Research Article

Short-Term Episodic Memory for Visual Textures

A Roving Probe Gathers Some Memory

Feng Zhou, Michael J. Kahana, and Robert Sekuler

Volen Center for Complex Systems, Brandeis University

ABSTRACT—*Cognition is shaped by the way that past experiences are represented in memory. To examine the representation of recent visual experiences, we devised a novel procedure that measures episodic recognition memory for synthetic textures. On each trial, two brief study stimuli were followed by a probe, which either replicated one of the study stimuli or differed in spatial frequency from both. The probe's spatial frequency roved from trial to trial, testing recognition with a range of differences between probe and study items. Repeated testing of recognition generated mnemometric functions, snapshots of memory strength's distribution. The distributional characteristics of the mnemometric functions rule out several hypotheses about memory representation, including the hypothesis that representations are prototypes constructed from previously seen stimuli; instead, stimuli are represented in memory as noisy exemplars.*

Important everyday tasks depend on the ability to recognize previously seen objects and events. In broad terms, this ability requires encoding and storing visual information for later retrieval and matching stored information against subsequent visual input. The nature and fidelity of the stored information has been the subject of continuing controversy, which we address with a novel approach that makes use of continuously varying probe items to characterize memory representation.

In order to characterize stimulus variables' influence over visual memory, we combined Sternberg's (1966) recognition procedure with a simple signal detection model. On each trial in Sternberg's task, a set, S, of n study stimuli is followed by a probe stimulus, P, which may or may not have been a member of S. P is characterized as a "target" if it had been a member of S, and is characterized as a "lure" otherwise. A subject judges whether P was or was not in S. The members of S vary from trial to trial, forcing subjects to use episodic information, that is, to compare P against some stored representation of the most recent S. We are especially interested in the character of that representation, which might comprise separate exemplars (e.g., Hintzman, 1988;

Address correspondence to Robert Sekuler, Volen Center, Mailstop 013, Brandeis University, Waltham, MA 02454; e-mail: vision@brandeis.edu.

Nosofsky, 1984), a prototype reflecting a weighted mean of items in S (Reed, 1972), or some combination of these alternatives. Our paradigm permits a choice among these possibilities; later, in the Discussion section, this choice is explained in more detail.

Sternberg's (1966) paradigm bridges sensory research, which emphasizes discriminal processes, and memory research, particularly research on recognition memory (Kahana & Sekuler, 2002). When the set size, n, is reduced to just one item, the paradigm reduces to the same/different procedure used to study sensory discrimination (Sternberg, 1998). Our stimuli were textures synthesized by summing horizontal and vertical sinusoidal gratings. These stimuli offer a number of advantages for studying memory (Kahana & Sekuler, 2002), including a reduced likelihood that verbal description will aid recognition (McIntosh et al., 1999), and an increased ability to manipulate interstimulus differences.

In the study we report here, on each trial, two brief, sequentially presented study textures, S_1 and S_2, were followed by a single probe texture, P. Samples are shown in Figure 1. The spatial frequencies present in S_1, S_2, and P varied trial by trial. Participants judged whether P's vertical spatial frequency matched that of either S_1 or S_2, or whether it differed from both. An important and novel aspect of our procedure is that from trial to trial the spatial frequency of the probe texture "roved," or varied in spatial frequency. The roving probe sampled memory strength at various points along the spatial-frequency continuum, sweeping out a probability function that is a unique snapshot of the distribution of memory strength. Such snapshots, which we call *mnemometric functions*, open windows onto the workings of memory. In particular, the central tendency and shape of mnemometric functions make it possible to test hypotheses about participants' representation of the stimuli.

After presenting the empirical results, we explain how mnemometric functions can be used to evaluate alternative hypotheses about the memorial representation of study stimuli. We then introduce a simple signal detection model of the visual and memory processes that are used in the experimental task.

METHOD

Participants

Participants were 10 paid volunteers whose ages ranged from 19 to 24 years. They had normal or corrected-to-normal Snellen acuity and
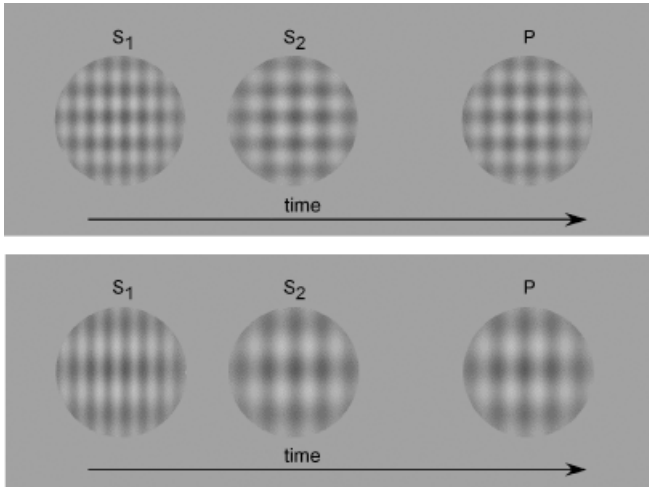
Fig. 1. Sample texture stimuli from two trials. In the upper panel, the probe stimulus (P) replicates neither of the study items, S_1 and S_2. In the lower panel, P replicates S_2. As noted in the text, horizontal spatial frequency varied randomly from trial to trial, but was constant within a trial.

normal contrast sensitivity as measured with Pelli-Robson charts (Pelli, Robson, & Wilkins, 1988).

Apparatus

Stimuli were generated and displayed using Matlab 5 and extensions from the Psychophysics and Video Toolboxes (Brainard, 1997; Pelli, 1997). Stimuli were presented on a 14-in. computer monitor with a screen resolution of 800 by 600 pixels, refreshed at 95 Hz. Mean screen luminance was maintained at $36 \, cd/m^2$.

Stimuli

In each stimulus, one vertical and one horizontal sinusoidal luminance grating were superimposed, generating a luminance profile, $L_{(x,y)}$, given by

$$L_{(x,y)} = L_{avg}[1 + a_1 \cos(\pi f x) + a_2 \cos(\pi g y)]$$

where L_{avg} is the mean luminance, f is the spatial frequency of the stimulus' vertical component (vertical frequency) in cycles per degree, g is the frequency of the horizontal component (horizontal frequency), and a_1 and a_2 are the Michelson contrasts of the two components. Each component's contrast was 0.4, which was well above threshold. The gratings subtended $6°$ of visual angle at the viewing distance, 82 cm. Stimuli were windowed by a circular two-dimensional Gaussian function, to eliminate edges.

Prior to memory testing, each participant's spatial-frequency discrimination was measured with an adaptive procedure (Wetherill & Levitt, 1965), which identified the spatial-frequency difference that permitted 79% correct discrimination between gratings. Participants' Weber fractions ranged from 0.06 to 0.12. Each participant's own Weber fraction was used to scale the stimuli for testing that participant's recognition memory. This minimized differences in visual encoding as a source of between-subjects differences in recognition performance.

During memory testing, the horizontal and vertical spatial frequencies of the study stimuli varied randomly over trials. This randomization minimized learning effects, and forced observers to base each trial's judgment on information obtained from that trial. Vertical frequencies were randomly drawn from a uniform distribution, which ranged from 0.5 to 3.0 cycles/deg; horizontal frequencies came from a slightly broader distribution, 0.0 (no variation along the horizontal) to 3.0 cycles/deg.

On each trial, the same randomly chosen spatial frequency was used for the horizontal component of all three gratings, S_1, S_2, and P; differences among the items were generated only by variation in their vertical frequencies. A new random phase shift was applied to the vertical component of each stimulus. These phase shifts, which ranged from 0 to $\pi/2$ radians, made it difficult to base recognition judgments on the spatial information contained in some residual, retinotopic image (Averbach & Coriell, 1961).

Within a block of trials, the difference between the study textures' vertical frequencies, $|S_1 - S_2|$, was set to 1, 2, 4, or 8 threshold units. (A threshold unit is the product of spatial frequency and an individual's Weber fraction.) The four values of $|S_1 - S_2|$ occurred randomly but equally often. On half the trials, S_1's vertical spatial frequency was higher than S_2's; on the remaining trials, the reverse was true. Moreover, on half the trials, P constituted a target stimulus; that is, it replicated either S_1 or S_2. Because a target stimulus was equally likely to replicate either one of the study items, observers had to attend to both. On the remaining trials, P constituted a lure; that is, it replicated neither S_1 nor S_2. On lure trials, P's vertical frequency was drawn from a uniform random distribution in the range of ± 6 threshold units relative to the mean of S_1's and S_2's vertical frequencies.

At the start of each trial, a fixation point was centered on the screen for 750 ms. The fixation point was followed by a 750-ms period of uniform luminance. S_1 and then S_2 followed, each for 1,000 ms. S_1 and S_2 were separated by an interstimulus interval of 1,000 ms, during which the display area was filled with uniform luminance. Finally, 1,000 ms after S_2, P was presented, remaining visible until the observer responded, but for no more than 1,000 ms.

Procedure

Participants served in four sessions of two 180-trial blocks each. Sessions were separated by between 24 hr and 72 hr. Each participant sat with head supported by a chin-and-forehead rest, viewing the computer display binocularly. A trial was initiated by the press of a key on the computer keyboard. Participants were instructed to respond as accurately and quickly as possible. By pressing computer keys representing "yes" and "no," the participants signaled their judgment whether or not P was identical to either S_1 or S_2 ("yes") or differed from both ("no"). The computer produced distinctly different tones after correct and incorrect responses, providing trial-wise knowledge of results.

RESULTS

We compared accuracy on the two types of probe-target trials: when P matched S_1 and when P matched S_2. Over all trials, recognition that P matched the second study item, S_2, was significantly higher than recognition that P matched the first study item, S_1; mean hit rates and standard errors of the mean were .61 (.01) and .73 (.01) for S_1 and S_2,
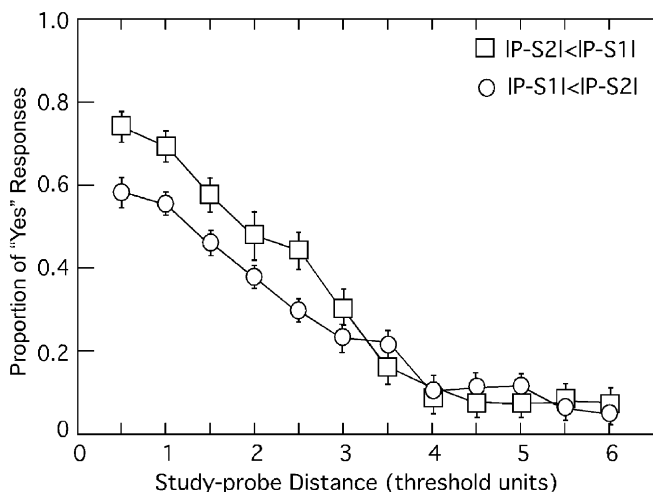
Fig. 2. Proportion of "yes" responses to lure trials (false alarms) as a function of the distance between the probe (P) and whichever study item (S_1 or S_2) was closer to the probe. Separate curves are shown for the case in which S_1 was the closer study item and the case in which S_2 was the closer study item. Values on the *x*-axis are threshold units. Error bars represent 95% confidence intervals (Loftus & Masson, 1994).

respectively, $t(9) = 4.52$, $p < .01$. This result is a recency effect, which is consistent with previous experiments using compound gratings (Kahana & Sekuler, 2002) and Sternberg-type experiments using verbal stimuli (Monsell, 1978).

For lure trials, parametric variation in P's spatial frequency relative to S_1 and S_2 generates information not otherwise available from recognition experiments. As an example, Figure 2 plots the proportion of incorrect "yes" responses (false alarms) to lures as a function of the difference between P's spatial frequency and the spatial frequency of whichever study item, S_1 or S_2, was closer to P's spatial frequency. Separate functions are shown for cases in which P more closely approximated S_1 and for cases in which P more closely approximated S_2. For either case, as P diverged from the more similar study item, the proportion of false alarms declined. The figure shows that for small study-probe distances, the rate of false alarms was higher when P was more similar to S_2 than when P was more similar to S_1. We return to this result later, in the context of a signal detection model.

To explore how interstimulus similarity influences episodic recognition, we tabulated hit and false alarm rates for various differences between P's vertical frequency and the vertical frequencies of S_1 and S_2. The resulting curves, which constitute mnemometric functions, are shown in Figure 3; empirical data are represented by filled and open symbols, for lures and targets, respectively. Each panel presents the mean mnemometric function for one signed value of $[S_1 - S_2]$. Note that P's spatial frequency has been normalized to zero relative to the mean of S_1's and S_2's spatial frequencies. The vertical axis in Figure 3 represents the proportion of "yes" responses, hits plus false alarms. The horizontal axis shows values of P, expressed in normalized threshold units. When a lure was very close to S_1 or S_2, represented by the vertical lines in each panel of the figure, the false alarm rate coincided well with the nearby hit rate. This coincidence suggests that under these conditions, the observers were likely using the same decision rule to generate responses to both lure and target probes.

The empirical mnemometric functions in Figure 3 allowed us to discriminate among alternative ways in which remembered spatial frequencies (S_1' and S_2') might relate to the frequencies actually presented (S_1 and S_2). For example, we wanted to determine whether memory involved some single prototype constructed from S_1 and S_2 or separate exemplars. In the following discussion, we illustrate how a simple signal detection approach can link empirical results with the character of representations in memory.

Imagine that the study series S consists of a single compound grating, S_1, and an observer must judge whether a subsequently presented probe, P, had been or had not been in S. Suppose that over trials, repetition of S_1 gives rise to a distribution of neural events. Suppose also that P and S_1 differ on only one stimulus dimension, say, vertical spatial frequency. S_1', the remembered value of S_1 at the time P is presented, can be written as $S_1' = S_1 + c_1 + \varepsilon_1$, where c_1 is a constant shift in remembered frequency, and ε_1 is a random error produced by additive noise in memory.

Figure 4a illustrates the probability distribution for S_1's representation in memory. In the case shown, $c_1 = 0$, and the probability density function of ε_1 is Gaussian. Figure 4b illustrates one plausible way that a response might be generated. Assume that P is presented at some spatial frequency, for example, the frequency pointed to by the arrow in the figure. To generate a response, an observer adopts a criterion, k, and when $|P - S_1'| < k$, the observer responds, "Yes, P was in the study series." When $|P - S_1'| > k$, the observer responds, "No, P was not in the study series." In the figure's lower panel, k is represented by the half width of the light-gray zone centered on the probe position. The proportion of "yes" responses, summing hits and false alarms, will vary with P's position relative to S_1. As the value of P varies over trials, the repeated calculation of proportion of "yes" responses sweeps out the curve that we designate a mnemometric function.[1] This account can be easily extended to study sets of more than one item.

With this description of recognition judgments and mnemometric functions, we can evaluate hypotheses about the relationship between the study stimuli and their representation in memory.

---

[1] In the text, mnemometric functions were motivated by a discrete sampling process represented by variation in P. More properly, mnemometric functions are generated by the process of convolution. In particular, with P positions distributed densely over spatial frequency, the function relating the proportion of "yes" responses to P's spatial frequency is given by the convolution of the observer's criterion and the probability density function for remembered values of the stimulus. Let both the distribution of memories, $\mathbf{m}$, and the criterion, $\mathbf{k}$, be vectors defined over values of spatial frequency, x. The equation for the convolution, $\mathbf{g}(x)$, of vector $\mathbf{m}(x)$ with the convolution kernel for the criterion, $\mathbf{k}(x)$, is

$$\mathbf{g}(x) = \mathbf{m}(x) * \mathbf{k}(x) = \sum_{\text{all-}k} \mathbf{k}(x - k)\mathbf{m}(k).$$

As is customary, values of the convolution kernel $\mathbf{k}$ are zero outside the region of support, which is the region (light gray bar in Fig. 4) in which kernel values are valid. The graph of $\mathbf{g}(x)$, the convolution output, constitutes a theoretical mnemometric function, which is a snapshot of the memory for S_1 averaged over many trials. As in all applications of convolution, linearity and shift invariance are assumed. Our application requires one other assumption, about the composition of $\mathbf{k}(x)$. If $\mathbf{k}(x)$ is constant over x, or if $\mathbf{k}(x)$ has a Gaussian distribution, so long as the values are not asymmetrical about the midpoint of $\mathbf{k}(x)$, the shape of $\mathbf{g}(x)$ will approximate the shape of $\mathbf{m}(x)$. As $\mathbf{k}(x)$ broadens, so too will $\mathbf{g}(x)$, but its shape will be preserved, as will the x value on which $\mathbf{g}(x)$ is centered.
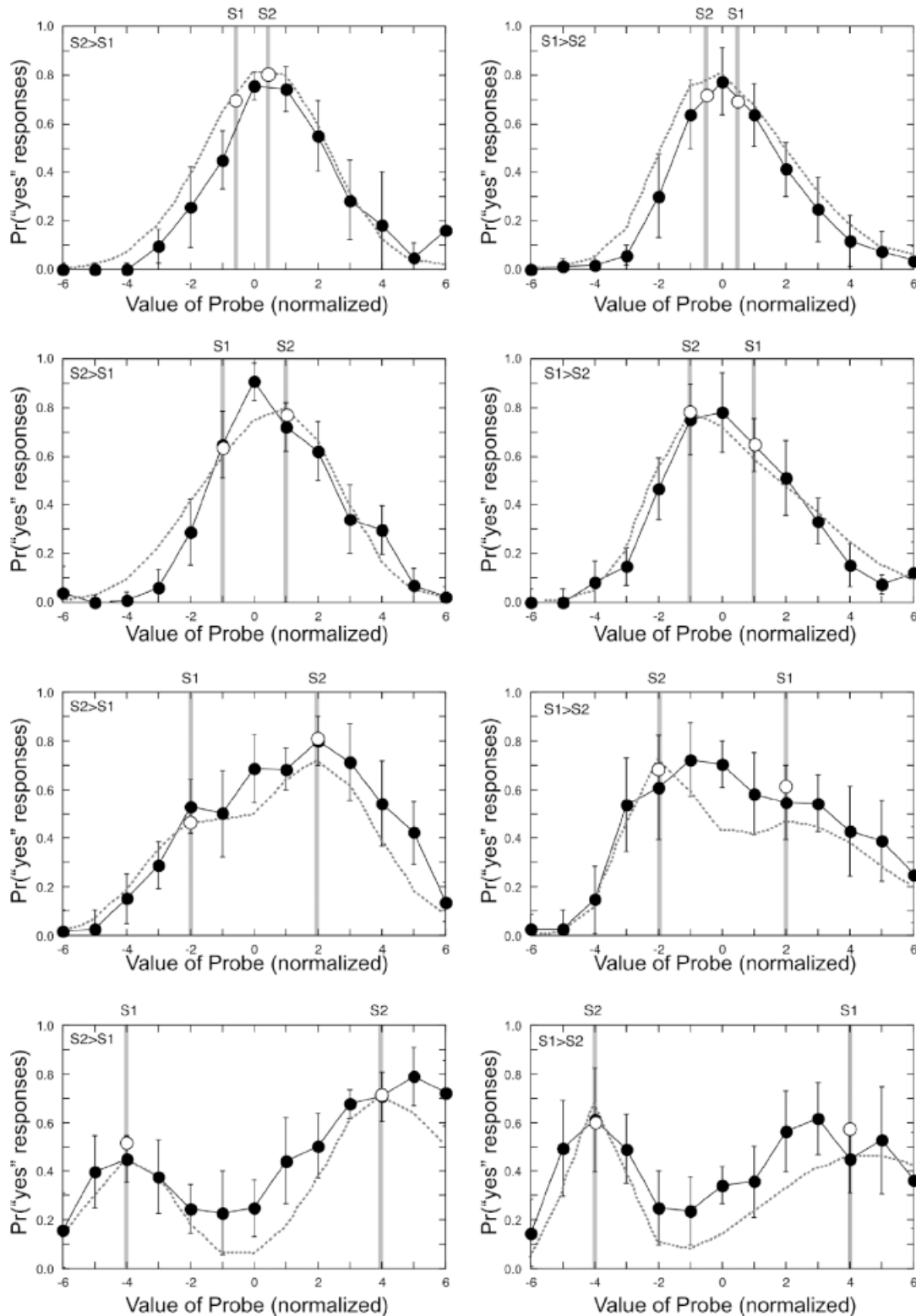
Fig. 3. Hit and false alarm rates as a function of spatial frequency of the probe stimulus (P). Data from each trial were normalized by setting the arithmetic mean of the frequencies of the study items (S_1 and S_2) to zero. P's position is plotted against multiples of threshold units. In the four graphs on the left, S_2's vertical spatial frequency was higher than S_1's; in the four graphs on the right, the reverse was true. In each panel, the normalized spatial frequencies of S_1 and S_2 are indicated by thick vertical lines. From top to bottom, $|S_1 - S_2| = 1$, 2, 4, and 8 threshold units, respectively. Results, represented by filled circles for lures and open circles for targets, were averaged across 10 subjects, and sorted into bins 1 threshold unit wide. Values predicted by the three-parameter model are shown by the dotted lines (parameters: $\varepsilon_1 = 2.0$, $\varepsilon_2 = 1.15$, and $k = 1.3$). Error bars show 95% confidence intervals.
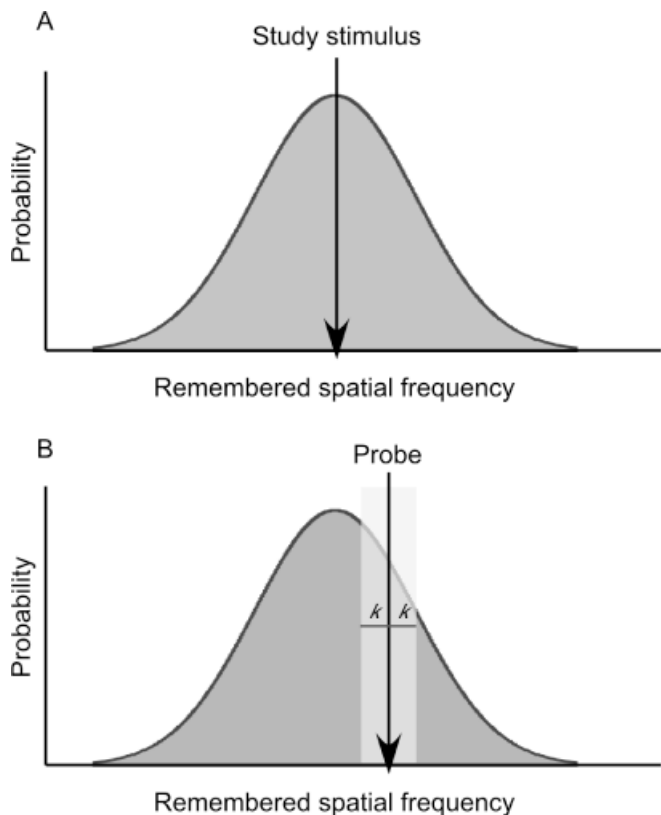
A Study stimulus

B Probe

Fig. 4. The generation of mnemometric functions. The graph in (a) shows a hypothetical distribution of remembered spatial frequencies generated by repeated presentation of a study stimulus whose spatial frequency is indicated by the arrow. The graph in (b) shows the presumed basis for deciding whether the probe matches the study stimulus. A probe is presented with spatial frequency represented by the arrow. The participant employs a criterion *k* units wide on either side of the probe's spatial frequency. If the remembered frequency on a trial falls within the criterion range (indicated by the width of the light gray vertical bar), the participant decides, "The probe matches the study stimulus."

Hypothesis 1: prototype: The remembered frequencies are represented on each trial by a single value, a prototype, whose distribution over trials is centered on the mean of S_1' and S_2'.

Hypothesis 2: shifted exemplars: The memory contains separate representations, S_1' and S_2', but these are systematically shifted relative to the presented frequencies, producing a constant error. Harvey (1986) reported such a result with long-term memory for scenes.

Hypothesis 3: unshifted exemplars: Memory comprises separate representations, S_1' and S_2', whose distributions over trials are centered on S_1 and S_2. This hypothesis is mute about the relative variability in the distributions for S_1' and S_2', which can be estimated from recognition judgments.

We devised a simple, three-parameter signal detection model to evaluate these hypotheses. Following signal detection theory, we treat the remembered spatial frequency of a study item (S') as a random deviate from a normal distribution centered on the spatial frequency of that study stimulus (S). This treatment of memory acknowledges that

Weber's law holds for spatial frequency, as demonstrated by Campbell, Nachmias, and Jukes (1970). Therefore, ε', the standard deviation of the distribution of S', is expressed as a proportion of S' (the coefficient of variation). On any trial, $S_1' \sim N(S_1, \varepsilon_1)$, and $S_2' \sim N(S_2, \varepsilon_2)$, where S' is a random sample from a Gaussian distribution, with a mean of S and a standard deviation of ε. For simplicity's sake, we assume that the covariance of S_1' and S_2' is zero.

The predicted response on any trial is determined by only three parameters, ε_1, ε_2, and *k*, the threshold criterion used in comparing P and S'. Over trials, the parameter set, ε_1, ε_2, and *k*, is constant, and S_1' and S_2' are samples from zero-mean Gaussian distributions. A decision is assumed to depend on the relationship between P and either S_1' or S_2'. When $|P - S_1'| < k$, or $|P - S_2'| < k$, the model predicts a "yes" response; otherwise, the model predicts a "no" response.

We fit the experimental data in Figure 3 using a downhill simplex search (Nelder & Mead, 1965). The resulting optimal parameters were $\varepsilon_1 = 2.0$, $\varepsilon_2 = 1.15$, and $k = 1.3$, $\chi^2(117) = 513.78$, $p > .90$. All parameter values are expressed in discrimination threshold units.

The optimal value for *k* was only 30% higher than the sensory discrimination threshold, suggesting that observers used a consistent criterion across the two types of judgments, discrimination and recognition. This agreement strengthens the claim that the Sternberg task can bridge sensory and memory domains.

Note that in the model's optimal parameter set, ε_1 is greater than ε_2, signifying that the variability of S_1' is greater than the variability of S_2'. This difference in parameters reflects the higher fidelity of memory for S_2 than for S_1. Figure 5 plots the obtained proportion of "yes" responses versus corresponding values predicted by the model. Using only three free parameters, the model captures 87% of the variance in the data.

Figure 3 shows the predicted values (dotted lines) superimposed on the empirical results (filled symbols for lures, open symbols for
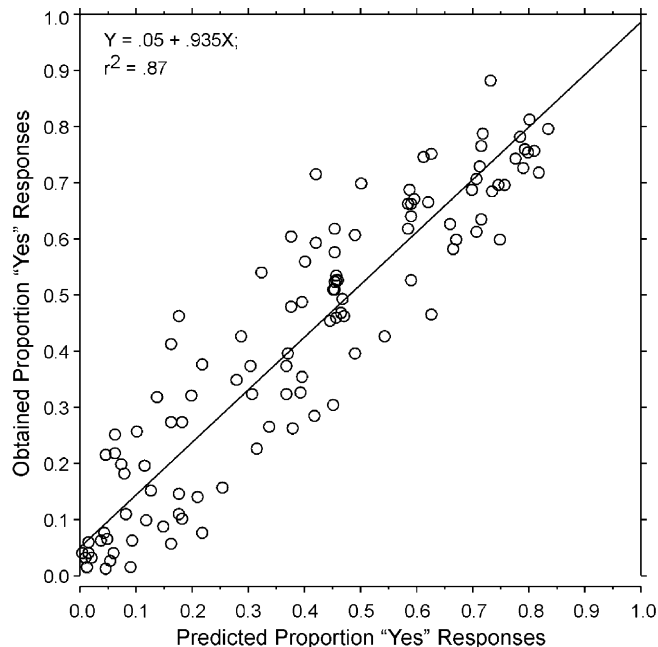

Y = .05 + .935X;
$r^2 = .87$

Fig. 5. Proportion of "yes" responses actually obtained plotted against the proportion predicted by the three-parameter model. Also shown is a line representing the empirical results regressed against the model's predictions. The equation for the line is shown in the graph.

targets). The model simulations reproduced most of the observed variation of hit and false alarm rates generated as P roved across the spatial-frequency dimension. For example, the model reproduced the difference between hit rates when P matched S_1 and when P matched S_2 (i.e., hit rate|S_1 < hit rate|S_2). The model also reproduced the observed difference between false alarm rates with lure frequencies near S_1 versus S_2 (as shown in Fig. 2): The empirical results and the model predictions agree that lures in the vicinity of S_1 attract fewer false alarms than do lures in the vicinity of S_2. From some perspectives, this result seems counterintuitive, as S_1 is associated with the noisier representation in memory but attracts fewer false alarms. But this result falls out naturally from the three-parameter model, so long as S_1' has a broader distribution than S_2', as our best-fitting parameters suggest, and the criterion for saying "yes" operates as described earlier.

The empirical mnemometric functions in Figure 3 are approximations to the underlying memory distributions. These approximations, which afford the first close-up views of representations in memory, allow us to evaluate the three hypotheses we introduced earlier. If, as Hypothesis 1 suggests, recognition decisions were made with respect to a prototype of the study series, the mnemometric curves would be unimodal. The spatial frequency of the predicted mode would depend on a precise specification of how the prototype was generated, for example, whether the prototype represented an equal or an unequal mixture of S_1 and S_2. For small values of $|S_1 - S_2|$, the empirical mnemometric curves are ambiguous with respect to the prototype hypothesis, but for large values of $|S_1 - S_2|$, the curves' clear bimodality is contrary to the hypothesis.

Hypothesis 2 proposes that decisions are based on separate exemplars in memory, but that these exemplars are systematically shifted in frequency relative to S_1 and S_2. Such frequency shifts are absent from the bimodal mnemometric functions, whose modes tend to lie at S_1 and S_2 (see Fig. 3). The data set contained some exceptions, but these tended to occur in conditions with relatively large standard errors. Although our data cannot conclusively rule out the possibility of very small systematic shifts, the data are consistent with Hypothesis 3: Stimuli are represented in memory not as prototypes, but as exemplars whose remembered spatial frequencies vary randomly and symmetrically around the frequencies of the corresponding study stimuli.

DISCUSSION

The mnemometric functions in Figure 3 reveal several important features of short-term episodic memory. First, the functions support an exemplar rather than a prototype account of memory. At least when the to-be-remembered textures are separated by several threshold units, each texture leaves a separate and distinct representation in memory. Second, trial-wise variation in a memorial representation is well described by a Gaussian or near-Gaussian function centered on the spatial frequency actually presented. Third, this random variation is smaller for the more recent study item than for the study item that preceded it.

The recency effect in our data was not caused by a time-dependent loss in memory, which could have arisen from the fact that S_2 was nearer in time to P than S_1 was. In fact, time alone does not change memory for spatial frequency for delays up to 10 s (Magnussen, 2000; Magnussen, Greenlee, Asplund, & Dyrnes, 1991) or more (Bennett

& Cortese, 1996). The experiments just cited used one-dimensional gratings, but we (Kahana & Sekuler, 2002) have reported a similar finding for textures synthesized from a trio of sinusoidal components.

On any trial in our experiment, the stimuli differed from one another only in vertical spatial frequency; all stimuli on a trial shared the same horizontal frequency. Our model assumes that observers process the textures' vertical component without interference from the accompanying horizontal component. Although the literature supports the validity of this assumption (Graham, 1989), we wanted to verify it for our data. To determine whether processing of vertical frequency was influenced by the co-occurring horizontal frequency, we sorted the trials into 10 equal-width bins of horizontal frequency. An analysis of variance showed no significant effect of horizontal frequency on the recognition of the vertical frequency component, $F(9, 81) = 0.51$, $p > .50$.

Summed-Similarity Exemplars

As noted earlier, our empirical results are not consistent with the proposition that recognition is based on a prototype constructed from the study stimuli on each trial. In fact, our results suggest that at the moment of the probe's presentation, a participant has access to a separate representation of each study stimulus. This does not guarantee, however, that recognition decisions are generated as our model assumes. For example, such decisions could reflect a summed-similarity computation (Clark & Gronlund, 1996; McKinley & Nosofsky, 1996; Nosofsky, 1984; Shiffrin & Steyvers, 1997). Summed similarity refers to the sum of similarities between the probe and each individual study item. In a previous study (Kahana & Sekuler, 2002), we successfully incorporated a summed-similarity operation in a model for short-term episodic recognition memory for visual textures. On its face, the three-parameter model used here seems to be able to account for the data without assuming summed similarity. But our data cannot entirely rule out a version of summed similarity in which pairwise differences contributing to the sum are weighted differentially, depending on each study item's own similarity to P. Choosing between the three-parameter model and a summed-similarity model will require enough additional data so that the mnemometric functions' x-axes could be sampled much more densely than in Figure 3. However, this unresolved question is not related to the distinction between prototype and exemplar memorial representations: A summed-similarity computation requires that each study stimulus be represented separately in memory at the time the computation is performed.

Extensions of the Paradigm

The stimuli used to test recognition memory were scaled to match each participant's own visual discriminative capacity. This form of stimulus scaling, which represents a novel contribution to the study of cognition, affords two related advantages. First, by compensating for individual differences in visual encoding, stimulus scaling tends to equate various participants' overall performance levels, thereby reducing between-subjects differences in recognition. Second, when data are pooled across participants, the use of scaled stimuli tends to preserve theoretically important distributional features of individual participants' data. Here, such distributional features include the locations and relative sizes of modes in the mnemometric functions (Fig. 3). In other domains, researchers have long recognized the dif-

ficulty of ensuring that any distribution derived from averaging over subjects is a fair representation of the subjects' individual distributions. When experimental results involve reaction time distributions, statistical adjustments are applied in order to preserve a resemblance between an average distribution and its component distributions (Ratcliff, 1979; van Zandt, 2002). By adjusting stimuli beforehand, our approach avoids the need for statistical correction.

The roving-probe procedure introduced here can be used in studies of recognition memory for other types of visual, as well as nonvisual, sensory stimuli. The limiting factor would be the ability to generate a useable stimulus metric that allows control of the psychological distance between study items, and allows the probe value to vary relative to those study items. Although our study stimuli varied along only one dimension, with horizontal frequency rendered irrelevant, there is no obstacle to using a roving-probe procedure with more complex, multidimensional stimuli, including naturalistic stimuli such as synthetic faces (Wilson, Wilkinson, & Loffler, 2002).

In the memory literature, most episodic recognition experiments in the Sternberg framework have used verbal or symbolic stimuli. It would be valuable to know how well our basic findings, including differences in noisiness for S_1' and S_2', hold up for verbal stimuli. Applying a roving-probe procedure to such stimuli would require controlling interitem similarities, which is not easy, but not impossible (Howard & Kahana, 2002). It may be that with verbal stimuli, or rehearsable stimuli in general, a roving-probe procedure would yield results appreciably different from those reported here. For one thing, some Sternberg experiments show both primacy as well as recency effects, but with compound grating stimuli, we (Kahana & Sekuler, 2002) found recency effects only. Moreover, there is evidence that the neural networks that support memory for certain types of visual stimuli may be distinct from the networks that support memory for symbolic material, such as words (Bennett, Sekuler, McIntosh, & Della-Maggiore, 2001; McIntosh et al., 1999). This hypothesis is consistent with behavioral demonstrations of striking differences between visual and verbal memory (e.g., Melcher & Kowler, 2001).

REFERENCES

Averbach, E., & Coriell, A.S. (1961). Short-term memory in vision. *Bell System Technical Journal*, *40*, 309–328.

Bennett, P.J., & Cortese, F. (1996). Masking of spatial frequency in visual memory depends on distal, not retinal, frequency. *Vision Research*, *36*, 233–238.

Bennett, P.J., Sekuler, A.B., McIntosh, A.R., & Della-Maggiore, V. (2001). The effects of aging on visual memory: Evidence for functional reorganization of cortical networks. *Acta Psychologia*, *107*, 249–273.

Brainard, D.H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Campbell, F.W., Nachmias, J., & Jukes, J. (1970). Spatial-frequency discrimination in human vision. *Journal of the Optical Society of America*, *60*, 545–559.

Clark, S.E., & Gronlund, S.D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review*, *3*, 37–60.

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

Harvey, L.O., Jr. (1986). Visual memory: What is remembered? In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities* (pp. 173–187). The Hague, The Netherlands: Elsevier.

Hintzman, D.L. (1988). Judgments of frequency and recognition in a multiple-trace memory. *Psychological Review*, *95*, 528–551.

Howard, M.W., & Kahana, M.J. (2002). When does semantic similarity help episodic retrieval? *Journal of Memory and Language*, *46*, 85–98.

Kahana, M.J., & Sekuler, R. (2002). Recognizing spatial patterns: A noisy exemplar approach. *Vision Research*, *42*, 2177–2192.

Loftus, G.R., & Masson, M.E.J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*, 476–490.

Magnussen, S. (2000). Low-level memory processes in vision. *Trends in Neurosciences*, *23*, 247–251.

Magnussen, S., Greenlee, M.W., Asplund, R., & Dyrnes, S. (1991). Stimulus-specific mechanisms of visual short-term memory. *Vision Research*, *31*, 1123–1219.

McIntosh, A.R., Sekuler, A.B., Penpeci, C., Rajah, M.N., Grady, C.L., Sekuler, R., & Bennett, P.J. (1999). Recruitment of unique neural systems to support visual memory in normal aging. *Current Biology*, *9*, 1275–1278.

McKinley, S.C., & Nosofsky, R.M. (1996). Selective attention and the formation of linear decision boundaries. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 294–317.

Melcher, D., & Kowler, E. (2001). Visual scene memory and the guidance of saccadic eye movements. *Vision Research*, *41*, 3597–3611.

Monsell, S. (1978). Recency, immediate recognition memory and reaction time. *Cognitive Psychology*, *10*, 465–501.

Nelder, J.H., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, *7*, 308–313.

Nosofsky, R.M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 104–114.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Pelli, D.G., Robson, J.G., & Wilkins, A.J. (1988). The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences*, *2*, 187–199.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, *86*, 446–461.

Reed, S.K. (1972). Pattern recognition and classification. *Cognitive Psychology*, *3*, 382–407.

Shiffrin, R.M., & Steyvers, M. (1997). A model for recognition memory: REM-retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*, 145–166.

Sternberg, S. (1966). High speed scanning in human memory. *Science*, *153*, 652–654.

Sternberg, S. (1998). Inferring mental operations from reaction-time data: How we compare objects. In D. Scarborough & S. Sternberg (Eds.), *Invitation to cognitive science: Vol. 4. Methods, models, and conceptual issues* (pp. 365–454). Cambridge, MA: MIT Press.

van Zandt, T. (2002). Analysis of response time distributions. In J. Wixted (Ed.), *Stevens' handbook of experimental psychology: Vol. 4. Methodology in experimental psychology* (3rd ed., pp. 461–516). New York: J. Wiley.

Wetherill, G.B., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical & Statistical Psychology*, *18*, 1–10.

Wilson, H.R., Wilkinson, F., & Loffler, G. (2002). Synthetic faces, face cubes, and the geometry of face space. *Vision Research*, *42*, 2909–2923.