

Squinting with the mind's eye: Effects of stimulus resolution on imaginal and perceptual comparisons

STEPHEN M. KOSSLYN and KATHERINE E. SUKEL
Harvard University, Cambridge, Massachusetts

and

BENJAMIN MARTIN BLY
Rutgers University, New Brunswick, New Jersey

Subjects either viewed or visualized arrays that were divided into four quadrants, with each quadrant containing a set of stripes. In two experiments, one array contained only relatively narrow (high-resolution) stripes, and one contained only relatively thick (low-resolution) stripes. The subjects compared sets of stripes in different quadrants according to their length, spacing, orientation, or width. When the subjects visualized the arrays, they required much more time to compare high-resolution patterns than low-resolution patterns; when the subjects saw the arrays, they evaluated both types of arrays equally easily. In addition, the results from the third experiment provide strong evidence that people use imagery in this task; in one condition, the subjects evaluated oblique sets of stripes, and in another condition, they evaluated vertical and horizontal stripes. In both imagery and perception, the subjects made more errors when evaluating oblique stimuli; in imagery, they also required more time to evaluate oblique stimuli. The results suggest that additional effort is required in imagery to represent visual patterns with high resolution. This finding demonstrates that, although imagery and perception may activate common brain regions, it is more difficult to represent high-resolution information in imagery than in perception.

The experience of visual mental imagery is usually described as *seeing with the mind's eye*. This characterization suggests that at least some of the same mechanisms used in visual perception are also used in visual mental imagery. And, in fact, there is ample evidence that the two functions draw on some of the same mechanisms (for reviews, see Craver-Lemley & Reeves, 1987, 1992; Farah, 1988; Finke, 1989; Kosslyn, 1980, 1994; Shepard & Cooper, 1982). However, it is clear that imagery does not rely on all the perceptual mechanisms. Indeed, only about two thirds of the cortical areas activated by imagery and by perception are activated in common by both functions (Kosslyn, Thompson, & Alpert, 1997). In this article, we provide evidence that even the shared mechanisms are not necessarily drawn upon equally easily by the two functions.

One of the most obvious ways in which images seem unlike percepts is that objects in images seem less vivid than real objects. Although this hypothesis has a long history (see, e.g., Hume, 1739/1969) and appears plausible, there is very little behavioral evidence that bears on it. The only example we have found was reported by Harvey (1986), who showed that people rate their visual mental images of faces as being similar to blurred photographs (i.e.,

photographs that have had high-spatial-frequency variations removed). However, even this evidence rests on subjective assessment and could reflect in part the subjects' concept of what an image is supposed to be like. If images are in fact less vivid than percepts, this fact could help to explain why people sometimes have difficulty seeing unexpected patterns in images (as is required to reinterpret an ambiguous figure; see, e.g., Chambers & Reisberg, 1985, 1992; Reisberg & Chambers, 1991; but see, also, Anderson & Helstrup, 1993; Finke, Pinker, & Farah, 1989; Peterson, Kihlstrom, Rose, & Glisky, 1992; Rouw, Kosslyn, & Hamel, 1997, 1998).

One way to conceptualize imagery vividness hinges on facts about the neuroanatomy of the visual system. Much is now known about the visual system of the macaque monkey, which performs similarly to humans in key psychophysical tasks (see, e.g., DeValois, Morgan, Polson, Mead, & Hull, 1974) and has a similar neurophysiology to that of humans; thus, it is reasonable to generalize findings about its visual system to humans, at least as a working hypothesis. Three facts about the neuroanatomy of the visual system are particularly relevant (see, e.g., Felleman & Van Essen, 1991). First, some of the visual areas are topographically organized. That is, the pattern of activation on the retina is physically laid out on the cortex of these areas (albeit with various distortions, such as larger representation of the foveal region). Images may be represented by patterns of activation in such areas (see, e.g., Kosslyn et al., 1993; Kosslyn, Thompson, Kim, & Alpert,

We thank Jennifer Shephard and William Thompson for technical assistance. Correspondence concerning this article should be sent to S. M. Kosslyn, 832 William James Hall, 33 Kirkland Street, Cambridge, MA 02138 (e-mail: smk@wjh.harvard.edu).

1995; Kosslyn, Thompson, Kim, Rauch, & Alpert, 1996). Second, about half of the visual areas of the monkey are topographically organized, but as signals travel deeper into the brain, the receptive field sizes of the areas increase. The larger receptive field sizes occur because each neuron in the higher areas receives inputs from a large number of neurons in the lower areas. Thus, the resolution of the spatial representation declines in the higher areas. Third, the connections among these areas are typically reciprocal—they run in both directions. This arrangement would allow, in principle, information stored in higher areas to flow backwards, inducing patterns of activation in the lower areas—and these patterns of activation would be the representation of the image itself (for a detailed discussion of such processing, see Kosslyn, 1994).

Within this context, high-resolution images would require top-down activation of low-level visual areas, which have the greatest resolution. This notion is consistent with Hebb's (1968) ground-breaking theory of visual mental imagery. Hebb believed that *cell assemblies* activated during perception could also be activated during imagery. Moreover, he believed that more vivid imagery resulted when cell assemblies that register details were activated. The key idea was that some of the cell assemblies that register details dealt with details, and others with coarser patterns. Translated into our present understanding, more vivid images would occur when stored information activated lower areas, which have higher spatial resolution than do higher areas.

Recent results from positron emission tomography have shown that when people visualize objects with their eyes closed, the first visual area (known as area V1, area 17, primary visual cortex, striate cortex, and area OC) can become activated (see Chen, Kato, Zhu, Ogawa, & Ugurbil, 1996; Damasio et al., 1993; Kosslyn et al., 1993; Kosslyn et al., 1995; Kosslyn et al., 1996; LeBihan et al., 1993; Menon et al., 1993; Putz et al., 1996; Sabbah et al., 1995). However, in some studies, this area was not activated (see, e.g., Charlot, Tzourio, Zilbovicius, Mazoyer, & Denis, 1992; Fletcher et al., 1995; Mellet et al., 1996; Mellet, Tzourio, Denis, & Mazoyer, 1995; Mellet et al., 1993; Roland, Erikson, Stone-Elander, & Widen, 1987; Roland & Gulyas, 1994). The studies that evoke activation in low-level areas often appear to have required subjects to use relatively high resolution in their images (but the studies differ in other ways as well, such as the use of different baseline conditions; see Kosslyn et al., 1995). It is tempting to hypothesize that the differences in results arise, at least in part, because one must work harder to form high-resolution images and, hence, subjects typically do not form such images unless they are necessary to perform the task. If so, we expect people to have more difficulty visualizing patterns that require relatively high spatial resolution than those that require relatively low spatial resolution. The present experiments directly test this hypothesis.

EXPERIMENT 1

In this experiment, we varied the widths of sets of stripes and asked subjects to visualize or view these stimuli and to compare sets of such patterns. Narrow stripes require greater resolution to resolve than do thicker stripes; hence, we refer to the stimuli as high- and low-resolution patterns. In addition, the subjects compared the patterns in different ways, which allowed us to observe the relative ease, in imagery and perception, of performing different discriminations. If the resolution of the imaged patterns differs, as compared with the perceptual patterns, we expect different orderings of difficulty for the discriminations for the high- and low-resolution patterns in the two conditions.

Method

Subjects

Thirty-two Harvard University students (with a mean age of 19.6 years) volunteered to be paid subjects in this study. Eight subjects were eliminated because they could not perform the task at better than 80% accuracy. Half of the remaining 24 subjects (mean age of 19.8 years) were male, and half were female; all were right-handed native speakers of English.

Materials and Apparatus

As is illustrated in Figure 1, we created two stimulus arrays, each of which included four sets of black-and-white stripes (one in each quadrant of the display). One set included only relatively narrow stripes (with stripes that averaged 0.2 cm), and one included only relatively thick stripes (with stripes that averaged 0.78 cm). In both displays, each set of stripes was approximately 5 in. high and 5 in. wide (subtending approximately 8° of visual angle). The patterns in each quadrant differed in length, width, orientation, and the amount of space between the bars. The quadrants were named by the letters A, B, C, and D. The stimuli were created with Aldus Superpaint.

We also created 48 sound files, using Soundedit 16. Each sound file included a stimulus of the form "A, B, Length," which indicated the two quadrants to be compared and the required discrimination (length, width, orientation, or spacing between the stripes). All cue words were a single syllable (thus, we used the word *bar* to cue orientation discriminations and *space* to cue spacing discriminations). Two sets of cues were created, one that required the subjects to compare length or width and one that required them to compare the spacing or orientation. All the tasks were created and presented using Psycscope software and were administered on a Macintosh Quadra 950, with a 15-in. black-and-white display monitor.

The auditory stimuli were arranged in a random order, except that no more than three responses of the same type, repetitions of the same letter (naming a quadrant), or type of discrimination could appear in a row, every combination of discrimination, quadrant, and response occurred once before any could occur twice, and all combinations occurred twice before any appeared a third time.

Procedure

Each subject was tested individually in two conditions, perception and imagery. The perception condition was always presented first; we were concerned that, if subjects were explicitly told to visualize in the first condition they received, they would then adopt that strategy in the perceptual condition (because the stimuli were presented briefly).

Perception condition. Instructions for each task were presented on the computer screen. After the subjects were informed that they

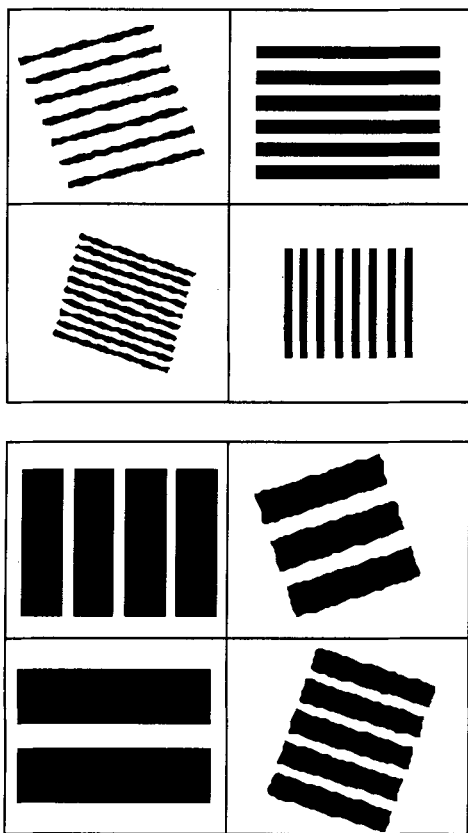


Figure 1. The two sets of gratings used as stimuli in Experiments 1 and 2.

would be making discriminations between quadrants of a complex pattern, a study display appeared that presented the name of each quadrant: A, B, C, and D. Immediately thereafter, a display appeared that defined the four discriminations: length, width, spacing, and orientation. The subjects were told to compare the patterns in the two named quadrants according to the named discrimination; they were to press a key labeled *true* on the computer's keyboard if the first-named quadrant had a pattern that was longer, wider, spaced farther apart, or more tilted (in the clockwise direction), as appropriate, than was the pattern in the second-named quadrant. Examples of the discriminations (using other stimuli) were also presented. As soon as the subjects indicated that they understood the nature of the discriminations, four practice trials were presented; these trials had the same format as the subsequent test trials.

Each trial consisted of the following series of events. A cue consisting of two letters (naming different quadrants) and the name of the discrimination was presented auditorily. For example, the subjects might have heard, "A, B, Space," which instructed them to determine whether the space between the bars in Quadrant A was greater than that between the bars in Quadrant B. Immediately following these cues, one of the stimulus arrays appeared for 500 msec. The subjects were asked to respond by pressing either the *true* (the "b" key on the keyboard) or the *false* (the "n" key on the keyboard) key, labeled appropriately on the keyboard, as quickly and accurately as possible. The subjects always responded with the dominant hand, using their index and middle fingers. A new trial began 1 sec after the subject responded.

The trials were blocked by array; so, after 24 trials with either the high- or low-resolution array, using only two of the possible dis-

criminations (length and width, or space and orientation), the subjects were instructed that a new pattern would be used. A new block of 24 trials was then presented, with the second array and the other pair of discriminations. The subjects were not aware which pattern would be used first or what discriminations they would be asked to make for each pattern. Counterbalancing over subjects ensured that each array appeared in each order equally often and that each discrimination appeared with each array in each order equally often.

Imagery condition. All the instructions again were presented on the screen. The subjects were reminded of the names of the quadrants and the definitions of the cues for the discriminations. Following this, the subjects were told that they would learn to visualize one of the arrays by breaking it down quadrant by quadrant. First, one of the two arrays appeared, and the subjects were asked to study it carefully. Next, only Quadrant A appeared (at the same size at which it had appeared in the full array), and the subjects were asked to study it carefully. When the subjects felt that they had memorized the pattern, they pressed a key, and a blank screen appeared. The subjects then were asked to visualize the set of stripes exactly as it had appeared. Once the pattern was visualized, the subjects pressed a key, and Quadrant A appeared again. The subjects were now asked to compare their mental image to the actual pattern and to make any necessary changes in the image so that it would correspond to the actual pattern. This procedure was repeated twice more, and then the subjects moved on to Quadrants B, C, and D, repeating this entire sequence of events with each. After the subjects had studied all four quadrants separately, the entire array appeared again, and the same type of image-learning technique was used with it; the subjects now visualized the entire array and compared it with the display a total of three times.

After completing the learning session, the subjects were instructed that they would be asked to visualize the array and compare pairs of quadrants, just as they had in the perception condition. However, they would now compare the quadrants in terms of the two discriminations that were not made in the perception condition. The subjects were asked to base their discriminations on "inspection" of the arrays in their mental images. After visualizing and evaluating one array, they went on to the other one. The subjects were not aware of which discriminations would be presented with each array. All other aspects of this procedure were the same as those in the perception condition.

Counterbalancing allowed us to ensure that half the subjects received the high-resolution array first in each condition and half received the low-resolution array first. Moreover, every discrimination was used equally often with each display in each condition in each order.

Results

The data were analyzed, using a yoked-subject design in which each subject was matched with another who received the exact opposite counterbalancing factors. This was necessary because only two discriminations were used in each condition, and hence we had to yoke subjects in order to nest all the variables properly. We included condition (imagery/perception), resolution (high/low), discriminations (height, width, orientation, spacing), and response (true/false) as independent variables. Response time and error rate data were analyzed in separate analyses of variance (ANOVAs); only times from trials on which the subjects made the correct responses were analyzed. Moreover, outliers were eliminated prior to analysis; an outlier was defined as a response time greater than 2.5 times the mean of the other times in that cell for a given person. This procedure resulted in our eliminat-

ing 3.9% of the data prior to analysis. All effects and interactions not noted were not significant ($p > .1$).

When we found a nonsignificant result of potential interest, we assessed the possibility of Type II error by conducting power analyses (Cohen, 1969; Glass & Hopkins, 1984). Power is here defined as one minus the probability of accepting the null hypothesis when false. Following the method outlined in Glass and Hopkins, we computed power by assessing the noncentrality inherent in the data, relative to the central distribution that would be expected of F scores with a particular number of degrees of freedom in the numerator and in the denominator. This noncentrality parameter can be used to compare the data with hypothetical, centrally distributed data. The greater the proportion of the data that lies beyond a critical alpha of .05 in the central distribution, the higher the power of the test (i.e., the lower the probability of falsely accepting the null hypothesis).

Response Times

As was predicted, we found that the relative ease of evaluating the two types of arrays was different in the imagery and perception conditions [$F(1,11) = 6.31, p < .03$] for the interaction of array type and condition. Planned comparisons revealed that the subjects required more time to evaluate the high-resolution arrays than the low-resolution arrays in the imagery condition [$F(1,11) = 16.17, p < .001$] but no more time to evaluate the two arrays in the perception condition ($F < 1$). An additional set of comparisons revealed that the subjects required more time for the high-resolution arrays in the imagery condition than in the perception condition [$F(1,11) = 7.36, p = .02$] but required comparable amounts of time

for the low-resolution array in the two conditions ($F < 1$). These results are presented in Table 1 and Figure 2.

In addition, we also found that the subjects required more time to evaluate the high-resolution arrays than the low-resolution ones [$F(1,11) = 8.33, p < .02$] and that this difference varied for different discriminations [$F(3,33) = 3.91, p < .02$]; this is not surprising, given that resolution is more or less important for the different discriminations. In general, the discriminations appeared to be of similar difficulty for imagery and perception [$F < 1$, power ($1 - \beta$) = 0.82] for the interaction between condition and discrimination. However, inspection of Figure 3 suggests that the discriminations had different relative difficulty in the imagery and perception conditions for the low-resolution patterns. Although the three-way interaction of condition, resolution, and discrimination was not significant ($F < 1$) and we had ample power [($1 - \beta$) = 0.965], we nevertheless performed separate analyses of only the high-resolution data and only the low-resolution data. As is illustrated in Figure 3, the relative time to make the different discriminations varied in imagery and perception for the low-resolution patterns [$F(3,33) = 5.40, p = .004$] for the interaction of condition and discrimination. In contrast, this interaction did not approach significance for the high-resolution data ($F < 1$). In addition, separate ANOVAs were conducted for the data for each discrimination, to examine the relative ease of evaluating the high- versus low-resolution stimuli in imagery and perception. Although the pattern of means was in the expected direction in each case, the interaction of condition and resolution was significant only when the subjects evaluated length [$F(1,11) = 8.22, p = .015$] and width [$F(1,11) = 6.60, p = .026$].

Error Rates

No significant differences were found in the comparisons of error rates ($p > .1$, in all cases); the mean error rates were 6.0% and 5.9% for the imagery and perception conditions, respectively. Therefore, there was no evidence of speed/accuracy tradeoffs.

Discussion

Two results from this experiment are noteworthy. First, when they used imagery, the subjects required more time to evaluate high-resolution patterns than to evaluate low-resolution ones; in contrast, they evaluated the two types of patterns equally easily during perception. The fact that the subjects required more time in this one condition—imagery with high-resolution patterns—is consistent with the hypothesis that imagery has a limited resolution and the imaged patterns were degraded and, hence, more difficult to evaluate. Second, the relative ease of the discriminations varied in imagery and perception. However, we only found such an interaction with the low-resolution patterns. This result may suggest that, when the necessary resolution is high enough, the processes used in making the evaluations operate comparably in imagery and perception (exhibiting what Finke, 1989,

Table 1
Response Times (in Milliseconds) and Error Rates, With the Corresponding Standard Errors of the Mean, in Experiment 1

	Response Times		Error Rates	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Imagery: High-Resolution Stimuli				
Length	1,692	235	0.08	0.02
Width	1,689	210	0.05	0.02
Space	2,336	273	0.07	0.02
Orientation	2,264	389	0.05	0.02
Imagery: Low-Resolution Stimuli				
Length	1,608	147	0.08	0.02
Width	1,442	130	0.06	0.01
Space	1,500	178	0.05	0.02
Orientation	1,445	168	0.08	0.02
Perception: High-Resolution Stimuli				
Length	1,203	77	0.06	0.01
Width	1,249	129	0.03	0.01
Space	1,878	194	0.06	0.02
Orientation	1,944	186	0.06	0.01
Perception: Low-Resolution Stimuli				
Length	1,994	176	0.09	0.02
Width	1,539	153	0.08	0.02
Space	1,096	77	0.06	0.01
Orientation	1,340	102	0.10	0.02

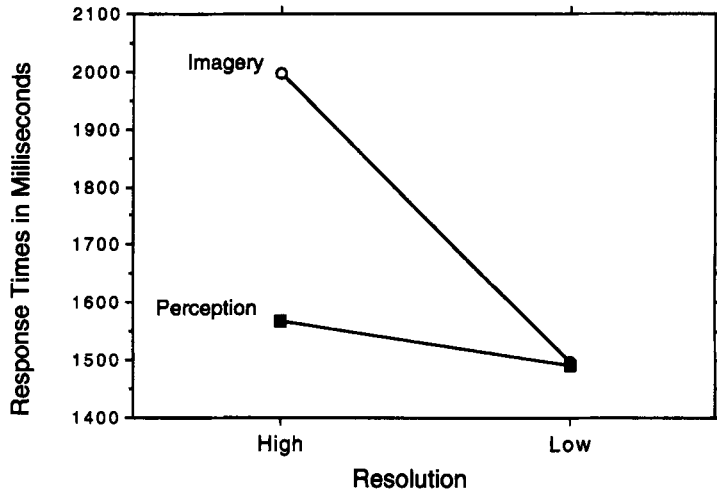


Figure 2. Effects of resolution on response times in imagery and perception in Experiment 1.

called *functional equivalence*); in contrast, when the necessary resolution is low, one may be able to use strategies in perception for some discriminations that cannot easily be used in imagery. For example, such strategies may rest on comparing parts that would be difficult to maintain simultaneously in an image. These strategies may seem easier to the subjects but need not actually be more efficient; thus, we cannot rule out this possibility on the basis of the comparable overall response times in the imagery and perception conditions with these stimuli.¹

EXPERIMENT 2

The results of Experiment 1 suggest that imagery has more limited resolution than does perception. We also found that the relative ease of making the discriminations with the low-resolution patterns was different for

imagery and perception. We were concerned that, by presenting the perceptual stimuli for only half a second, the subjects may have developed special strategies for the low-resolution patterns in perception, which could only be used with those relatively simple displays. If so, performance should be different if the subjects have more time to evaluate the perceptual displays and, hence, have the opportunity to use such strategies for both types of stimuli. In this experiment, we replicated Experiment 1 but now presented the perceptual displays for 1 sec; this seemed to be enough time to ensure that the discriminations were based on perception but not so much time that the subjects could begin to memorize possible comparisons. We also modified the materials and procedure slightly to eliminate problems encountered in Experiment 1 that sometimes produced poor performance (and led us to discard data from some of the subjects).

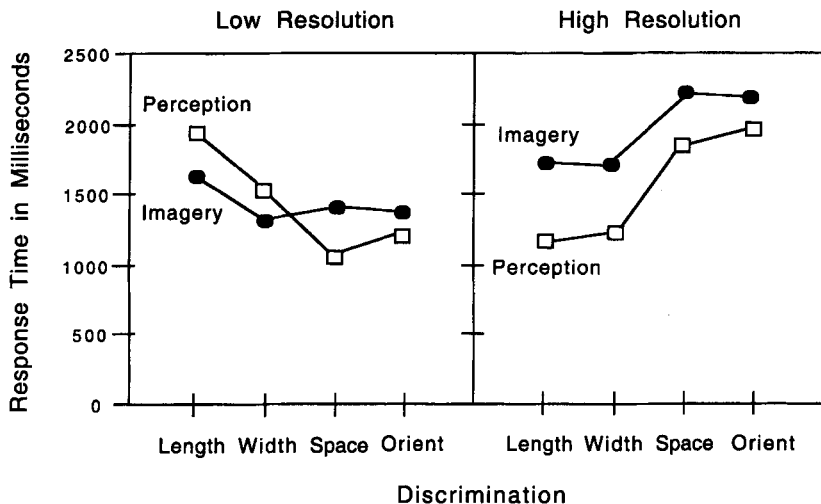


Figure 3. The relative ease of the four discriminations in imagery and perception in Experiment 1.

Method

Subjects

Sixteen Harvard University students (10 women, 6 men; mean age, 18.7 years) volunteered to participate as paid subjects in this study. None of these people had participated in the previous experiment. All the subjects were right-handed native speakers of English. Our methodological changes resulted in all the subjects' making fewer than 20% errors; hence, we did not eliminate any subjects from these analyses.

Materials and Procedure

The materials of Experiment 1 were used here, with four changes: (1) The quadrants were now referred to by the numbers 1, 2, 3, and 4 (because the letters B, C, and D were sometimes confused by the subjects in Experiment 1); (2) the *bar* cue was replaced by the word *tilt*; (3) the subjects no longer saw the definitions of all possible discriminations at the outset, so they would not be able to memorize relevant comparisons; and (4) in the perception condition, the stimulus arrays were presented for 1 sec, twice the time they were presented in Experiment 1.

Results

These data were analyzed in the same way that the data were analyzed in Experiment 1; 3.1% of the data were eliminated as outliers prior to analysis.

Response Times

As is evident in Figure 4, we again found an interaction between resolution and condition [$F(1,7) = 6.28, p = .04$], with differing response times for high- and low-resolution arrays in the imagery and perception conditions. Planned contrasts revealed that the subjects required more time to make discriminations among the high-resolution patterns than among the low-resolution patterns in the imagery condition [$F(1,7) = 7.06, p < .04$] but required comparable amounts of time for the two types of patterns in the perception condition ($F < 1$). Other planned contrasts revealed that the subjects required more

time in the imagery than in the perception condition for the high-resolution stimuli [$F(1,7) = 31.16, p < .001$] but required only marginally more time in the imagery condition for the low-resolution stimuli [$F(1,7) = 4.15, p = .08$]. Finally, the subjects required more time to respond in the imagery condition than in the perception condition [with means of 2,317 vs. 1,410 msec; $F(1,7) = 31.05, p < .001$]. The relative ease of the different discriminations is illustrated in Figure 5. There was no interaction between discrimination and any other factor [$p > .15$, in all cases; power ($1 - \beta$) for the interaction of condition and discrimination = 0.82, and for the three-way interaction of condition, resolution, and discrimination, power = 0.64]. We again performed separate analyses of only the low-resolution data and only the high-resolution data and now found no hint of interactions with condition and judgment [$F < 1$ and $F(3,21) = 1.54, p = .23$, respectively]. We also performed separate analyses for data from each discrimination and found an interaction between resolution and condition only for the space discrimination [$F(1,7) = 6.92, p = .034$], but the pattern of means was as expected for the other three discriminations.

Error Rates

As in Experiment 1, no effect or interaction was significant in this analysis, thereby providing evidence against speed/accuracy tradeoffs. Error rates were 10.1% for imagery and 9.95% for perception. Table 2 presents the data from all comparisons.

Discussion

The major result of Experiment 1 was replicated, despite the fact that the stimuli in the perception condition were visible for 1 sec and that the subjects did not know the discriminations that they would need to make prior to studying the stimuli. The additional time necessary for

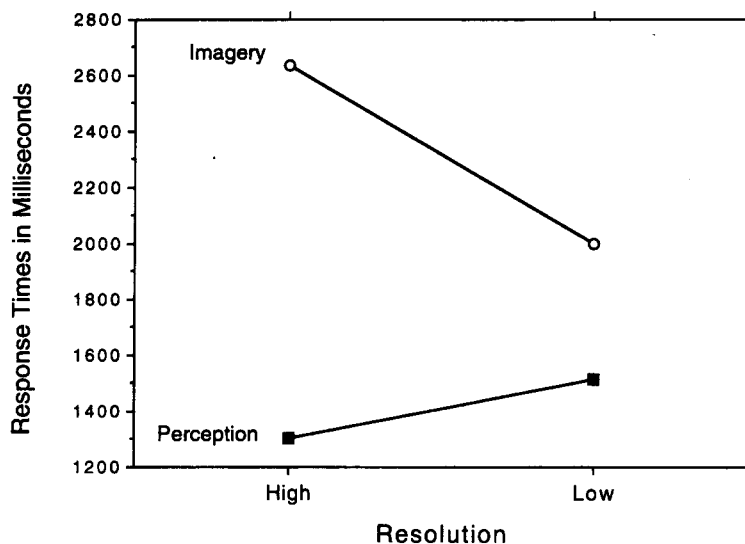


Figure 4. Effects of resolution on response times in imagery and perception in Experiment 2.

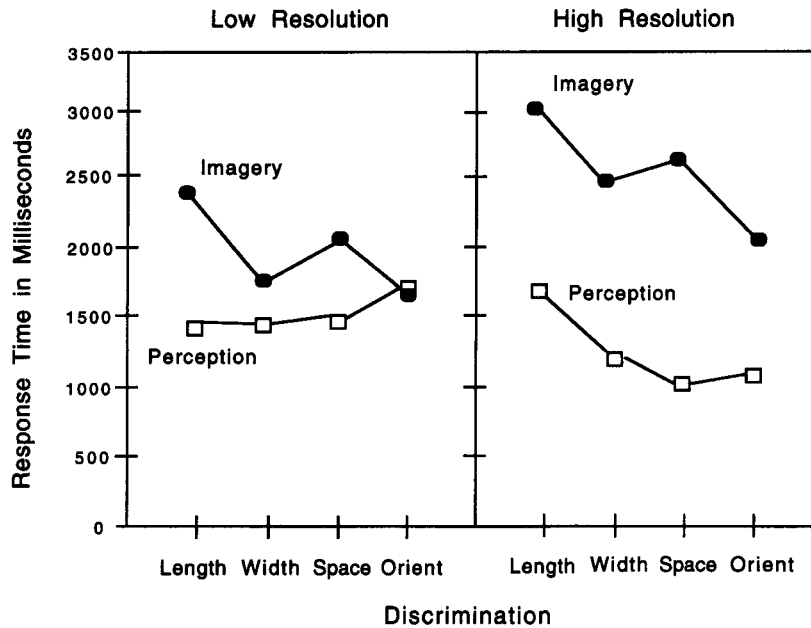


Figure 5. The relative ease of the four discriminations in imagery and perception in Experiment 2.

high-resolution patterns in the imagery condition presumably reflects the relatively poor quality of the representation. That is, we hypothesize that, when generating images, the subjects could not easily impose a pattern of activation on low-level visual areas, which have the highest spatial resolution. Thus, the representation of the high-resolution patterns was degraded in imagery, rela-

tive to the representation formed when bottom-up processing activates these areas in perception.

In addition, in this study we found no evidence that the ease of performing the discriminations was different in imagery and perception. However, the error variance associated with these means is high and, thus, we cannot be confident in drawing conclusions about the relative ease of performing the different comparisons in imagery and perception.

Table 2

Response Times (in Milliseconds) and Error Rates, With the Corresponding Standard Errors of the Mean, in Experiment 2

	Response Times		Error Rates	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Imagery: High-Resolution Stimuli				
Length	3,052	555	0.12	0.02
Width	2,545	544	0.09	0.03
Space	2,742	397	0.07	0.02
Orientation	2,196	335	0.11	0.04
Imagery: Low-Resolution Stimuli				
Length	2,446	364	0.10	0.02
Width	1,785	235	0.13	0.02
Space	2,020	198	0.09	0.03
Orientation	1,755	223	0.10	0.02
Perception: High-Resolution Stimuli				
Length	1,733	363	0.10	0.05
Width	1,254	199	0.07	0.02
Space	1,077	103	0.07	0.02
Orientation	1,155	136	0.14	0.05
Perception: Low-Resolution Stimuli				
Length	1,410	217	0.11	0.02
Width	1,412	265	0.12	0.01
Space	1,487	167	0.06	0.02
Orientation	1,756	226	0.12	0.02

EXPERIMENT 3

We have interpreted the results of the previous experiments as indicating that the same neural structures are used in imagery and perception. However, one could argue that these results actually reflect the ease of holding an abstract representation in short-term memory, which is more difficult with high-resolution gratings. If perceptual mechanisms are in fact used in imagery in our task, we should find perceptual effects in imagery that are like those found in vision proper. In this experiment, we examined one such effect, which has the virtue of being counter-intuitive. Appelle (1972) reviewed considerable evidence that oblique lines are perceived more poorly than vertical or horizontal ones, and Pennington and Kosslyn (described in Kosslyn, 1983) found such an effect in imagery; however, this effect was not found by Kosslyn, Brunn, Cave, and Wallach (1984), but their methodology differed and was arguably less precise.

In this study, we investigated whether the current paradigm could detect an oblique effect in imagery. If so, this is good evidence that perceptual mechanisms are indeed being used in this task. In addition, if we succeed

in reducing the resolution further by using exclusively oblique stimuli, we should see clear differences in the relative difficulty of the different discriminations in this condition, relative to when vertical and horizontal lines are visualized.

Method

Subjects

Twenty four Harvard University undergraduates volunteered to participate as paid subjects in this study (11 males and 13 females; mean age, 19.7 years). All the subjects were right-handed, and none had participated in any of the previous experiments reported in this article. Data from all the subjects were included in the analyses.

Materials

As is illustrated in Figure 6, two new stimulus patterns, both of high resolution, were used in this experiment. The first consisted of only vertical or horizontal stripes, whereas the second contained only oblique stripes. The same stripes were used in both stimuli, with only their location, orientation, and (in 2 cases) number varying (no subject reported having noticed this, when queried during debriefing). In addition, we replaced the orientation discrimination with one (cued by the word *fill*) that asked the subjects to decide whether the pattern in the quadrant named first filled the area of the quadrant more than did the pattern in the quadrant named second. All other materials were identical to those of Experiment 2.

Procedure

The procedure was identical to that of Experiment 2.

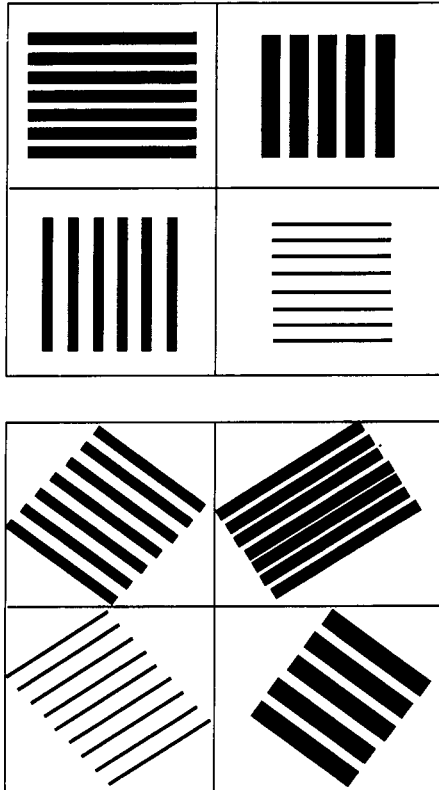


Figure 6. The two sets of gratings used as stimuli in Experiment 3.

Results

The data were analyzed in a yoked-subjects design, as was used in the previous experiments; 3.7% of the response times were eliminated as outliers prior to analysis.

Response Times

The most interesting result is illustrated in Figure 7, which shows that the time the subjects required to compare the oblique versus the nonoblique stimuli varied, depending on whether they visualized or perceived the stimuli [$F(1,11) = 4.84, p = .05$], for the interaction of orientation and condition. Contrasts revealed that the subjects required more time for the oblique stimuli than for the nonoblique stimuli only during imagery [$F(1,11) = 11.28, p = .006$, for the contrast on the imagery data, and $F < 1$ for the contrast on the perception data]. Pooling over conditions, there was only a trend toward an oblique effect [$F(1,11) = 3.70, p < .08$]. In addition, the subjects were faster in the perception condition than in the imagery condition [1,255 vs. 2,055 msec; $F(1,11) = 23.12, p < .001$]. No other effect or interaction approached significance ($p > .33$, in all cases). We found no interaction between imagery versus perception and the type of discrimination [$F < 1$, with power $(1 - \beta) = 0.85$] or between these factors and orientation [$F < 1$, with power $(1 - \beta) = 0.64$]. Separate analyses of the oblique and nonoblique stimuli revealed that the discriminations had comparable relative difficulty in imagery and perception in both cases ($p > .33$, for both interactions). And again, separate analyses were conducted for each discrimination type but now found only one trend toward an interaction between condition and orientation, for the width discrimination [$F(1,11) = 3.42, p = .09$].

Error Rates

We found that the subjects made more errors with the oblique stimuli [$F(1,11) = 11.0, p < .007$]. This effect was comparable in both imagery and perception [$F(1,11) = 1.92, p > .19$, for the interaction]; in addition, contrasts revealed that subjects made more errors for oblique patterns in imagery [$F(1,11) = 11.60, p = .006$] and in perception [$F(1,11) = 28.77, p = .0002$].

In this analysis, we found that the discriminations produced different numbers of errors [$F(3,33) = 31.00, p < .001$]. Moreover, as is illustrated in Figure 8, the differences in the difficulties of the discriminations varied for the oblique versus nonoblique stimuli [$F(3,33) = 14.32, p < .0001$]. As is evident, the *fill* discrimination was more difficult for oblique stimuli, but the other discriminations were comparable for the two stimuli. Finally, as is evident in Figure 9, the difference in discriminations for the two stimuli in turn varied in imagery and perception [$F(3,33) = 3.48, p < .03$]. Separate analyses of data from each discrimination revealed interactions of condition and stimulus type for the length [$F(1,11) = 5.03, p < .05$] and fill discriminations [$F(1,11) = 5.42, p = .04$]. Table 3 presents the results from all the comparisons.

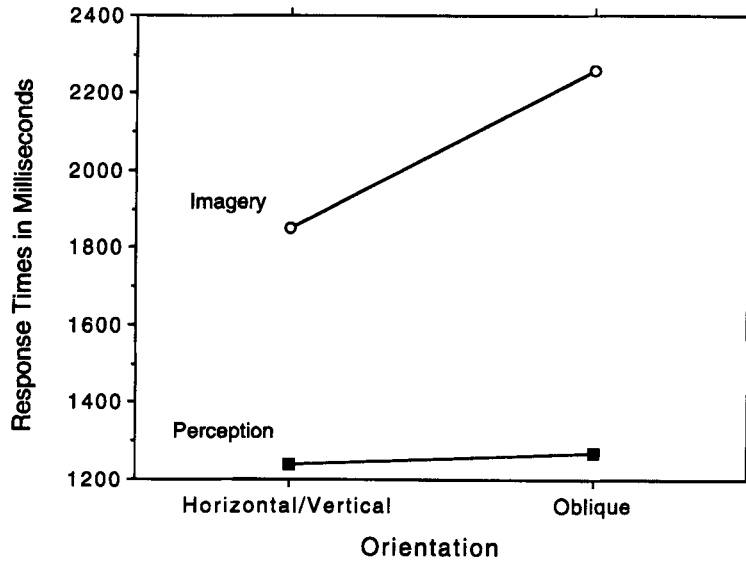


Figure 7. Effects of orientation on response times in imagery and perception in Experiment 3.

Discussion

As was predicted, we found an oblique effect in both imagery and perception. Indeed, if anything, the effect was more pronounced in imagery, appearing in both the response time and error rate data. These results not only provide convergent evidence that the subjects did in fact use imagery in this task, but also that visual mental images are processed by mechanisms used in vision. Despite the fact that the subjects performed the perceptual discrimination first, they apparently did not memorize the responses—which would have been almost impossible, because different discriminations were used with the stimuli in the imagery and perceptual conditions.

In addition, in this experiment we found clear-cut differences in the relative ease of performing the different discriminations with the oblique versus the nonoblique stimuli. Apparently, as the resolution decreased with the oblique stimuli, at least one of the discriminations became more difficult.

GENERAL DISCUSSION

The results clearly indicate that people expend more effort to compare patterns that require high resolution in imagery than patterns that require less resolution. We can explain these results if the resolution limits in imagery essentially blur high-resolution stimuli, making them less vivid. That is, the subjects require more time to evaluate high-resolution patterns in imagery because the processes that interpret imaged patterns have to work on degraded representations.

We also found that the relative ease of the four discriminations in imagery and perception varied in the different experiments. This result suggests that, at least in

some conditions, resolution differences may have affected processing. However, given the relatively small amount of data for each of the separate discriminations, we must be cautious in drawing strong conclusions from these results. Note, however, that for oblique stimuli we found that one of the discriminations was clearly more difficult in imagery. Apparently, the oblique stimuli were difficult enough to evaluate in imagery that relative difficulties in the discriminations were exacerbated. The ease of performing a process depends on the quality of the representation, and, apparently, oblique images are degraded enough to affect processing.

These results are consistent with much other literature that implicates shared mechanisms in imagery and perception. For example, Craver-Lemley and Reeves (1987)

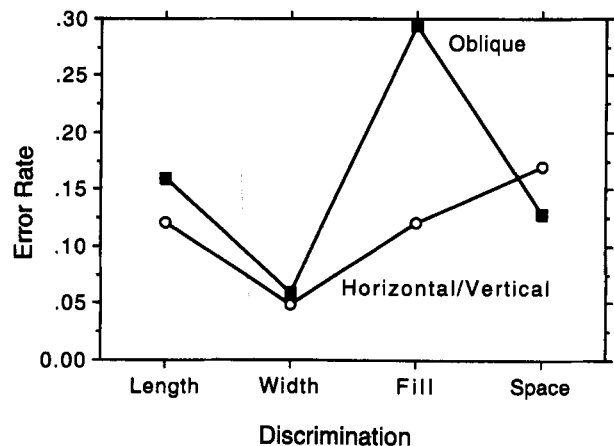


Figure 8. The relative ease of the four discriminations for oblique versus horizontal/vertical stimuli in Experiment 3.

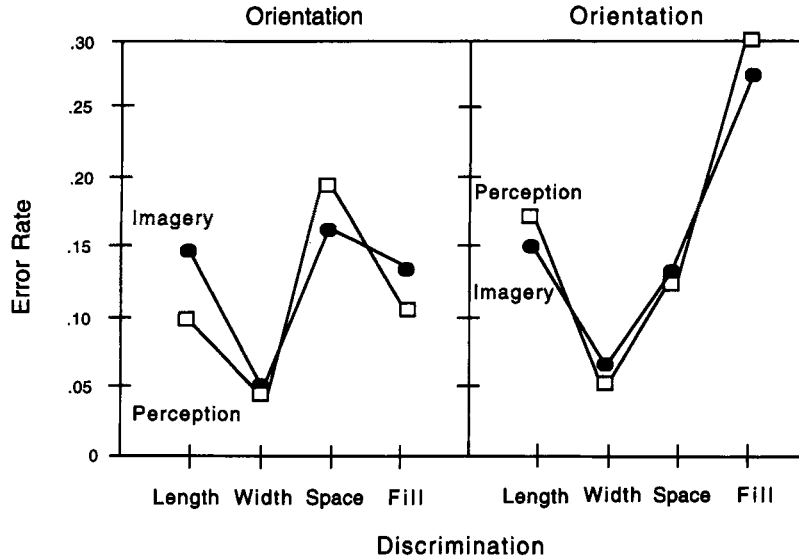


Figure 9. The relative ease of the four discriminations for stimuli of different orientations during imagery and perception in Experiment 3.

found that forming a visual image can reduce visual perceptual acuity but that this effect occurs only when the image is positioned over the to-be-discriminated stimulus (see, also, Craver-Lemley, Arterberry, & Reeves, 1997). In addition, Finke, Johnson, and Shyi (1988), Intraub and Hoffman (1992), and Johnson and Raye (1981) have found that people confuse having visualized a stimulus for actually having seen it. All such results can be interpreted if imagery relies on neural structures also used in perception.

The present results are of interest in part because they may help us to understand the neural basis of image representations. If image representations occur in topographically organized parts of the visual cortex (see Kosslyn, 1994), the difference in the effort required to process finely detailed images may result in locally enhanced regional cerebral blood flow. Thus, the brain activation measured in neuroimaging studies of mental imagery may depend critically on the spatial properties of the imaged stimuli. This hypothesis might explain why some brain-scanning studies of imagery have found activation of the topographically organized cortex (see the introduction), whereas others have not. The tasks and stimuli differ in many ways among the studies, but a review of the articles (Thompson & Kosslyn, in press) suggests that the topographically organized cortex (specifically, Areas 17 and 18) is activated only when subjects need to use high-resolution images. For example, the Mellet et al. (1995; Mellet et al., 1993) and Roland et al. (1987) experiments used imagery tasks that required subjects to visualize very large differences in directions, which do not require high resolution—and these studies did not produce evidence that imagery activates the topographically organized cortex. In contrast, LeBihan et al. (1993) required subjects

to visualize relatively high-resolution patterns of stripes; LeBihan et al. used binocular goggles fitted with red light-emitting diodes. Although they do not specify the resolution of their stimuli, the manufacturer of the goggles estimated that the display subtends approximately 3°–4° of visual angle; thus, their stimuli were definitely of a higher resolution than our low-resolution patterns. Kosslyn et al. (1993; Kosslyn et al., 1995; Kosslyn et al., 1996) required subjects to visualize letters or line draw-

Table 3
Response Times (in Milliseconds) and Error Rates, With the Corresponding Standard Errors of the Mean, in Experiment 3

	Response Times		Error Rates	
	M	SE	M	SE
Imagery: Oblique				
Length	2,228	434	0.15	0.03
Width	2,584	653	0.07	0.02
Space	2,277	341	0.13	0.02
Fill	1,954	358	0.27	0.02
Imagery: Horizontal/Vertical				
Length	1,816	225	0.15	0.02
Width	1,564	192	0.05	0.02
Space	2,119	445	0.16	0.02
Fill	1,896	374	0.13	0.02
Perception: Oblique				
Length	1,169	217	0.17	0.02
Width	1,175	247	0.05	0.02
Space	1,258	197	0.12	0.03
Fill	1,480	155	0.31	0.02
Perception: Horizontal/Vertical				
Length	1,186	178	0.10	0.02
Width	1,150	135	0.04	0.02
Space	1,346	271	0.18	0.02
Fill	1,278	285	0.11	0.02

ings and then to evaluate them in ways that required relatively high resolution—and these studies did produce evidence that imagery activates topographically organized cortex.

The present results, then, lead us to take seriously the possibility that at least some of the disparities in findings with brain scanning reflect the need to form high-resolution images. If such images are more difficult to form and use, we expect people not to use them unless they must do so in order to perform the task.

In addition, the present findings may bear on a controversy about the ability to reinterpret images (see, e.g., Reisberg, 1998; Rouw et al., 1998). Reisberg and Chambers (1991) and Chambers and Reisberg (1985, 1992) originally reported that people could not reinterpret ambiguous shapes in images. However, later researchers showed that this could be done, if the stimuli were not overly complex or proper cues were provided (e.g., Anderson & Helstrup, 1993; Finke et al., 1989; Peterson et al., 1992; Rouw et al., 1997). One reason why such cues may be useful is that the stimuli required relatively high resolution and, thus, the images were not very vivid.

In summary, the results of this study suggest that mental imagery is more difficult when subjects must form representations that allow discrimination of high-resolution properties of stimuli. Our finding of increasing difficulty associated with high-resolution discriminations suggests that additional effort is required in imagery to represent visual patterns with high resolution. This finding is consistent with the notion that imagery and perception activate common brain regions but that it is more difficult to represent high-resolution information in imagery than in perception.

REFERENCES

- ANDERSON, R. E., & HELSTRUP, T. (1993). Visual discovery in mind and on paper. *Memory & Cognition*, *21*, 283-293.
- APPELLE, S. (1972). Perception and discrimination as a function of stimulus orientation: The "oblique effect" in man and animals. *Psychological Bulletin*, *89*, 266-273.
- CAMPBELL, F. W., & ROBSON, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, *197*, 551-566.
- CHAMBERS, D., & REISBERG, D. (1985). Can mental images be ambiguous? *Journal of Experimental Psychology: Human Perception & Performance*, *11*, 317-328.
- CHAMBERS, D., & REISBERG, D. (1992). What an image depicts depends on what an image means. *Cognitive Psychology*, *24*, 145-174.
- CHARLOT, V., TZOURIO, N., ZILBOVICIUS, M., MAZOYER, B., & DENIS, M. (1992). Different mental imagery abilities result in different regional cerebral blood flow activation patterns during cognitive tasks. *Neuro-psychologia*, *30*, 565-580.
- CHEN, W., KATO, T., ZHU, X.-H., OGAWA, S., & UGURBIL, K. (1996). Primary visual cortex activation during visual imagery in human brain: A fMRI mapping study. *NeuroImage*, *3*, S204.
- COHEN, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- CRAVER-LEMLEY, C., ARTERBERRY, M. E., & REEVES, A. (1997). Effects of imagery on vernier acuity under conditions of induced depth. *Journal of Experimental Psychology: Human Perception & Performance*, *23*, 3-13.
- CRAVER-LEMLEY, C., & REEVES, A. (1987). Visual imagery selectively reduces vernier acuity. *Perception*, *16*, 533-614.
- CRAVER-LEMLEY, C., & REEVES, A. (1992). How visual imagery interferes with vision. *Psychological Review*, *99*, 633-649.
- DAMASIO, H., GRABOWSKI, T. J., DAMASIO, A., TRANEL, D., BOLES-PONTO, L., WATKINS, G. L., & HICHA, R. D. (1993). Visual recall with eyes closed and covered activates early visual cortices. *Society of Neuroscience Abstracts*, *19*, 1603.
- DEVALOIS, R. L., MORGAN, H. C., POLSON, M. C., MEAD, W. R., & HULL, E. M. (1974). Psychophysical studies of monkey vision: I. Macaque luminosity and color vision tests. *Vision Research*, *14*, 53-67.
- FARAH, M. J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review*, *95*, 307-317.
- FELLEMAN, D. J., & VAN ESSEN, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*, 1-47.
- FINKE, R. A. (1989). *Principles of mental imagery*. Cambridge, MA: MIT Press.
- FINKE, R. A., JOHNSON, M. K., & SHYI, G. C.-W. (1988). Memory confusions for real and imagined completions of symmetrical visual patterns. *Memory & Cognition*, *16*, 133-137.
- FINKE, R. A., PINKER, S., & FARAH, M. (1989). Reinterpreting visual patterns in mental imagery. *Cognitive Science*, *13*, 51-78.
- FLETCHER, P. C., FRITH, C. D., BAKER, S. C., SHALLICE, T., FRACKOWIAK, R. S. J., & DOLAN, R. J. (1995). The mind's eye—precuneus activation in memory-related imagery. *NeuroImage*, *2*, 195-200.
- GLASS, G. V., & HOPKINS, K. D. (1984). *Statistical methods in education and psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- HARVEY, L. (1986). Visual memory: What is remembered? In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities* (pp. 173-187). The Hague: Elsevier.
- HEBB, D. O. (1968). Concerning imagery. *Psychological Review*, *75*, 466-477.
- HUME, D. (1969). *A treatise in human nature*. Baltimore: Pelican. (Original work published 1739)
- INTRAUB, H., & HOFFMAN, J. E. (1992). Reading and visual memory: Remembering scenes that were never seen. *American Journal of Psychology*, *105*, 101-114.
- JOHNSON, M. K., & RAYE, C. L. (1981). Reality monitoring. *Psychological Review*, *88*, 67-85.
- KOSSLYN, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.
- KOSSLYN, S. M. (1983). *Ghosts in the mind's machine*. New York: Norton.
- KOSSLYN, S. M. (1994). *Image and brain*. Cambridge, MA: MIT Press.
- KOSSLYN, S. M., ALPERT, N. M., THOMPSON, W. L., MALJKOVIC, V., WEISE, S. B., CHABRIS, C. F., HAMILTON, S. E., & BUONANO, F. S. (1993). Visual mental imagery activates topographically organized visual cortex: PET investigations. *Journal of Cognitive Neuroscience*, *5*, 263-287.
- KOSSLYN, S. M., BRUNN, J. L., CAVE, K. R., & WALLACH, R. W. (1984). Individual differences in visual imagery: A computational analysis. *Cognition*, *18*, 195-243.
- KOSSLYN, S. M., THOMPSON, W. L., & ALPERT, N. M. (1997). Neural systems shared by visual imagery and visual perception: A positron emission tomography study. *NeuroImage*, *6*, 320-334.
- KOSSLYN, S. M., THOMPSON, W. L., KIM, I. J., & ALPERT, N. M. (1995). Topographic representations of mental images in primary visual cortex. *Nature*, *378*, 496-498.
- KOSSLYN, S. M., THOMPSON, W. L., KIM, I. J., RAUCH, S. L., & ALPERT, N. M. (1996). Individual differences in cerebral blood flow in Area 17 predict the time to evaluate visualized letters. *Journal of Cognitive Neuroscience*, *8*, 78-82.
- LEBIHAN, D., TURNER, R., ZEFFIRO, T. A., CUNOD, C. A., JEZZARD, P., & BONNEROT, V. (1993). Activation of human primary visual cortex during visual recall: A magnetic resonance imaging study. *Proceedings of the National Academy of Sciences*, *90*, 11802-11805.
- MELLET, E., TZOURIO, N., CRIVELLO, F., JOLIOT, M., DENIS, M., & MAZOYER, B. (1996). Functional anatomy of spatial mental imagery generated from verbal instructions. *Journal of Neuroscience*, *16*, 6504-6512.
- MELLET, E., TZOURIO, N., DENIS, M., & MAZOYER, B. (1995). A positron emission tomography study of visual and mental spatial exploration. *Journal of Cognitive Neuroscience*, *7*, 433-445.

- MELLET, E., TZOURIO, N., PIETRZYK, U., RAYNAUD, L., DENIS, M., & MAZOYER, B. (1993). Visual perception vs. mental imagery: A PET activation study. *Journal of Cerebral Blood Flow & Metabolism*, *3*(Suppl. 1), S536.
- MENON, R., OGAWA, S., TANK, D. W., ELLERMANN, J., MERKELE, H., & UGURBIL, K. (1993). Visual mental imagery by functional brain MRI. In D. L. Bihan, R. Turner, M. Mosley, & J. Hyde (Eds.), *Functional MRI of the brain: A workshop presented by the Society of Magnetic Resonance in Medicine and the Society for Magnetic Resonance Imaging* (p. 252). Arlington, VA: Society of Magnetic Resonance in Medicine.
- PETERSON, M. A., KIHLMSTROM, J. F., ROSE, P. M., & GLISKY, M. L. (1992). Mental images can be ambiguous: Reconstruals and reference-frame reversals. *Memory & Cognition*, *20*, 107-123.
- PUTZ, B., MIYAUCHI, S., SASAKI, Y., TAKINO, R., OHKI, M., & OKAMOTO, J. (1996). Activation of the visual cortex by imagery used in mental calculations. *NeuroImage*, *3*, S215.
- REISBERG, D. (1998). Constraints on image-based discovery: A comment on Rouw et al. (1997). *Cognition*, *66*, 95-102.
- REISBERG, D., & CHAMBERS, D. (1991). Neither pictures nor propositions: What can we learn from a mental image? *Canadian Journal of Psychology*, *45*, 336-352.
- ROBSON, J. G. (1993). Contrast sensitivity: One hundred years of clinical measurement. In R. Shapley & D. K. T. Lam (Eds.), *Contrast sensitivity: Proceedings of the Retina Research Foundation symposia* (Vol. 5, pp. 253-267). Cambridge, MA: MIT Press.
- ROLAND, P. E., ERIKSON, L., STONE-ELANDER, S., & WIDEN, L. (1987). Does mental activity change the oxidative metabolism of the brain? *Journal of Neuroscience*, *7*, 2373-2389.
- ROLAND, P. E., & GULYAS, B. (1994). Visual imagery and visual representation. *Trends in Neurosciences*, *17*, 281-296.
- ROUW, R., KOSSLYN, S. M., & HAMEL, R. (1997). Detecting high-level and low-level properties in percepts and mental images. *Cognition*, *63*, 209-226.
- ROUW, R., KOSSLYN, S. M., & HAMEL, P. (1998). Aspects of mental images: Is it possible to get the picture? *Cognition*, *66*, 103-107.
- SABBAH, P., SIMOND, G., LEVRIER, O., HABIB, M., TRABAUD, V., MURAYAMA, N., MAZOYER, B. M., BRIANT, J. F., RAYBAUD, C., & SALAMON, G. (1995). Functional magnetic resonance imaging at 1.5T during sensorimotor and cognitive task. *European Neurology*, *35*, 131-136.
- SHEPARD, R. N., & COOPER, L. A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.
- THOMPSON, W. L., & KOSSLYN, S. M. (in press). Neural systems activated during visual mental imagery: A review and meta-analyses. In J. Mazziota & A. Toga (Eds.) *Brain mapping II: The applications*. New York: Academic Press.

NOTE

1. One could be concerned about our use of stripes with sharp edges, which produced square wave gratings. The sharp edges imply that even the thick stripes have a high-resolution component. We have three responses to this concern: First, Robson (1993) reviews applications of research on contrast sensitivity measurements and notes (p. 261) that "the threshold contrast for square wave gratings at high spatial frequencies was in reality just slightly less than that for sine wave gratings" (see, e.g., Campbell & Robson, 1968). Following this, he observes that the relative visibility of sine and square wave gratings is "very much the same until the resolution was made very low indeed." Given that we did not use very low frequencies, our use of square wave gratings would not seem to be a problem. Second, these results all apply to near-threshold stimuli. Our stimuli were well above threshold. Third, the subjects did not need to use the highest resolution, required to resolve the sharp edges, to perform the task. Indeed, the interaction between resolution and mode of stimulus presentation was driven by a difference in the high-resolution stimuli, despite the fact that both sets of stimuli had sharp edges.

(Manuscript received April 25, 1997;
revision accepted for publication February 16, 1998.)