, 1997). A critical strength of the manipulations used in these types of studies is the ability to experimentally manipulate and thus examine the separate contributions of ranges of spatial frequency information in a stimulus.[1]

A possible application of this ability is by way of the suggestion that perception of the more detailed and local aspects of the scene (what are commonly referred to as its *features*[2]) may depend preferentially on the high-frequency components. In addition, it has also been suggested that perception of the configural aspects of the scene may depend preferentially on the low-frequency components. The latter could be referred to as *configural*

or *Gestalt characteristics* of the image. In particular, Ginsburg (e.g., 1978, 1986) has championed the view that such Gestalt relationships are captured almost entirely by the low-frequency channels in the visual system. Although in principle configural characteristics need not be confined to the low frequencies, particularly if configuration can be assessed by the geometric relations among the constituent features, it is not unreasonable to think that configural information in some way requires those attributes of the stimulus patterns that have been hypothesized to be preferentially carried by the lower spatial frequencies. Indeed, this notion has been stated, at least in heuristic form, in the literature on face perception and memory, (Bruce, 1988; Harvey, 1986; Sergent, 1984, 1989) and appears in contemporary textbooks (e.g., Goldstein, 1996). This hypothesis is compatible with, although not identical to, the notion that configural processing implies that either limited or no featural decomposition occurs (see also discussions in Ellison & Massaro, 1997; Tanaka & Farah, 1991, 1993; Tanaka & Sengco, 1997).

Yet this heuristic, in spite of its wide use (see reviews in Uttal, 1988; Uttal et al., 1995a, 1995b, 1997), runs into some distinct problems. First, and possibly most important for present purposes, the validity and coherence of the mapping between ranges of spatial frequencies and those aspects of the stimulus that support performance indicative of configural, holistic, featural, and so on, processing is compromised by a lack of definitional precision with respect to the latter constructs. For example, consider the proposal that low spatial frequencies transmit configural information. Now further consider the possibility that

such configural information may be operationally defined in terms of (or functionally equivalent to) the geometric relations among elements within the stimulus. If these proposals are valid, then a change in the relative locations of elements within the stimulus (e.g., flaring of the nostrils, widening of the mouth or eyes, etc.) would, by these assumptions, constitute configural changes—changes that by the hypothesis of the preceding paragraph should be best detected on the basis of information in the low spatial frequencies. However, in this scenario, the high spatial frequencies (which should, by extension of the hypothesis of the preceding paragraph, preferentially capture local or featural information) may be well suited to allowing detection of such a change (see, Parker, Lishman, & Hughes, 1996) because of their ability to support the evaluation of distances between feature edges. The result of this application of the heuristic is, thus, a contradiction. Second, the heuristic oversimplifies the distinction between global and local processing (see, e.g., Hughes, Nozawa, & Kitterle, 1996; Lamb & Yund, 1993, 1996). Third, in some applications it overlooks the degree to which various spatial frequency ranges might function to support performance in task-specific ways (see, e.g., Uttal et al., 1995a, 1995b, 1997).[3] One goal of the present effort was to document some of the critical problems (including the definitional problems associated with terms such as *configural*, *relational*, *featural*, *global*, *local*, etc.) associated with three forms of the heuristic represented schematically in Figure 1.

To begin, consider the possibility that all or even most face perception, and later memory, might be driven by con-
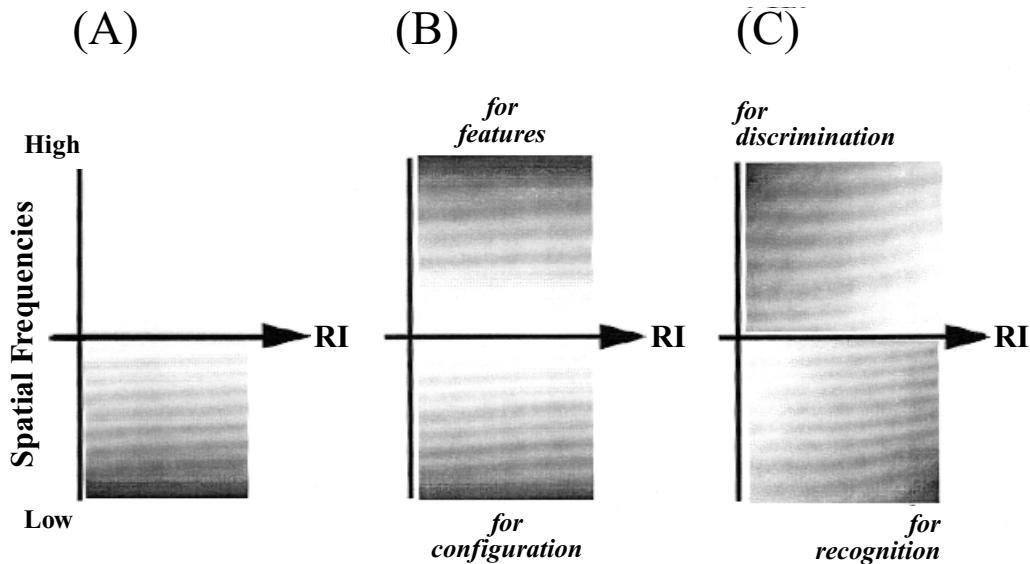
Figure 1. Schematic representations of the three versions of the heuristic relating ranges of spatial frequency to different types of psychological information: **(A)** low-frequency dominance hypothesis, **(B)** distinct informational roles hypothesis, **(C)** task-dependent information hypothesis. Darkness of the shading indicates the importance of the particular frequency range at different points in the retention interval (RI), for different aspects of the stimulus (panel B), and different tasks (panel C).

figural information. If configural characteristics were associated with low frequencies, as one version of the heuristic would suggest, then the immediate inference would be that low-frequency processing would be responsible for the bulk of face perception and, probably, memory as well. The possibility that much of real-world visual perception is governed by low frequencies is bolstered by the fact that spectral power of the various frequencies in visual scenes is proportional to $1/f$ (see, e.g., Sergent, 1989). That is, low frequencies dominate the high frequencies in the kinds of things we actually look at, for any sector or scale of an image. If this is so, then performance on tasks requiring use of this information (where sector or scale has been suitably controlled, or possibly matched, from initial exposure to later recognition test), such as a decision about whether a presented face has been seen before or a judgment about the similarity of that face to one seen previously should indicate a clear and general (i.e., task-independent) dominance of the low-frequency information (see panel A of Figure 1). For instance, if a target face (either in the task display or in memory) must be matched against a distractor or identical face that has been filtered to pass either the high or low frequencies, then the latter should be more effective, independent of the task and independent of the retention interval (RI). We will refer to this assertion—the first of three we test in the experiments below—as the *low frequency dominance* version of the heuristic.

An alternative version of the heuristic is the hypothesis that the higher frequencies are primarily responsible for encoding featural information whereas the lower frequencies capture configural information. In this case, the relative importance of the particular frequency range should be a function of the type of information that is available at either encoding (initial perception) or later retrieval. For example, according to this version of the heuristic, one would expect that featural aspects should survive a high-pass filtering of the test stimuli, leading to the prediction that if test items involve a change to features (as we are using the term here), then they should be accurately rejected in a recognition task and accurately be judged to have low similarity to the original target, independent of the length of the RI (see panel B of Figure 1). Lowpass filtering of these types of stimuli would support much lower levels of recognition accuracy and higher (less accurate) levels of judged similarity, with this effect again being independent of the length of the RI. However, lowpass filtered stimuli should support accurate rejection and low perceived similarity for probes involving a change in the configuration of features. We refer to this second hypothesis as the *distinct informational roles* version of the heuristic.[4]

Another version of the heuristic, one that can include both the low-frequency dominance and distinct informational roles hypotheses, is the suggestion that there is some "critical band" of frequencies that handles the bulk of face processing. This concept is potentially more difficult to falsify than the two hypotheses it includes. For example, the critical band could be defined such that it encompasses particular frequencies that might otherwise be labeled as "low" or "high." Indeed, the results of careful attempts to decide this issue are somewhat mixed (but see Costen et al., 1996, for a recent examination and update of this approach). The present work cannot rule out all band intervals as particularly suited to faces, but the regular, yet not simplistic, empirical results that follow will suggest that a strong version of such a hypothesis may have limited generality.

Despite the continuing exertions to find simple solutions in terms of categorical commitment to certain frequency bands in face cognition, there is a growing sense that distinct spatial frequency ranges may be called upon preferentially in different cognitive tasks. This general idea represents the third of the three versions of the heuristic that we consider and will be referred to as the *task-dependent information* hypothesis. One of the first studies to strongly suggest such a possibility was that of Sergent (1986); lower spatial frequencies appeared to support age and sex discrimination but not identification. Since then, other papers supporting the notion that the particular pattern of results may reflect critical task dependencies have appeared (e.g., Fiorentini, Maffei, & Sandini, 1983; Hayes, Morrone, & Burr, 1986). For example, Abdi, Valentin, Edelman, and O'Toole (1995) noted that in the principal component analysis approach to segregating different types of information in a facial stimulus, the largest valued eigenvectors of an image (those having a correspondence to low-frequency information) may support general category-level decisions, whereas the lowest valued eigenvectors (those having a correspondence to high-frequency information) may support decisions specific to particular previously learned stimuli.[5] And, notably pertinent to our study, Uttal et al. (1995a, 1995b, 1997) performed experiments whose analyses decisively pointed to low-frequency information being critical in recognition and high-frequency information being most important in discrimination tasks.[6]

This type of finding suggests that, unlike the two preceding hypotheses, the task-dependent information hypothesis may imply a dependency on the length of the RI (Figure 1, panel C). For example, when two stimuli (target and test probe) are presented simultaneously, the RI is effectively 0, and the task is one that we would commonly think of as discrimination. In contrast, when the two stimuli are separated in time, the RI is greater than 0 and the task is commonly one that we would think of as recognition. We should note that this simple example conflates RI and task instruction; we clarify and adhere to a more precise distinction as our presentation develops below.

Thus, this third hypothesis is distinct from the preceding two in a critical way, one that places important bounds on the use of spatial frequency manipulations in examining issues of facial perception and memory. That is, evidence for task dependencies such as those just noted points to a critical issue that can be lost when one bases experimental exploration solely on aspects of stimulus in-

formation. Specifically, task dependency highlights the need to consider both aspects of the stimulus information and how that information is used. Moreover, the importance of considering both of these aspects has been critical in revealing the incompleteness of a variety of simple dichotomies in the memory literature (e.g., Blaxton, 1989; Weldon & Coyote, 1996; Weldon, Roediger, Beitel, & Johnston, 1995) and it is at the foundation of suggestions that memory performance is a function of the interaction of presented (perceptual) information and retained (mnemonic) information (e.g., "synergistic ecphory," Tulving, 1983).

The interaction of aspects of stimulus information and aspects of task requirements was manipulated in the experiments presented here. Specifically, manipulations of both RI and response instruction were used to construct response situations that were most similar to either typical discrimination or recognition tasks. For the former, we used a 0-sec RI and instructions requiring the comparison of the simultaneously present target and probe faces. For the latter, we used 1- and 20-sec RIs and instructions requiring a yes/no recognition response. In addition, we included instructions for comparison at these two longer RIs. The general prediction derived from the task-dependent information hypothesis was that if, as indicated by the results of Sergent (1984, 1989) and Uttal et al. (1995a, 1995b) high-frequency information is best suited to support discrimination while low-frequency information is best suited to support recognition, then performance with stimuli that preserve the low-frequency components should show an increasing advantage (relative to stimuli that preserve the high-frequency components) with increases in RI.

However, there are at least two factors that could influence the general predictions of each of these three hypotheses. First, there is a well-documented and somewhat intriguing interaction between target–distractor similarity and RI that could influence the relative salience of high- and low-frequency stimulus information. Generally, intuition suggests and numerous experimental results have documented that increasing RI will lead to lower levels of accuracy or reduced levels of precision in memory tasks (e.g., Brown, 1958; Nairne, 1991, 1992; Peterson & Peterson, 1959; Wixted & Ebbesen, 1991). However, with distractors and targets that have a high degree of perceived similarity, recognition performance has been shown to improve across short RIs (e.g., Read, Vokey, & Hammersley, 1990). Specifically, when targets and distractors are presented in the same perceptual modality, recognition performance can be reliably disrupted by highly similar distractors, with this effect being most pronounced at short RIs and falling off as the RI increases (Glenberg, 1984). This interaction between stimulus similarity and RI has been shown in verbal materials (Chandler, 1993) and with facial stimuli (Shepherd, Gibling, & Ellis, 1991). With adults, the interacting effects of similarity and RI have been shown to disrupt not only memory for item information (e.g., item recognition) but also memory for order information (Anderson, 1980). With faces, the effect can be

observed with adults (Shepherd et al., 1991), school-age children (Ellis & Flin, 1990), and even infants (Fagan, 1977; Valenza, Simion, Cassia, & Umiltà, 1996).

Thus, increases in target–distractor similarity can potentially function to increase the degree to which discrimination processes are required. In this case, it is possible that stimuli that preserve high-frequency information could show an advantage relative to stimuli that preserve low-frequency information even when the response instructions emphasize recognition. However, given the regularity of the interaction between target–distractor similarity and RI, it becomes an intriguing empirical question whether one would expect any advantage for high-frequency information (relative to low-frequency information) to be restricted to the shortest RIs. This mandates that some measure of perceived similarity be obtained at each of the RIs in each of the experiments.

A second factor that could alter the general prediction for superior performance with low-frequency information with increases in RI is the degree of perceptual match that might exist between the internal representation (itself a function of the characteristics of the originally presented stimulus) and the probe stimulus. In studies involving spatial frequency manipulations, it is common to present unfiltered target items and filtered test probes; this was the approach used in our Experiment 2. Although it has been documented that recognition is a task that appears to benefit preferentially from high-level or conceptually driven processing at encoding (e.g., Goshen-Gottstein & Moscovitch, 1995; Roediger, Weldon, & Challis, 1989), there is also evidence that the low-level perceptual match between study and test stimuli exerts an influence on test performance (e.g., Rajaram, 1996; Weldon et al., 1995). Consequently, it might be possible to obtain higher levels of performance in a recognition task with high-frequency probe stimuli if the originally encoded stimuli were also high-pass frequency filtered. This possibility was tested in Experiment 3.

Our general goals in the present investigation, then, were to explore the three versions of the heuristic that relates ranges of spatial frequencies to various aspects of psychological information: the low-frequency dominance, distinct informational roles, and task-dependent information hypotheses. Since each of these hypotheses concerns the use of specific ranges of spatial frequencies, it was necessary to have basic measures of discriminability (d') and perceived similarity for the stimulus sets in the absence of filtering. This was provided by Experiment 1. In order to test the low-frequency dominance and distinct informational roles hypotheses, it was necessary to have probe stimuli in which either low- or high-frequency ranges were preserved. This was done in both Experiments 2 and 3. Tests of the distinct informational roles hypothesis also required stimulus manipulations involving distractors formed by changing the featural properties alone, the configural properties (specifically, the physical relations among the constituent features) alone, both of these, or total distractors (i.e., completely novel faces). This ma-

nipulation corresponds to standard approaches to investigating the roles of featural and configural information in facial cognition (e.g., Tanaka & Farah, 1991, 1993).

We should note that critical to the tests of each of these hypotheses is the provision of an operational definition of high- and low-frequency bands. We obtained this definition by determining a midrange frequency cutoff that equated overall similarity ratings for specific stimuli across low and high levels of spatial frequencies. This was accomplished in the calibration experiment reported in the Appendix. This strategy was intended to perhaps offset the advantage that low frequencies may have in faces, mentioned earlier, and to more or less level the playing field in general. We should also note that one crucial dilemma has confronted all studies with which we are familiar that have attempted to answer the above and similar questions involving faces and frequencies (and many other studies in general perception as well). The dilemma we are referring to concerns the thorny distinction between absolute frequencies (e.g., cycles/centimeter) and relative frequencies (e.g., cycles/face). Adding to the perplexity is the idea that if configural information is important, we surely must be able to process it at different scales, within limits, and the same would hold for featural content. Of course, it could happen that the weighting of configural versus local features could change as the overall scale is varied. Since the present work was not designed to adjudicate this controversy, we report our filtering manipulations (see the Appendix) in both cycles/ degree and cycles/face.

GENERAL METHOD

In all three experiments reported here (along with the calibration study reported in the Appendix), we used a similar set of materials and procedures. Consequently, we report the general approach here and note specific modifications in the presentation of each experiment.

Participants

All participants were recruited from introductory psychology courses at Indiana University and received course credit for their participation. All participants reported normal or corrected-to-normal vision.

Materials

Stimuli were digitized black-and-white photographs of faces taken from an alumni yearbook. All faces were frontal views of young, Caucasian males and females and lacked any extraneous features (such as facial hair, glasses, ornamental jewelry, or hats). Stimuli were captured via a Sony CCD/RGB video camera and were digitized using the ATT Sable image manipulation software.

A total of 32 faces were used to construct 16 sets of stimuli. Each set of stimuli consisted of a target (unmodified) face, a distractor face (of a different individual of the same sex), and four modified faces. The four modified faces were constructed from the target and distractor faces by changing the features, the relative position of features (their configuration), or both the features and their relative position. The design of each stimulus set is summarized in Table 1, and examples of each stimulus type are presented in Figure 2.[7]

One modified stimulus in each set was constructed by swapping the mouth of the target face with that of its paired distractor face. Another two modified stimuli in each set were constructed by

Table 1
Design of Each Stimulus Set

Change in Feature Relations?	Change in Features?	
	Yes	No
Yes	new (distractor) (A)	mouth up (C)
	internal swap (B)	mouth down (D)
No	mouth swap (E)	target (F)

Note—See text for details. Letters in parentheses refer to the specific panels of Figure 2. Note that the change in configuration refers to change in the relative position of facial features.

changing only the configuration of a feature within the target face. The first was constructed by moving the mouth of the target face down 3 pixels and the second was constructed by moving the mouth of the target face up 3 pixels.[8] The final modified stimulus was constructed by swapping the internal portion of the target face (eyebrows, eyes, nose, and mouth) with that of the paired distractor face.

All stimuli were presented using a VGA display controlled by a PC-compatible microcomputer, and participants responded using the keyboard of this PC. Participants were seated at an unconstrained distance of approximately 100 cm from the screen. At this distance, stimuli subtended approximately 2.75º of visual angle for width and approximately 3.80º of visual angle for height. The filtering procedures used in the calibration study and in Experiments 2 and 3 are reported in the Appendix. Examples of the filtered stimuli are presented in Figure 3.

Design

The basic design for each experiment was a 5 (stimulus manipulation: target, configuration change, feature change, configuration and feature change, distractor; see Table 1) $\times$ 3 (retention interval: immediate, short, long) factorial, with both factors manipulated within participants. This design was supplemented with a between-participants manipulation of stimulus filtering in Experiments 2 and 3.

Procedure

Each trial in each experiment was initiated by participants and began with a .5-sec presentation of a fixation point in the center of the screen. The fixation point was replaced with the target stimulus, which remained on the screen for 1.5 sec in the short and long RI conditions. In the immediate condition, the target stimulus was simultaneously presented alongside one of the six possible probe stimuli for a total of 5 sec. Following presentation of the target in the short RI condition, the screen was blanked for 1 sec, after which one of the six possible probe stimuli was presented. In the long RI condition, a 20-sec period followed the presentation of the target stimulus, during which participants were shown a digit and were then required to count the number of occurrences of that digit in a random sequence of digits.[9] This distractor task was presented on the screen and replaced the target stimulus. Following the presentation of the last digit in the sequence, participants were required to provide the total number of occurrences of the target digit using the numeric keypad.

In the immediate condition, participants were required to rate the similarity of the probe stimulus to the target using a scale ranging from 0 (*completely dissimilar*) to 9 (*identical*). Ratings were made using the numeric keypad of the computer. There was no speed stress for this response; participants were instructed to rate the magnitude of the difference (if any) they perceived between the target and probe faces. Participants were asked to respond within 8 sec of stimulus onset (5-sec stimulus presentation and no more than 3-sec delay following).

Following the RI in the short and long RI conditions, the probe was presented and participants made two responses—a yes/no

(A)   (B)   (C)   (D)

(E)   (F)
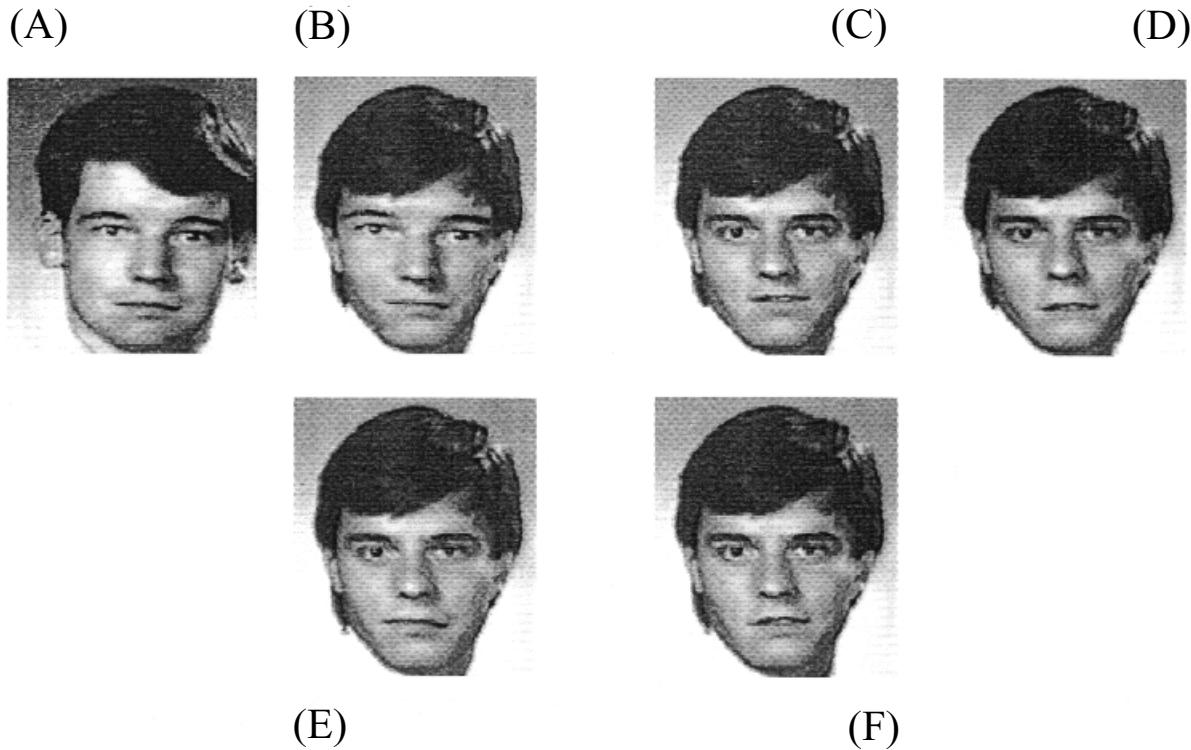
Figure 2. Examples of the various stimulus types (see text and Table 1 for details). Panel F presents a target face; panel A presents its accompanying distractor. Panels C and D present shifts (up or down, respectively) of the mouth of the target face. Panel E presents a swap of mouths between the target and distractor faces. Panel B presents a swap of the entire interior portions of the target and distractor faces.

recognition decision followed by a similarity rating. Participants were instructed to respond "yes" if the comparison stimulus was the same as the target stimulus and "no" if it was different. Participants responded using keys in the upper left and right corners of the keyboard for "yes" and "no" responses, respectively. Participants were instructed to respond as quickly and as accurately as possible; responses with latencies of greater than 3 sec were discarded prior to analyses. After making their recognition decision, participants in the

short and long RI conditions were asked to rate the similarity of the comparison and target stimuli using the procedure described above.

For each participant, eight sets of stimuli were randomly assigned to each of the short and long RI conditions. All 16 sets were presented to each participant in the immediate condition. Every possible pairing of a target with a comparison stimulus from the set was used. The short and long RI conditions were always performed first, with the specific ordering determined randomly for each participant.

(A)   (B)   (C)
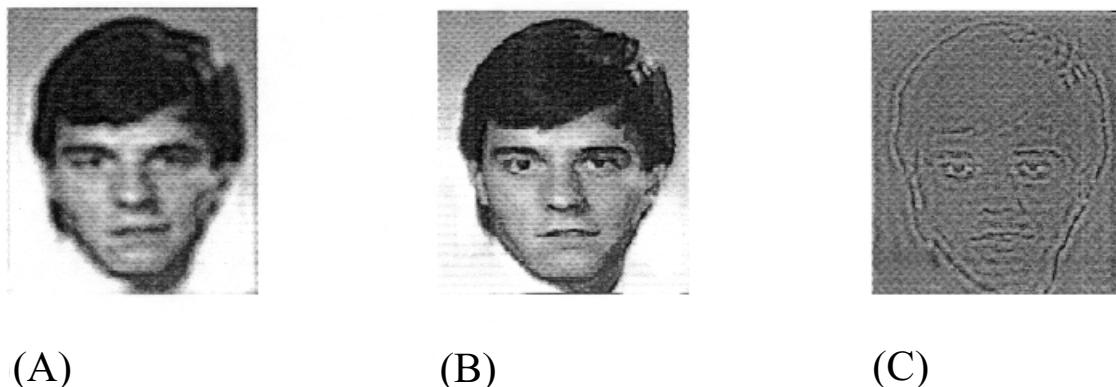
Figure 3. Examples of lowpass (panel A) and highpass (panel C) filtered stimuli, in comparison with unfiltered (panel B) stimuli.

Within each of the three RI conditions, order of pairs was randomized for each participant.

EXPERIMENT 1

Experiment 1 was conducted to provide a baseline for the effects of filtering used in Experiments 2 and 3.

Method

A total of 21 participants (14 females and 7 males) participated in Experiment 1. All stimuli used were unfiltered and all other aspects of the design and procedure were as described in the General Method.

Results

All results reported here and throughout the paper were significant at $p < .05$ unless otherwise noted. We begin with data relevant to perceived similarity and then consider discriminability.[10]

Figure 4 presents the mean similarity ratings obtained for each of the stimulus manipulations as a function of RI. These similarity ratings were analyzed using a 5 (stimulus manipulation: target, configuration change, feature change, configuration and feature change, distractor; see Table 1) $\times$ 3 (RI: immediate, short, long) mixed-factorial analysis of variance (ANOVA). As anticipated, this analysis revealed a main effect of stimulus manipulation [$F(4,80) = 295.78$, $MS_e = 2.04$]. Planned comparisons indicated that each of the probes was rated as less similar to the target than was the target stimulus. The main effect for RI was marginal [$F(2,40) = 2.67$, $MS_e = 0.84$, $p = .06$], with the highest similarity ratings being provided in the
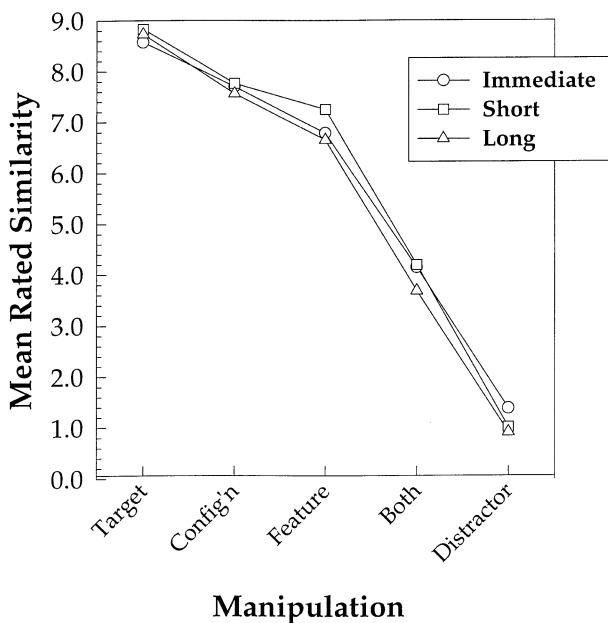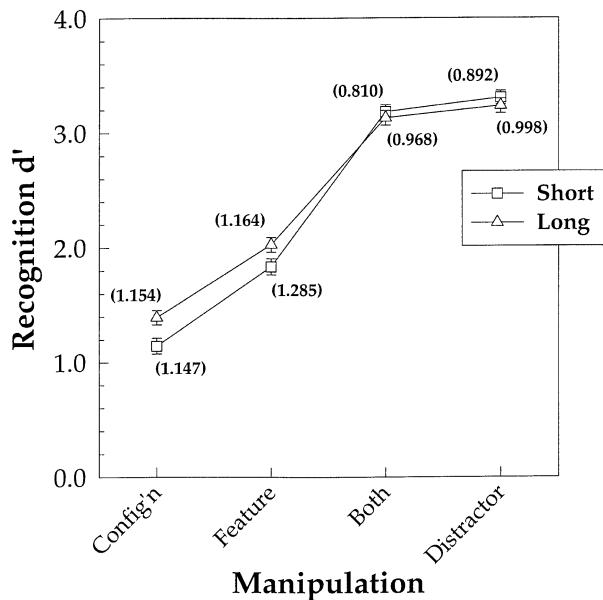

Figure 5. Experiment 1: Mean discriminability (in d') for each of the manipulations relative to the target stimulus as a function of retention interval. The numbers in parentheses next to each data point indicate the β value of for the plotted value of d'.

short RI condition (6.13), followed by the immediate (6.04) and the long (5.85) RI conditions. These higher ratings for the short RI condition were most pronounced for the feature manipulation (Figure 4).

On the basis of these similarity ratings, we anticipated that the discriminability of stimuli would be lower in the short RI than in the long RI conditions. Mean discriminability data (in d') for the target stimulus with respect to each of the manipulations are presented in Figure 5, with the values in parentheses next to the data points being the value of β associated with that of d'. Following Macmillan and Creelman (1991), we calculated 95% confidence intervals for these means (see also Gourevitch & Galanter, 1967; Miller, 1996), and used these confidence intervals as the basis for our analysis. As can be seen in Figure 5, discriminability improved as a function of RI for the configuration and feature manipulations, with no difference due to RI being observed for the stimuli involving a change in both configuration and features or for the distractor stimuli. Thus, impairment of performance due to high similarity at short RIs that has been observed in previous work with both faces and nonfacial stimuli (e.g., Chandler, 1993; Glenberg, 1984; Shepherd et al., 1991) was obtained in Experiment 1.

With these results as a baseline, we can now move to considering tests of the three versions of the heuristic relating spatial frequencies to psychological information in this task. Recall that if the low-frequency dominance hypothesis is correct, then lowpass filtering should result in lower similarity ratings and higher d's for all distractors, relative to highpass filtering, with these effects being in-

Figure 4. Experiment 1: Mean rated similarity for the various stimulus manipulations as a function of retention interval.

dependent of RI. If the distinct roles hypothesis is correct, then highpass filtering should produce the best levels of performance (highest levels of discriminability and lowest levels of rated similarity for distractors) when the stimuli involve featural changes, whereas lowpass filtering should produce the best levels of performance when the stimuli involve changes in featural configuration, with these effects again being independent of RI. Finally, if the task-dependent information hypothesis is correct, then the best levels of performance should be observed with highpass filtered stimuli when the task preferentially requires discrimination processes (e.g., at short RIs), whereas the best levels of performance with the lowpass filtered stimuli should be observed when the task preferentially requires recognition processes (e.g., at long RIs).

In order to perform the filtering manipulations, it was first necessary to determine a spatial frequency that would allow the spatial frequency spectrum to be divided into a high and low range. Our goal was to determine a frequency cutoff point that would allow for equivalent performance (as measured by immediate similarity ratings) when frequencies above or below were removed from stimuli. The Appendix reports in detail the method and results of the experiment conducted to determine this cutoff frequency. This cutoff frequency was then used in the construction of the stimuli involved in Experiments 2 and 3.

EXPERIMENT 2

Method

Materials. The 16 sets of stimuli used in Experiment 1 were used to create two new sets of stimuli consisting of 16 sets each. The first of the two sets was created by filtering the five distractors in each set

using a highpass filter with a cutoff frequency of 12 cycles/face (4.2 cycles/degree). The second was created by filtering the five distractors in each set using a lowpass filter with a cutoff frequency of 12 cycles/face (4.2 cycles/degree). The target stimuli in each set were not filtered. For reference, examples of the filtered stimuli are presented in Figure 3.

Participants, Design, and Procedure. A total of 24 participants (15 female, 9 male) were recruited from the source described in the General Method. The experiment was conducted as a 2 (filtering condition: lowpass, highpass) × 3 (RI: immediate, short, long) × 5 (stimulus manipulation: target, configuration change, feature change, configuration and feature change, distractor; see Table 1) mixed factorial, with filtering condition being a between-participants factor and RI and stimulus manipulation being within-participants factors. All aspects of the procedure were identical to those described in the General Method.

Results and Discussion

The mean similarity ratings given to each of the stimulus manipulations, as a function of both RI and filtering conditions, are presented in the three panels of Figure 6. These data[11] were analyzed using a 2 (filtering condition: lowpass, highpass) × 3 (RI: immediate, short, long) × 5 (stimulus manipulation) mixed-factorial ANOVA. Importantly, there was no main effect for filtering condition ($F < 1$), suggesting that our choice of cutoff frequency succeeded in equating similarity across the high- and low-frequency ranges. The analysis did reveal a main effect for stimulus manipulation [$F(4,88) = 61.92$, $MS_e = 1.36$]. Although Figure 6 (see the leftmost and rightmost panels) suggests that the stimuli involving either a feature or configuration change alone might have been rated as less similar to the target than was the target (as in Experiment 1), these differences were not significant (both $Fs < 1.50$, $ps > .20$). This lack of a significant difference may have been
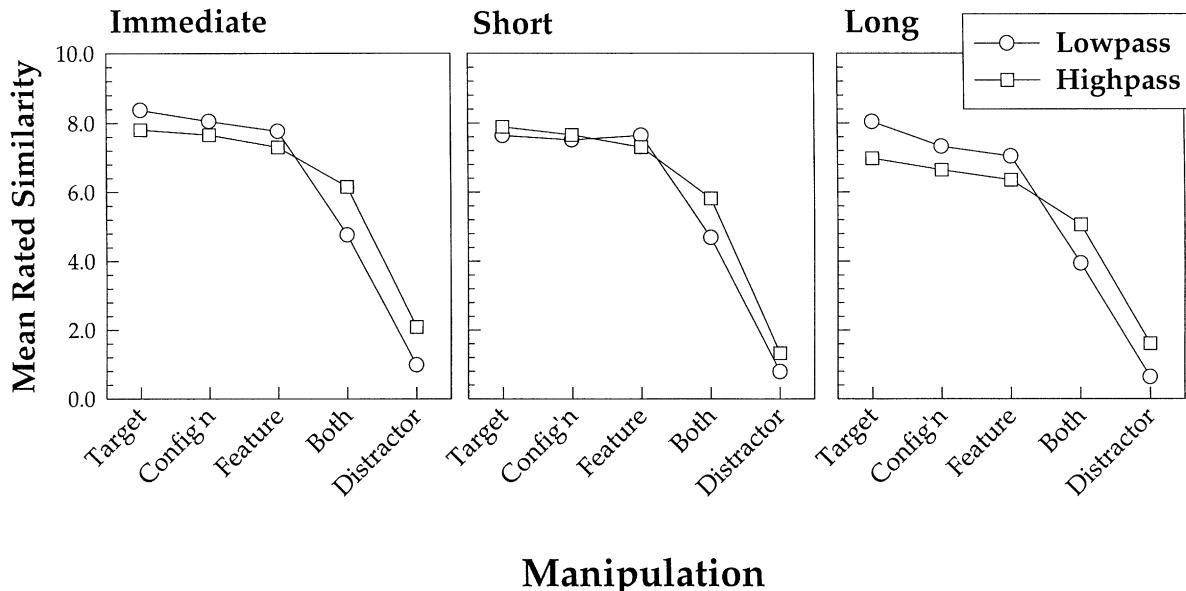

Figure 6. Experiment 2: Mean rated similarity for the various manipulations as a function of filtering conditions in the immediate (left panel), short (center panel), and long (right panel) retention interval conditions.

due in part to the fact that targets in Experiment 2 were given lower overall similarity ratings than were those in Experiment 1. (This, in turn, may be due to a lack of perceptual match between the test probes and the originally encoded, unfiltered, targets, a possibility we discuss below and test in Experiment 3.) In contrast, and consistent with the results of Experiment 1, the stimuli involving a change in both feature and configuration and the distractor stimuli were rated as less similar to the target than was the target [$F(1,22) = 40.87$, $MS_e = 11.74$, and $F(1,22) = 114.90$, $MS_e = 11.28$, respectively].

There was also an interaction between stimulus manipulation and filtering condition [$F(4,88) = 8.44$, $MS_e = 1.36$]. This interaction can be seen in the crossover of the curves representing the lowpass and highpass filtering data in each of the three panels of Figure 6. Essentially, the lowpass filtering condition produced higher similarity ratings (relative to the highpass filtering condition) for the stimuli involving featural or configural changes alone, but produced lower similarity ratings for the stimuli involving both types of change and for the distractor face. This crossover would seem to provide evidence contradicting the low-frequency dominance hypothesis, which would predict lower similarity ratings for all of the lowpass filtered stimuli relative to the highpass filtered stimuli.

According to the distinct informational roles hypothesis, one would expect that the lowpass filtered stimuli involving configural changes alone would be rated lower in similarity than the highpass filtered versions of these stimuli. In addition, one would expect that the highpass filtered stimuli involving featural changes would be rated lower than the lowpass filtered versions of these stimuli. As is apparent in Figure 6, the latter was true across RI conditions, but the former was not. The lowpass filtered stimuli involving a change in configuration (i.e., a change in the relative position of the anatomical features) were perceived as more similar than were the same stimuli when highpass filtered. Although it could be the case that the lowpass filtering introduced a unique source of noise, one would expect that this noise would be independent of type of stimulus manipulation, meaning that the observed crossover would not be predicted.

The similarity ratings do, however, appear to generally offer mixed of support for the task-dependent information hypothesis. If we assume that the task of providing a similarity rating is the same at each RI (though the memory information used may differ as a function of RI), then one would expect that highpass filtered stimuli would result in lower similarity ratings than would lowpass filtered stimuli. This is because this task was designed to be maximally dependent on discrimination processes that, according to the task-dependent information hypothesis, rely preferentially on high-frequency information. This pattern was observed for targets and for the configural and featural changes alone. It was not observed for distractors involving both types of changes or for new distractors. In addi-
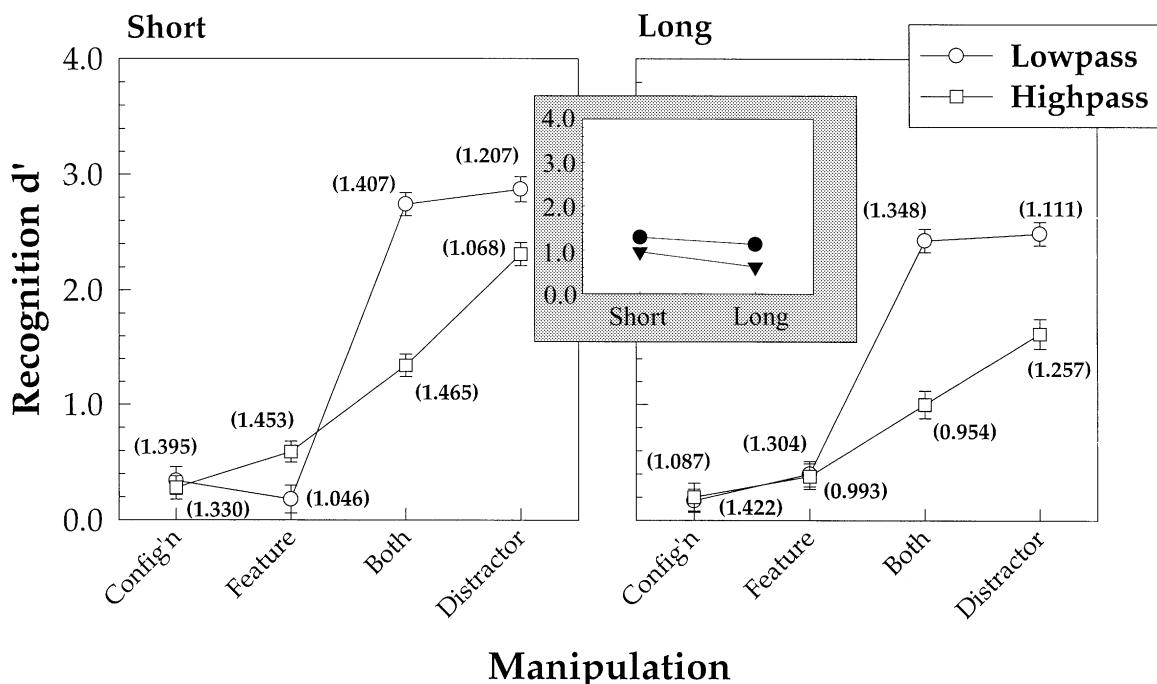
Figure 7. Experiment 2: Mean discriminability (in *d′*) for each of the manipulations relative to the target stimulus as a function of filtering condition in the short (left panel) and long (right panel) retention interval (RI) conditions. Error bars indicate 95% confidence intervals. The numbers in parentheses next to each data point indicate the value *β* of for the plotted value of *d′*. The inset plots mean *d′*s for the lowpass (circles) and highpass (inverted triangles) filtered stimuli at the two RIs.

tion, the advantage of highpass filtered stimuli that was present for the target, configural change, and featural change stimuli with simultaneous (0-sec RI) presentation was attenuated or absent at the 1-sec RI and reappeared at the 20-sec RI. This suggests that the interacting effects of RI and similarity in the recognition task might also be a factor in the similarity rating task. Still, we should note that at least for three stimulus types, the proposal that a task designed to be functionally similar to discrimination showed a marginal advantage for high-frequency information.

Figure 7 presents the discriminability data (in d') for each of the stimuli relative to the target in the 1-sec and 20-sec RI conditions. As before, we calculated 95% confidence intervals to guide our analysis of these data. Recall that the low-frequency dominance hypothesis predicts higher d's for all distractors in the lowpass filtering condition (relative to highpass filtered stimuli), whereas the distinct informational roles hypothesis predicts superior performance for the highpass filtered stimuli (relative to the lowpass filtered versions) involving featural changes and superior performance for the lowpass filtered stimuli (relative to the highpass filtered versions) involving configural changes. In addition, the task-dependent information hypothesis suggests that these differences, particularly the latter, should increase with increasing RI, as the task demands shift to become similar to those of a recognition task.

An initial observation of interest in the discriminability data presented in Figure 7 is the general agreement between these data and the similarity ratings presented in Figure 6. Note that higher similarity ratings for all distractors should correspond to lower discriminability, and indeed this is what we observed. In particular, note that probes involving only configuration or feature changes were given slightly higher similarity ratings in the lowpass filtering condition than in the highpass filtering condition (but with some qualification regarding the nonsignificant differences in the long RI condition). In contrast, probes involving both types of change (or distractors) were given lower similarity ratings in the lowpass filtering conditions than in the highpass filtering condition. A corresponding crossover (left panel) or separation (right panel) is apparent in the discriminability data, with a pronounced advantage in discriminability for the lowpass filtering condition relative to the highpass filtering condition for the probes involving both feature and configuration changes and for the distractors. This consistency between our two dependent variables is pleasing from a methodological standpoint and underscores the failure of both the low-frequency dominance and the distinct informational roles hypotheses.

It does appear that the highpass filtered stimuli involving featural changes were more discriminable than the corresponding lowpass filtered stimuli at the short RI (where the task could be functionally similar to a discrimination task), but this difference disappeared at the long RI. This would seem to be generally consistent with the task-dependent information hypothesis. Moreover, the superior discriminability of the lowpass filtered stimuli relative to the highpass filtered stimuli actually increased with increasing RI. For the lowpass filtered stimuli, overall d' was 1.29 in the short RI condition and 1.13 in the long RI condition, with the 95% confidence intervals for these means overlapping (see the inset to Figure 7). In contrast, for the highpass filtered stimuli, the overall d' was .96 in the short RI condition and .62 in the long RI condition, with no overlap in the 95% confidence intervals for these means. Thus, it appears that the ability of the highpass filtered stimuli to cue memory decreased as a function of increases in RI, whereas the ability of the lowpass filtered stimuli did not change appreciably across RI.

This apparent difference in the robustness of these two types of cues may be the source of the effect documented by Harmon and Julesz (1973) and is consistent with the findings and conclusions reported by Uttal et al. (1995a, 1995b, 1997). Although it is true that low-frequency information contains more power than high-frequency information (see Costen et al., 1996; Edwards, 1967), there is at present no principled way to predict how this difference would produce effects at different RIs. We return to this point in the General Discussion.

In sum, the results of Experiment 2 appear to generally contradict the predictions of both the low-frequency dominance and the distinct informational roles versions of the heuristic relating ranges of spatial frequency to different types of psychological information, while being generally consistent with the task-dependent information hypothesis (e.g., Uttal et al., 1995a, 1995b, 1997). Relative to the low-frequency dominance hypothesis, lowpass filtering did not produce generally lower levels of perceived similarity or higher levels of d' for distractors than did highpass filtering. Relative to the distinct informational roles hypothesis, it was not the case that featural manipulations were detected or discriminated from originally learned faces better with highpass filtered probes or that configural changes were detected or discriminated better with lowpass filtered probes. Instead, lowpass filtered stimuli seemed to permit higher similarity ratings when features or configuration were manipulated alone but supported lower (implying higher accuracy in d') similarity ratings when both features and configuration were changed. And in terms of discriminability, lowpass filtered stimuli appeared to be become more effective in their ability to probe memory with increases in RI.

Still, there remains the fact that highpass filtered stimuli did produce higher levels of performance (lower rated similarity and higher d') relative to the lowpass filtered stimuli, with the stimuli that involved only a change in features at the short RI. This is congruent with the distinct informational roles hypothesis and, in spite of the other evidence contradicting that hypothesis, requires additional exploration. There is at least one possible explanation for this superior performance that does not require accepting

the distinct informational roles hypothesis. Specifically, consider the fact that at the short RI, the test task closely resembles a discrimination task, in that a 1-sec unfilled RI was used. This RI is short enough, and faces are generally memorable enough, that participants might functionally treat the task at this RI as a discrimination task. Now consider the fact that whereas the probes were filtered, the targets were not. Thus, the highpass filtered stimuli could allow for the detection of the types of local discrepancies (present for both featural and configural manipulations) noted as critical in the performance of discrimination tasks by Uttal (1995a, 1995b). A similar alternative hypothesis is that the general degree of perceptual match between the studied stimulus and the comparison probe was lower in the highpass filtering condition for stimuli involving only one type of change. This would be similar to the notion of encoding specificity (see Tulving, 1983; Tulving & Thomson, 1973) and consistent with the findings of perceptual effects in the conceptually driven task of recognition (Rajaram, 1996; Weldon et al., 1995).

Fortunately, it is straightforward to test these alternatives. Specifically, if the general perceptual match between the originally studied item and the probe is responsible for the superior discriminability of the stimuli involving feature changes alone (at the short RI), rather than aspects of processing pertinent to the correspondence hypothesis, then filtering the studied stimuli (the targets) should remove this performance advantage. In contrast, if certain aspects of processing relevant to the correspondence hypothesis are of primary importance, then filtering the tar-

get stimuli should have no effect: The obtained performance advantage should remain.

EXPERIMENT 3

As noted, it is possible that the evidence obtained in Experiment 2 in partial support of the distinct informational roles hypothesis was due to the ability in the highpass filtered conditions to detect the lower degree of perceptual match, due to the fact that the targets were unfiltered whereas the probe stimuli were filtered. In order to check this possibility, targets in Experiment 3 were filtered using the same filters used for the probes in each condition.

Method
A total of 12 participants (8 females and 4 males) were recruited from the source described in the General Method. All other aspects of Experiment 3 were identical to those of Experiment 2, with one exception. In Experiment 3, targets were filtered at the same level as the corresponding probes.

Results and Discussion
The mean similarity ratings obtained for each of the manipulations in each of the filtering and RI conditions are presented in the three panels of Figure 8. These data[12] were analyzed using a 2 (filtering condition: lowpass, highpass) $\times$ 3 (RI: immediate, short, long) $\times$ 5 (stimulus manipulation) mixed-factorial ANOVA. In contrast to Experiment 2, this analysis did reveal a main effect for filtering condition [$F(1,10) = 4.91$, $MS_e = 7.60$], with mean similarity being rated higher in the highpass filtering con-
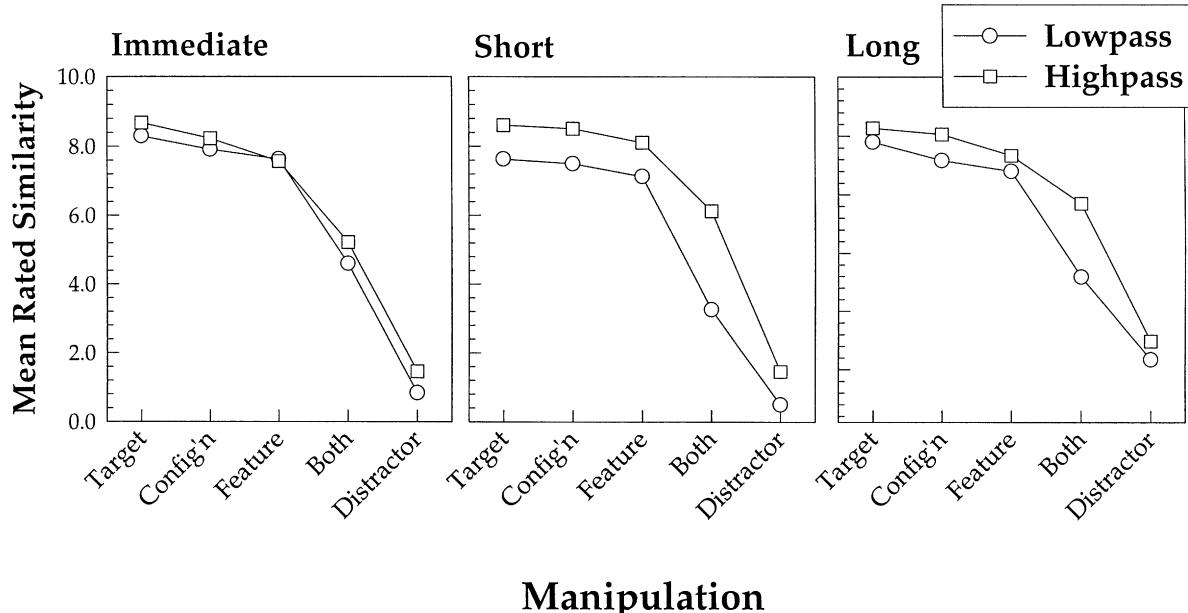

Figure 8. Experiment 3: Mean rated similarity for the various manipulations as a function of filtering conditions in the immediate (left panel), short (center panel), and long (right panel) retention interval conditions.

ditions (6.62) than in the lowpass filtering conditions (5.73). There was also a marginal main effect for RI [$F(2,20)$ = 2.71, MS_e = 1.61, p = .07], with rated similarity being higher in the immediate (6.38) and short (6.24) RI conditions than it was in the long RI condition (5.91). Finally, there was a main effect of stimulus manipulation [$F(4,40)$ = 2.146, MS_e = 11.77]. As in Experiment 2, and in contrast to Experiment 1, planned comparisons indicated that the stimuli that involved only featural or configural changes were rated as being as similar to the target as the target itself (both Fs < 1, ps > .49), while the stimuli that involved both featural and configural changes and distractor stimuli were rated as less similar to the target than the target itself [$F(1,10)$ = 18.41, MS_e =16.71, and $F(1,10)$ = 34.48, MS_e = 15.46, respectively].

With regard to the primary question, in contrast to Experiment 2, there were no crossovers in the similarity ratings: Highpass filtered stimuli were always rated as more similar than were lowpass filtered stimuli in each of the RI conditions. This would seem to indicate that increasing the perceptual match between targets and comparison stimuli had the most deleterious effect on the highpass filtered stimuli. We should note that this should not be in-

terpreted as support for the low-frequency dominance hypothesis, since that hypothesis is enunciated with respect to the ability of filtered probes to cue unfiltered targets. On the basis of these similarity data, then, we would expect that discriminability should be superior in the lowpass filtered conditions for all manipulations. If obtained, this would imply that the evidence for the distinct informational roles hypothesis obtained in Experiment 2 (with respect to the featural manipulation at the short RI) was due primarily to a lack of perceptual match.

The discriminability of each of the stimulus types (relative to the targets) in each of the RI and filtering conditions is presented in Figure 9, with error bars noting the 95% confidence interval for each plotted value. As is immediately apparent in this figure, there is no evidence for superior discriminability of the highpass filtered stimuli, for any of the manipulations at either of the RIs. Providing a perceptual match between the target and the comparison stimuli removed the observed superiority for the highpass filtered stimuli involving featural changes at the short RI observed in Experiment 2. This suggests that evidence that might have been interpreted as support for the distinct informational roles hypothesis was due to the lack
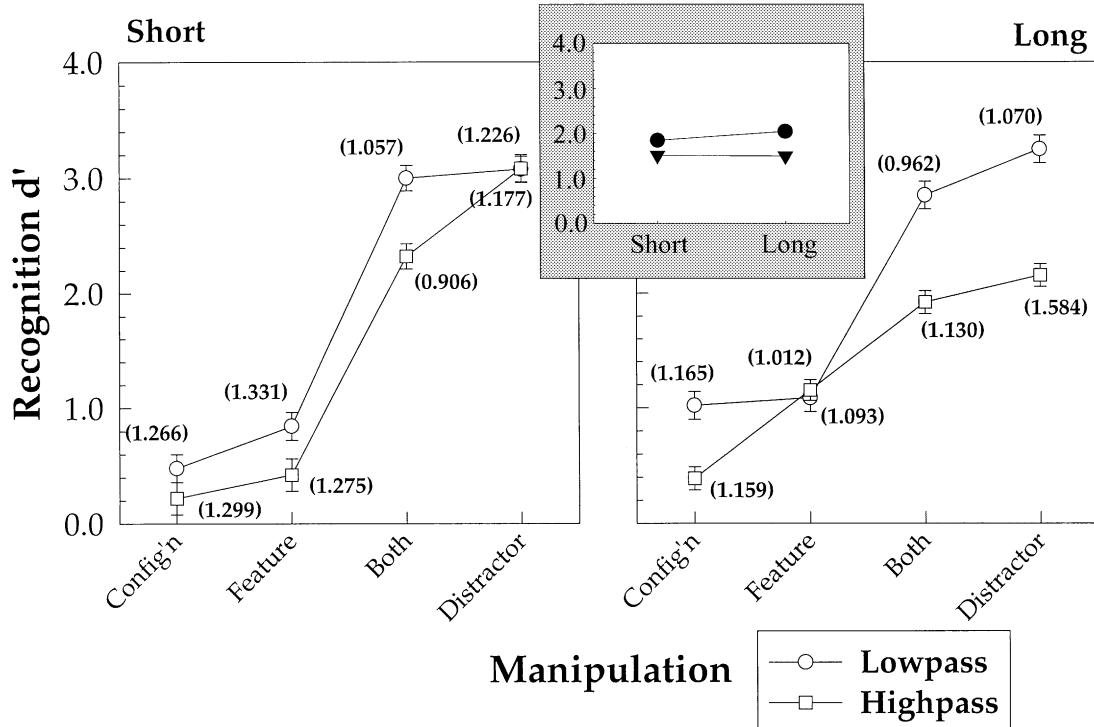
Figure 9. Experiment 3: Mean discriminability (in d') for each of the manipulations relative to the target stimulus as a function of filtering condition in the short (left panel) and long (right panel) retention interval (RI) conditions. Error bars indicate 95% confidence intervals. The numbers in parentheses next to each data point indicate the value of β for the plotted value of d'. The inset plots mean d's for the lowpass (circles) and highpass (inverted triangles) filtered stimuli at the two RIs.

of perceptual match that existed in Experiment 2 as a function of having the comparison stimuli filtered and the targets unfiltered.

For the short RI, discriminability for the lowpass filtered stimuli was higher than that for the highpass filtered stimuli for every manipulation except the distractor. (The confidence interval for the lowpass filtered stimulus involving a configural change extends down to .36 whereas the confidence interval for the corresponding highpass filtered stimulus extends up to .35, a difference that is impossible to discern in the figure.) For the long RI, the differences in discriminability between the lowpass and highpass filtered stimuli were significant for all manipulations (including the distractor) except the feature manipulation. Given the general patterns observed in these data, we are inclined to interpret this lack of a difference as an anomaly, subject to further evidence. Again, we should note that this cannot be unambiguously interpreted as support for the low-frequency dominance hypothesis, given the fact that both targets and probes were filtered.

Importantly, it appears that the superior discrimination performance obtained with the lowpass filtered stimuli was actually enhanced with increasing RI. The overall d's for the lowpass and highpass filtered stimuli at the short RI were 1.85 and 1.51, respectively, while the corresponding d's at the long RI were 2.05 and 1.50, respectively (see the inset in Figure 9). The overall increase for the lowpass filtered stimuli was significant, according to a lack of overlap of the appropriate confidence intervals. Essentially, the robustness of lowpass filtered stimuli as probes of memory, relative to highpass filtered stimuli, actually increased across the range of RIs used here, whereas the ability of highpass filtered probes to cue memory did not, even with a perceptual match between the targets and the probes. If the recognition task has a functional similarity to a discrimination task at the 1-sec RI, then this would provide additional evidence supporting the general predictions of the task-dependent information hypothesis (e.g., Uttal et al., 1995a, 1995b, 1997).

GENERAL DISCUSSION

The results of the three experiments reported here suggest three important conclusions. First, there was scant support from our experiments for either the low-frequency dominance or the distinct informational roles versions of the heuristic relating spatial frequency to psychological information. Second, it appears that low-frequency information in facial stimuli can reliably cue recognition memory; this cuing ability of low-frequency information (relative to high-frequency information) appears to be reasonably immune from or even enhanced by the effects of increasing RI. Finally, the interacting effects of similarity and RI observed in both face and non-face stimuli in certain previous studies was observed here as well. We consider each of these conclusions in turn.

The first of our major conclusions has to do with the lack of support in our data for either the low-frequency dominance or distinct informational roles hypotheses. Recall that the low-frequency dominance hypothesis (see panel A of Figure 1) predicts that low-frequency information should support performance to a greater degree than high-frequency information and that it should do so for all the stimulus manipulations considered here. We found little evidence to support this prediction. Experiment 2 provides the most direct information pertinent to this point, and as can be seen in Figures 6 (similarity ratings) and 7 (discriminability), lowpass filtered stimuli did not support lower overall levels of rated similarity (for nontargets) or higher overall levels of discriminability (relative to highpass filtered stimuli). Lower perceived similarity co-occurring with lowpass filtering did appear for two of the five stimulus types in the similarity ratings and for two of the four stimulus types in the discriminability data. Although the superior levels of performance observed for the lowpass filtered stimuli in these cases may be due to differences in the power contained in the stimulus as a function of filtering (see, e.g., Costen et al., 1996; Edwards, 1967), this hypothesis would not explain the differences that we observed as a function of RI (see below).

The distinct informational roles hypothesis (see panel B of Figure 1) predicts that high-frequency information should preferentially support the use of facial features whereas low-frequency information should preferentially support the use of their configuration. In the two experiments in which filtering manipulations were employed, we found only one bit of evidence supporting this prediction: This was the observation (in Experiment 2) of lower rated similarity and higher discriminability for probes involving only featural manipulations in the highpass filtering condition relative to the lowpass filtering condition. However, results of Experiment 3 strongly suggest that this evidence supporting the correspondence hypothesis may have been due to a simple degradation of the perceptual match between the target and comparison stimuli. This is because filtering the targets and probes at the same levels in Experiment 3 removed the advantage for the highpass filtering condition (relative to the lowpass filtering condition) for the probes involving only a change in features (observed in Experiment 2). In addition, lowpass filtering did not differentially support higher levels of performance with a configuration change in Experiment 2 in terms of either perceived similarity or discriminability. Thus, the results obtained in these two experiments fail to confirm either the low-frequency dominance or the distinct informational roles hypothesis.

It is appropriate to ask at this point whether, in fact, the low-frequency dominance and distinct informational roles hypotheses are in any way falsifiable. Indeed, there is a level of discourse at which either is tautologically true. That is, for a given retinal size of a face, if features are defined, as they usually are, in terms of finer grained geo-

metrical aspects (e.g., lip size), then it follows that such information must be associated with higher spatial frequencies (though not necessarily processed by associated filters) than, say, coarser properties such as distance between the ears, or relational aspects that rely on widely scattered information. Nevertheless, a hypothesis of such a dichotomy in terms of specific frequency bands in cycles per degree (i.e., considered as a psychological processing hypothesis) runs into trouble with different viewing distances and therefore retinal sizes of the entire object. On the other hand, such critical bands figured in terms of cycles per face, although perhaps not as theoretically problematic, demand justification via some type of computational model as to how, say, spatial frequency analyzers might turn their mechanisms into this kind of relativistic operator. These problems being noted, however, it is still the case that it is both important and possible to ask questions about the relative importance of high- and low-frequency information in tasks such as those presented here. And, in that context, the low-frequency dominance hypothesis predicted a powerful and task-independent superiority for low frequency information in natural facial patterns, whereas the distinct informational roles hypothesis predicted uniform roles for high- and low-frequency information across conditions (manipulations) and tasks. Both of these predictions were decisively falsified.

With regard to the distinct informational roles hypothesis, a reviewer suggested the possibility that the configural change (involving the movement of the mouth; see Table 1 and Figure 2) could be viewed as occurring at about the same scale as the featural change. If configurality requires a larger scale of change than featural processing or information, then perhaps one would not predict a difference in the results obtained for our configural change and a similar featural change. The configural manipulation certainly altered the relation between our operationally (and intuitively) defined features.

Suppose the conjecture is true. Then one would predict that our small, local configural change would produce effects very similar to those obtained with a change to the feature (the mouth). Yet the data (particularly the recognition d' data) suggest that this was not the case. It appears that moving the mouth alters the relation between it (and various combinations of it with other features) and other features in the face and with global aspects such as anatomical physiognomy (e.g., head shape, etc.). Thus, it is entirely possible—and perhaps probable, if those who stress the importance of such configural relations are correct—that even our relatively small configural alterations engendered much more massive changes on a configural level (see also note 8). Also, the truth or falsity of the hypothesis in no way impacts the validity or import of our third major conclusion, to be discussed shortly. Finally, we should note that there are a number of examples in the literature of manipulations labeled or interpreted as configural that are equivalent in type and magnitude to those used here (e.g., Bruce, Doyle, Dench, & Burton, 1991;

Haig, 1984; Hosie, Ellis, & Haig, 1988; Tanaka & Sengco, 1997).

On a related note, it must be acknowledged and emphasized that our design cannot answer the question of whether there exists *any* critical spatial frequency band for faces. For instance, the band of frequencies pointed to by Costen et al. (1996) overlaps the high as well as the low frequencies in our experiments, as defined by cycles per face, but lies entirely in the low end of the spectrum when the frequencies are defined by cycles per degree. Still, we must admit to some skepticism regarding the hypothesis of a critical band. That is because the various principled relationships between psychophysical and cognitive factors observed here would seem to suggest that such a hypothesis might be relevant only for the lowest level sensory phenomena. That is, such a hypothesis may be assayable in the simplest of psychophysical tasks, but apparently can be overshadowed by pervasive cognitive influences at other, higher, levels of task complexity.

The experimental results did reveal evidence pertinent to the second of our three major conclusions—specifically, the relative efficacy of low frequencies in facial recognition. That is, the data exhibited a striking robustness of low-frequency information (relative to high-frequency information) as a probe of recognition memory, with this robustness being either static in, or even improved by, the effects of increasing RI. This conclusion finds support in the comparison of discriminability across the three experiments (Figures 5, 7, and 9). As can be seen in Figure 5 (the unfiltered stimuli), the effect of RI on discriminability was to reduce the overall level slightly, but to maintain the general shape of the functional relationship between type of manipulation and level of performance. Looking now at Figures 7 and 9, one can see that the same finding is obtained for the lowpass filtered stimuli, both without (Experiment 2) and with (Experiment 3) a match in the level of filtering between the targets and probes. In contrast, highpass filtered stimuli showed pronounced deleterious effects of RI, far greater than those obtained when the stimuli were lowpass filtered or unfiltered.

Note that this finding does not a contradict our conclusion regarding the falsification of the low-frequency dominance hypothesis (our first major conclusion, above). That hypothesis holds that low-frequency information should optimally support performance, independent of task. Our finding suggests that low-frequency information may be more robust with respect to the effects of RI but solely in the context of a recognition task. Examination of the similarity data from Experiments 2 (Figure 6) and 3 (Figure 8) shows that low-frequency information did not show either a consistent advantage over high-frequency information in similarity ratings or increasing advantage with increasing RI. In fact, the data from Experiment 2 (Figure 6) suggest that the advantage for low-frequency over high-frequency information was greatest at the 1-sec RI and actually diminished as the RI was increased to 20 sec. Consequently, the robustness of low-frequency in-

formation was specific to the recognition task, a prediction that is at odds with the task-independent predictions of the low-frequency dominance hypothesis.

This robustness of low-frequency information across RI could well have been a contributing element to the phenomenon originally studied by Harmon and Julesz (1973). In that oft-cited study, a picture of Abraham Lincoln was degraded, by first blocking (or quantizing) and then removing selected high-frequency components. The noteworthy finding from that study was that recognition seemed, paradoxically, to be improved when the second degradation was combined with the first. Our finding of a robustness for low-frequency information in a recognition task is compatible with the observations of Uttal et al. (1995a, 1995b, 1997), who noted that the ability of low-frequency information to support performance was dependent on the degree to which the task resembles recognition rather than discrimination. For the experiments reported here, that task resemblance was a function of task type and RI and, in both Experiments 2 and 3, the effect of RI was to increase the superiority of lowpass filtered stimuli relative to highpass filtered stimuli. This interaction of perceptual information (the spatial frequency ranges) with memory information (revealed as a function of RI) returns us to the point made earlier. Specifically, our findings reinforce that it is critical to consider both the psychophysical and cognitive factors that are at work in tasks such as ours.

In fact, this last statement brings us to the third of our major conclusions. One of our critical manipulations involved shifting response demands between a task requiring discrimination and a task requiring recognition. On this dimension, our study parallels (in spirit) the explorations of Uttal et al. (1995a, 1995b, 1997) and, in fact, our results are quite coherent with the findings from those and other studies (Fiorentini et al., 1983; Hayes et al., 1986). Relevant to this finding, we noted in the introduction that the detrimental impact of high levels of similarity on performance is reduced with increases in RI. This has been observed when the target and probe stimuli are presented in the same modality (Glenberg, 1984) and has been observed with both facial and nonfacial stimuli. Our findings demonstrate this same relationship: For the two manipulations that were rated as most similar to the target (those involving featural or configural changes alone), the effect of RI was to increase discriminability. For the two manipulations that were rated as least similar to the target (the distractor and the stimuli that involved both featural and configural changes), there was little if any effect of increasing RI. As might be inferred from Figure 5, there may have been little need (or room) for any improvement with these stimuli.

The evidence presented here suggests to us that the processing of faces (or, to be precise, the processing of pictures of faces) appears to be subject to the same constraints and effects as other stimuli within the general literature. Recent work on other aspects of facial cognition (e.g., Campbell & Massaro, 1997; Ellison & Massaro, 1997) produced similar conclusions, suggesting that the hypothesis that faces are privileged stimuli (e.g., Bruce, 1988; Diamond & Carey, 1986; Farah, 1992) deserves additional and more precise attention. Finally, we believe that the useful confluence of results from both the perceptual and the memorial literature points to the need and utility for more top-down and bottom-up integration of facts and theory in both perception and memory.

REFERENCES

ABDI, H., VALENTIN, D., EDELMAN, B., & O'TOOLE, A. J. (1995). More about the difference between men and women: Evidence from linear neural networks and the principal-component approach. *Perception*, **24**, 539-562.

ANDERSON, N. S. (1980). Coding strategies in short-term memory. In F. L. Denmark (Ed.), *Psychology, the leading edge* (Annals of the New York Academy of Sciences, Vol. 340, pp. 1-8). New York: New York Academy of Sciences.

BLAXTON, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 657-668.

BRACEWELL, R. N. (1986). *The Fourier transform and its applications*. New York: McGraw-Hill.

BROWN, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, **10**, 12-21.

BRUCE, V. (1988). *Recognizing faces*. London: Erlbaum.

BRUCE, V., DOYLE, T., DENCH, N., & BURTON, M. (1991). Remembering facial configurations. *Cognition*, **38**, 109-144.

BRUCE, V., & YOUNG, A. (1986). Understanding face recognition. *British Journal of Psychology*, **77**, 305-327.

BUSEY, T. A., & LOFTUS, G. R. (1994). Sensory and cognitive components of visual information acquisition. *Psychological Review*, **101**, 446-469.

CAMPBELL, C. S., & MASSARO, D. W. (1997). Visible speech perception: Influence of spatial quantization. *Perception*, **26**, 627-644.

CAMPBELL, F. W. C., & ROBSON, J. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, **197**, 551-566.

CHANDLER, C. C. (1993). Accessing related events increases retroactive interference in a matching recognition task. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 967-974.

COSTEN, N. P., PARKER, D. M., & CRAW, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception Psychophysics*, **58**, 602-612.

DIAMOND, R., & CAREY, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, **115**, 107-117.

DILLON, R. F., & REID, L. S. (1969). Short term memory as a function of information processing during the retention interval. *Journal of Experimental Psychology*, **81**, 261-269.

EDWARDS, R. E. (1967). *Fourier series: A modern introduction*. New York: Holt, Rinehart & Winston.

ELLIS, H. D., & FLIN, R. H. (1990). Encoding and storage effects in 7-year-olds' and 10-year-olds' memory for faces. *British Journal of Developmental Psychology*, **8**, 77-91.

ELLISON, J. W., & MASSARO, D. W. (1997). Featural evaluation, integration, and judgment of facial affect. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 213-226.

FAGAN, J. F. (1977). Infant recognition memory: Studies in forgetting. *Child Development*, **48**, 68-78.

FARAH, M. J. (1992). Is an object an object an object? Cognitive and neuropsychological investigations of domain specificity in visual object recognition. *Current Directions in Psychological Science*, **1**, 164-169.

FIORENTINI, A., MAFFEI, L., & SANDINI, G. (1983). The role of high spatial frequencies in face perception. *Perception*, **12**, 195-201.

GINSBURG, A. P. (1978). *Visual information processing based on spatial filters constrained by biological data* (Tech. rep.). Dayton, OH: Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base.

GINSBURG, A. P. (1986). Spatial filtering and visual form perception. In K. R. Boff & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 2: Cognitive processes and performance* (pp. 34.1-34.41). New York: Wiley.

GLENBERG, A. M. (1984). A retrieval account of the long-term modality effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 16-31.

GOLDSTEIN, E. B. (1996). *Sensation and perception*. Pacific Grove, CA: Brooks/Cole.

GOSHEN-GOTTSTEIN, Y., & MOSCOVITCH, M. (1995). Repetition priming for newly formed and preexisting associations: Perceptual and conceptual influences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1229-1248.

GOUREVITCH, V., & GALANTER, E. (1967). A significance test for one parameter isosensitivity functions. *Psychometrika*, **32**, 25-33.

GRAHAM, N. V. S. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

HAIG, N. D. (1984). The effect of feature displacement on face recognition. *Perception*, **13**, 104-109.

HARMON, L. D., & JULESZ, B. (1973). Masking in visual recognition: Effects of two-dimensional filtered noise. *Science*, **180**, 1194-1197.

HARVEY, L. O. (1986). Visual memory: What is remembered? In F. Klix H. Hagendorf (Eds.), *Human memory and cognitive abilities* (pp. 180-206). Amsterdam: Elsevier.

HAYES, T., MORRONE, M. C., & BURR, D. C. (1986). Recognition of positive and negative bandpass-filtered images. *Perception*, **15**, 595-602.

HOSIE, J. A., ELLIS, H. D., & HAIG, N. D. (1988). The effect of feature displacement on the perception of well-known faces. *Perception*, **17**, 461-474.

HUGHES, H. C., NOZAWA, G., & KITTERLE, F. (1996). Global precedence, spatial frequency channels, and the statistics of natural images. *Journal of Cognitive Neuroscience*, **8**, 197-230.

LAMB, M. R., & YUND, E. W. (1993). The role of spatial frequency in the processing of hierarchically organized stimuli. *Perception Psychophysics*, **54**, 773-784.

LAMB, M. R., & YUND, E. W. (1996). Spatial frequency and interference between global and local levels of structure. *Visual Cognition*, **3**, 193-219.

LOFTUS, G. R., & RUTHRUFF, E. (1994). A theory of visual information acquisition and visual memory with special application to intensity–duration trade-offs. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 33-49.

MACMILLAN, N. A., & CREELMAN, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.

MASSARO, D. W., & LOFTUS, G. R. (1996). Sensory and perceptual storage: Data and theory. In E. L. Bjork & R. A. Bjork (Eds.), *Memory: Handbook of perception and cognition* (pp. 67-99). San Diego: Academic Press.

MCGILLEM, C. D., & COOPER, G. R. (1974). *Continuous and discrete signal and system analysis*. New York: Holt, Rinehart & Winston.

MILLER, J. (1996). The sampling distribution of d'. *Perception Psychophysics*, **58**, 65-72.

NAIRNE, J. S. (1991). Positional uncertainty in long-term memory. *Memory & Cognition*, **19**, 332-340.

NAIRNE, J. S. (1992). The loss of positional certainty in long-term memory. *Psychological Science*, **3**, 199-202.

PARKER, D. M., LISHMAN, J. R., & HUGHES, J. (1996). Role of coarse and fine spatial information in face and object processing. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 1448-1466.

PETERSON, L. R., & PETERSON, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, **58**, 193-198.

PETERSON, L. R., PETERSON, M. J., & MILLER, A. (1961). Short term retention and meaningfulness. *Canadian Journal of Psychology*, **15**, 143-147.

PINKER, S. (1985). *Visual cognition*. Cambridge, MA: MIT Press.

RAJARAM, S. (1996). Perceptual effects on remembering: Recollective processes in picture recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 365-378.

READ, J. D., VOKEY, J. R., & HAMMERSLEY, R. (1990). Changing photos of faces: Effects of exposure duration and photo similarity on recognition and the accuracy–confidence relationship. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 870-882.

ROEDIGER, H. L., III, WELDON, M. S., & CHALLIS, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger, III, & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3-41). Hillsdale, NJ: Erlbaum.

SCHADE, O. H. (1956). Optical and photoelectric analog of the eye. *Journal of the Optical Society of America*, **46**, 721-739.

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.

SELWYN, E. W. H. (1948). *Photographic Journal*, Sec. B, 88B.

SERGENT, J. (1984). An investigation of component and configural processes underlying face recognition. *British Journal of Psychology*, **75**, 221-242.

SERGENT, J. (1986). Methodological constraints on neuropsychological studies of face perception in normals. In R. Bruyer (Ed.), *The neuropsychology of face perception and facial expression* (pp. 69-81). Hillsdale, NJ: Erlbaum.

SERGENT, J. (1989). Structural processing of faces. In A. W. Young & H. D. Ellis (Eds.), *Handbook of research on face processing* (pp. 57-91). Amsterdam: Elsevier.

SHEPHERD, J. W., GIBLING, F., &. ELLIS, H. D. (1991). The effects of distinctiveness, presentation time, and delay on face recognition. *European Journal of Cognitive Psychology*, **3**, 137-145.

TANAKA, J. W., & FARAH, M. J. (1991). Second-order relational properties and the inversion effect: Testing a theory of face perception. *Perception Psychophysics*, **50**, 367-372.

TANAKA, J. W., & FARAH, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, **46**A, 225-245.

TANAKA, J. W., & SENGCO, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, **25**, 583-592.

TULVING, E. (1983). *Elements of episodic memory*. Oxford: Oxford University Press.

TULVING, E., & THOMSON, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, **80**, 352-373.

UTTAL, W. R. (1988). *On seeing forms*. Hillsdale, NJ: Erlbaum.

UTTAL, W. R., BARUCH, T., & ALLEN, L. (1995a). Combining image degradations in a recognition task. *Perception & Psychophysics*, **57**, 682-691.

UTTAL, W. R., BARUCH, T., & ALLEN, L. (1995b). The effect of combinations of image degradations in a discrimination task. *Perception & Psychophysics*, **57**, 668-681.

UTTAL, W. R., BARUCH, T., & ALLEN, L. (1997). A parametric study of face recognition when image degradations are combined. *Spatial Vision*, **11**, 179-204.

VALENZA, E., SIMION, F., CASSIA, M., & UMILTÀ, C. (1996). Face preference at birth. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 892-903.

WELDON, M. S., & COYOTE, K. C. (1996). Failure to find the picture superiority effect in implicit conceptual memory tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 670-686.

WELDON, M. S., ROEDIGER, H. L., III., BEITEL, D. A., & JOHNSTON, T. R. (1995). Perceptual and conceptual processes in implicit and explicit tests with picture fragment and word fragment cues. *Journal of Memory & Language*, **34**, 262-285.

WIXTED, J. T., & EBBESEN, E. B. (1991). On the form of forgetting. *Psychological Science*, **2**, 409-415.

NOTES

1. The theory of Fourier analysis provides the traditional means of implementing spatial frequency studies, and since the work presented below falls in this tradition, we offer some preliminary comments on the relation of the theory to empirical efforts such as the one described here.

It is worth noting that Fourier theory, now close to 200 years old, has had an impact in science that is difficult to exaggerate. As has frequently been the case, it received its rigorous mathematical underpinning only within the last 50 years. It is critical to understand, and is a point that is sometimes overlooked, that such mathematical theories are in and of themselves, devoid of empirical content. Only when converted into testable models of empirical phenomena do they gain that status (see, in particular, discussions by Pinker, 1985; Uttal, 1988). Fourier analysis is closely associated with linear systems theory (see, e.g., Bracewell, 1986; McGillem & Cooper, 1974). As an *empirical* model, the consequences of this connection are severe. Although linearity is accepted as a good approximation for near-threshold simple visual psychophysics (Graham, 1989), it may be suspect for higher processes (although, see Busey & Loftus, 1994; Loftus & Ruthruff, 1994; Massaro & Loftus, 1996). Nevertheless, it continues to influence interpretations of visual perception and cognition in a number of ways.

2. Note that by *features* we mean those gross anatomical aspects of the face, such as the nose, eyes, mouth, and so on, and by *configuration* (or *configural* or *relational*) we mean the general arrangement of those features as well as general aspects such as the shape of the facial surround. Although it is the case that the first term has been used as a referent for any information extracted from the stimulus pattern (see, e.g., Diamond & Carey, 1986) and that the second term has been used with equal generality (and lack of definition), we have chosen to adhere to the modal use of the terms in oder to aid in our illustration of the problems with the various hypotheses under consideration.

3. Our thanks to Denis Parker for reinforcing the importance of these issues.

4. Before continuing, we should note that Tanaka and Farah (1991, 1993) and Diamond and Carey (1986) discussed the notion that configural processing could be based on relations among features, possibly subsequent to early featural establishment. Even if this is the case, it is quite conceivable that the information requisite for relational properties could reside in the lower frequency data, whether or not the featural aspects themselves are strategically involved in face processing. Thus, the concept of second-order relational properties does not necessarily invalidate the above spatial frequency implications.

5. Abdi et al. (1995) have also discussed the relationship and differences between the principal component analysis approach and the spatial frequency analysis approach.

6. Schyns and Oliva (1994) have offered evidence for a similar perspective in scene perception, and Campbell and Massaro (1997) have suggested conclusions coherent with this idea in their work with perception of visible speech.

7. Although it is true that a change in facial features produces a change in configuration, we present these changes as independent for the purposes of exposition.

8. The magnitude of this configural change may seem small, putting it outside the limits of what one would consider to be a valid change to the "global" characteristics of the stimuli. However, two sources of evidence suggest otherwise. First, the magnitude of the change (3 pixels) is within the range of magnitudes labeled by other investigators as configural. For example, Bruce, Doyle, Dench, and Burton (1991) used configural shift manipulations as small as 3 pixels on images that were larger than those used in the present study. Hosie, Ellis, and Haig (1988) used a configural manipulation that involved moving the mouth up by 3 or 6 pixels. Finally, Tanaka and Sengco (1997) used a change in feature configuration involving a change of approximately 1 pixel /1° of visual angle (6 pixels on an image subtending 6° of visual angle horizontally), approximately the same magnitude of change used in our stimuli. Second, the 3-pixel change represents an increase or decrease of approximately 25% of the original distance between the center of the mouth and the bottom of the nose, a change certainly reflected in the spatial frequency characteristics of the altered images.

9. Although faces would seem to be stimuli that should be difficult if not impossible to rehearse, it is conceivable that participants could strategically rehearse an observation regarding some aspect of a target face (e.g., "big nose"). The intent of the distractor task was thus to be maximally effective in reducing the opportunity for rehearsal while being minimally interfering with respect to memory for the target faces (see, e.g., Dillon & Reid, 1969; Peterson, Peterson, & Miller, 1961).

10. Note that since participants were making both recognition and similarity judgments on the same stimuli in the short and long RI conditions, it was possible for them to provide inconsistent recognition and similarity judgments. For purposes of analyzing the recognition data, we defined a cross-error as a trial on which a participant gave a "yes" response to the recognition probe and a 0 to the accompanying similarity judgment, or a "no" response to the recognition probe and a 9 to the accompanying similarity judgment. These cross-errors were very infrequent in Experiment 1 (a total of 19 responses, 0.2% of the total responses, with no single participant making more than 2) and were excluded from all analyses.

11. Total cross-errors in Experiment 2 were 24, 0.3% of total responses, with no single participant making more than 3. As in Experiment 1, these cross-errors were excluded from all analyses.

12. Total cross-errors in Experiment 3 were 12, 0.3% of total responses, with no single participant making more than 2. As in the previous experiments, these cross-errors were excluded from all analyses.

APPENDIX

In this Appendix we describe the methods and results of the calibration experiment conducted to determine the cutoff frequency for the filtering used in Experiments 2 and 3.

Subjects, Materials, Design, and Procedure

A total of 12 participants (6 females and 6 males) were recruited from the source described in the General Method. Three of the 16 sets of stimuli described in the General Method were selected at random for this experiment. Those three sets (18 stimuli in all) were filtered through six different spatial frequency filters: three low-pass filters with cutoff frequencies of 2.8, 4.2, and 5.6 cycles/degree (4, 8, or 12 cycles/face), and three high-pass filters with cutoff frequencies of 1.4, 2.8, and 4.2 cycles/degree (8, 12, or 16 cycles/face). Spatial frequency filtering was accomplished using Fast Fourier Transforms and circular ideal filters.[A1] This produced six groups of 18 stimuli each. In addition, the unfiltered stimuli from each of the original three sets were used as control comparison stimuli, producing a total of 126 stimuli.

The experiment was conducted as a 7 (filtering condition: 3 lowpass, 3 highpass, unfiltered control) $\times$ 6 (probe type: see Table 1) factorial, with both factors manipulated within subjects. Since the critical dependent measure was rated similarly, the recognition task was not included. Details of procedure were identical to those in the immediate condition of Experiment 1.

Results

Mean similarity ratings for each probe type in each of the filtering conditions are presented in Table A1. Pairwise comparison of mean similarity ratings for each probe type across all the pairs of filtering conditions showed no differences for any of the probes as a function of filtering conditions when 4.2 cycles/degree was used as the cutoff frequency. Therefore, 4.2 cycles/degree was used as the cutoff frequency for defining the high- and low-frequency filtering conditions used in Experiments 2 and 3.

NOTE

A1. It has been noted (L. Thibos, personal communication, 1996) that to represent filtering as a convolution in the space or time domain would require a convolution kernel defined at an infinite number of points. Since this is impossible in practice, non-ideal filters are commonly used, allowing only a finite number of values in the convolution kernel. However, in the present work, images were transformed into the frequency domain, multiplied by the ideal lowpass filter, and transformed back into the space domain. Since frequency was implemented as multiplication in the frequency domain, the compromise required by the convolution representation is no longer needed.

Table A1
Mean Similarity Ratings as a Function of Probe Type for
Each of the Filtering Levels

| Probe | c/d c/f | Highpass Filtering* | | | Lowpass Filtering† | | | |
		2.80 8	4.20 12	5.60 16	1.40 4	2.80 8	4.20 12	Ctrl
Target		8.58	7.52	7.58	7.18	7.81	8.39	8.52
Mouth down		8.06	7.91	7.11	6.88	7.58	7.52	8.20
Mouth up		7.97	7.45	7.76	6.72	7.55	7.79	8.48
Mouth swap		7.50	7.28	7.00	6.58	7.06	7.27	7.17
Internal swap		4.16	5.31	5.39	4.12	4.78	4.67	4.31
Distractor		0.63	1.19	0.78	0.72	0.52	0.86	1.15

Note—Ctrl, control (unfiltered) stimuli; c/d, cycles per degree; c/f, cycles per face. *The values for the highpass filtering represent the lower cutoff; that is, frequencies at or above the reported level were preserved. †The values for the lowpass filtering represent the upper cutoff; that is, frequencies at or below the reported level were preserved.