

Distortions in recall from visual memory: Two classes of attractors at work

Jie Huang

Volen Center for Complex Systems and Department of Psychology, Brandeis University, Waltham, MA, USA



Robert Sekuler

Volen Center for Complex Systems and Department of Psychology, Brandeis University, Waltham, MA, USA



In a trio of experiments, a matching procedure generated direct, analogue measures of short-term memory for the spatial frequency of Gabor stimuli. **Experiment 1** showed that when just a single Gabor was presented for study, a retention interval of just a few seconds was enough to increase the variability of matches, suggesting that noise in memory substantially exceeds that in vision. **Experiment 2** revealed that when a pair of Gabors was presented on each trial, the remembered appearance of one of the Gabors was influenced by: (1) the relationship between its spatial frequency and the spatial frequency of the accompanying, task-irrelevant non-target stimulus; and (2) the average spatial frequency of Gabors seen on previous trials. These two influences, which work on very different time scales, were approximately additive in their effects, each operating as an attractor for remembered appearance. **Experiment 3** showed that a timely pre-stimulus cue allowed selective attention to curtail the influence of a task-irrelevant non-target, without diminishing the impact of the stimuli seen on previous trials. It appears that these two separable attractors influence distinct processes, with perception being influenced by the non-target stimulus and memory being influenced by stimuli seen on previous trials.

Keywords: visual short-term memory, selective attention, prototype effect, perceptual averaging

Citation: Huang, J., & Sekuler, R. (2010). Distortions in recall from visual memory: Two classes of attractors at work. *Journal of Vision*, 10(2):24, 1–27, <http://journalofvision.org/10/2/24/>, doi:10.1167/10.2.24.

Introduction

Research on visual working memory has been dominated by a focus on questions of *quantity*, such as the number of stimuli that can be remembered, or for how long. Although this focus has produced valuable insights (Vogel, Woodman, & Luck, 2005; Zhang & Luck, 2008, 2009), it is important to recognize the value of a parallel focus, namely, on the *quality* of visual memory. This parallel focus attempts to characterize factors that promote systematic discrepancies between a stimulus as presented and that same stimulus as remembered. Over the years, researchers have identified myriad memory-distorting factors. For visual working memory (vWM), these factors include the assimilative effect of verbal labels (Bruner, Busiek, & Minturn, 1952), changes associated with normal aging (Bennett, Sekuler, & Sekuler, 2007), and the time available for encoding a stimulus (Bays, Catalao, & Husain, 2009), to name just a few.

Here we examine a pair of potential influences on the quality of vWM for a target stimulus. One potential influence is residual memory for stimuli that a subject had seen on previous trials; the second potential influence is a task-irrelevant stimulus that accompanies the trial's target stimulus. Anticipating that these influences might be subtle and therefore easily lost in experimental noise, we optimized the yield and sensitivity of our experiments by focusing on subjects' *recall* of what had been presented,

rather than mere *recognition* of that previous stimulus. Memory researchers have argued that recall and recognition draw upon distinct forms of information, and that a recall task sometimes yields the more sensitive index of memory (Kahana, *in press*). In addition, gains in sensitivity that might accrue from a recall task are amplified if each trial's stimulus set is not a random selection from a pool of potential stimuli but, instead, is deliberately constructed, trial by trial, to serve well-defined, theoretical aims (e.g., Visscher, Kahana, & Sekuler, 2009; Viswanathan, Perl, Visscher, Kahana, & Sekuler, 2010).

Our experiments exploit one of psychophysics' most venerable tools, direct matches made by the subject.¹ This method generates an analogue assay of stimulus' appearance on each trial. We assumed that analogue measures would afford sensitivity sufficient to measure important effects that might have been too subtle to show up in the less sensitive, binary (e.g., “Yes” and “No”) responses generated by other methods. The stimuli we used, Gabors, have proven to be valuable test materials in previous studies of visual working memory and are well suited to our aims because they lend themselves to direct matching, and because they facilitate analysis of visual memory's dependence on similarity, a variable that influences numerous aspects of memory (Sekuler & Kahana, 2007).

Many studies and models of perception and of vSTM treat each experimental trial as an encapsulated epoch, with memory completely zeroed out or reset after each trial. Recently, Zhang and Luck (2009) presented a

mechanistic, theoretical justification for such treatment and presented data suggesting that within several seconds, short-term memory terminates abruptly and completely. Although it would be a considerable theoretical convenience if each experimental trial were completely encapsulated and if vSTM were reset at each trial's end, suggestions that this might not be so can be found as far back as the earliest systematic study of memory. In that study, Ebbinghaus (1885/1913) noted that

The vanished mental states give indubitable proof of their continuing existence even if they themselves do not return to consciousness at all, or at least not exactly at the given time, ... Most of the experiences remain concealed from consciousness and yet produce an effect which is significant and which authenticates their previous existence. (p. 2)

Mindful of empirical support for Ebbinghaus' assertion that memories' influence lingers well after the memories seem to have vanished (Nelson, 1985; Visscher et al., 2009), we used direct matching to quantify the impact that previously seen stimuli might have on the remembered appearance of the current trial's target stimulus. Our results reveal precisely such an influence.

As mentioned earlier, we are interested also in how the remembered appearance of one, *target*, stimulus might be altered by the presence of a second, *non-target*, task-irrelevant stimulus. That such an influence is possible is suggested by demonstrations that the visual system integrates and averages information over distinct stimuli if those stimuli are presented close together, in time and/or in space (e.g., Alvarez & Oliva, 2008; Haberman, Harp, & Whitney, 2009; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). To preview our result, a task-irrelevant, non-target stimulus acts as a strong attractor on the recalled appearance of an accompanying target stimulus. By controlling the similarity relationships between target and non-target stimuli, we show that this within-trial attractor effect depends critically upon the perceptual similarity of target and non-target stimuli to one another. Interestingly, a cue presented in a timely fashion before the stimuli allows selective attention to substantially curtail this influence of a trial's non-target stimulus. Finally, the two attractors demonstrated in our experiments—one operative within a trial, and the other operative across multiple trials—seem to be additive.

Experiment 1: Perception-limited and memory-dependent recall of single items

This first part of [Experiment 1](#) used direct matching to measure the accuracy with which a Gabor's spatial frequency was perceived; the experiment's second part

used the same technique to measure the accuracy of memory-based short-term recall of a Gabor's spatial frequency. For the initial, perception-limited measurements, Target and Comparison Gabors were shown side by side on a display screen and remained visible while a subject adjusted the spatial frequency of the Comparison Gabor to match that of the Target. As both stimuli remained visible, dependence upon memory was minimized, and the outcome can be taken as essentially perception-limited (Cohen & Bennett, 1997). In the experiment's second part, a subject saw a single, briefly presented Gabor, and then after some delay, reproduced its spatial frequency from memory, by adjusting the spatial frequency of a Comparison Gabor. In both phases of the experiment, each final setting of the Comparison Gabor was converted to a measure of error, given by the difference between (1) the spatial frequency of the Target Gabor, and (2) the spatial frequency of the adjusted Comparison Gabor. We used the distributional properties of the errors to characterize the properties of the perceptual representation and of the representation in memory.

Subjects

The same eight subjects, two males, participated in both parts of [Experiment 1](#) and in [Experiment 2](#) as well. Their ages ranged from 19 to 24 years. All had normal or corrected-to-normal vision as measured with Snellen targets and normal contrast sensitivity as measured with Pelli–Robson charts (Pelli, Robson, & Wilkins, 1988). Subjects were naive to the purpose of the experiments, and all were paid for their participation.

Apparatus

MATLAB 7 and extensions from the Psychophysics Toolbox (Brainard, 1997) were used to generate and display stimuli, which were shown on a 14-inch CRT monitor (refresh rate of 95 Hz; screen resolution of 1024 × 768 pixels). Display luminance was linearized by means of software adjustments, and the screen's mean luminance was maintained at 32 cd/m². During testing, a subject sat with head supported in a chin rest, viewing the computer display binocularly from a distance of 57 cm.

Stimuli

Gabor stimuli were unidimensional, vertically oriented sinusoidal luminance gratings windowed by a circular Gaussian carrier whose space constant was 1.29° visual angle. Each Gabor subtended 5.15° visual angle, and its sinusoidal component had a fixed contrast of 0.2, a value

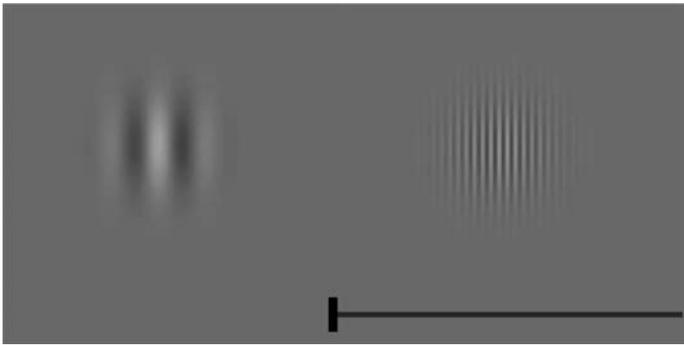


Figure 1. Stimulus arrangement in the first part of [Experiment 1](#). Target and Comparison Gabors were presented to the left and right of the display's center. To adjust the Comparison Gabor's spatial frequency to match that of the Target Gabor, a subject clicked on a horizontal adjustment bar located immediately below the Comparison stimulus. After each click, the spatial frequency of the Comparison Gabor changed, tracking changes in the location clicked. The short vertical bar changed position with each click, marking the clicked location (see text for details).

well above detection threshold. To undermine subjects' ability to base adjustments on local correspondences between stimuli, the absolute phase of each Gabor's sinusoidal component was shifted on each trial by a random value ranging from 0 to $\pi/2$. Before the actual experiment, we measured each subject's spatial frequency discrimination threshold using two-alternative forced-choice trials embedded in an adaptive psychophysical procedure. On each trial, two Gabors were presented sequentially, each for 500 ms, with an inter-stimulus interval of 500 ± 100 ms. The subjects' task was to identify the Gabor, either first or second, whose spatial frequency was the higher one. The higher spatial frequency Gabor was equally likely to appear first or to appear second. After a trial, a distinctive tone gave information about response correctness.

From trial to trial, the lower spatial frequency of the two stimuli was chosen from a uniform random distribution ranging from 0.5 to 5 cycles/deg. This distribution spanned and exceeded the range of spatial frequencies that would be used subsequently in our actual experiments. The QUEST algorithm (Watson & Pelli, 1983) controlled the difference in the spatial frequencies of each trial's two Gabors. As implemented here, QUEST estimated the difference between the paired spatial frequencies that produced correct judgments 79% of the time. Each subject's discrimination threshold was determined in three separate QUEST runs. The lowest Weber fraction of the three was taken as the subject's discrimination threshold.

Vision-limited direct matches. On each trial, a subject viewed a Target Gabor, then reproduced its spatial frequency by adjusting a Comparison Gabor. A trial began with a fixation point presented for 300 ms at the

display's center. Then, 300 ms later, the two Gabors were presented simultaneously side by side, each 4.15° visual angle from the display's center. Both Gabors remained visible until the subject was satisfied that the Comparison Gabor's adjusted spatial frequency matched that of the Target. This satisfaction was signaled by a key press. As [Figure 1](#) suggests, the Target Gabor was always presented on the left, and the Comparison Gabor was presented on the right. Below the Comparison Gabor was a dark horizontal adjustment bar, 9.64° visual angle long. Subjects adjusted the Comparison Gabor's spatial frequency by clicking at different locations along the adjustment bar; the computer read the coordinates of the point clicked and changed the Comparison Gabor's frequency accordingly. In addition, as an aid to the subjects, the location of each click was marked on the adjustment bar by a vertical line whose position tracked the clicked location. This process continued for a number of iterations, with each click producing a corresponding change in spatial frequency—until the subject was satisfied that the Comparison Gabor matched the Target Gabor.

Each trial's Target stimulus had a spatial frequency that was drawn from a continuous uniform random distribution ranging from 0.5 to 5 cycles/deg. The Comparison Gabor's spatial frequency was adjustable from 0.1 to 6.0 cycles/deg, a range that exceeded the range of spatial frequencies of Target stimuli. A single pixel change along the horizontal adjustment bar changed spatial frequency by $2.28 \times e^{-5}$ cycles/deg, a small, sub-threshold change. For half the subjects, the Comparison stimulus' initial spatial frequency was 0.1 cycles/deg and increased with rightward movements along the adjustment bar. For the remaining subjects, the Comparison stimulus' initial spatial frequency was 6 cycles/deg and decreased with rightward movements along the adjustment bar. The left-right variation and assignment of initial frequency were kept constant for individual subjects. At the beginning of a trial, the cursor appeared at the left end of the adjustment bar, and the Comparison stimulus' initial spatial frequency was fixed at either 0.1 cycles/deg (for half the subjects) or 6 cycles/deg (for the remaining subjects). Subjects had no time limit on their adjustments of the Comparison stimulus but were encouraged to finish each trial within 4 s, a time that was signaled by a computer-generated beep. After making a satisfactory adjustment, the subject pressed a key to start the next trial. By the end of the practice trials, all subjects were able to complete the adjustment process comfortably in advance of the reminder beep, although we did not record those adjustment times. As the Target stimulus remained visible during the adjustment, this simultaneous matching task had minimum reliance on memory, any discrepancy between the Target frequency and the final frequency of the Comparison stimulus likely arose from some other source(s), e.g., inaccuracies of perception, or imprecision in controlling the computer mouse to adjust the Comparison stimulus.

Assessing short-term memory for a single Gabor. The second part of [Experiment 1](#) used a delayed matching task, which assessed vSTM for a single Gabor that had just been seen but was now no longer visible. Each trial began with a fixation point presented for 300 ms at the center of the display. This was followed 300 ms later by a single study Gabor (the Target stimulus), which was presented for 500 ms, also at the center of the display. One of two retention intervals, 1400 ± 100 or 2400 ± 100 ms, followed the Gabor's disappearance. As explained below, these two values were chosen in order to match the effective retention intervals that would be used in [Experiment 2](#). Following the retention interval, the Comparison Gabor was presented at the center of the display, along with the dark, horizontal adjustment bar below it.

A subject reproduced the Target's spatial frequency from memory using the matching method described earlier. For each subject, the mapping of spatial frequency onto the adjustment bar and the Comparison Gabor's initial frequency preserved the mapping that the subject had used in the experiment's first part. This consistency made a subject's task easier and eliminated the interference and confusion that might have come from a change in the mapping assigned to a subject. Trials at each of the two retention intervals occurred with equal frequency and were randomly intermixed within a block of trials. The comparison between shorter and longer retention intervals was an index of any rapid decay in memory quality. As explained below, the particular values chosen for the two retention intervals provided baseline memory controls for test conditions that would be used in [Experiment 2](#).

Procedure

Both parts of [Experiment 1](#) were completed in a single session, which comprised six blocks of trials. In the first three blocks, QUEST measured the subject's spatial frequency discrimination threshold, with each block comprising 75 trials. The fourth block of trials used direct matching to measure visual perception of a Gabor's spatial frequency; the final two blocks measured short-term memory for a single Gabor. Blocks devoted to measuring perception or memory each consisted of 168 trials, with the first 24 trials treated as practice and excluded from data analysis.

Results and discussion

For each trial, the reproduction error was defined by the difference between (1) the actual spatial frequency of the Target Gabor and (2) the spatial frequency produced by the subject's manipulation of the Comparison Gabor. The magnitude of this difference was normalized relative to an individual subject's Weber fraction for spatial frequency

and was therefore expressed in units of just noticeable difference (JNDs). In addition to equalizing values across subjects, this transformation allowed the aggregation of measurements over trials, despite trial-wise variation in the Target Gabor's spatial frequency.

Let x be some subject's Weber fraction; let f_T be the spatial frequency of the Target Gabor on some trial; and let f_R be the Comparison Gabor's final, adjusted frequency. The raw error in reproduction, $(f_R - f_T)$, is normalized by the subject's Weber fraction to produce a Transformed Reproduction Error (*TRE*), i.e.,

$$TRE = \log_{(1+x)}\left(\frac{f_R}{f_T}\right) = \frac{\ln f_R - \ln f_T}{\ln(1+x)}. \quad (1)$$

When there is no error in reproduction, $TRE = 0$. Note that *TRE* preserves the sign of the raw error, $(f_R - f_T)$, so that $TRE > 0$ indicates that the reproduced spatial frequency was higher than that of the Target Gabor, and $TRE < 0$ indicates the reverse.

Vision-based matches. We examined the distribution of each subject's *TRE* values when Target and Comparison Gabors were both visible, alongside one another. Although based on just 144 trials per condition, each subject's distribution appeared to be approximately bell-shaped with a peak near zero, that is, near a value corresponding to zero error. To characterize the distributions more systematically, we applied the Anderson–Darling (1952) test. The result was that the distributions from four of the eight subjects did not differ significantly from a normal distribution ($p > 0.05$), but four other subjects' distributions showed small but statistically reliable departures from normality ($p < 0.05$). [Figure 2A](#) shows the distribution of *TRE* values summed over all eight subjects.

In order to reduce possible effects of non-normality of error distributions, and to minimize the possible influence of extreme values, robust statistics were used to summarize the key properties of distributions of *TRE*. A distribution's location (central tendency) was represented by its median value, and a distribution's scale (variability) was represented by MAD, the median absolute deviation from the median (Wilcox, 2005). To facilitate comparisons among conditions, the boxplots (Tukey, 1977) in [Figure 3A](#) summarize individual subjects' median *TRE*s for each condition. The leftmost boxplot represents median *TRE*s from vision-based matches. As can be seen in that boxplot, on average, subjects' median reproduction error for vision-based performance was -0.151 JNDs. That is, subjects' reproductions tended to lie very close to the spatial frequency that they were trying to reproduce. This small constant error did not result from the Comparison Gabor's starting spatial frequency, or from the way that spatial frequency was mapped onto location along the adjustment bar. When we compared the

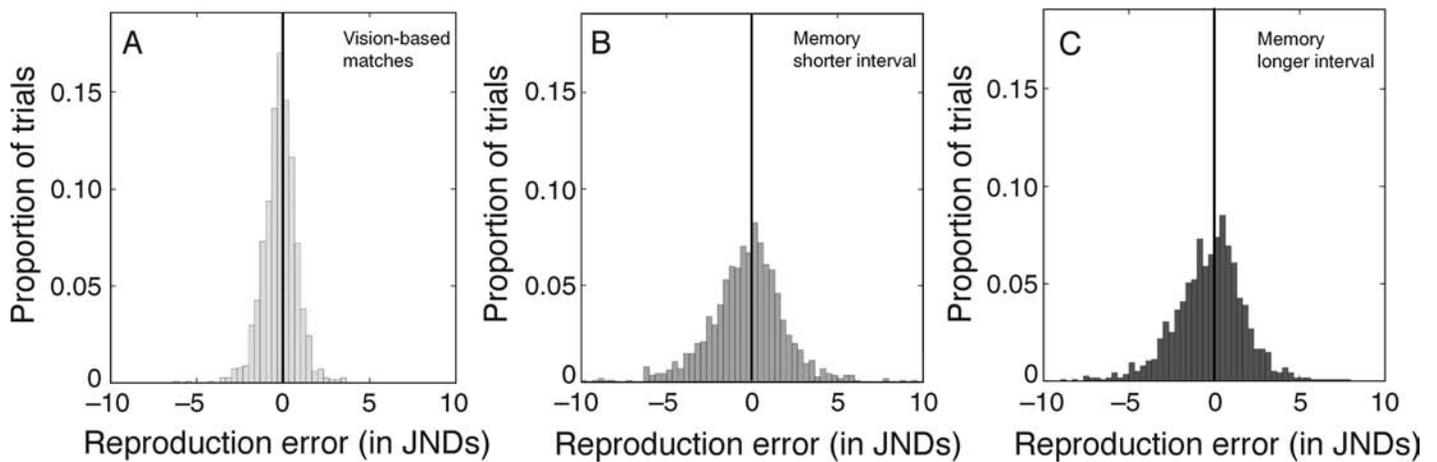


Figure 2. Frequency histograms showing transformed reproduction errors (*TREs*) from the three test conditions in [Experiment 1](#). Data are aggregated over all eight subjects. (A) Distribution of transformed reproduction errors based on simultaneous matching of Target and Comparison Gabors. (B, C) Distributions of transformed reproduction errors produced when the Target Gabor's spatial frequency was reproduced from memory. Results in (B) were taken with a post-stimulus retention interval of 1400 ms; results in (C) were taken with a post-stimulus retention interval of 2400 ms. In (A–C), black vertical lines indicate zero error.

reproduction errors collected from subjects tested with a mapping in which spatial frequency increased left–right along the adjustment bar against errors from subjects tested with a mapping that increased right–left, the median reproduction errors of these two groups were not reliably different from one another, $t(6) = 1.230$, $p = 0.265$. Note that although negative values of *TREs* were most common, the span of the box's whiskers shows that some subjects' median *TRE* values were positive.

Next, we examined vision-based matches' variability, aggregating values of MAD from individual subjects into the leftmost boxplot in [Figure 3B](#). The averaged

value of MAD across eight subjects was 0.546 *JNDs*. As visual support was available throughout the matching process, we believe that this small but genuine variability most likely reflects perceptual variability (Cohen & Bennett, 1997) and/or imprecision of motor output. The task relies minimally on memory, although some trans-saccadic memory could have been recruited when subjects looked back and forth between Target and Comparison Gabors (Demeyer, De Graef, Wagemans, & Verfaillie, 2009; Melcher & Colby, 2008).

Memory-based matches. To assess the fidelity of a single Gabor's representation in short-term memory,

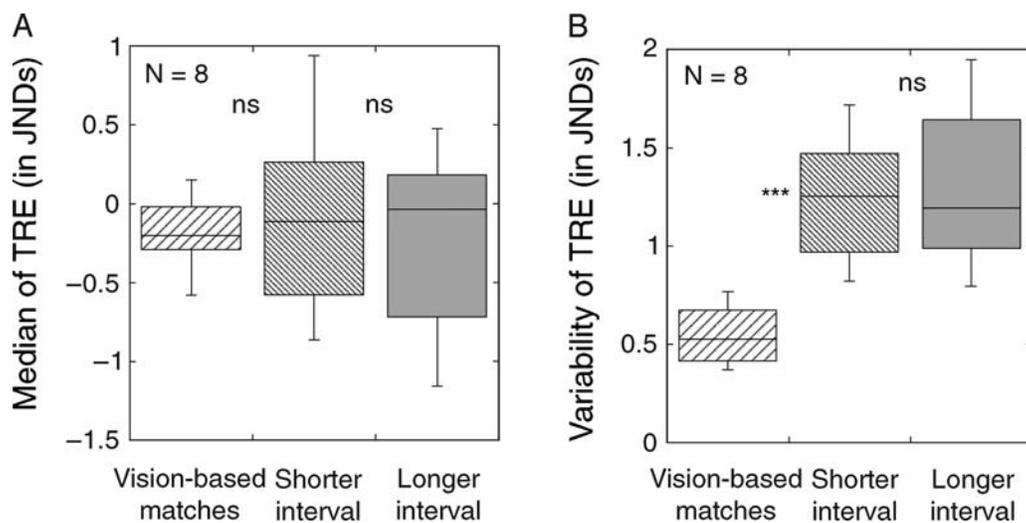


Figure 3. (A) Boxplots showing the median transformed reproduction errors (*TREs*) in the three test conditions of [Experiment 1](#). (B) Boxplots for the median absolute deviation (MAD) of transformed reproduction errors for the three conditions. A non-significant difference is shown as *ns*; a difference significant at $p < 0.001$ is shown as ***.

reproductions were separated according to retention interval, the shorter (1400 ± 100 ms) and the longer one (2400 ± 100 ms). Histograms of all subjects' reproduction errors in both retention intervals are given in [Figures 2B](#) and [2C](#). The two rightmost boxplots in [Figure 3A](#) present the subjects' median reproduction errors for the two retention intervals. Neither the median for the shorter interval ($M = -0.167$, $SD = 0.591$), nor the median for the longer interval ($M = -0.130$, $SD = 0.585$) differed significantly from zero (one sample t -tests, $t(7) = -0.486$, $p = 0.654$ and $t(7) = -1.106$, $p = 0.305$, respectively). Additionally, the medians of all three conditions (the vision-based condition and the two memory conditions) did not differ significantly from one another, $F(2,14) = 0.477$, $p = 0.630$. These results demonstrate that when a brief retention interval separates a trial's study and reproduction phases, the distributions of memory-based reproductions remain centered at or near the spatial frequency that had actually been seen. Moreover, this result is essentially unchanged from the result obtained when reproductions were vision-based.

Having found no reliable differences between the central tendencies of vision-based and memory-based errors in reproduction, we examined the variability of reproductions under these same conditions. This examination was prompted by studies that demonstrated an increase in noise (variability) during the time that representations are maintained in visual short-term memory (e.g., Gold, Murray, Sekuler, Bennett, & Sekuler, 2005), though usually with intervals longer than those used here.

As explained earlier, the variability of reproductions was quantified by MAD, the median absolute deviation from the median *TRE*. [Figure 3B](#) shows MAD for vision-based reproductions (leftmost box: $M = 0.546$) and also for memory-based reproductions (two rightmost boxes: $M = 1.299$ and 1.240 , for the shorter and longer retention intervals, respectively).

Given the very small difference between the respective durations of our two retention intervals, it is not surprising that reproduction errors are no more variable in one interval than in the other, $t(7) = 0.998$, $p = 0.352$. However, note that the boxes for the two memory-based conditions in [Figure 3B](#) are shifted upward relative to the box for the vision-based condition. This upward displacement means that memory-based reproduction errors are more variable with either retention interval than when reproduction was vision-based, $t(7) = 7.659$, $p < 0.001$ and $t(7) = 5.632$, $p < 0.001$, for comparisons of vision-based MAD values against those from the shorter and the longer retention intervals, respectively. These differences are consistent with the overall trends in the spread of the histograms in [Figures 2A–2C](#).

We wondered if this difference between conditions might have been caused by the difference between the display geometries used for the two conditions. In particular, for memory-based matches, Target and Comparison Gabors were at and around the fovea, but stimuli

for the vision-based matches were presented several degrees away from the fovea. Although this difference cannot be entirely discounted, it is worth noting that contrary to what one might expect from a distinction between foveal and non-foveal stimulus presentations, the foveal, memory-based matches are significantly more variable than the non-foveal, vision-based matches. So, a retention interval of only 1400 ms seems to be sufficient to significantly increase the variability of matches relative to the variability with vision-based reproductions. This difference in variability between memory- and vision-based reproductions suggests that some process associated with memory increases the noise of the stimulus representation, and perhaps begins to do so as soon as the stimulus disappears from view. That the onset of such a process would be rapid is consistent with previous reports (Gold et al., 2005; Zhang & Luck, 2008).

Biases in matching. It has long been known that the psychophysical responses elicited by one stimulus can be influenced by stimuli that have been presented to the subject, on previous trials (e.g., Helson, 1948; Hollingworth, 1910; Poulton, 1989). Moreover, research on memory has demonstrated analogous, between-trial effects, the best known of which may be proactive interference (Visscher et al., 2009). We therefore examined subjects' responses in both vision- and memory-based conditions for evidence of such an influence. For each condition, the match produced by the subject on trial n was regressed against the stimulus seen on that same trial and against stimuli seen on trials $n - 1$, $n - 2$, ... $n - 5$.

[Figure 4A](#) displays the coefficients from the linear multiple regression for vision-based matches. Each value is the mean coefficient taken over all subjects' individual regression analyses. Not surprisingly, trial n 's own stimulus was a potent determinant of the match made on that trial ($b = 0.95$); more to the point, the coefficient representing the influence of the stimulus on trial $n - 1$ was indistinguishable from zero. Thus, with vision-based matches, the stimulus on the current trial is essentially the sole determinant of the match made on that trial. However, the case is quite different when matches are memory-based. [Figure 4B](#) displays the coefficients from the linear multiple regression on the memory-based matches. Again, each value is the mean coefficient taken over all subjects' individual regression analyses. Note first that the stimulus on trial n is a less potent influence on that trial's match than it was for vision-based matches (here, $b = 0.71$). In addition, for all eight subjects, the stimulus on trial $n - 1$ had a small, but statistically significant influence on the match made on the next trial, n . For three subjects, the same held true for trial $n - 2$'s influence. So, unlike the case for vision-based matches, memory-based matches reflect a lingering effect of stimuli presented at least one, perhaps two trials earlier.

[Figure 4B](#) shows the influence of some single prior trial on the memory-based reproduction of the stimulus most recently seen. Each coefficient isolates the relationship

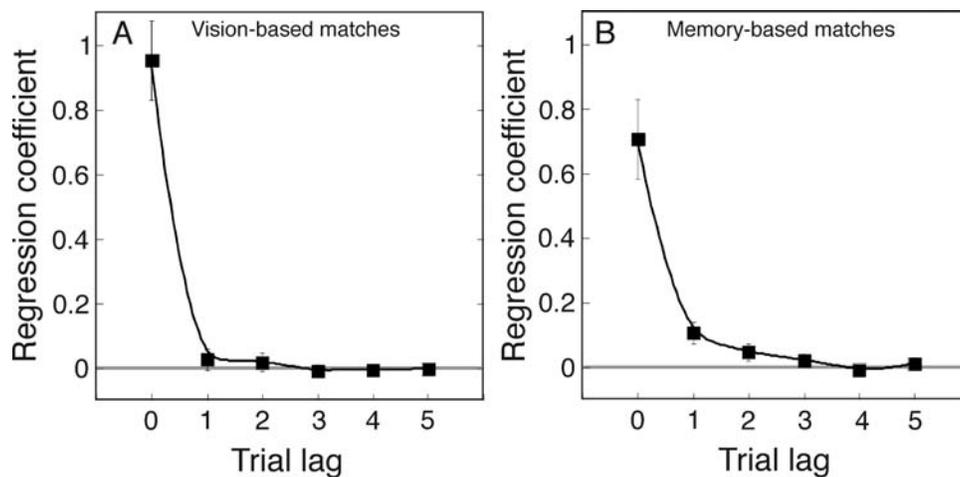


Figure 4. Regression coefficients for the dependence of the n th trial's reproduction on that trial's stimulus and on the stimuli presented on each of the five preceding trials. (A) Results from vision-based matches. (B) Results from memory-based matches. Error bars are within-subject standard errors. The horizontal gray line indicates a value of zero.

between trials taken one at a time, for example the relationship between the match on one trial and the stimulus on another trial. There is reason to suspect that Figure 4B actually understates the aggregate influence of multiple, preceding trials. The fact that stimuli were generated randomly for each trial also randomized the relationship between stimuli on any pair of trials. Suppose that the influence of one trial's stimulus on another trial's match depended upon the similarity of the stimuli on the two trials. The randomization of stimulus values over trials could cause the inter-trial effect to be underestimated when examined one trial at a time. To circumvent this problem, we evaluated each trial's reproduction relative to the mean of all the preceding stimuli in that same block of trials. We reasoned that because the mean stimulus was representative of all the stimuli that a subject had seen, comparing each trial's reproduction against that mean stimulus would smooth out trial-wise variation in the relationship between stimuli on any pair of trials. In order to capture the influence of the mean stimulus on each trial's reproduction of spatial frequency, we retained the absolute value of each trial's reproduction error but transformed the sign of each error. As a result, the sign of the error corresponded to the *direction* of the influence, that is, either toward or away from the mean stimulus. Under this convention, a positive value signifies that a reproduction was shifted *toward* the mean value of the frequencies that had been presented in preceding trials within that block; a negative value signifies that a reproduction was displaced *away* from the mean of previously seen values. If subjects' reproductions were not biased by the mean spatial frequency from preceding trials within that block, the expected median of sign-adjusted errors should be zero. As the mean stimulus from preceding trials comprises a prototypical stimulus for all

the stimuli that had been seen, we will refer to its influence as a "Prototype effect."

When Target Gabors remained visible throughout the process of matching, the median of the sign-adjusted reproduction errors did not differ significantly from zero ($M = 0.027$, $SD = 0.137$), one sample t -test, $t(7) = 0.555$, $p = 0.596$. So once again, the stimuli seen on the preceding trials had no systematic effect on subjects' reproductions of what they were currently seeing. However, the outcome was different when reproductions were made from memory. Here, the median of the sign-adjusted reproduction errors was significantly greater than zero, and was so for both the shorter retention interval ($M = 0.679$, $SD = 0.482$), $t(7) = 3.979$, $p < 0.006$, and the longer retention interval ($M = 0.612$, $SD = 0.522$), $t(7) = 3.316$, $p < 0.013$. Finally, the magnitude of this Prototype effect did not differ between the two retention intervals, $t(7) = -1.278$, $p = 0.242$, with trials in each being shifted by an average of more than six-tenths of a JND toward the mean stimulus.

In summary, after a very brief retention interval, subjects' average memory for a Target's spatial frequency preserves the spatial frequency that had been seen, although that memory representation very quickly becomes more variable, as vision-based matching gives way to memory-based matching. This time-dependent increase in the variability of representation is consistent with Zhang and Luck's (2008) report of time-dependent changes in subjects' reproductions of remembered color. Although Zhang and Luck collected no responses with a still-visible memorandum, they did find a substantial increase in the variability of recalled color between one short retention interval and another slightly longer retention interval. It has not escaped our notice that the shorter of the retention intervals we used in probing

memory is not so different from inter-stimulus intervals that are commonly used in psychophysical studies of vision. That such a short interval could increase variability of matches—in our experiment and in Zhang and Luck’s—suggests caution in interpreting results from such psychophysical studies; in fact, the result, even with very short inter-stimulus intervals, might well reflect some mixture of influences from vision and influences from memory.

Additionally, the shift in our experiment from vision-limited judgments to memory-limited ones allowed the remembered, prototypical stimulus to influence subjects’ matches. So, overall, memory-based matches are shifted toward other stimuli that had been seen, but vision-based matches are not significantly influenced by stimuli outside the current trial. This substantial cross-trial influence demonstrated that memory is not completely expunged after each trial. At first, this result appears to be inconsistent with Zhang and Luck’s (2009) recent claim that without active support, short-term memories terminate abruptly and completely between some 4 and 8 s after the stimulus has disappeared. However, this apparent inconsistency between Zhang and Luck’s claim and our result is easily resolved by harkening back to the first systematic study of human memory. In that pioneering work, Ebbinghaus (1885/1913) demonstrated that when a memory seemed to have disappeared and could not be recalled, an appropriate indirect measure demonstrated that the trace of the recall-resistant memory remained active. In Ebbinghaus’ paradigm, a non-recallable memory facilitated relearning of the “forgotten” material, a phenomenon that has been confirmed many times since (e.g., Nelson, 1985). Our subjects probably could not have recalled and accurately reproduced the stimuli they had seen on the preceding few trials, but that failure alone would not prove that the memories of those stimuli past had actually suffered complete, irreversible death. As Ebbinghaus demonstrated, a properly sensitive indirect assay can reveal signs of life in memories that otherwise would have been mistaken for dead.

Experiment 2

Experiment 1 confirmed that direct matching can produce reliable trial-by-trial measures of subjects’ memory, at least when only a single Gabor stimulus is being held in memory. Although memory on any trial was clearly dominated by the characteristics of the stimulus seen on that trial, the experiment revealed a second modulator of vSTM, namely, the characteristics of stimuli presented on previous trials (a Prototype effect). The stimuli contributing to this influence were separated temporally from the to-be-matched stimulus and were also clearly irrelevant to the task of matching that

stimulus. To isolate the temporal component of this influence, Experiment 2 inserted a task-irrelevant stimulus directly into each trial, thereby shrinking the temporal separation between task-relevant and task-irrelevant stimuli. We asked how a trial’s task-irrelevant study item would affect the remembered appearance of the accompanying, task-relevant Target, and how any such effect would interact with the Prototype effect seen in Experiment 1. It is possible that the within-trial effect of the task-irrelevant stimulus could diminish the Prototype effect altogether. The close temporal proximity of the stimulus responsible for any within-trial effect might allow it to override any effect from sources that were temporally further removed.

In this second experiment, two study Gabors were presented in rapid succession on each trial. These were followed by a visual cue that indicated which of two study items, the first or second, was to be reproduced from memory. On randomly interleaved trials, the cue corresponded to the first stimulus or to the second stimulus. Regardless of its serial position, first or second, the study item to be reproduced will be referred to as the Target stimulus. As the other study item was *not* reproduced, it is designated the Non-Target stimulus. Despite its task-irrelevant status, the spatial frequency of the Non-Target could well influence the reproduction of the Target stimulus. In order to isolate possible within-trial influence of the Non-Target from the Prototype effect demonstrated earlier, we manipulated the difference in the spatial frequencies of the two Gabors presented on a single trial. This manipulation made it possible to determine how the similarity of the study stimuli to one another affected the Non-Target’s influence upon recall of the Target stimulus.

Subjects

The eight subjects from Experiment 1 participated here and, again, were paid for their participation.

Apparatus and stimuli

Except as specified below, Experiment 2 used the same methods as the previous experiment. Figure 5 shows the sequence of events on a typical trial, which began with the display of a fixation point, for 300 ms at the center of the display screen. Then, after a 300-ms delay, two study Gabors, s_1 and s_2 , were presented sequentially, each for 500 ms, with a 500-ms inter-stimulus interval between the two. Next, after 750 ± 50 ms, a numerical cue was presented for 300 ms. The cue indicated which one of the two study stimuli was Target Gabor: “1” indicated that s_1 was the Target, and “2” indicated that s_2 was the Target. s_1 and s_2 were equally likely to be the Target Gabor, with

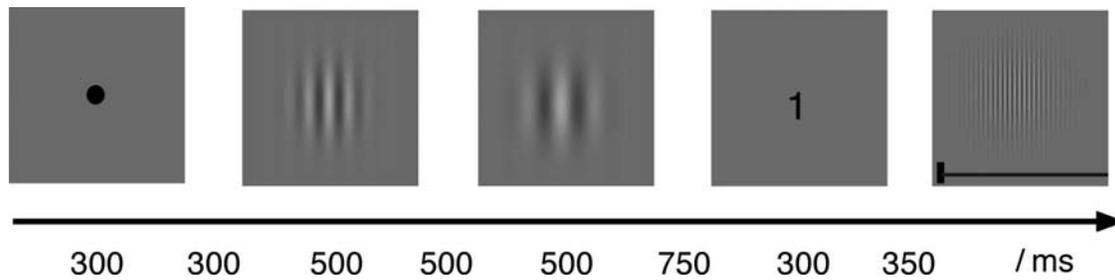


Figure 5. The timeline of events comprising one of [Experiment 2](#)'s trials. Subjects viewed two Gabors (s_1 and s_2) presented in rapid succession, followed by a cue, which was either the digit “1” or the digit “2”. In the example shown, the cue was “1”, meaning s_1 was Target Gabor and was to be reproduced; s_2 was Non-Target Gabor. Subjects reproduced the Target's spatial frequency by adjusting the Comparison Gabor to match the remembered frequency of Target. See text for additional details.

the two types of trials randomly alternating within a block of trials. The black cue had sufficient contrast that the cue was easily and quickly recognized. Finally, after a 350 ± 50 ms interval, a Comparison Gabor was presented at the center of the display. As in the previous experiment, this Gabor was accompanied by an adjustment bar that controlled the Gabor's spatial frequency. The subject attempted to reproduce the cued, Target stimulus (either s_1 or s_2) from memory, using the same method as in [Experiment 1](#). There was no limit on the time that could be taken to produce the reproduction, but subjects invariably finished the task within 4 s. The Comparison Gabor and adjustment bar remained visible until the subject finished the reproduction and signaled satisfaction with the reproduction by pressing a key. This key press initiated the next trial.

Note that the timing of s_1 , s_2 , and the Comparison stimulus means that when s_2 had to be reproduced the retention interval was 1400 ± 100 ms, which replicates the shorter retention interval in [Experiment 1](#); additionally, when s_1 had to be reproduced, the retention interval for that stimulus was 2400 ± 100 ms, which replicates the longer retention interval in [Experiment 1](#). These equivalences between experiments made it possible to use the results from the two retention intervals in [Experiment 1](#) as memory baselines for performance with s_1 and s_2 in [Experiment 2](#).

The term “inter-item difference” refers to the difference, in *JND* units, between the spatial frequencies of s_1 and s_2 . With larger values of inter-item difference, s_1 and s_2 are perceptually less similar. We tested subjects with three inter-item differences: 2, 4, and 6 *JND*s. These occurred with equal frequency and were randomly interleaved. In addition, s_1 and s_2 were equally likely to have the lower of the two study items' spatial frequencies. On each trial, the lower of the two spatial frequencies was determined by the sum of (1) a random sample from the set $\{0, 1, 2, 3, 4\}$ *JND*s, and (2) a base frequency drawn from a uniform random distribution ranging from 0.5 to 1 cycles/deg. This randomization of spatial frequencies greatly diminished

the likelihood that the studied spatial frequencies would be repeated on successive trials.

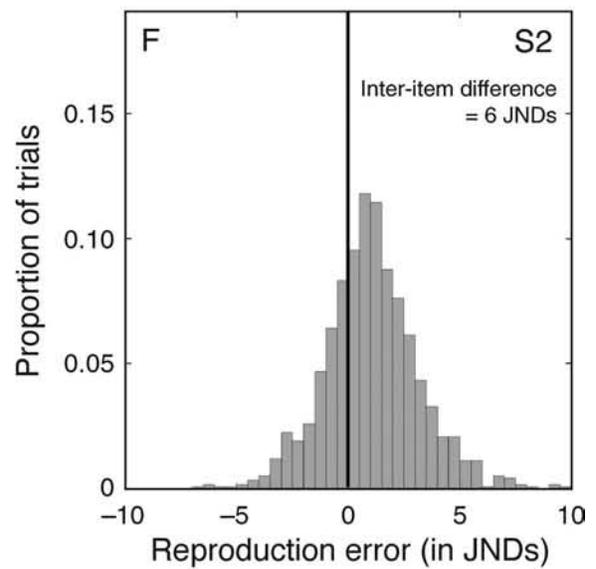
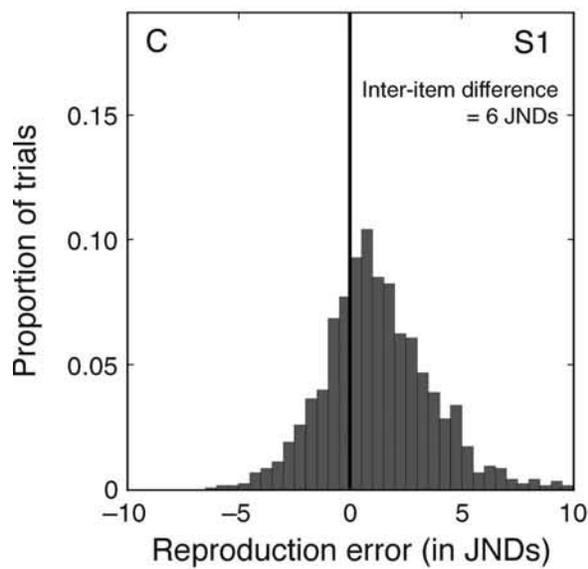
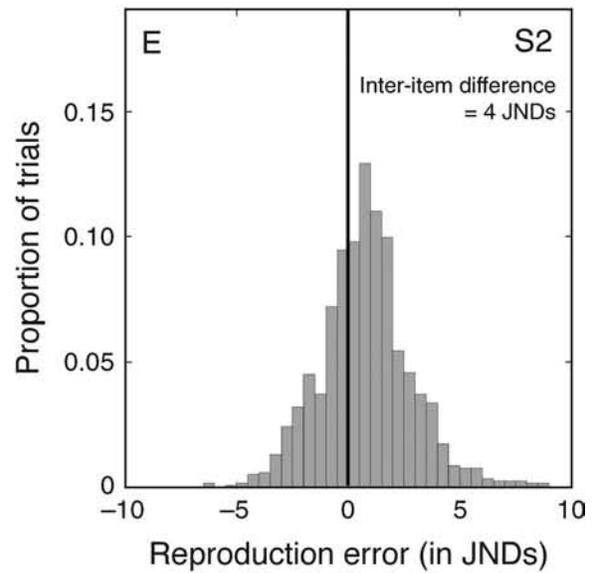
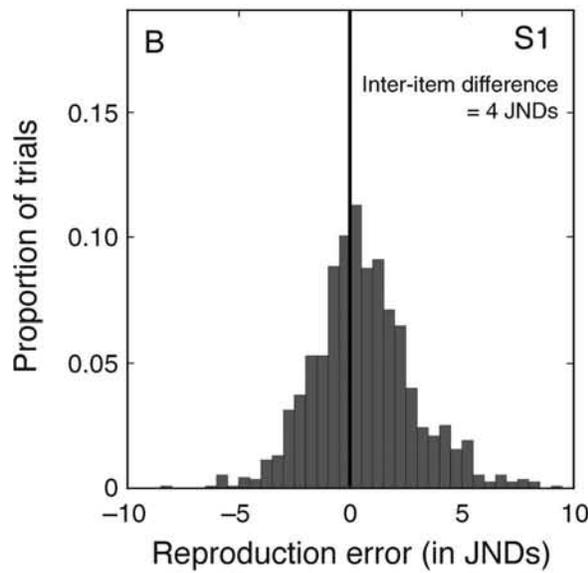
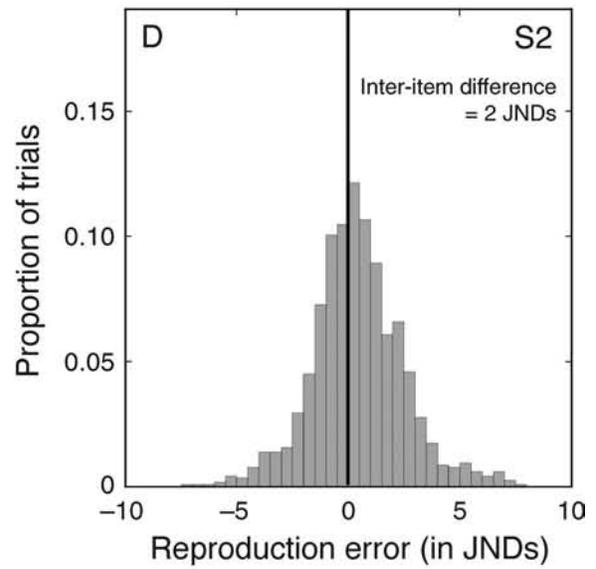
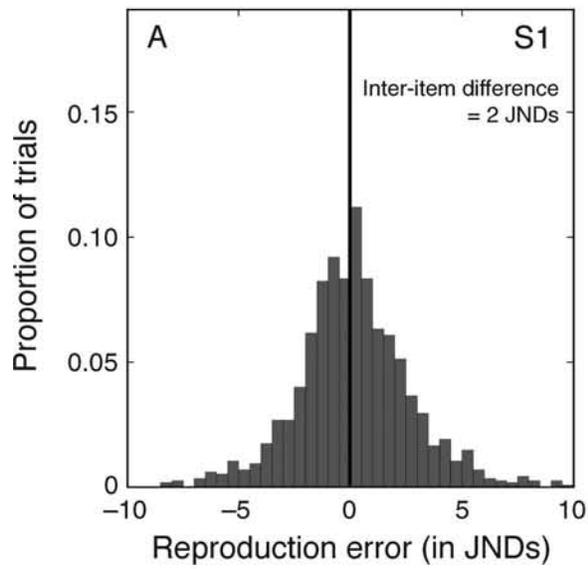
Procedure

Each subject served in three blocks, all within a single session. A block comprised 312 trials, with the first 24 comprising practice trials that were excluded from all data analyses. Of the 864 experimental trials from each subject, 288 were devoted to each inter-item difference. The experimental design called for reproduction of s_1 on half the trials and s_2 on the other half. These two types of trials, as well as trials from the three different inter-item differences, were randomly intermixed within a block.

Results and discussion

In order to evaluate the Non-Target stimulus' influence on the reproduction of the Target stimulus, each trial's reproduction error was algebraically transformed further in order to reflect the relationship between the spatial frequencies of the two stimuli on that trial. Recall that trials on which Target's frequency was the higher one were randomly intermixed with trials on which Non-Target's frequency was the higher one. As a result, if errors were to be algebraically summed with no further transformation, any putative effect of the Non-Target stimulus would be obscured, as on some trials the effect would be toward a higher frequency, while on other trials the effect would be toward a lower frequency.

We therefore transformed each error made in reproducing the Target stimulus' spatial frequency so that the transformed error's sign corresponded to the direction of the error relative to the Non-Target's spatial frequency. Specifically, when the Target's spatial frequency was higher than that of the Non-Target, the sign of the reproduction error was inverted; no change was made to the error's sign when the Target's spatial frequency was



lower than that of the Non-Target. With this manipulation, a positive reproduction error means that the reproduced spatial frequency was shifted *toward* the Non-Target, and a negative reproduction error means that the reproduced spatial frequency was shifted *away* from the Non-Target. With this transformation, if reproductions of the Target were biased toward the Non-Target, the distribution of errors would be centered on some value >0 .

In order to determine if the presence of the Non-Target affected the recall of the Target stimulus, we examined the distributions of errors associated with reproductions in each of the six experimental conditions (3 inter-item differences \times 2 Target serial positions, either s_1 or s_2). Figures 6A–6F show the results for all six conditions. Distributions in the left column represent the errors made when s_1 was the Target; distributions in the right column represent errors made when s_2 was the Target. Proceeding from the top row of panels (Figures 6A and 6D) to the bottom row (Figures 6C and 6F), the frequency difference between Target and Non-Target increases from 2 to 6 JNDs. The black vertical line in each panel marks the value corresponding to zero error.

Note that in Figure 6 as the difference between the spatial frequencies of Target and Non-Target increases, the sign-adjusted reproductions shift systematically rightward. This rightward shift suggests that the Target's remembered spatial frequency shifts increasingly toward the spatial frequency of the Non-Target. Figure 6 reveals another dimension of the results. Comparing the three distributions in the right column to their counterparts in the left column, it seems that the dispersion of errors when Target was s_1 is somewhat greater than when Target was s_2 . We postpone until later a discussion of this phenomenon, focusing for now on the measures of central tendency.

Reproduction errors: Central tendency

Figure 7A shows the mean shift associated with varying inter-item difference in spatial frequency. Values shown are means of individual subjects' median reproduction errors. Separate curves are shown for s_1 as Target (\square) and for s_2 as Target (\blacksquare). Note first that for every condition the reproduction error is positive, showing that the reproduced spatial frequency of Target was consistently shifted toward the spatial frequency of the Non-Target study item. Note also that the median reproduction error increases systematically with the inter-item difference

Figure 6. Frequency histograms of reproduction errors made in various conditions in Experiment 2. In each histogram, the solid black vertical line marks the value of zero, that is, no error. (A–C) Histograms of errors made in reproducing s_1 , at inter-item differences of 2, 4, and 6 JNDs, respectively. (D–F) The errors for s_2 , with inter-item differences of 2, 4, and 6 JNDs, respectively.

between study items. This result was confirmed in an ANOVA by the significant linear component of the effect of inter-item difference, $F_{\text{Linear}}(1,7) = 52.375$, $p < 0.001$.

To facilitate comparisons with results from Experiment 1, Figure 7A also replots from an earlier figure the median reproduction errors for Experiment 1's vision-based matches and memory-based matches. In that experiment, only a single Gabor was presented on each trial. In all six of Experiment 2's conditions, reproduction errors were significantly larger than those in either the vision-based condition or the memory-based conditions in Experiment 1, $p < 0.05$.

The results shown in Figure 7A are consistent with the proposition that the Non-Target affects memory for that trial's accompanying Target. When a second study item is present, even though it is task-irrelevant, subjects' reproductions of the Target shift systematically toward the spatial frequency of that task-irrelevant Non-Target stimulus. Moreover, this bias grows with the spatial frequency difference between Target and Non-Target, with larger differences giving rise to more substantial bias.

However, before concluding that the observed bias reflects only the influence of that trial's Non-Target, an alternative possibility must be considered. Recall that Experiment 1 revealed the existence of a Prototype effect: the memory for a single Gabor stimulus was shifted toward the average spatial frequency of all previously seen study items (see also Lu, Williamson, & Kaufman, 1992). Thus, it might be that this Prototype effect, either alone or in combination with the influence of a trial's remembered Non-Target, was responsible for the bias shown in Figure 7A. To evaluate the possible effect of the prototypical stimulus, we divided the trials within a block of trials into two sets. Specifically, for each subject we separated trials (1) on which the two possible effects would work in the same direction, from trials (2) on which the two effects would oppose one another. In the first set of trials, each trial's Non-Target's frequency and the mean frequency of all the items seen on preceding trials bore the same ordinal relationship to the Target stimulus' spatial frequency on that trial (e.g., both values were higher than Target's frequency, or both were lower); in the other set of trials, the Non-Target's frequency and the mean frequency had opposite ordinal relationships to the Target stimulus' spatial frequency (e.g., one value was higher, and the other was lower than Target's frequency).

Equations 2 and 3 present more formally the rationale behind our division of trials into two sets. In those equations, $\text{Error}_{\text{Sdir}}$ and $\text{Error}_{\text{Odir}}$ are reproduction errors from trials in the first and second sets of trials, respectively; $\text{Bias}_{\text{Non-Target}}$ is the magnitude of reproduction bias toward the Non-Target, and $\text{Bias}_{\text{Proto}}$ is the magnitude of the Prototype effect, that is, the bias toward the mean spatial frequency seen on preceding trials. On trials of the first type, because the two components' influences operate in the same direction, the observed reproduction error reflects a sum of the two separate

influences; on trials of the second type, the two components' influences operate in opposite directions, so the resulting reproduction error reflects the difference between the two influences. As the sign of reproduction errors had been previously adjusted so that a positive-signed error represents a reproduction bias toward Non-Target, the difference between the two components' influences in Equation 3 is found by subtracting $Bias_{Proto}$ from $Bias_{Non-Target}$. Thus,

$$Error_{Sdir} = Bias_{Non-Target} + Bias_{Proto}, \quad (2)$$

$$Error_{Odir} = Bias_{Non-Target} - Bias_{Proto}. \quad (3)$$

If $Bias_{Proto}$ influenced subjects' reproductions, errors in the two sets of trials should differ, which turns out to be the case. Figure 7B shows the averaged medians of subjects' reproduction errors on trials of the two types, as a function of inter-item difference. Separate curves are shown for trials on which s_1 was Target and for trials on which s_2 was Target. The upper pair of curves in Figure 7B represent trials on which the two influences worked in the same direction; the lower pair of curves come from trials on which the two influences worked against each other. The mean value from the upper pair of curves, $Error_{Sdir}$ ($M = 0.976$, $SD = 0.194$) was significantly larger than $Error_{Odir}$ ($M = -0.114$, $SD = 0.112$), as confirmed by a significant main effect of trial type, $F(1,14) = 18.024$, $p < 0.004$ (three-way repeated-measures ANOVA). Thus, the average stimulus on preceding trials does indeed influence the reproduction errors, as shown in Figure 7A.

Having established that memory-based matches are affected both by the Non-Target stimulus on a trial, and by the average or prototypical stimulus, we quantified each of these components separately. To do this, we rearranged the terms in Equations 2 and 3, producing Equations 4 and 5. Referring back to the components comprising Equations 2 and 3, one can see that Equation 4 isolates the contribution of the Non-Target, nulling the contribution from the prototypical stimulus, while Equation 5 does the opposite, nulling the contribution of Non-Target stimulus:

$$Bias_{Non-Target} = 0.5(Error_{Sdir} + Error_{Odir}), \quad (4)$$

$$Bias_{Proto} = 0.5(Error_{Sdir} - Error_{Odir}). \quad (5)$$

Applying these new equations to individual subjects' results allows us to estimate both influences, $Bias_{Non-Target}$ and $Bias_{Proto}$, for each subject in all six conditions. The

average bias estimates are shown in Figure 7C. The left panel shows the errors attributable to the Non-Target stimulus; the right panel shows the influence of the prototypical stimulus. Both are plotted against inter-item difference, with separate curves for trials on which s_1 was Target and for trials on which s_2 was Target. The data points at the rightmost side of Figure 7C are estimates of the Prototype effect derived from Experiment 1's memory-based matches. In that experiment, just a single study item was presented on each trial, which by definition rules out the possibility of an influence from another study item on that trial. As explained earlier, the longer retention interval (2400 ms) and the shorter retention interval (1400 ms) used for Experiment 1's memory-based matches were identical to Experiment 2's retention intervals for s_1 and s_2 , respectively.

The influence of the Non-Target stimulus. The Non-Target stimulus' influence on subjects' reproductions of the remembered Target is shown in the left-hand section of Figure 7C. Estimates of bias toward the Non-Target were entered into a two-way repeated-measures ANOVA. The significant main effect for the serial position of the Target stimulus, $F(1,7) = 9.314$, $p < 0.019$, confirms that the bias toward Non-Target in reproducing s_1 ($M = 0.339$, $SD = 0.072$) was generally smaller than the error made in reproducing s_2 ($M = 0.514$, $SD = 0.103$). In addition, for either serial position the reproduction bias toward the Non-Target showed a significant linear increase with inter-item difference, $F_{Linear}(1,7) = 17.069$, $p < 0.001$, and $F_{Linear}(1,7) = 6.268$, $p < 0.011$, for s_1 and s_2 , respectively. These results demonstrate that the presence of a task-irrelevant study item influences the reproduction of the Target, and also that the influence grows with the difference in spatial frequency between Target and Non-Target.

The influence of the prototypical stimulus. The right-hand panel in Figure 7C represents the prototypical stimulus' influence on the reproduction of a Target stimulus held in short-term memory. There, this influence is plotted as a function of inter-item difference, with separate curves shown for trials on which s_1 was Target, and for trials on which s_2 was Target. The Prototype effect that we extracted by algebraic isolation was comparable in size to its directly measured counterparts in Experiment 1's memory-based matches. Neither for s_1 nor for s_2 were the differences significant, $F(3,21) = 0.257$, $p = 0.855$ for s_1 and $F(3,21) = 1.174$, $p = 0.343$ for s_2 (one-way ANOVAs). This similarity validates the logic, embodied in Equations 4 and 5, used to separate sources of bias in Experiment 2. Furthermore, neither the serial position of the Target nor inter-item difference significantly influenced the Prototype effect, $F(1,14) = 1.636$, $p = 0.242$ for the effect of Target's serial position and $F(2,14) = 1.122$, $p = 0.353$ for the effect of inter-item difference (two-way repeated-measures ANOVA). These results show that when subjects attempt to match the spatial frequency of a remembered Target, their matches shift toward the

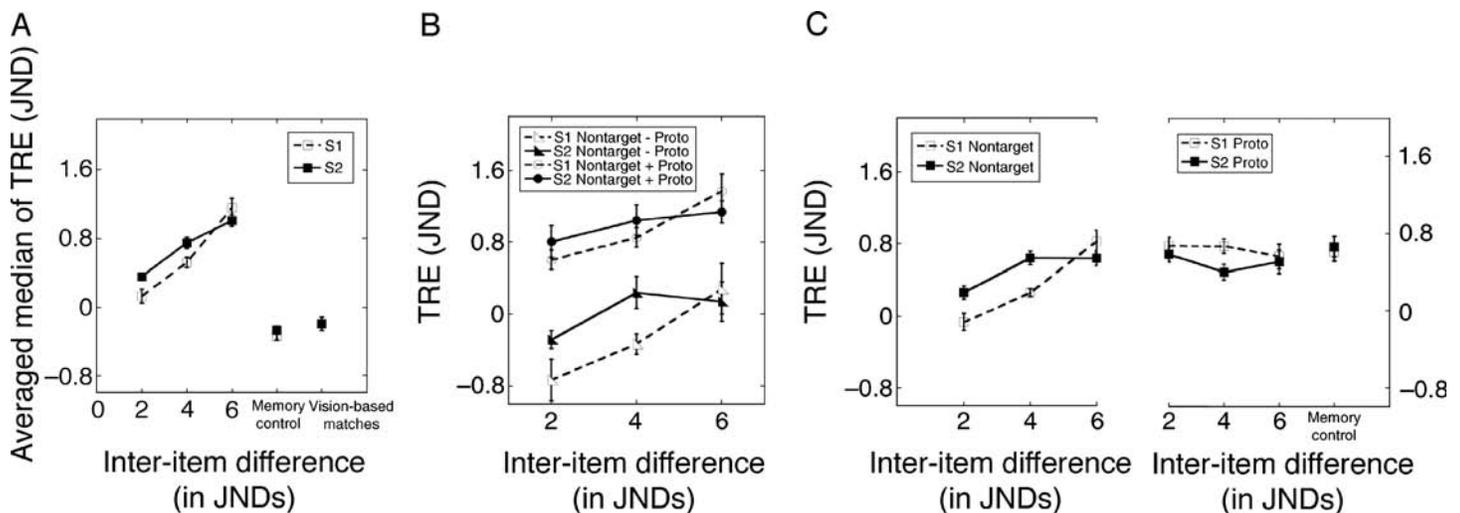


Figure 7. (A) Averaged median reproduction errors for s_1 and s_2 as a function of inter-item difference. Data are for all eight subjects. The median reproduction errors from vision-based and memory-based (with a shorter interval or with a longer interval) performances in Experiment 1 are plotted as perceptual and memory controls. (B) Averaged median reproduction errors for s_1 and s_2 from two different sets of trials plotted separately as a function of inter-item difference. In one set of trials, the bias toward Non-Target and the bias toward the prototypical stimulus were in the same direction; in the second set, these two biases were toward different directions. (C) Left shows the average bias toward Non-Target for s_1 and s_2 as a function of inter-item difference. Right: the average bias toward the mean frequency of all preceding items within that block. The bias toward the mean frequency obtained in Experiment 1 when studying a single stimulus was plotted as memory control. Error bars are ± 1 within-subject standard errors of the average value.

mean frequency seen on previous trials. Note that the size of this Prototype effect was essentially the same when subjects studied just a single stimulus on each trial (in Experiment 1) or studied two stimuli on each trial (in Experiment 2). As the Prototype effect presumably reflects an influence aggregated over many trials, it is not surprising that it is unaffected by the characteristics of the task-irrelevant stimulus on individual trials.

Reproduction errors: Variability

Figure 6 suggests that dispersions of errors made in reproducing s_1 are larger than the corresponding dispersions with s_2 . To quantify this effect and to identify its possible relationship to the sources of bias discussed in the preceding paragraphs, we began with MAD, the median absolute deviation from the median error, for each subject and condition. Values of MAD were averaged over subjects, producing the results in Figure 8A, where average MAD values are shown as a function of the inter-item difference between study items. Results are presented separately for trials on which s_1 was reproduced and trials on which s_2 was reproduced. The results from vision- and memory-based matches in Experiment 1 are also shown.

As suggested earlier in connection with Figure 6, reproductions of the Target were significantly more variable when s_1 was Target than when s_2 was Target

($M = 1.446$, $SD = 0.132$, and $M = 1.232$, $SD = 0.105$, respectively, $F(1,14) = 7.622$, $p < 0.028$, two-way repeated-measures ANOVA). This difference in variability is consistent with the idea that s_1 's representation in memory is more variable than that of s_2 . Moreover, the difference in variability was not modulated by inter-item difference, indicating that reproduction variability is independent of the similarity of Target and Non-Target.

Recall that Experiment 1 used retention delays of both 1400 and 2400 ms. Each of these delay intervals was equivalent to one of the effective pre-reproduction delays in Experiment 2: a shorter delay when a subject was cued to reproduce the more recent study item, s_2 (1400 ms), and a longer delay when the subject was cued to reproduce the earlier of the two study items, s_1 (2400 ms). These equivalences allow us to discount effects associated with the passage of time per se, isolating as a residual any effect of having seen two rather than one study item. There was no difference in the variability produced by the four conditions that shared a 1400-ms delay (three inter-item difference conditions in Experiment 2 and one in Experiment 1), $F(3,21) = 0.396$, $p = 0.757$, one-way ANOVA. In addition, there was no difference among the four conditions that share the longer 2400-ms delay ($F(3,21) = 1.967$, $p = 0.15$). So the presentation of a second stimulus has little influence on the variability of memory for s_1 or s_2 .

In summary, Experiment 2 shows that the central tendency of errors made in reproducing a Target stimulus

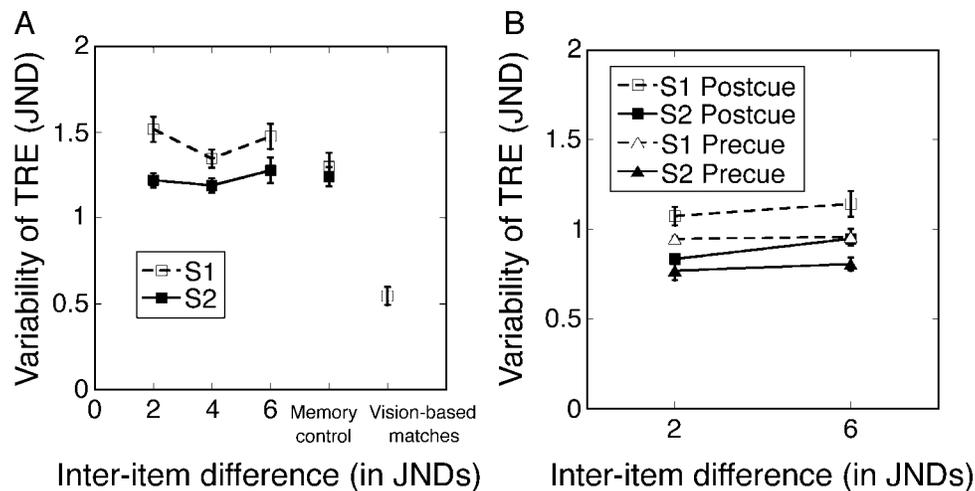


Figure 8. (A) Mean variability of reproduction errors across subjects in Experiment 2. Separate curves are for s_1 and s_2 , both as a function of inter-item difference. Results from Experiment 1's memory and perceptual control conditions are also shown. Error bars are ± 1 within-subject standard errors of the mean value. (B) Mean variability of reproduction errors for pre-cue and post-cue conditions in Experiment 3, separately for trials on which s_1 was Target and trials on which s_2 was Target.

from memory was influenced by two distinct sources: the average characteristics of the stimuli on many different trials, and the characteristics of the task-irrelevant, Non-Target stimulus on the current trial. The influence of the average or prototypical stimulus was unaffected by the inter-item difference, which varied randomly over trials. In contrast, the other influence was strongly modulated by the spatial frequency difference between Target and Non-Target. In addition, when subjects retained two stimuli in memory, the representation of the stimulus (s_1) held longer in memory was more variable than that of the stimulus (s_2) that had been in memory for less time. That this difference between conditions could be resolved despite the very small difference (only 1 s) in the retention intervals for the two stimuli suggests that some factor other than the passage of time was at work, e.g., some interaction between stimuli in memory.

Experiment 3

The preceding experiments revealed two distinct, task-irrelevant influences on recall: the characteristics of a particular trial's Non-Target stimulus and the average properties of stimuli on previous trials. In order to test our interpretation of these empirical effects, we manipulated subjects' selective attention, adapting Yotsumoto and Sekuler's (2006) strategy of exploiting selective attention as way to diminish a task-irrelevant stimulus' impact on recognition memory. Based on Yotsumoto and Sekuler's findings, we hypothesized that if an attentional-directing cue were timely, selective attention could modulate

processing of a Non-Target stimulus. In turn, this modulation would reduce the biasing effect of that Non-Target stimulus. This same attention-driven modulation, however, should leave the Prototype effect untouched. Over trials, Target and Non-Target stimuli had been drawn from the same distribution, so even if selective attention effectively shut out every Non-Target stimulus, the mean frequency of stimuli that were attended to would closely match that of all stimuli, those that were attended and those that were ignored. As a result, the Prototype effect would be unaffected by attention-driven modulation of Non-Target stimuli.

Additionally, by assessing selective attention's impact on short-term memory we addressed a question raised by Yotsumoto and Sekuler's (2006) study of short-term visual memory. Because their results depended upon binary recognition responses ("Yes" or "No") rather than analogue, direct matching responses like those used here, Yotsumoto and Sekuler could not gauge the impact that selective attention might exert on individual trials. In their study, the measurement of any attentional effects depended upon averaging binary responses over many trials. As a result, Yotsumoto and Sekuler acknowledged that they could not distinguish between (1) some relatively modest effect that operated uniformly across all trials, and (2) a strong effect that operated probabilistically, affecting only some random subset of trials. In addition, their dependent measure was the proportion of correct recognition responses, which made it difficult to translate the observed effect from units more meaningful, stimulus-based units, such as *JNDs*.

To manipulate the attentional resources given to one Gabor in a pair of rapidly presented, successive Gabors, we varied the timing of a numerical cue. On some trials a

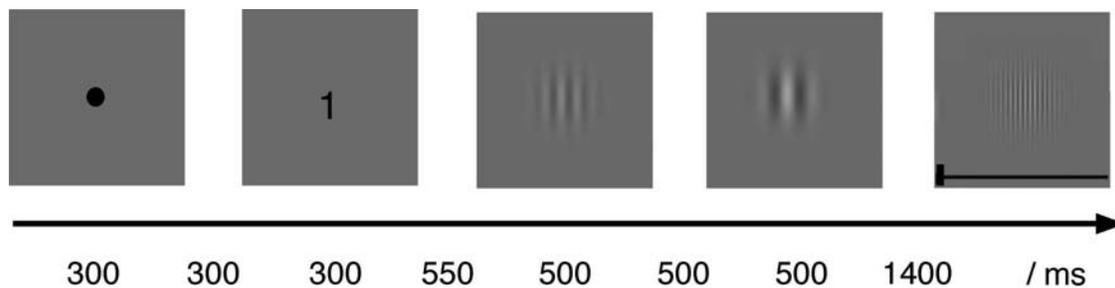


Figure 9. The sequence of events on a trial in [Experiment 3](#). A sample trial from the pre-cue condition is shown. A subject saw a numerical cue (“1” or “2”), which was followed by two Gabors presented in rapid succession. In the example illustrated, the cue was “1”, which signified that s_1 was the Target Gabor to be remembered and reproduced; s_2 was the Non-Target stimulus and could be ignored. Here, the subject reproduced s_1 by adjusting the Comparison Gabor’s spatial frequency until it matched the remembered frequency of s_1 .

digit (“1” or “2”) was presented *after* both study items. The digit cued the subject as to which stimulus, s_1 or s_2 , was the Target stimulus that was to be reproduced from memory. That post-stimulus cue (a post-cue) replicated the cue condition in [Experiment 2](#). In another condition, a digit *preceding* the study items cued the subject which stimulus would be the Target. The digit cue in this pre-cue condition defined a Non-Target item as task-irrelevant before that item was presented and, therefore, potentially allowed a subject to selectively attend to and remember just one stimulus, s_1 on some trials, s_2 on others. So, although the same number of stimuli was presented on every trial and although the timing of the stimuli was constant, the pre- and post-cue conditions differed in the number of items in memory.

Subjects

Eight subjects participated in [Experiment 3](#); five were male. One subject had participated in the previous experiments. Subjects ranged in age from 18 to 22 years; all had normal or corrected-to-normal vision as measured with Snellen targets and normal contrast sensitivity as measured with Pelli–Robson charts (Pelli et al., 1988). Subjects were naive to the purpose of the experiments and were paid for their participation.

Apparatus and stimuli

[Experiment 3](#) used the same apparatus and stimuli as [Experiment 2](#). At the outset, each subject’s spatial frequency discrimination threshold was measured by the same method that had been used previously. After a subject’s Weber fraction for spatial frequency was known, it was used to scale the spatial frequencies of stimuli presented to that subject in the experiment. These values were also used to normalize subjects’ reproduction errors.

Methods

The post-cue and pre-cue conditions were presented in a randomized block design. As in [Experiment 2](#), a subject’s task was to reproduce, from memory, the stimulus (either s_1 or s_2) that the numerical cue designated as the Target. These memory-based matches were carried out as in the preceding experiments. Four subjects were tested with an increasing mapping of location on the adjustment bar to spatial frequency; the other four subjects were tested with the reverse mapping. Generation of the Gabor stimuli followed the method used in [Experiment 2](#), except that only two inter-item differences, 2 and 6 *JNDs*, were tested rather than three. Within a block of trials, the two values of inter-item difference were randomly intermixed over trials. Finally, within a condition, pre-cue or post-cue, the Target (s_1 or s_2) varied randomly from trial to trial. Other details are given below.

Post-cue and pre-cue conditions. As the post-cue condition replicated a condition from [Experiment 2](#), the sequence of events for a post-cue trial has already been described (see [Figure 5](#)). [Figure 9](#) illustrates the sequence of trial events in the pre-cue condition. The numerical cue (“1” or “2”) was presented at the start of each trial, that is, before the first of the study items. As subjects knew ahead of time, due to the cue, which study item would be the Target item, they could attend to and remember just that one study item. Subjects were explicitly instructed to view but ignore the task-irrelevant Non-Target. As the figure shows, a trial began with a fixation point presented for 300 ms at the center of the display. This was followed 300 ms later by a numerical cue that indicated whether s_1 or s_2 would be that trial’s Target item. The cue was presented for 300 ms, followed 550 ± 50 ms later by the sequence of study stimuli, s_1 and s_2 , each displayed for 500 ms. Finally, after a delay of 1400 ± 100 ms, a Comparison Gabor appeared at the center of the display, together with an adjustment bar just below it. The Comparison Gabor and adjustment bar remained visible until the subject

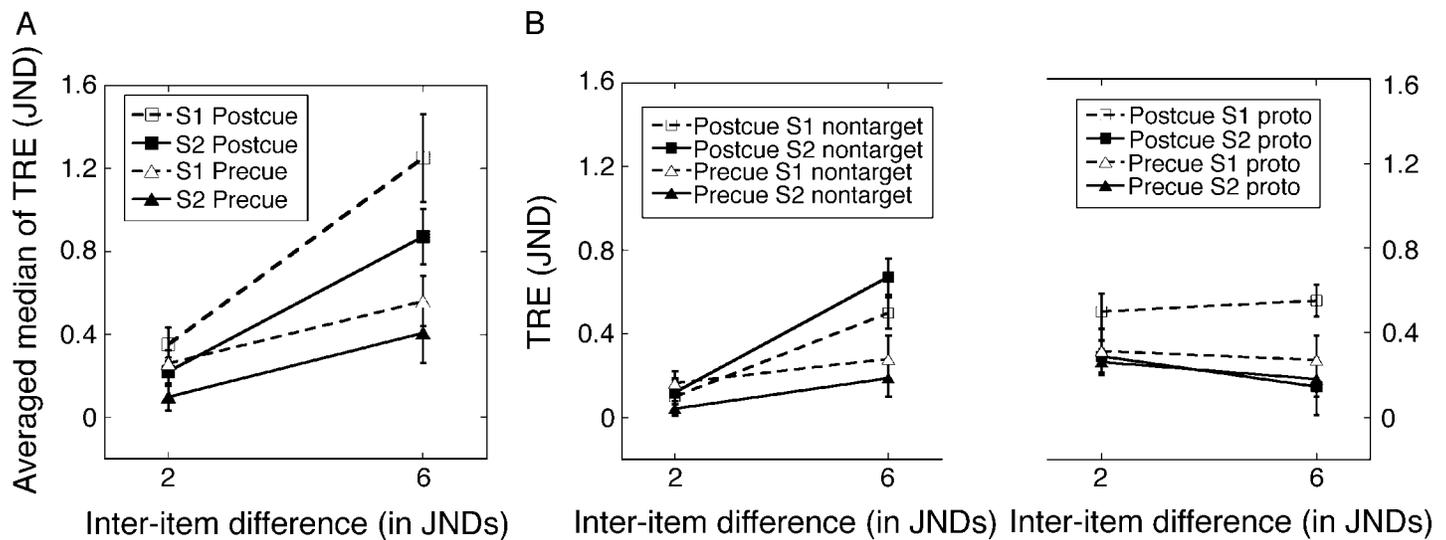


Figure 10. (A) Averaged median reproduction errors for s_1 and s_2 as a function of inter-item difference. Results from post-cue and pre-cue conditions are shown separately. (B) The left-hand section shows the average bias toward Non-Target for s_1 and s_2 , reflecting the influence of the Non-Target stimulus; the right-hand section shows the average bias toward the mean frequency of all observed items on previous trials. In each panel, error bars represent ± 1 within-subject standard errors of the average value.

finished reproducing the Target stimulus and had pressed the designated key to initiate the next trial. The intervals in the pre-cue condition for which s_1 and s_2 had to be retained in memory matched the corresponding intervals in the post-cue condition. This equated the conditions on the basis of the times that separated the presentation of the Target from the start of the matching process.

Procedure

Each subject was tested in four 312-trial blocks. The first 24 trials of each block constituted practice and were excluded from data analysis. Two of the blocks of trials were devoted to the pre-cue condition, and two were devoted to the post-cue condition. The order of blocks was randomized anew for each subject. Each combination of inter-item difference, cue condition, and Target serial position was presented equally often among a subject's 1152 trials, which resulted in 144 trials per combination.

Results and discussion

Using the same methods that were applied to data from Experiment 2, reproduction errors were normalized to individual subjects' Weber fractions for spatial frequency, and the signs of individual reproduction errors were adjusted to represent the direction of the Non-Target's influence. In each condition, the central tendency and

variability of the transformed reproduction errors for each subject were represented by the errors' median and MAD, respectively.

Reproduction errors: Central tendency

Figure 10A shows the mean of subjects' median reproduction errors in post-cue and pre-cue conditions. Results for s_1 and s_2 are shown separately, plotted against the inter-item difference between a trial's stimuli. Of particular interest is any difference between the post-cue and pre-cue conditions. Figure 10A suggests that overall errors are reduced in the pre-cue condition. Recognizing that the values of errors in Figure 10A probably reflect some mixture of two influences, one from the Non-Target and the other from prototypical stimulus on preceding trials, we set out to examine each of these influences separately. To isolate the influences from one another, each value of the errors was algebraically transformed using the procedure introduced for Experiment 2. Equations 4 and 5 were then used to decompose each reproduction error into the two components of bias. Figure 10B shows the outcome of this decomposition. Values in the left panel represent the bias toward a trial's Non-Target stimulus; values in the right panel represent the bias toward the mean spatial frequency of stimuli presented on preceding trials within that block. In each panel, values are plotted against inter-item difference.

The influence of the Non-Target. The numerical cue presented before the study items could have allowed

selective attention to filter out that trial's task-irrelevant Non-Target stimulus. Therefore, we hypothesized that in the pre-cue condition, the bias toward the Non-Target stimulus would be diminished. If selective attention had filtered out the Non-Target entirely, there would be no bias whatever toward a trial's Non-Target in the pre-cue condition.

As expected, the Non-Target's influence was significantly affected by the timing of the cue: the Non-Target's influence in the pre-cue condition was significantly less than in the post-cue condition ($M = 0.168$, $SD = 0.035$ and $M = 0.348$, $SD = 0.063$, respectively). This pattern was confirmed by a significant main effect of cue timing in a three-way repeated-measures ANOVA, $F(1,7) = 6.366$, $p < 0.04$. Moreover, the curves in the left panel in [Figure 10B](#) show that the effect of inter-item difference differed between pre-cue and post-cue conditions ($F(1,7) = 8.742$, $p < 0.021$ for the interaction between inter-item difference and cue timing). Specifically, with a post-cue, the shift toward the Non-Target increased significantly with inter-item difference; but with a pre-cue, there was no such effect ($F(1,7) = 56.019$, $p < 0.001$, and $F(1,7) = 1.384$, $p = 0.278$, respectively). Although selective attention successfully diminished the Non-Target's impact by approximately $2\times$, selective attention's influence was incomplete. As the left-hand panel in [Figure 10B](#) shows, when the cue preceded both study items, the Non-Target's influence on s_1 was greater than zero for each of the inter-item differences, $t(7) = 2.867$, $p < 0.024$ for 2 JNDs, and $t(7) = 2.478$, $p < 0.042$ for 6 JNDs.

The influence of the prototypical stimulus. Previously, [Experiments 1](#) and [2](#) showed that memory-based reproductions of the Target stimulus were biased toward the mean spatial frequency of stimuli on preceding trials. Here we ask whether that bias was affected by variation in selective attention to one of the two stimuli on a trial. As explained earlier, even if selective attention effectively eliminated one stimulus on each trial, over trials, the mean of the remaining stimuli would closely approximate the mean of all the stimuli. Therefore, we hypothesized that the Prototype effect would be unaffected by selective attention.

Prior to comparing the prototypical stimulus' influence in the pre-cue and post-cue conditions, we sought confirmation of the effect's presence. In right-hand section of [Figure 10B](#), reproduction errors tend to be positive, a sign of the Prototype effect. Next, we compared reproduction errors that came from the Prototype effect in the post-cue and pre-cue conditions. In right-hand panel of [Figure 10B](#), curves for s_1 and curves for s_2 suggest that the bias toward the mean stimulus with s_1 was generally larger than with s_2 ($M = 0.409$, $SD = 0.07$ and $M = 0.219$, $SD = 0.074$, respectively). This difference showed up as a significant main effect of Target serial position, $F(1,7) = 13.806$, $p < 0.007$ (in a three-way repeated-measures ANOVA). Moreover, this difference between s_1 and s_2

seemed to come mainly from the post-cue condition. Confirming this observation, although the post-cue condition produced a significant difference between s_1 and s_2 , no such difference was seen in the pre-cue condition ($F(1,7) = 40.869$, $p < 0.001$ and $F(1,7) = 1.017$, $p = 0.347$, for post-cue and pre-cue, respectively). As expected, the change in cue timing had no appreciable influence on the bias toward the prototypical stimulus, $F(1,7) = 0.582$, $p = 0.471$ for the main effect of cue timing.

The results confirm that on any trial, the prototypical stimulus' influence was unaffected by the deployment of selective attention. In the pre-cue condition, the cue's timing made it possible for a subject to focus attention on the one, Target, stimulus that would be task relevant. In that condition, the prototypical stimulus affected memory-based matches to s_1 and s_2 equally, despite the fact that the sequential presentation of the stimuli means that one would have been in memory longer than the other. The absence of an effect replicates the null effect of delay in [Experiment 1](#), when only a single stimulus had to be remembered. However, in the post-cue condition of [Experiment 3](#), when subjects' attention presumably was divided between the two study stimuli, the prototypical stimulus' influence was larger on s_1 , the Gabor that had been held in memory longer.

Reproduction errors: Variability

Next, we considered how the deployment of selective attention affected the variability of memory-based matches. [Figure 8B](#) shows the mean variability of reproduction errors for s_1 and for s_2 , in both pre-cue and post-cue conditions. Reproductions of s_1 ($M = 1.030$, $SD = 0.143$) were consistently more variable than those of s_2 ($M = 0.839$, $SD = 0.108$), $F(1,7) = 24.183$, $p < 0.002$. In addition, although there is a hint in [Figure 8B](#) that the timing of the cue affects variability, that effect is small—pre-cue: $M = 0.869$, $SD = 0.100$, and post-cue: $M = 0.999$, $SD = 0.152$ —and fails to reach statistical significance $F(1,7) = 4.558$, $p = 0.07$.

In summary, [Experiment 3](#) confirmed that an informative cue preceding the study stimuli could successfully induce a redistribution of attention, allowing subjects to focus on the task-relevant Target stimulus, while filtering out the Non-Target. This filtering, however, seems to have been incomplete. Although selective attention significantly diminished the Non-Target's influence, under some conditions a small but distinct residue of the Non-Target's influence remained. This result replicates previous results on selective attention's modulation of a task-irrelevant stimulus' effect (Yotsumoto & Sekuler, 2006). As expected, the influence of the prototypical stimulus was unaffected by the cued redistribution of attention. However, when subjects' attentional resource had to be shared between both study stimuli, the prototype effect was more

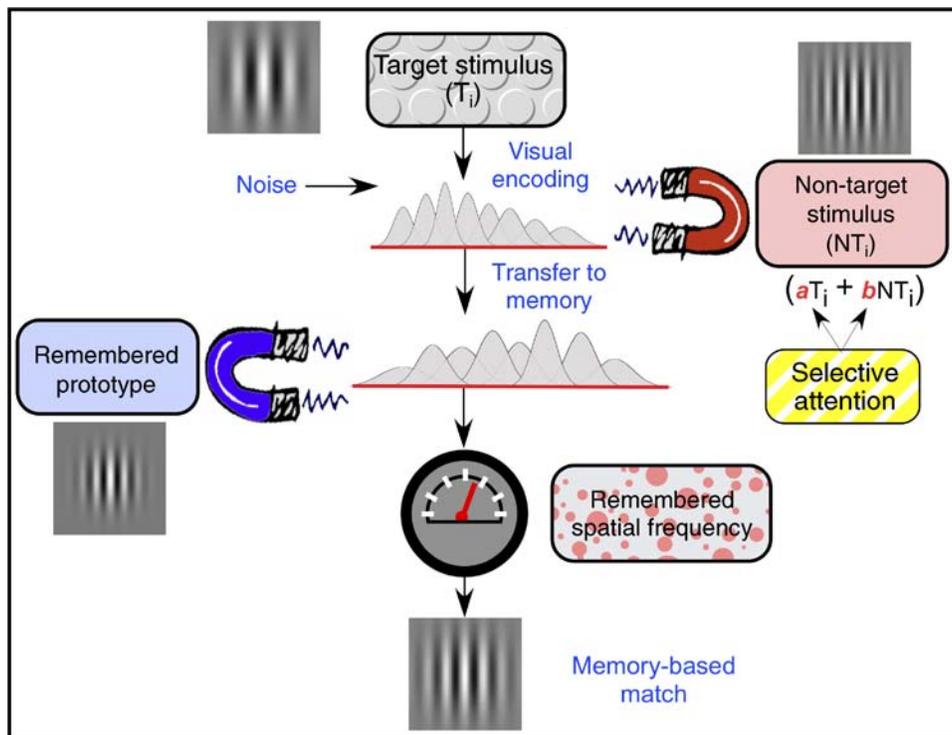


Figure 11. Graphical summary of processes that may be implicated in the present study's results. On the i^{th} trial, the spatial frequency of the Target stimulus (T_i) is encoded by an ensemble of frequency-tuned mechanisms. The spatial frequency of the Non-Target stimulus (NT_i) acts as an attractor on this representation of T_i . In this process, selective attention can bias the relative influences of Target (a) and of Non-Target (b) stimuli. After being transferred to memory, the resulting representation is attracted toward the remembered prototypical stimulus. The outcome of this attraction is read out (indicated by the meter) and directs the subject's reproduction of T_i 's spatial frequency (the memory-based match). Note that the two classes of attractors are shown as influencing distinct stages of processing. As the diagram suggests, our working hypothesis is that (1) the Non-Target effect operates in early visual processing, whereas (2) the Prototype effect operates on the Target stimulus' memory representation.

prominent on the stimulus that had been held in memory longer, and whose reproduction therefore had become more variable.

General discussion

Our experiments confirm the value of analyzing subjects' matches to stimuli, ones that remained visible as well as ones that are remembered. Our experiments also confirm the value of parametric variation and control of the stimulus attributes on each trial. Figure 11 summarizes our main findings in diagram form and identifies some of the processing stages and influences that we believe are responsible for our results. That figure is meant as a guide and graphical supplement to the following discussion.

Experiment 1 demonstrated that vision-based reproductions of spatial frequency were essentially unbiased, clustering as they did around a stimulus' actual spatial frequency. However, as expected, perceived spatial frequency varied from trial to trial, likely reflecting noise in the neural mechanisms that encode the stimulus

(Fredericksen, Bex, & Verstraten, 1997). The activation evoked by a Gabor in each spatial-frequency-selective neural mechanism is represented in Figure 11 by the height of the response functions of mechanisms tuned for spatial frequency. Although it is not crucial for our interpretation of recall from short-term memory, we assume that perceived, and later, remembered, spatial frequency is read out from a distributed, population code (Davis, Yager, King, & Kirkland, 1995; Olshausen & Field, 2004; Tynan & Sekuler, 1974). The figure represents this read-out operation by a meter whose output gives the perceived or remembered spatial frequency. When a Gabor stimulus disappears, visual support for a match is withdrawn, which shifts the basis for any match to short-term memory. Although reproductions continue to cluster around the actual spatial frequency of the just-seen stimulus, those reproductions grow more variable, even with a retention interval of only a second or two (Sperling, 1960). Zhang and Luck (2008) showed that the variability of a color stimulus representation also increased rapidly following stimulus offset. So although the noisiness of a representation in memory may well be constant over the course of several seconds (Zhang & Luck, 2009), the

transition from vision-based to memory-based representation initiates a significant, rapid reduction in fidelity. This abrupt change in fidelity affirms the importance of including a vision-based, perceptual control in studies of visual memory and suggests that it would be useful to examine more closely what happens during the period immediately around the transition from vision to memory. The diagram in [Figure 11](#) portrays the decrease in fidelity as an increase in the variability (spread) within spatial-frequency-tuned mechanisms. Note that [Figure 11](#) depicts the short-term store of remembered spatial frequency as activity within an ensemble of frequency-tuned mechanisms rather than as some single, scalar value. This depiction was chosen because working memory for simple attributes such as spatial frequency seems to depend upon a perceptual memory store situated at an early, retinotopically organized stage of the visual system, possibly including Area V1 (Baumann, Endestad, Magnussen, & Greenlee, 2008; Klein et al., 2004; Kosslyn, Ganis, & Thompson, 2001).

When just a single Gabor was presented and remained visible ([Experiment 1](#)'s vision-based matches), subjects' reproductions showed no bias. However, the situation changed substantially when a pair of Gabor stimuli was presented ([Experiment 2](#)). With a post-stimulus cue designating either the first or the second Gabor as the Target for reproduction, two distinct systematic biases distort subjects' reproductions. One bias appears to operate within the bounds of a single trial; the other operates across trials and reflects a residual from previous trials. Both biases function as attractors, each pulling the Target's remembered spatial frequency toward itself. These two distinct sources of bias are represented in [Figure 11](#) by a pair of magnets. The first of these biases comes from the task-irrelevant Non-Target stimulus that accompanies a trial's Target stimulus. This bias, indexed by a systematic shift in subjects' matches, increases with the difference between the spatial frequency of a trial's Target and that of the Non-Target stimuli. The formula given in [Figure 11](#) portrays the resultant representation of the Target stimulus from this bias as a function of some averaging of the i th trial's Non-Target and Target stimuli, where each type of stimulus carries a role-specific weight (We return later to this intra-trial influence and to the significance of the weights in the formula.) [Experiment 2](#)'s second bias mirrored the prototypical spatial frequency that had been seen on some number of preceding trials. This bias was insensitive to the relationship between Target and Non-Target stimuli within any particular trial. Although our data do not permit a precise description of the way that the two biases are integrated, it appears that each makes its own contribution to subjects' memory-based reproductions. In various conditions, we succeeded in making the two influences either work in the same direction or oppose one another, which suggests that the two influences can sum. The results of [Experiment 3](#) offer yet another reason for treating these two biases as distinct.

In that experiment, when a timely cue preceded both Gabors, selective attention to the task-relevant stimulus curtailed the influence of the other, task-irrelevant stimulus, but had no effect on the Prototype effect. The diagram of [Figure 11](#) depicts selective attention as operating on the representation in visual perception, a view that is consistent with behavioral and physiological demonstrations that selective attention can operate in early visual processing (e.g., Liu, Larsson, & Carrasco, 2007; Martinez-Trujillo & Treue, 2004; Pestilli, Ling, & Carrasco, 2009). Note, however, that our behavioral results are compatible also with the proposition that selective attention has some effect at a post-perceptual stage. It would be valuable to resolve this ambiguity about locus of selection, but as various researchers have noted, attempts at disambiguation can be fraught with difficulty and are contingent upon multiple task-related parameters (Vogel et al., 2005).

Before analyzing further each of two sources of memory bias, we should note the potential implications of our results for the controversy about the nature of visual working memory. This controversy centers on two competing models of memory: a slot-based model that describes working memory as comprising a fixed number, ~ 4 , "slots" (Zhang & Luck, 2008), and an alternative model that treats working memory as resource-limited, rather than slot-limited (Bays et al., 2009; Bays & Husain, 2008). The fact that our experiments never entailed more than two study stimuli per trial means that rarely if ever would the number of items in working memory approach the number of "slots" hypothesized by the fixed "slot" model. However, even with a distinctly sub-span memory load of two items, sensitive measures revealed distinct imperfections in memory. That the memory of one item was so much influenced by the memory of an accompanying item suggests that rather than a series of independent, isolated "slots", visual working memory includes some more flexible process that permits an interaction between representations in memory.

Characteristics of the prototype effect

Before considering further the two biases that affected matches in our experiments, we should explain one of the assumptions made in estimating the value of the variable we have called the Prototype effect. As described earlier, the Prototype effect is defined by a shift in a trial's reproduction toward or away from the mean value of the preceding stimuli in a block of trials. One aspect of this definition needs explanation: the number of trials taken into account when quantifying the value of the prototypical stimulus. The aggregation of many trials in quantifying the reference value for the prototypical stimulus does not mean we believe that the stimuli seen on all those trials are retained in memory, and therefore all actually influence a subject's match. For one thing, regressing trial n 's memory-based match against stimulus

values on trials $n - 1$, $n - 2$, etc., showed that only a few, immediately preceding trials impacted the match to a statistically reliable degree (see [Figure 4B](#)). Of course, the results of the regression analysis do not rule out the possibility that more than just two or three preceding trials exerted some >0 influence, though, as explained earlier, that influence may not have been powerful enough to reach statistical significance.

To appreciate this last point, consider a recent study of the classic psychophysical technique known as the “Method of Single Stimuli” (Woodworth & Schlosberg, 1954). With that method, a subject judges a stimulus against some implicit standard that is built up over trials and held in memory. The memory-based implicit standard is analogous to the prototypical stimulus whose influence was revealed in our experiments. By systematically varying the set of stimuli that a subject sees and judges against the implicit standard, it is possible to track the development of the implicit standard and to estimate the number of trials that are required to build up and maintain a stable implicit standard. Because the stimuli in our experiments varied randomly from trial to trial, our data do not allow a detailed analysis of how many trials actually contributed to the representation of the prototypical stimulus. However, using Monte Carlo simulations to supplement their empirical study, Morgan, Watamaniuk, and McKee (2000) examined the development of subjects’ implicit standard for judgments of the separation between a pair of lines. Although their methods and task differed from ours, their results may bear on subjects’ representation of the prototypical stimulus in our experiments. Morgan et al.’s simulations led them to conclude that a subject’s implicit standard incorporated contributions from as many as 10–20 prior stimuli. We should note that their simulations incorporated one dubious assumption, namely, that all previous stimuli, no matter where in the sequence they occurred, made the same contribution to the implicit standard. As a result, it may be that Morgan et al.’s simulations misestimated the number of trials that actually affected their subjects’ implicit standard. Also relevant to the present work is a signal detection analysis that was carried out by Gorea and Sagi (2000) on “Yes”–“No” data collected in a detection task. Gabor targets of varying contrast were presented either singly, or simultaneously, in pairs, one slightly to the left of fixation and one slightly to right. In the conditions that are most relevant to our study, Gorea and Sagi unbalanced the probabilities with which test targets would appear in the two possible locations. A signal detection analysis demonstrated that instead of maintaining separate and distinct decision criterion for each of the two target locations, subjects used a single, fused criterion to which each stimulus location contributed in proportion to its probability of occurrence. Finally, using a matching task with random dot pattern stimuli, Busemeyer and Myung (1988) were able to generate trial-by-trial estimates of the prototype that was being built up from exposure to noisy

exemplars of a stimulus. Their results show a clear recency weighting of the contribution made to the prototype by the pattern on each trial, with an influence that reached back just three to five trials. Taking account of these results, as well as those of our own regression analysis (presented earlier), our working hypothesis going forward is that memory-based matches for spatial frequency probably reflect a recency-weighted influence of somewhere between three and five preceding stimuli.

The integration of prototype information and vSTM

In [Experiment 3](#), the impact of the prototypical stimulus in the post-cue condition varied with the Target stimulus’ serial position. Specifically, the prototypical stimulus affected memory-based matches to the first stimulus more than it did matches to the second stimulus. A previous study in which direct matches were used to measure memory for spatial frequency may help explain the association between serial position and the magnitude of the Prototype effect. In that study, Wilken and Ma (2004) showed subjects briefly presented arrays of either 2, 4, 6, or 8 Gabors. The Gabors within an array differed from one another in spatial frequency. After the array had disappeared from view, a spatial cue instructed the subject to reproduce the spatial frequency of the Gabor that had occupied the cued location. Because the spatial frequencies represented in their arrays were randomized from trial to trial, it was impossible to extract from the data any systematic effect that the Non-Target Gabors might have had on matches to the Target Gabor. However, Wilken and Ma were able to extract the bias produced by their prototypical stimulus. Figure 8 in their paper showed a striking bias in subjects’ matches. This bias, which constitutes a shift toward the mean of the entire set of spatial frequencies, grew with the number of Gabors in the study array. As the variability in reproductions of spatial frequency also grew with the number of Gabors in a stimulus array, the more variable or unreliable memory was, the greater was the influence of the prototypical stimulus. This connection between variability of memory, on one hand, and memory’s vulnerability to the Prototype effect, on the other, has a counterpart in our study (see post-cue condition in [Experiment 3](#)).

To explain the association between variability in memory and the influence of the prototypical stimulus, Wilken and Ma hypothesized that “observers judge spatial frequencies by combining observed values with an internal template based on prior experience, perhaps through a weighted sum in which the weights increase with increasing variance” (pp. 1130–1131). This suggestion is consistent with reports that when memory for some item is weak or uncertain, that memory seems to be augmented by the “gist” or general characteristics of other stimuli (e.g., Brady & Alvarez, 2010; Koriati, Goldsmith, &

Pansky, 2000; Maryott & Sekuler, 2009). Although the exploitation of gist has often been portrayed as a deliberate, conscious process, Ma, Beck, and Pouget (2008) argued that such behavior may also be the natural product of variability within populations of neurons. These researchers demonstrated that the neural code instantiated in such populations was able to represent simultaneously not only the value of a stimulus, spatial frequency in our experiments, but also the degree of uncertainty associated with that estimate of stimulus value. Ma et al. noted that this dual representation could be the neural basis for Bayes-optimal or near-optimal cue integration that has been found in many psychophysical studies (e.g., Cheng, Shettleworth, Huttenlocher, & Rieser, 2007; Ernst, 2007; Ernst & Banks, 2002; Knill, 2007). The design of our experiments prevents a detailed description of the integration of item and prototype information, but the association in our results between variability and bias is qualitatively consistent with a Bayesian framework.

Whether the integration of prototype and item information proves ultimately to be Bayes-optimal or not, our working hypothesis is that prototype information serves an adaptive function, compensating for imperfections in memory for the Target stimulus. A corollary is that the prototype's influence should be inversely proportional to the degree of those imperfections in memory. This hypothesis is supported by the fact that in [Experiment 1](#), the Prototype effect was seen only with the more variable, memory-based matches and was absent when matches were of the less variable, vision-based variety. Additional support for our working hypothesis comes from [Experiment 3](#)'s post-cue condition. There, matches to s_1 were more variable and showed a correspondingly larger Prototype effect than did matches to s_2 . However, this relationship between variability and Prototype effect was not evident in [Experiment 2](#), suggesting that the predicted relationship between variability and the effect of the prototypical stimulus is not some direct, obligatory product of population coding. Instead, that relationship may prove to be susceptible to cognitive control. As different groups of subjects served in [Experiments 2](#) and [3](#), it may be that individual differences in strategy modulate the use of available prototype information.

Characteristics of the non-target stimulus effect

It has long been recognized that the visual system can integrate information over space and over time. In fact, such integration is among the most-studied attributes of vision (Blake & Sekuler, 2006). Although early studies focused on spatio-temporal integration of low-level attributes such as luminance, integration extends also to higher level stimulus attributes, including motion direction (Watamaniuk, Sekuler, & Williams, 1989) and speed

(Watamaniuk & Duchon, 1992), line length (Morgan et al., 2000), size (Albrecht & Scholl, [in press](#); Brady & Alvarez, 2010; Chong & Treisman, 2003), and emotion of facial expressions (Haberman et al., 2009; Haberman & Whitney, 2007). When multiple stimuli are presented within the receptive fields of different neurons, spatio-temporal activation patterns sometimes exhibit what can be described as “neural averaging,” with the resulting activation approximating an average of the neural responses evoked by the stimuli individually (Chelazzi, Duncan, Miller, & Desimone, 1998; Kastner et al., 2001; Zoccolan, Cox, & DiCarlo, 2005). The influence of a Non-Target stimulus in our experiments can be thought of perceptual averaging, a behavioral correlate of neural averaging, on each trial, of responses evoked by the spatial frequency of the Target Gabor and by the spatial frequency of the Non-Target Gabor. In fact, perceptual averaging has been demonstrated with a range of stimuli and tasks (e.g., Albrecht & Scholl, [in press](#); Chong & Treisman, 2005; Parkes et al., 2001). Sweeny, Grabowecky, Paller, and Suzuki (2009) proposed that perceptual averaging depends upon the relationships among the receptive fields within which stimuli fall. In the case at hand, if Target and Non-Target Gabors were widely separated in space and therefore fell within distinct receptive fields, visual averaging would be foreclosed. As a result, the Non-Target stimulus would exert no influence on the Target stimulus. Some preliminary data from our laboratory are consistent with this hypothesis (Huang & Sekuler, 2010). Those data demonstrate that the influence of the Non-Target stimulus disappears when the Target and Non-Target are presented simultaneously but within different hemifields ($\sim 5^\circ$ apart).

Adopting the perspective of neural and perceptual averaging makes possible an informative, quantitative dissection of the effect of the Non-Target stimulus. To facilitate that dissection, [Figure 12](#) replots the Non-Target effects from [Experiments 2](#) and [3](#). The size of the effect, which in this case is based on [Equation 4](#), is plotted against the *JND* difference between Target and Non-Target stimuli. In addition, for a theoretical benchmark, we have plotted what would be expected if the Non-Target effect were the outcome of an average of Target and Non-Target spatial frequencies, and if each of those spatial frequencies contributed equally to the result. This theoretical benchmark is shown by the solid line through the \blacktriangle symbols. Assuming the total effect from both sources summed to 1.0, the Non-Target's contribution would grow with slope $b = 0.5$ as the two spatial frequencies diverged. All the curves that represent empirical results lie well below this theoretical curve, demonstrating that any perceptual averaging is very far from complete. In addition, the slopes of the empirical curves are well short of $b = 0.5$, demonstrating that in any perceptual averaging in the present study, the Target's contribution dominates the representation on which a subject's match is based, with the Non-Target stimulus

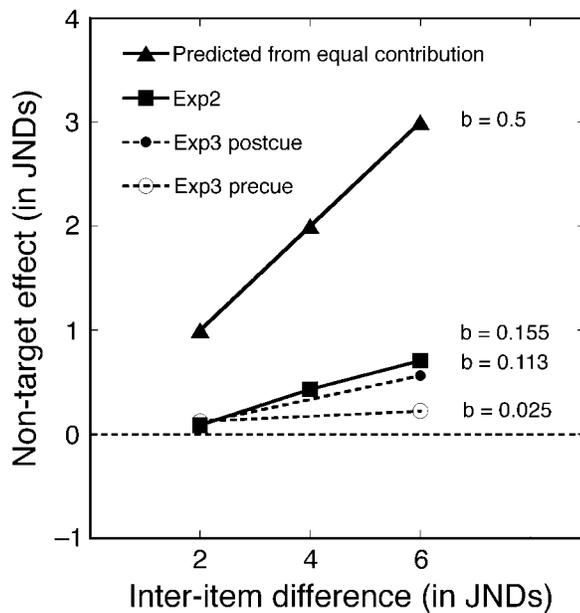


Figure 12. Data points represent the Non-Target's effect as a function of the inter-item difference in spatial frequency between Target and Non-Target stimuli. Data from trials when s_1 was Target and those when s_2 was Target are aggregated. Different lines show data in Experiment 2 (■) and Experiment 3's post-cue (•) and pre-cue (○) conditions. The values of Non-Target's effect predicted for a process in which the reproduction of the Target stimulus was determined by an equal-weighted visual averaging of s_1 and s_2 's spatial frequencies (▲) are also shown. The slope coefficient of each line, b , is shown as an index of the Non-Target's weight in perceptual averaging. Note that b is the weight assigned to the Non-Target stimulus by the formula in Figure 11.

making a far smaller contribution. Although the curves for the post-cue conditions of Experiments 2 (■) and 3 (•) have slopes well below $b = 0.5$, their slopes (0.113–0.155) still exceed the slope ($b = 0.025$) from the pre-cue condition of Experiment 3. Presumably, the timing of the pre-cue makes it possible to nearly nullify any contribution from the task-irrelevant, Non-Target stimulus. Note that this non-zero weight suggests that selective attention is incompletely effective in filtering out the Non-Target stimulus, as discussed in the next section.

Alternative explanations

Figure 11 depicts the bias toward the Non-Target stimulus as coming from some attractor (hence, the magnet) that operates on the internal representation. However, before accepting this depiction, we must consider an alternative account of the underlying behavioral result. Imagine that on some proportion of trials, a subject simply confused the identities of the two study stimuli, and therefore reproduced the Non-Target instead of the Target. Over trials, such confusion would shift the

central tendency of matches toward to the Non-Target, just as was observed. However, there is reason to doubt a connection between confusion and the shift demonstrated in Experiments 2 and 3. The distributions of matches made by individual subjects are inconsistent with the hypothesis that subjects confused Target and Non-Target stimuli. We began an examination of these distributions by dividing the eight subjects from Experiment 2 into two groups, based on the skewness of their matches in the six different conditions (3 levels of inter-item difference \times 2 levels of Target's serial position). If the observed bias in matches toward the Non-Target resulted from confusion between Non-Target and Target, subjects whose matches were more positively skewed, indicating that they had mistakenly reproduced Non-Target more frequently, should have shown a larger bias toward the Non-Target. We averaged each subject's bias toward the Non-Target across the six conditions and compared the two groups' mean bias. The mean bias toward the Non-Target from the group whose distributions were more skewed did not differ from the bias seen in the group whose distributions were less skewed ($M = 0.456$, $SD = 0.211$, with 95% confidence interval $[-0.3506, 0.792]$, and $M = 0.361$, $SD = 0.295$, with 95% confidence interval $[-0.109, 0.852]$, $t(6) = 0.520$, $p > 0.1$, respectively).

At a more-detailed level, individual subjects' distributions of matches showed no sign of a secondary mode in the vicinity of the Non-Target spatial frequency. Although, such a secondary mode would have been difficult to detect with small inter-item differences in spatial frequency, with differences as large as 6 JNDs, the mode should have been quite distinctive. Of course, had the Non-Target's bias come from confusions on only a small proportion of trials, the expected secondary mode could have eluded detection, even with the largest inter-item difference. To assess how many confusions would have been needed to produce our results, we simulated different mixture models, allowing confusions between stimuli to affect varying proportion of trials. For each proportion of mixtures we examined the magnitude of the shift in the resulting distribution of matches. The simulations generated mixtures of two normal distributions, which represented for the distributions of Target and Non-Target memories. The spread of the sampled distributions was set to the empirical value of MAD from subjects' memory-based matches in Experiment 1. To simulate the largest inter-item difference from our experiments, we set the means of the two sampled distributions to 0 and 6. Matches on trials that entailed no confusion between Non-Target and Target were simulated by randomly sampling the memory distribution of Target; matches on trials that entailed confusion of Non-Target and Target were simulated by sampling from the memory distribution of Non-Target. For each of the resulting mixed distributions we used the median to represent the shift produced by the simulated confusions. The results showed that a shift in the distribution of matches as large

as 0.8 *JND*s, the shift observed with an inter-item difference of 6 *JND*s in [Experiment 2](#), would have required confusions on at least 30% of the trials. Confusions as frequent as this would have produced a distinct, readily detectable secondary mode in the distribution of matches. Therefore, our empirical result is inconsistent with the notion that the bias toward a trial's Non-Target mostly resulted from a subject's mistakes about which Gabor, first or second, was actually the Target.

The operation and limits of selective attention

As mentioned earlier, to measure selective attention's effect on visual short-term memory, Yotsumoto and Sekuler (2006) used a basic task and stimuli similar to our own. However, in their study, a subject made one of just two possible responses on each trial: "Yes," signifying that the probe stimulus was recognized, or "No," signifying that the probe stimulus was not recognized. Yotsumoto and Sekuler's demonstration of attentional selectivity required that binary recognition responses be averaged over dozens of trials. Consequently, they could not distinguish between two modes in which selective attention might have operated on any single trial. In one mode, subjects might have been extremely successful in filtering out the task-irrelevant Non-Target, but only on a fraction of all trials; then on the other trials, filtering would have failed. In the second mode, the effect of attentional filtering was partial, but fairly consistent over trials. Due to the analogue, direct matching responses collected in our experiments, these competing accounts can be evaluated. To select between them, we examined the distributions of reproduction errors that subjects made in the pre-cue condition. We divided the eight subjects who participated in [Experiment 3](#) into two groups based on the mean skewness of their matches across four pre-cue conditions (2 levels of inter-item difference \times 2 levels of Target serial position). If the first explanation above were true, that is, if Non-Target's residual influence reflected a failure of selective attention on only a small fraction of trials, there should be an association between (1) the magnitude of the bias toward the task-irrelevant Non-Target item in pre-cue condition, and (2) the number of trials on which attention had completely failed. This latter variable would be reflected in the skewness of the subject's distribution of errors.

To evaluate this hypothesized association, we averaged each subject's bias toward the Non-Target across the four pre-cue conditions and compared the two groups' mean bias toward the Non-Target. The mean bias toward Non-Target from the group whose distributions were more skewed ($M = 0.208$, $SD = 0.114$, with 95% confidence interval [0.0262, 0.390]) did not differ significantly from the bias seen in the group whose distributions were less skewed ($M = 0.139$, $SD = 0.110$, with 95% confidence interval [-0.036, 0.314]), $t(6) = -0.865$, $p > 0.1$. This

result supports the second hypothesis described above, namely that selective attention operates at a fairly steady, if incomplete level, over trials.

As can be seen in [Figure 10B](#), changing the timing of the numerical cue—from post-cue to pre-cue—greatly enhanced the effectiveness of selective attention. In summarizing the operation of selective attention, [Figure 11](#) depicts attention as modulating the relative weights, a and b , given to Target and Non-Target stimuli, respectively. Modulation of the weights alters the outcome of the averaging operation, $aT_i + bNT_i$, and therefore alters the Non-Target stimulus' influence on perception and memory. However, our behavioral results are equivocal with respect to some essential details of how this process would work. When a cue allows selective attention to be focused on the Target, b , the weight associated with the Non-Target stimulus is diminished, which reduces the Non-Target's influence. This account of attentional selectivity parallels ones in the domains of perceptual learning and spatial cuing. For example, Doshier and Lu (2000) showed that in some circumstances spatial cuing enhances perception largely by excluding or diminishing task-irrelevant external noise (see also Zanto & Gazzaley, 2009). This would correspond to a reduction, in our framework, in the value b . However, because we constrained b to be $1 - a$, our behavioral results are equally compatible with a case in which selective attention operates via an increase in a , the weight of the attended Target stimulus. This formulation would be consistent with demonstrations by Carrasco et al. that selective attention increases the gain of neural mechanisms tuned to a cued location or a cued feature, such as spatial frequency (e.g., Liu et al., 2007).

As mentioned earlier, [Experiment 3](#) demonstrated that selective attention curtailed, but did not eliminate, the influence of the task-irrelevant, Non-Target stimulus. This result confirmed what Yotsumoto and Sekuler (2006) found using a recognition paradigm, rather than the matching procedure used here. Specifically, Yotsumoto and Sekuler showed that selective attention was diverted away from a Gabor that occupied a particular serial position in a study list, the influence of that item was reduced. In their study, this reduction in influence was inferred from improved recognition for the attended item. In both studies, the present one and the one by Yotsumoto and Sekuler, selective attention's influence proved to be substantial but decidedly incomplete. In the present study, the partial failure of selective attention could have come from the timing of the numerical pre-cue. In particular, was the pre-cue not delivered sufficiently far in advance of the study items to allow selective attention to have its full effect? In our experiments, the Target Gabor's serial position varied from one trial to the next. In contrast, Yotsumoto and Sekuler varied the Target Gabor's serial position between one entire block of trials to another. That constraint allowed subjects to know well in advance of each trial which item would be task relevant. Our third

experiment presented each trial's pre-cue 550 ± 50 ms prior to the presentation of the first study item. We reasoned that this lead time would have been sufficiently long to allow the full deployment of selective attention (see Ball & Sekuler, 1981). Despite what we think was adequate pre-stimulus readiness, in Yotsumoto and Sekuler's experiments and in our own, selective attention was only partially effective.

This limitation on the effect of selective attention might reflect some irreducible delay in initiating the neural processes that filter out the task-irrelevant stimulus or enhance response to the task-relevant one. For example, if that neural process were triggered by the onset of the task-irrelevant stimulus, any appreciable latency associated with that process would permit some early portion of the stimulus to escape suppression. If this unfiltered portion of the stimulus entered into memory, attentional selectivity would be only partially effective, as we and Yotsumoto and Sekuler observed. Precisely such a delay was recently seen in electroencephalographic recordings made while subjects performed a short-term memory task and attempted to limit interference from a task-irrelevant stimulus attribute (Zanto & Gazzaley, 2009). In that study, the stimulus had already been visible for about 200 ms before the neural signal of selective attention reached its maximum. Although we draw encouragement from Zanto and Gazzaley's results, future research is needed to determine whether the latency of the suppressive process is actually the cause of the imperfect selective attention in our task.

To conclude, a trio of experiments demonstrated that subjects' recall of a stimulus' spatial frequency depends upon an integration of three factors: the actual spatial frequency of the Target stimulus on that trial; the spatial frequency of the accompanying, but task-irrelevant Non-Target stimulus; and the prototypical spatial frequency derived from previously seen stimuli. Each of the last two, memory-distorting influences attracts the remembered Target stimulus toward itself. Our working hypothesis is that the prototypical stimulus influences a stored representation in memory of a trial's Target stimulus. In contrast, we believe that the Non-Target stimulus operates on an earlier, visual stage of processing, perhaps through some form of perceptual averaging. Whatever the fate of these hypotheses proves to be, it is clear from the complex network of elements and interactions postulated by Figure 11 that much work is needed to produce a full understanding of how visual memory can be distorted, in a task like ours, and perhaps in many other tasks as well.

Acknowledgments

This work was supported by the National Institutes of Health Grant MH-068404. A portion of this project was

presented at the 2009 Meeting of the Vision Sciences Society.

Commercial relationships: none.

Corresponding authors: Jie Huang and Robert Sekuler.

Email: jiehuang@brandeis.edu; sekuler@brandeis.edu.

Address: Volen Center for Complex Systems and Department of Psychology, Brandeis University, Waltham, MA, USA.

Footnote

¹Over the years, this method has taken on many names, including "method of reproduction," "method of adjustment," "method of average error," and "method of equivalent stimuli."

References

- Albrecht, A. R., & Scholl, B. J. (in press). Perceptually averaging in a continuous world: Extracting statistical summary representations over times. *Psychological Science*. doi:10.1177/0956797610363543.
- Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science*, *19*, 392–398. [PubMed] [Article]
- Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. *Annals of Mathematical Statistics*, *23*, 193–212.
- Ball, K., & Sekuler, R. (1981). Cues reduce direction uncertainty and enhance motion detection. *Perception & Psychophysics*, *30*, 119–128. [PubMed]
- Baumann, O., Endestad, T., Magnussen, S., & Greenlee, M. W. (2008). Delayed discrimination of spatial frequency for gratings of different orientation: Behavioral and fMRI evidence for low-level perceptual memory stores in early visual cortex. *Experimental Brain Research*, *188*, 363–369. [PubMed]
- Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10):7, 1–11, <http://journalofvision.org/9/10/7/>, doi:10.1167/9.10.7. [PubMed] [Article]
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*, 851–854. [PubMed] [Article]
- Bennett, P. J., Sekuler, R., & Sekuler, A. B. (2007). The effects of aging on motion detection and direction

- identification. *Vision Research*, *47*, 799–809. [[PubMed](#)]
- Blake, R., & Sekuler, R. (2006). *Perception* (5th ed.). New York: McGraw-Hill.
- Brady, T. F., & Alvarez, G. A. (2010). Ensemble statistics of a display influence the representation of items in visual working memory. *Visual Cognition*, *18*, 114–118.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 443–446. [[PubMed](#)]
- Bruner, J. S., Busiek, R. D., & Minturn, A. L. (1952). Assimilation in the immediate reproduction of visually perceived figures. *Journal of Experimental Psychology*, *44*, 151–155. [[PubMed](#)]
- Busemeyer, J. R., & Myung, I. J. (1988). A new method for investigating prototype learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 3–11.
- Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, *80*, 2918–2940. [[PubMed](#)] [[Article](#)]
- Cheng, K., Shettleworth, S. J., Huttenlocher, J., & Rieser, J. J. (2007). Bayesian integration of spatial information. *Psychological Bulletin*, *133*, 625–637. [[PubMed](#)]
- Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*, 393–404. [[PubMed](#)]
- Chong, S. C., & Treisman, A. (2005). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*, *67*, 1–13. [[PubMed](#)] [[Article](#)]
- Cohen, D. J., & Bennett, S. (1997). Why can't most people draw what they see? *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 609–621. [[PubMed](#)]
- Davis, E. T., Yager, D., King, R. A., & Kirkland, B. A. (1995). A labeled lines explanation of the perceived spatial frequency of moderate-, near-threshold- and zero-contrast spatial patterns. *Vision Research*, *35*, 1025–1040. [[PubMed](#)]
- Demeyer, M., De Graef, P., Wagemans, J., & Verfaillie, K. (2009). Transsaccadic identification of highly similar artificial shapes. *Journal of Vision*, *9*(4):28, 1–14, <http://journalofvision.org/9/4/28/>, doi:10.1167/9.4.28. [[PubMed](#)] [[Article](#)]
- Dosher, B. A., & Lu, Z. L. (2000). Mechanisms of perceptual attention in precuing of location. *Vision Research*, *40*, 1269–1292. [[PubMed](#)]
- Ebbinghaus, H. (1885/1913). *On memory: A contribution to experimental psychology*. New York: Teachers College, Columbia University.
- Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision*, *7*(5):7, 1–14, <http://journalofvision.org/7/5/7/>, doi:10.1167/7.5.7. [[PubMed](#)] [[Article](#)]
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433. [[PubMed](#)]
- Fredericksen, R. E., Bex, P. J., & Verstraten, F. A. (1997). How big is a Gabor patch, and why should we care? *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, *14*, 1–12. [[PubMed](#)]
- Gold, J. M., Murray, R. F., Sekuler, A. B., Bennett, P. J., & Sekuler, R. (2005). Visual memory decay is deterministic. *Psychological Science*, *16*, 769–774. [[PubMed](#)]
- Gorea, A., & Sagi, D. (2000). Failure to handle more than one internal representation in visual detection tasks. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 12380–12384. [[PubMed](#)] [[Article](#)]
- Haberman, J., Harp, T., & Whitney, D. (2009). Averaging facial expression over time. *Journal of Vision*, *9*(11):1, 1–13, <http://journalofvision.org/9/11/1/>, doi:10.1167/9.11.1. [[PubMed](#)] [[Article](#)]
- Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, *17*, R751–R753. [[PubMed](#)]
- Helson, H. (1948). Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychological Review*, *55*, 297–313. [[PubMed](#)]
- Hollingworth, H. L. (1910). The central tendency of judgment. *The Journal of Philosophy, Psychology and Scientific Methods*, *7*, 461–469.
- Huang, J., & Sekuler, R. (2010, May). *α -Oscillations and the fidelity of visual memory*. Paper presented at the meeting of the Visual Sciences Society, Naples, FL.
- Kahana, M. J. (in press). *Foundations of human memory*. Oxford University Press.
- Kastner, S., De Weerd, P., Pinsk, M. A., Elizondo, M. I., Desimone, R., & Ungerleider, L. G. (2001). Modulation of sensory suppression: Implications for receptive field sizes in human visual cortex. *Journal of Neurophysiology*, *86*, 1398–1411. [[PubMed](#)] [[Article](#)]
- Klein, I., Dubois, J., Mangin, J.-F., Kherif, F., Flandin, G., Poline, J.-B., et al. (2004). Retinotopic organization

- of visual mental images as revealed by functional magnetic resonance imaging. *Cognitive Brain Research*, 22, 26–31. [PubMed]
- Knill, D. C. (2007). Robust cue integration: A Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *Journal of Vision*, 7(7):5, 1–24, <http://journalofvision.org/7/7/5/>, doi:10.1167/7.7.5. [PubMed] [Article]
- Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology*, 51, 481–537. [PubMed]
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, 2, 635–642. [PubMed]
- Liu, T., Larsson, J., & Carrasco, M. (2007). Feature-based attention modulates orientation-selective responses in human visual cortex. *Neuron*, 55, 313–323. [PubMed] [Article]
- Lu, Z. L., Williamson, S. J., & Kaufman, L. (1992). Behavioral lifetime of human auditory sensory memory predicted by physiological measures. *Science*, 258, 1668–1670. [PubMed]
- Ma, W. J., Beck, J. M., & Pouget, A. (2008). Spiking networks for Bayesian inference and choice. *Current Opinion in Neurobiology*, 18, 217–222. [PubMed]
- Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology*, 14, 744–751. [PubMed]
- Maryott, J., & Sekuler, R. (2009). Age-related changes in imitating sequences of observed movements. *Psychology and Aging*, 24, 476–486. [PubMed]
- Melcher, D., & Colby, C. L. (2008). Trans-saccadic perception. *Trends in Cognitive Science*, 12, 466–473. [PubMed]
- Morgan, M. J., Watamaniuk, S. N., & McKee, S. P. (2000). The use of an implicit standard for measuring discrimination thresholds. *Vision Research*, 40, 2341–2349. [PubMed]
- Nelson, T. O. (1985). Ebbinghaus's contribution to the measurement of retention: Savings during relearning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 472–479. [PubMed]
- Olshausen, B. A., & Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14, 481–487. [PubMed]
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, 4, 739–744. [PubMed]
- Pelli, D. G., Robson, J. G., & Wilkins, A. J. (1988). Designing a new contrast sensitivity. *Clinical Vision Science*, 2, 187–199.
- Pestilli, F., Ling, S., & Carrasco, M. (2009). A population-coding model of attention's influence on contrast response: Estimating neural effects from psychophysical data. *Vision Research*, 49, 1144–1153. [PubMed]
- Poulton, E. C. (1989). *Bias in quantifying judgments*. Hillsdale, NJ: Erlbaum.
- Sekuler, R., & Kahana, M. J. (2007). A stimulus-oriented approach to memory. *Current Directions in Psychological Science*, 16, 305–310.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74, 1–30.
- Sweeny, T. D., Grabowecky, M., Paller, K. A., & Suzuki, S. (2009). Within-hemifield perceptual averaging of facial expressions predicted by neural averaging. *Journal of Vision*, 9(3):2, 1–11, <http://journalofvision.org/9/3/2/>, doi:10.1167/9.3.2. [PubMed] [Article]
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley Publishing.
- Tynan, P., & Sekuler, R. (1974). Perceived spatial frequency varies with stimulus duration. *Journal of the Optical Society of America*, 64, 1251–1255.
- Visscher, K. M., Kahana, M. J., & Sekuler, R. (2009). Trial-to-trial carry-over in auditory short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 46–56. [PubMed] [Article]
- Viswanathan, S., Perl, D. R., Visscher, K. M., Kahana, M. J., & Sekuler, R. (2010). Homogeneity computation: How inter-item similarity in visual short-term memory alters recognition. *Psychonomic Bulletin & Review*, 17, 59–65. [PubMed]
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2005). Pushing around the locus of selection: Evidence for the flexible-selection hypothesis. *Journal of Cognitive Neuroscience*, 17, 1907–1922. [PubMed]
- Watamaniuk, S. N., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research*, 32, 931–941. [PubMed]
- Watamaniuk, S. N., Sekuler, R., & Williams, D. W. (1989). Direction perception in complex dynamic displays: The integration of direction information. *Vision Research*, 29, 47–59. [PubMed]
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33, 113–120. [PubMed]
- Wilcox, R. R. (2005). *Introduction to robust estimation and hypothesis testing* (2nd ed.). Amsterdam, The Netherlands: Elsevier/Academic Press.

- Wilken, P., & Ma, W. (2004). A detection theory account of change detection. *Journal of Vision*, 4(12):11, 1120–1135, <http://journalofvision.org/4/12/11/>, doi:10.1167/4.12.11. [[PubMed](#)] [[Article](#)]
- Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology* (2nd ed.). New York: Holt.
- Yotsumoto, Y., & Sekuler, R. (2006). Out of mind, but not out of sight: Intentional control of visual memory. *Memory & Cognition*, 34, 776–786. [[PubMed](#)] [[Article](#)]
- Zanto, T. P., & Gazzaley, A. (2009). Neural suppression of irrelevant information underlies optimal working memory performance. *Journal of Neuroscience*, 29, 3059–3066. [[PubMed](#)] [[Article](#)]
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453, 233–235. [[PubMed](#)] [[Article](#)]
- Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science*, 20, 423–428. [[PubMed](#)]
- Zoccolan, D., Cox, D. D., & DiCarlo, J. J. (2005). Multiple object response normalization in monkey inferotemporal cortex. *Journal of Neuroscience*, 25, 8150–8164. [[PubMed](#)] [[Article](#)]