Varieties of Size-Specific Visual Selection

Kyle R. Cave
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology

Stephen M. Kosslyn
Harvard University

Compared time to evaluate stimuli of varying sizes. When Ss expect an upcoming stimulus to be a certain size, response time increases with the disparity between expected and actual size. There are, however, 2 size adjustment processes, and they reflect 2 types of visual selection. In the first, a shape-specific image representation is used to separate a visual object from a superimposed distractor. These representations require the type of slow size scaling demonstrated in imagery experiments. The second size scaling process is faster and not shape-specific. At any given time the visual system is set to process information at a particular scale, and that scale can be adjusted to match an object's size. Because both selection mechanisms depend on size, they probably occur at a relatively low, spatially organized processing level. These findings lead to a new explanation for results that had been taken as evidence for attentional selection at the level of object representations.

Visual attention entails making quick decisions about which aspects of the input are likely to be important and then selecting those aspects for further processing. It is now clear that many mechanisms subserve visual attention (Parasuraman & Davies, 1984), but it is not clear why so many are needed. Our experiments implicate two distinct mechanisms that allow one to attend to different-sized stimuli and also delineate important functional differences between the two mechanisms. We argue that although the end result may be the same—attention to a specific size—the different functional properties of the mechanisms make them useful for performing different kinds of tasks.

Larsen and Bundesen (1978) demonstrated that humans can set themselves to attend to different-sized stimuli. Their subjects responded faster when a stimulus appeared at an expected size than at an unexpected size. Indeed, as the disparity between the expected size and the unexpected size increased, the response times increased as well. Presumably the subject was set to view objects at a particular size. When the stimulus was a different size, an adjustment was necessary, and larger adjustments required more time.

The selection mechanism at work in Larsen and Bundesen's (1978) experiment seems to apply in general to all stimuli at a particular size, regardless of shape. The subjects responded more quickly to the expected size without knowing the identity of each letter before it appeared. Larsen and Bundesen contrasted this *perceptual scale adjustment* process with an *image size adjustment* process, which they measured in visual imagery experiments. In these experiments they used tasks that were designed to require a visual image, such as comparing a complex shape with another similar shape that was presented earlier. These imagery experiments were similar to a number of others (Bundesen & Larsen, 1975; Bundesen, Larsen, & Farrell, 1981; Kubovy & Podgorny, 1981; Larsen, 1985; Larsen & Bundesen, 1978; Sekuler & Nash, 1972). Like the perceptual scale adjustment process, the image size adjustment process observed in these experiments requires more time with larger disparities in the sizes of the original and test shapes. Presumably, subjects must adjust the represented size of the imaged shape in order to match it to the test shape.

Larsen and Bundesen (1978) claimed that these two size-scaling processes can be distinguished by the shape of the size-scaling function. They claimed that the time necessary for perceptual scale adjustment increases with the ratio of the expected and actual sizes, whereas the time necessary for image size adjustment increases with the log of the size ratio. However, as we demonstrate in the Appendix, both functions fit both sets of data very well, leaving little room to choose between them. In addition, however, we have observed in the literature that size scaling in imagery experiments always takes more time than the size scaling in Larsen and Bundesen's perceptual scale adjustment experiment. This difference suggests that the two kinds of size scaling can be distinguished by the rates at which they occur. However, the existing data do not allow one to rule out the possibility that subjects are using only a single mechanism, which simply is slower when more complex tasks are performed. In our experiments we show that the slower times for image size adjustment do reflect the operation of a distinct mechanism, and we also demonstrate that this process is used in visual parsing.

Experiment 1

In order to measure image scale adjustment, it is necessary to use a task that requires imagery, or at least one that is made easier by the use of an image. One task for which an image representation could be useful is visual parsing. An image must be a representation of one specific shape, and shape-specific representations of this sort have proved useful for parsing in computer vision systems, such as that developed by Lowe (1987a, 1987b). Lowe's system identifies objects by matching stored object shapes top-down against the visual input. Lowe's system uses three-dimensional models to generate two-dimensional templates. These templates are used to parse a complex visual scene into separate objects, even though the objects are jumbled together and partly occlude one another. In Lowe's system the templates (which are shape-specific representations) are adjusted in size and orientation until they match the input pattern. The part of the input that matches the template can then be parsed as a distinct object. Lowe's work raises the possibility that shape-specific object representations can be used to select particular objects from complex scenes for further processing, provided that those representations are properly adjusted in size and orientation to match the input.

Thus we hypothesize that the two types of size scaling identified by Larsen and Bundesen (1978) are associated with two kinds of visual selection. The more general type of selection mechanism simply picks a region of space or a scale of resolution for further processing. This mechanism does not depend on shape, and it results in more efficient processing for any stimulus that appears at the chosen size. Experiments in which this general selection process is manipulated should show patterns of relatively fast size scaling. The other mechanism, which should be useful in parsing figure from ground, is shape specific and may be related to visual imagery. This mechanism should produce relatively slow size scaling.

Experiment 1 was designed to differentiate between these two types of size-scaling processes. In this experiment, subjects began each trial with the expectation that the stimulus would appear at a particular size, just as the subjects did in Larsen and Bundesen's (1978) Experiment 2. The size of the stimuli varied from trial to trial, as did the size expected by the subject. We manipulated this expectation by using the same method that Larsen and Bundesen used: Usually the upcoming stimulus would be the same size as the preceding stimulus.

Each stimulus consisted of two superimposed objects. The same two objects were used in every trial, although from trial to trial each object could differ in the relative lengths of its sides. In each trial, one object was drawn in heavy lines and the other was drawn in light lines. The correct response was determined by whether all sides of the object drawn in light lines were of equal length. The object drawn in heavy lines was to be ignored. Therefore, it was necessary to parse the target object from the overlapping distractor object. Usually the target object in the upcoming trial was the same object as in the preceding trial, and so subjects could prepare for a particular object just as they could for a particular size. Thus the subject could begin each trial with an object expectation, as well as a size expectation.

If the visual system can be adjusted so that stimuli at a particular size are processed optimally, then responses should be faster when a stimulus appears at the expected size. If size adjustments are done gradually (either continuously or in small steps), then response times should increase steadily as the ratio between expected and actual sizes increases. Furthermore, if a perceptual scale adjustment is used, then there should be a similar effect of size ratio whether the target object is the expected or the unexpected object. On the other hand, if a size-specific image scale adjustment is used, then size ratio should affect response time only when the expected object appears. We should be able to compare the two types of size scaling processes by examining the size ratio effect for expected and unexpected objects.

Method

Subjects. Fourteen people (6 men and 8 women) served as paid subjects. Most were Massachusetts Institute of Technology (MIT) undergraduates, and all had normal or corrected vision. None knew the purpose of the experiment or the expected results beforehand.

Apparatus. The experiment was controlled by a Terak microcomputer running the RT-11 operating system. The stimuli were drawn with the Terak's graphics system and displayed on a small black-and-white monitor. Two microswitches served as response keys.

Stimuli. Each stimulus item was composed of two rectangles. For each one, all four sides could be equal, forming a square, or two parallel sides could be longer than the other two. One rectangle, which we call the *diagonal* object, was made up of diagonal lines, whereas the other, which we call the *orthogonal* object, was made up of horizontal and vertical lines (i.e., lines that are orthogonal to the frame of reference). The two were superimposed, as illustrated in Figures 1 and 2. In each stimulus one rectangle, the target, was drawn with relatively light lines, whereas the other, the distractor, was drawn with relatively heavy lines. In the following discussion, the orthogonally and diagonally oriented rectangles are referred to as different objects. Even though they are different rotations of the same shape, they are generally described as different figures (Mach, 1914). A square and a nonsquare rectangle, on the other hand, are for the purposes of discussion considered to be two instances of the same object if they are both orthogonally oriented or both diagonally oriented. Thus in this terminology each stimulus was made up of two superimposed objects, one diagonal and one orthogonal.

The stimulus items appeared in four sizes with ratios of 1:2:6:9. Because orthogonal distances between pixels are smaller than diagonal distances, orthogonal squares at the smallest size were made up of 19 pixels on each side, whereas diagonal squares had 14 on each side. The visual angles for these four sizes were approximately 1.9°, 3.8°, 11.4°, and 17.1°. Within any one stimulus, the two rectangles were always the same size. If one was a square and the other not, the sides of the square were equal in length to the longer sides of the other object.

Each subject viewed an equal number of stimuli at each of the four sizes. In half of the stimuli, the orthogonal object was the target and the diagonal object was the distractor, and in the other half, vice versa. These two types of stimuli were distributed evenly among the four sizes. Within each group of stimuli of a particular size and type, half of the target objects were square. Of those that were not, half were longer along one axis, and half were longer along the other. The distractor objects were varied in the same way, and this variation was counterbalanced against the variation in the target objects. Light lines were one pixel across. Heavy lines were two pixels across for the two smaller sizes and three pixels across for the two larger sizes. The

position of the orthogonal object was randomly shifted slightly to the left or right in order to reduce clues from the relative positions of the two objects. The amount of the shift was proportional to the size of the stimulus; the shift for the smallest size was a single pixel.

A different trial order was generated for each subject. The orders were constructed so that 75% of the trials matched the preceding trial according both to size of objects and to which object (either diagonal or orthogonal) was the target. Within this constraint, the orders were random. Within the 25% of the trials with either a size difference or an object difference from the preceding trial, there were equal numbers of every possible combination of previous size, current size, previous object, and current object. This constraint dictated that 10.7% of the trials had both a size and an object different from those in the previous trial, 10.7% had only a different size, and 3.6% had only a different object. Each subject received a total of 1,792 trials. Subjects were informed that "most of the time" the size would remain the same from trial to trial and that the target objects would "usually remain the same" as well.

Procedure. The subject sat in front of a video display monitor. A single stimulus (two superimposed objects) was displayed on the screen for 267 ms. The duration of the display was limited to discourage subjects from contemplating their responses for too long. The task was the same, regardless of whether the object to be judged was the orthogonal or the diagonal one. The subject was to decide whether all four sides of the light-lined object were the same length. This task was used to ensure that subjects examined a large part of the object and not just an isolated local feature.

A buzzer sounded after an incorrect response in order to provide feedback. Immediately after the response, the next trial began. A break occurred after every 50 trials. A filler trial occurred at the beginning of the experiment and after each break in order to provide a cue for the size and the target object of the actual trial to follow.
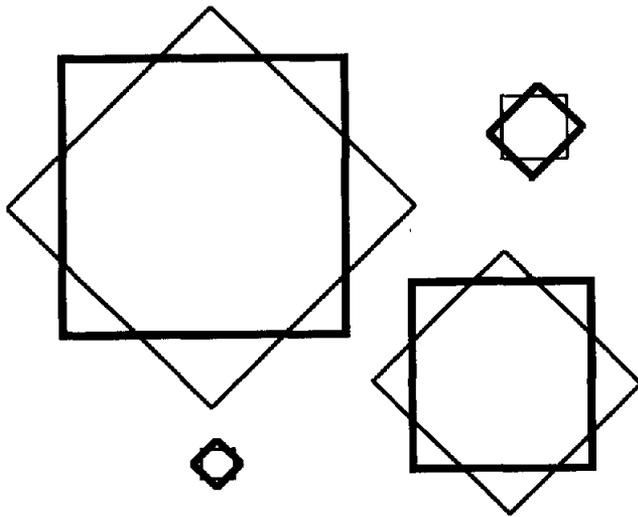
Figure 1. Examples of the stimuli in Experiment 1, illustrating the difference between the four sizes. (The difference between squares and nonsquares was adjusted according to each subject's performance. The stimuli in this figure are the most difficult stimuli to judge. The next figure shows those from the easiest end of the scale. In the upper two drawings in both figures, the orthogonal object is a square, whereas in the lower two, it is not. Likewise, the diagonal objects on the left are square, whereas those on the right are not. Subjects were to attend to the light-lined object and ignore the heavy-lined object.)
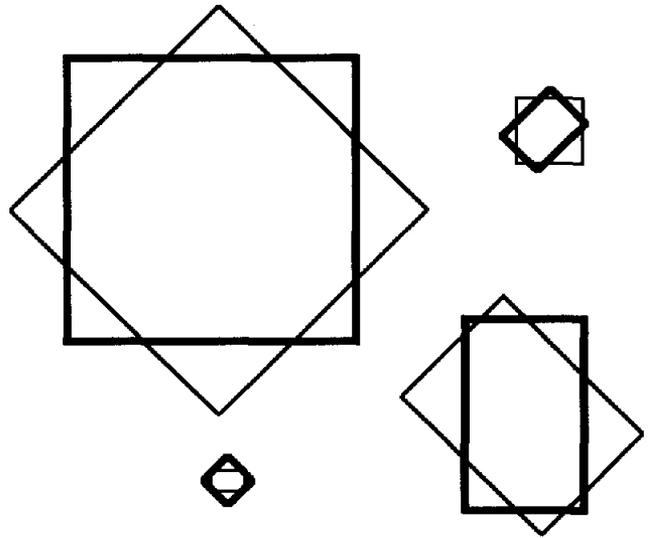
Figure 2. The easiest possible stimuli from Experiment 1.

Each subject performed all the trials in a single session, which typically lasted between 60 and 90 min.

We attempted to make the task equally challenging for all subjects, regardless of differences in their abilities to distinguish squares from nonsquares. After every 40 trials, the computer calculated the percentage of correct responses. Percentages were computed separately for the diagonal and the orthogonal stimuli that had occurred within the last 40 trials. If more than 97% of the responses for either type of stimulus were correct, the short sides on the nonsquare rectangles of that type were lengthened by one pixel on future trials, which made these objects more similar to squares and thus made the discrimination more difficult. If performance dropped below 90%, the short sides were shortened by one pixel. The difference in length between short and long sides never exceeded seven pixels and could be as low as a single pixel. This adjustment was performed automatically by the computer, and there was no mention of it in the instructions. The stimuli displayed in Figure 1 are the most difficult possible, and those in Figure 2 are the least difficult. The beginning level for each subject was determined by performance on practice trials.

Results

All data for incorrect responses were discarded. The highest error rate for an individual subject was 13.6%, and the mean error rate was 8.2%. For each subject, the response times were sorted into 64 groups, each with a different combination of expected size, actual size, expected object, and actual object. Within each group, all response times more than three standard deviations from the mean for that group were discarded. On the average, 1.4% of the correct trails were discarded for each subject. The largest percentage of discarded trials for a single subject was 1.9%.

Response times. The response times (measured in milliseconds) from both the expected- and unexpected-object conditions were organized according to the ratio between expected size and actual size. (We always made the larger term the numerator so that all ratios were greater than or equal to 1.) An analysis of variance (ANOVA) was performed with size

ratio, actual object, object expectation (expected object vs. unexpected object), and correct response as factors. The effects of size ratio and object expectation are considered first because they are the most theoretically interesting and because the effects associated with them were generally the largest.

As is evident in Figure 3, responses generally were much faster when subjects were prepared for the correct target object, $F(1, 13) = 18.78, p < .002, MS_e = 198,008$. Response times also varied greatly with the size ratio, $F(6, 78) = 17.50, p < .001, MS_e = 21,506$, and this variation differed between expected and unexpected targets, $F(6, 78) = 11.94, p < .001, MS_e = 10,810$.

When the target object was the one the subject expected, response times increased linearly with the size ratio, $r = .80$, $F(1, 78) = 140.64, p < .001, MS_e = 16,915$ (see Figure 3). There was also a linear increase in the unexpected-object condition, $r = .38, F(1, 78) = 13.34, p < .001, MS_e = 15,401$. A significant interaction between the linear contrast and target expectation revealed that the linear increase in the expected-object condition was much larger, $F(1, 78) = 55.07, p < .001$.

Even though the linear increase for unexpected objects was highly significant, the effect was small, and one might be tempted to believe that there was not a steady increase with size ratio but merely a difference between those stimuli that were exactly the expected size and all those that were not. To test this conjecture, we performed the linear contrast without the trials in which the stimulus appeared at the expected size (size ratio = 1). The result was still highly significant, $F(1, 65) = 10.19, p < .005, MS_e = 15,671$.

The remaining effects did not appear to be directly relevant to our questions about size scaling and selection. First, sub-

jects responded more quickly when the correct response was "equal" or, in other words, when the designated object was square, $F(1, 13) = 6.99, p < .03, MS_e = 29,674$. This effect is not surprising, partly because "equal" responses were made with the dominant hand. Subjects also responded faster when the object to be examined was oriented diagonally than when it was oriented orthogonally, $F(1, 13) = 20.45, p < .002, MS_e = 30,611$. Many subjects reported that discriminations were easier with the diagonal rectangles. Some claimed that they compared two vertices on opposite corners to determine whether they both fell on the same horizontal or vertical line.

In addition, subjects apparently took longer for size ratios of 1.5 and 4.5 when the expected object was orthogonal, which resulted in a significant interaction of size ratio, object expectation, and actual object, $F(6, 78) = 3.90, p < .003, MS_e = 12,163$. All trials in both of these size-ratio groups had either expected sizes or actual sizes of 9, the largest size used. Some subjects reported that stimuli at this size were especially difficult. Apparently this difficulty was more pronounced when the subject was prepared for an orthogonal stimulus. When a separate analysis was done without the data from trials in which either the expected size or the actual size was the largest size, there was no such interaction ($F < 1$).

Also, in the expected-object condition, responses to equal-sided targets were faster than responses to unequal-sided targets, but only for the larger size ratios. A different pattern emerged in the unexpected-object condition, in which the advantage for equal-sided targets applied to all size ratios except 1.5 and 4.5. Trials with these ratios involved either an expected or an actual size of 9, as noted earlier. These effects were revealed by a significant interaction of size ratio, object
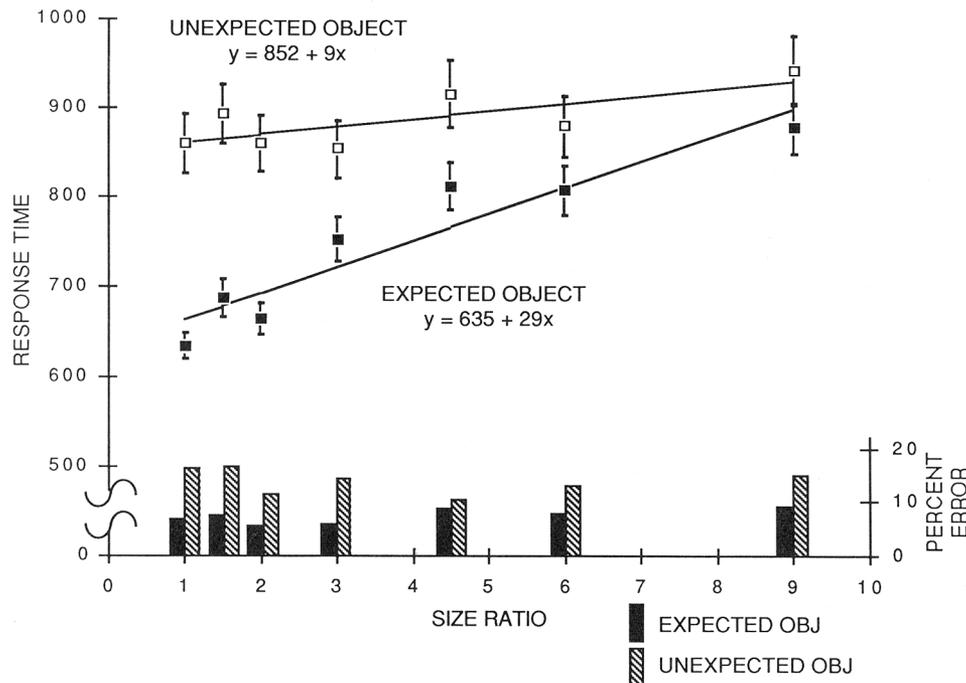

Figure 3. The results of Experiment 1, organized by size ratio. (At the bottom are the percentage of errors made for expected and unexpected object conditions for each size ratio.)

expectation, and correct response, $F(6, 78) = 2.28$, $p < .05$, $MS_e = 11,044$. No other effects or interactions were significant ($ps > .05$ in all cases).

We performed another analysis to examine the effect of adjusting the size upward versus adjusting it downward. In this analysis, we omitted all trials in which the actual size was the expected size because no size scaling was necessary in these trials. The direction of size scaling made no difference ($F < 1$). The only significant interactions involving scaling direction were among direction, size ratio, and actual object, $F(5, 65) = 3.15$, $p < .02$, $MS_e = 27,511$, and among direction, size ratio, actual object, and object expectation, $F(5, 65) = 2.37$, $p < .05$, $MS_e = 18,964$. Inexplicably, response time did not increase with size ratio when an unexpected diagonal target was larger than expected. All other effects and interactions that were significant in the original analysis were also significant in this analysis, except for the interaction of size ratio, object expectation, and equal/unequal response.

Error rates. Subjects made many more errors when they were prepared for the wrong object, $F(1, 13) = 51.58$, $p < .001$, $MS_e = 159$. They also made more errors when the correct response was "unequal," $F(1, 13) = 10.15$, $p < .01$, $MS_e = 226$, and when the indicated object was orthogonal, $F(1, 13) = 7.76$, $p < .02$, $MS_e = 321$. These effects all correspond to similar effects in the response time data. There were also different error rates for different size ratios, $F(6, 78) = 2.34$, $p < .05$, $MS_e = 89$, but there was no linear trend ($F < 1$). In fact, the differences in error rates for different size ratios were small (see the bottom of Figure 3). There is no evidence of a speed–accuracy trade-off. No other main effects or interactions in the error data were significant ($ps > .05$ in all cases).

Discussion

In Figure 3, the y intercept is clearly much higher for unexpected objects, and the slope is clearly much higher for expected objects. The disparity in y intercepts reveals that subjects required extra time to process an object for which they were not prepared. The disparity in slopes reflects differences in the size scaling for expected and unexpected objects. If the same size-scaling process were at work in both conditions, then the effects of expected/unexpected object and of size ratio should have been additive (Sternberg, 1969). Instead, the strong interaction between these factors suggests that different size scaling processes are used in the two conditions.

Subjects were best prepared for the upcoming stimulus when they knew which shape *and* which size to expect. This result would be forthcoming if subjects used this information to construct a mental representation to use in a comparison against the perceptual representation. When this prepared representation matched the stimulus in size and shape, the comparison could be performed quickly, and the response was fast. When the shape was correct but the size was not, the represented size could be adjusted. These adjustments were relatively slow; larger adjustments took more time. Even so, subjects were still able to respond faster if they adjusted the

size of the current representation than if they did the discrimination without it.

The representation that the subjects used here must be shape specific because it apparently cannot be used when the subjects prepare for the wrong object. Also, the size apparently cannot be changed all at once; it must be adjusted gradually. This result suggests that size is an integral part of the organization of the representation and argues against a completely "conceptual" representation in which spatial properties such as size have been factored out and coded symbolically (see Kosslyn, 1980; Pinker, 1984). Given these observations, it is reasonable to label the representation used in this condition as a *visual image* and to assume that the size scaling in the expected-object condition involves image size adjustment.

In the unexpected-object condition, the image prepared in advance of the target stimulus is of no use, and subjects must use a different strategy. The data from this condition exhibit a shallower slope but a higher y intercept. The slope in this case might reflect not the adjustment of an image but a perceptual scale adjustment. Setting the perceptual scale requires knowledge of the size of the upcoming stimulus, but not knowledge of its shape. If all higher level visual processing depends on the proper setting of the perceptual scale, then this adjustment must be fairly fast so that processing is not delayed too long. Even though this fast size scaling produces a shallower slope, responses are still relatively slow because of the higher y intercept. This extra time is necessary because the task must be done without the benefit of an image generated in advance. This method is slower than the image comparison used with expected objects, but it can be successfully applied to any shape, regardless of whether it is the expected shape. (It is possible, however, that an image is used in this case as well. The increase in y intercept might reflect the time necessary to build an image of the target object at the correct size after the stimulus appears.)

The perceptual scale adjustment takes time and is necessary only when the image shape does not match the target shape. Thus there might be good reason for subjects to delay this scale adjustment until after the image size scaling is completed and they can be sure that it is necessary. If they followed this strategy, however, the slope resulting from image size adjustment would be added to the unexpected-object slope, which would make it much larger than it actually was. Instead, subjects can apparently start the perceptual scale adjustment without first waiting for the image size adjustment to finish.

Perhaps both size-scaling operations are performed in parallel. Alternatively, subjects might be able to identify the target object before adjusting the image size, perhaps by choosing a light line and determining whether its orientation is orthogonal or diagonal. Only if the target object is the expected object would the subjects proceed with the image adjustment and use the image to perform the task.

There exists an important parallel between this experiment and the first of the mental-rotation experiments reported by Cooper and Shepard (1973). In both experiments, subjects were able to respond much more quickly if they had all the information necessary to prepare before the stimulus was presented. They had to know not only the amount of transformation required but also the shape that had to be trans-

formed. Without accurate information about the specific shape, information about size or orientation was much less useful.

Experiment 2

Experiment 2 served two purposes. First, we tested an alternative account of the results of Experiment 1 that does not posit that shape-specific representations were used in this task. Second, we explored the role of imagery in both the parsing of the target object and the square/nonsquare judgment. A visual image might be useful for either or both of these subtasks.

The alternative explanation tested here is built on the claim that the same sort of size scaling is used for both expected and unexpected objects, and it is always slow. In this case, the slow size-scaling process would be started for both objects as soon as the stimulus appears, regardless of whether the target is the expected or unexpected object. As this size scaling is going on, another process checks to determine whether the expected object is the target. If not, processing is switched from the expected to the unexpected object. This switch takes time, but it can be done in parallel with size scaling. If the size scaling finishes quickly, however, it is still necessary to wait for the switching process to finish. Thus response times in the unexpected-object condition are always slow, even for small size ratios, because the size-scaling time is obscured by the switching time. If this account is correct, then size scaling should be slow in the expected-object condition even when no distractor object is present.

In addition, Experiment 2 was designed to indicate how the image was used in Experiment 1. The decision in Experiment 1 was based only on the light-lined object, and the heavy-lined object that overlaps it had to be ignored. Somehow, information about the light-lined object must be extracted while information about the overlapping heavy-lined object is filtered out. We hypothesize that a shape-specific image representation helps in this selection of the target object. There is, however, another reason why subjects might find an image to be useful in this task: An image might be used in a template comparison to make the square/nonsquare decision easier. In fact, our task is very similar to the ratio comparison task used by Sekuler and Nash (1972), and their subjects apparently used images for just such a comparison.

On the other hand, there is good reason to believe that a task as simple as our square/nonsquare discrimination should not require an image. From previous experiments we can conclude that mental images are not always necessary to identify familiar shapes. For example, Kubovy and Podgorny's (1981) subjects needed no extra time to discriminate between two familiar shapes of different sizes, and Cooper (reported in Cooper & Shepard, 1973) found that subjects were able to identify letters quickly regardless of their orientation (although Jolicoeur, 1985, reported evidence that there is a small effect of orientation in identifying familiar shapes). It appears that images are most likely to be used to distinguish shapes from other shapes, such as reflections and rotations, that share many of the same simple visual features.

The role of imagery in parsing can be easily investigated in an experiment that requires the same discrimination as Experiment 1 but does not require the parsing. Thus Experiment 2 was like Experiment 1 in every respect but one: Each stimulus consisted of only a single object. There was no overlapping object to be ignored. If subjects in Experiment 1 used a visual image to distinguish squares from nonsquares, then Experiment 2 would yield similar results. If, on the other hand, the image was used to separate the target object from the overlapping distractor, then it would not be used here.

Method

Subjects. Fifteen people (9 men and 6 women) served as paid subjects. As before, most were MIT undergraduates, and all had normal or corrected vision. None had participated in Experiment 1, and none knew the purpose of the experiment or the expected results beforehand.

Stimuli. The stimuli were the same objects used in Experiment 1. Only the target object was displayed in each trial, however, and as before, it was always drawn in light lines. Each trial held the same predictive value for the following trial as in Experiment 1, and the instructions concerning this predictive value were the same as before.

Apparatus and procedure. The experiment was performed in the same room and with the same equipment used in Experiment 1. The instructions were the same, except that references to the overlapping heavy-lined object were removed. As before, the difficulty was adjusted according to performance. As before, each subject performed all the trials in a single session, which usually lasted between 60 and 90 min.

Results

As in Experiment 1, all data for incorrect responses were discarded. The mean error rate was 8.4%, and the highest individual error rate was 17.0%. On the average, 1.0% of the correct response times were more than three standard deviations from the corresponding trial group mean and were thus discarded. The highest amount of discarded data from an individual subject was 1.5%. After these outliers were removed, 1 subject still had a single response time of 34,866 ms. (This trial was not eliminated earlier because an even larger response time raised this subject's mean response time substantially.) This trial was also removed. All the remaining response times from all subjects were less than 3,000 ms.

Response times. For both conditions and all size ratios, responses were fast. As before, an overall ANOVA was performed with size ratio, object expectation, actual object, and equal versus unequal response as factors. As is evident in Figure 4, responses were faster when the indicated object was the one expected, $F(1, 14) = 5.80, p < .04, MS_e = 5,029$. As before, response times varied with size ratio, $F(6, 84) = 15.39, p < .001, MS_e = 4,121$, and as before, this variation was different for the expected and the unexpected objects, $F(6, 84) = 4.20, p < .002, MS_e = 3,421$. As in Experiment 1, these overall effects were tested with linear contrasts. We tested for a linear increase in size ratio separately for the expected- and unexpected-object conditions. The effect was very large with expected objects, $r = .711, F(1, 84) = 85.92, p < .001, MS_e =$
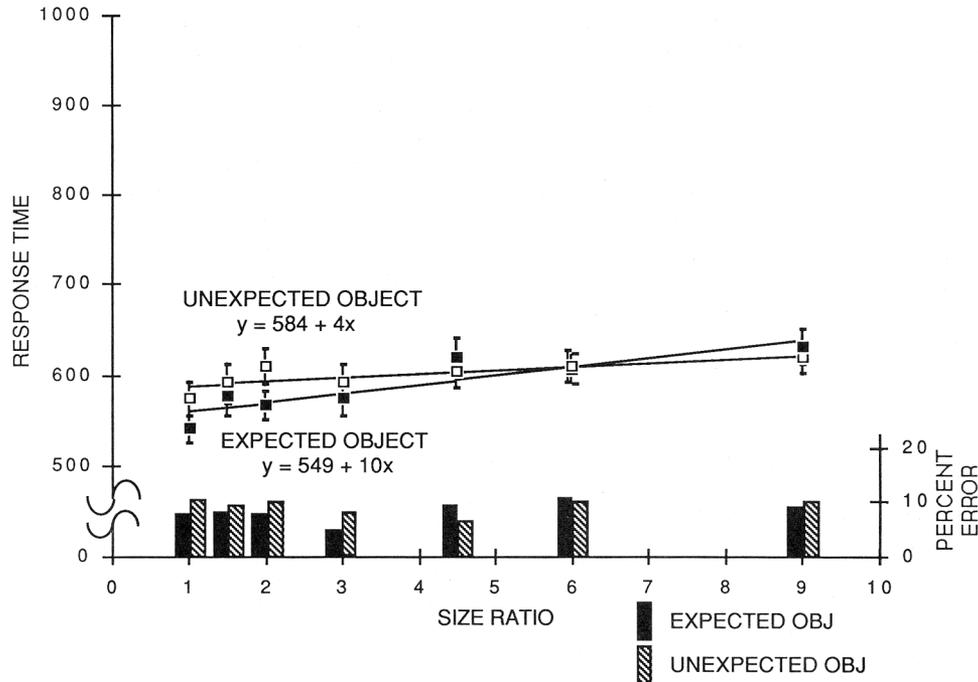
Figure 4. The results of Experiment 2, organized by size ratio. (Error rates are at the bottom.)

3,527, but still quite robust with unexpected objects, $r = .366$, $F(1, 84) = 13.03$, $p < .001$, $MS_e = 4,015$. To test for a difference in the linear increase between these two conditions, we returned to the overall analysis and crossed the linear contrast with object expectation. The result was significant, $F(1, 84) = 14.95$, $p < .001$. Thus there was a pattern somewhat similar to that of Experiment 1. Response time increased linearly with size ratio, and this increase was steeper when the target object was the expected one.

We tested the possibility that the significant linear contrasts were entirely due to differences between stimuli at exactly the expected size and those at all other sizes. For both the expected and unexpected object conditions, we performed linear contrasts after removing all trials with stimuli at the expected size. The effect persisted for both the expected-object condition, $F(1, 70) = 45.17$, $p < .001$, $MS_e = 3,893$, and the unexpected-object condition, $F(1, 70) = 5.36$, $p < .025$, $MS_e = 4,003$. Response times not only were higher when an unexpected size appeared but also increased steadily as the ratio between expected size and actual size increased.

In addition, there was an advantage for "equal" responses in trials with larger size ratios, as was reflected in the significant interaction between size ratio and correct response, $F(6, 84) = 3.61$, $p < .004$, $MS_e = 4,030$. No other interactions approached significance ($ps > .05$ in all cases).

As before, a separate analysis was done to compare trials in which the actual object was larger than the expected object with trials in which the actual object was smaller. There was no significant effect of this factor or of any interaction including this factor, and no other effects that were not significant in the original analysis were significant for this experiment ($ps > .05$ in all cases).

Error rates. The error rates are displayed at the bottom of Figure 4. An analysis of the error rates revealed differences for different size ratios, $F(6, 84) = 2.22$, $p < .05$, $MS_e = 92$, although, as in Experiment 1, the range of differences was small. There was some hint of a upward slope in error rate with size ratio, but it was very small and not significant, $F(1, 84) = 1.82$, $p > .1$. There was no speed–accuracy trade-off.

There were a number of other effects, and some were unexpected. The number of errors for each size ratio depended on the response type, $F(6, 84) = 2.24$, $p < .05$, $MS_e = 82$. When all four sides were equal, subjects made fewest errors with size ratios of 3 and 4.5. When the sides were not equal, subjects did best with the small ratios (1, 1.5, 2, and 3) and worst with the large ratios (6 and 9). Subjects made many more errors when attending to the orthogonally oriented stimulus, $F(1, 14) = 41.38$, $p < .001$, $MS_e = 55$, and in general made more errors when the sides were not all of equal length, $F(1, 14) = 36.87$, $p < .001$, $MS_e = 86$. They did worse with orthogonally oriented unequal targets than with others, $F(1, 14) = 5.14$, $p < .05$, $MS_e = 215$. No other effects approached significance ($ps > .05$ in all cases).

Comparison with Experiment 1. In Experiment 1, response times increased steeply with size ratio in the expected-object condition and increased much less steeply in the unexpected-object condition. Experiment 1 produced a large size-ratio effect with expected objects and a smaller size-ratio effect with unexpected objects. The purpose of Experiment 2 was to determine whether this pattern depended on the presence of a distractor object superimposed over the target object. Like the first experiment, this one revealed an interaction between the linear effect of size and object expectation. Such an interaction might at first glance be taken as evidence that

the same pattern of response times exists regardless of whether there is a superimposed object to filter out. However, a comparison of Figures 3 and 4 suggests that there is a big difference between the results from the two experiments. To test this impression, we combined data from both experiments in a single ANOVA with experiment as a between-subjects factor. Because there was 1 more subject in Experiment 2 than in Experiment 1, we chose a single subject at random from Experiment 2 and excluded that person's data from this analysis to make calculations easier.

The main motivation behind this analysis was to compare the effects of preparing for the correct object in the two experiments. Thus the most important interaction is that among size ratio, object expectation, and experiment (i.e., single vs. overlapping objects), and it was highly significant, $F(6, 156) = 6.07, p < .001, MS_e = 6,942$. To test whether the two experiments differed in the *linear* effect of size ratio, we performed a contrast in which we crossed the linear effect of size ratio, object expectation, and experiment. The result was highly significant, $F(1, 156) = 23.03, p < .001$. This result is confirmation that the linear increase in time with an increasing size ratio differed more between the expected- and unexpected-object conditions of Experiment 1 than it did between these two conditions in Experiment 2.

For the most part, the other results of this analysis were not surprising. Responses overall were slower in Experiment 1 than in Experiment 2, $F(1, 26) = 14.92, p < .002, MS_e = 1,391,951$, and the expected-object advantage was larger in Experiment 1, $F(1, 26) = 16.54, p < .001, MS_e = 99,942$. Also, of course, the general effect of size ratio was more pronounced in Experiment 1 than in Experiment 2, $F(6, 156) = 6.62, p < .001, MS_e = 12,711$.

In addition, subjects were faster for diagonal than for orthogonal targets in Experiment 2 but not in Experiment 1, $F(1, 26) = 16.30, p < .001, MS_e = 17,316$. Another difference between the two experiments involves the interaction among size ratio, object expectation, and actual object. The previous analysis of Experiment 1 showed this interaction, whereas the analysis of Experiment 2 did not. Thus it is not surprising that the combined analysis revealed that this interaction differed between the two experiments, $F(6, 156) = 2.89, p < .02, MS_e = 7,639$. In addition, the advantage of "equal" responses over "unequal" responses was large when subjects prepared for a diagonal stimulus and saw an orthogonal stimulus, but it was small when they prepared for an orthogonal stimulus and saw a diagonal one, $F(1, 26) = 4.88, p < .04, MS_e = 8,949$. This effect was not large enough to reach significance in the separate analyses for Experiments 1 and 2. All other interactions reflected effects already seen in the analyses performed on the two experiments separately.

Discussion

The results from Experiment 2 show that as long as the target appears unobscured, size scaling can be done very quickly. The fast size scaling that we observed allows us to argue against the single-mechanism account described earlier because that explanation predicts that size scaling will always be slow, for both expected and unexpected targets. The results from Experiment 2 also suggest that the primary role of imagery involves parsing and not recognition per se. The slow size-scaling process, which we attribute to image size adjustment, is evident only in the expected-object condition of Experiment 1, when the target must be parsed from the overlapping distractor and an image with the appropriate shape is ready for the job. The unexpected-object condition of Experiment 1 and both conditions of Experiment 2 showed very fast size scaling, which we take to reflect perceptual scale adjustment.

Although perceptual scale adjustment itself can be fast, the parsing operation required in Experiment 1 requires additional time, as can be seen when the overall response times from Experiment 2 are compared with the response times in the unexpected-object condition of Experiment 1. Subjects responded more slowly in the presence of a second overlapping object that was irrelevant to the decision. However, the subjects could save most of this extra parsing time if they knew in advance which object to observe *and* the size at which it will appear. When the expected size is incorrect, the time to adjust depends on how much the size must be adjusted. This combination of shape specificity and analog size scaling suggests that a visual image is involved in parsing.

Thus image-based parsing apparently lies behind the very steep slope observed in the expected-object condition of Experiment 1. However, in Experiment 2 there was still a small but significant difference in slopes between the expected- and unexpected-object conditions. No distractor objects were present in either condition, and so image-based parsing should have been unnecessary in either case. Why was the expected-object slope slightly steeper?

One possibility is that some or all of the subjects in Experiment 2 tried to use an image on a few trials, not to filter out the distractor, but to compare against the stimulus as a template. If they prepared an image before an expected-object trial, they would still need to adjust the image size if it did not match the stimulus size, and so these few trials would have a steep slope. When these few trials were averaged in with the rest in the expected-object condition, the average slope would be somewhat higher than it would be if no images were used on any trials. If subjects prepared an image for an unexpected-object trial, they would abandon it once the stimulus appeared, rather than adjust its size, because it would be the wrong shape. Thus the average slope for unexpected objects would be somewhat lower than that for expected objects, just as Figure 4 shows.

In short, Experiment 2 supports the claim that imagery was involved in Experiment 1 and that its primary role was to help in extracting the target object from the overlapping distractor. However, there is another alternative explanation that must be considered.

Experiment 3

In Experiment 1, subjects were confronted with two overlapping objects and had to decide which was the target, separate it from the other, and decide whether it was square. In Experiment 2 they saw only a single object, and they responded faster. Because it was not necessary to parse the target from a distractor, we inferred that the visual parsing required in Experiment 1 accounted for the slower perform-

ance. However, Experiment 1 also included a decision that Experiment 2 did not: Subjects in Experiment 1 had to decide which of the two objects was the target. If this decision took longer when the object's size was further from the expected size, then the steep slope in Experiment 1 might reflect decision time and not image size adjustment. The stimuli in Experiment 3 consisted of two overlapping rectangles as in Experiment 1, but in this experiment the trials were blocked so that the subjects could always be certain of the object that would be the target. Thus even though two objects were present, no decision was necessary. Any delay in responses in this experiment would have to be due to the parsing operation itself.

Method

Subjects. Fourteen people (7 men and 7 women) served as paid subjects and completed the experiment. Most were MIT undergraduates, and all had normal or corrected vision. None had participated in either of the previous experiments, and none knew the purpose of the experiment or the expected results beforehand.

Stimuli, apparatus, and procedure. The stimuli were the same as those used in Experiment 1. The experiment was performed in the same room with the same equipment as in the previous experiments. Each subject participated in two sessions on different days. In one session, the target was always orthogonal, whereas in the other, it was always diagonal. Half of the subjects participated in the orthogonal session first, and half participated in the diagonal session first. Each session usually lasted less than 1 hr. As before, the size of the previous stimulus served as a cue to the size of the current stimulus. In all other respects, the method and procedure were the same as in Experiment 1.

Results

As in the previous experiments, all response times for incorrect responses were discarded. The highest error rate for an individual subject was 6.9%, and the mean error rate was 5.5%. As in Experiment 1, response times for each subject were sorted into groups according to the combination of expected size, actual size, expected object, and actual object. In this case, because there was no unexpected-object condition, there were only 32 groups, rather than 64. Within each group, all values more than three standard deviations from the mean were discarded. On the average, 1.5% of the correct response times from each subject were discarded. The largest amount of discarded trials for a single subject was 2.3%.

Response times. As is evident in Figure 5, the presence of a distractor object clearly still made the task more difficult for subjects, even when they knew in advance which object would be the target. The ANOVA revealed that response times varied for different size ratios, $F(6, 78) = 32.79$, $p < .001$, $MS_e = 4,437$, with a significant linear trend, $r = .840$, $F(1, 78) = 187.22$, $p < .001$.

The remaining results are not obviously relevant to the effect of the distractor object. Not surprisingly, "equal" responses were faster, $F(1, 13) = 8.64$, $p < .02$, $MS_e = 8,496$, although this advantage apparently did not hold for size ratios of 1.5 and 6, $F(6, 78) = 3.26$, $p < .008$, $MS_e = 1,930$. Both of these ratios involve stimuli of size 6. The orientation of the target had no effect $(F < 1)$, nor did any of the other interactions $(ps > .05$ in all cases).

We performed another analysis in which all of the trials with the target at the expected size were removed and direction
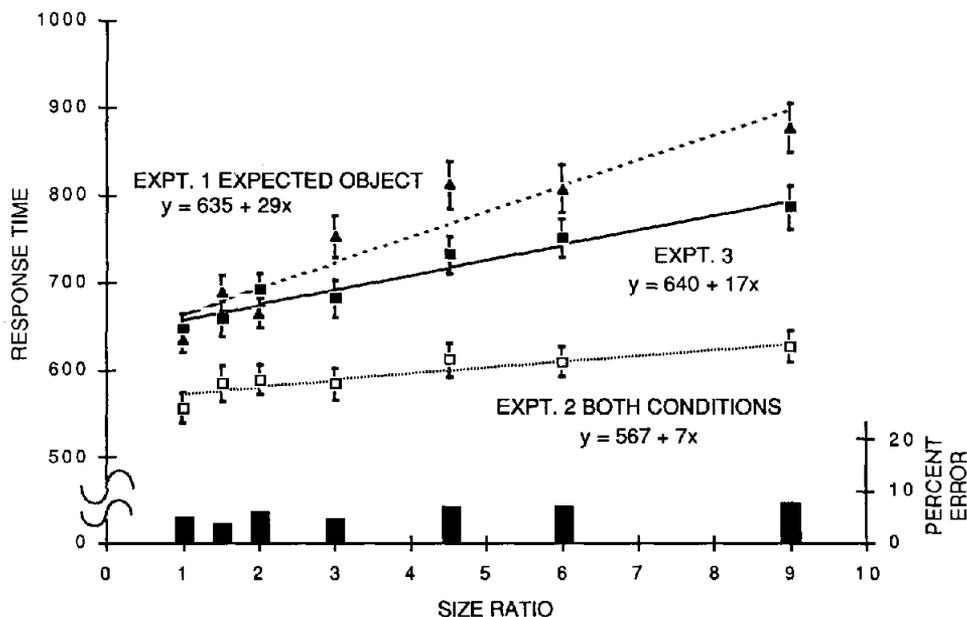

Figure 5. The results of Experiment 3, arranged by size ratio. (Also included are the data from the expected figure condition of Experiment 1 and the combined data from both conditions of Experiment 2. At the bottom are error rates from Experiment 3.)

of size adjustment was added as a factor. Surprisingly, subjects were faster when adjusting the size up than when adjusting it down in this experiment, $F(1, 13) = 7.10$, $p < .02$, $MS_e = 18,842$. Subjects had difficulty when the target unexpectedly appeared at the smallest size, which is reflected in an interaction between size ratio and adjustment direction, $F(5, 65) = 3.97$, $p < .004$, $MS_e = 7,436$. Given how small these stimuli were, it is perhaps surprising that this interaction did not appear in the earlier experiments. There were no other significant effects that did not also appear in the original analysis ($ps > .05$ in all cases).

Error rates. The error rates are displayed at the bottom of Figure 5. The analysis revealed differences in errors for the different size ratios, $F(6, 78) = 3.46$, $p < .005$, $MS_e = 37$, with a significant linear trend, $F(1, 78) = 13.53$, $p < .001$. Thus there was no speed–accuracy trade-off. There were no other significant effects ($ps > .05$ in all cases).

Comparison with Experiments 1 and 2. In order to compare the linear response time increase with size ratio in Experiment 3 with the similar patterns in Experiments 1 and 2, we performed two more analyses. In the first, the data from Experiment 3 were combined with the data from the expected-object condition of Experiment 1 in a single analysis, with experiment as an added factor. Although there was no significant difference in mean response times between the two experiments ($F < 1$), they did differ in the effect of size ratio, $F(6, 156) = 5.58$, $p < .001$, $MS_e = 10,676$. The linear increase in response time with greater size ratios was smaller in Experiment 3, when subjects could be certain of which target object to evaluate. This was demonstrated with a contrast in which we crossed the linear size-ratio effect with experiment, $F(1, 156) = 19.00$, $p < .001$. The interaction between experiment and target object was also significant, $F(1, 26) = 5.52$, $p < .03$, $MS_e = 64,322$, as was the interaction between target object and size ratio, $F(6, 156) = 3.30$, $p < .005$, $MS_e = 5,300$. All other effects reflected effects seen in earlier analyses.

In the second analysis, the data from Experiment 3 were combined with the data from the expected-object condition of Experiment 2. We chose a single subject from Experiment 2 at random and dropped that person's data to equate the sample sizes. Overall, responses were faster in Experiment 2, $F(1, 26) = 6.00$, $p < .03$, $MS_e = 502,381$, and once again the relation between response time and size ratio varied between the two experiments, $F(6, 156) = 4.64$, $p < .001$, $MS_e = 3,894$. As expected, response time increased with size ratio more sharply in Experiment 3 than in Experiment 2, $F(1, 156) = 21.32$, $p < .001$. All other significant effects reflected similar effects from earlier analyses.

Discussion

Two important results emerged from this experiment. The first is that size scaling was faster in Experiment 3 than it was in the expected-object condition of Experiment 1. Apparently the time to decide which object is the target does increase with size ratio. Part of the steep slope in the expected-object condition of Experiment 1 may be due to this decision time. One reason why decision time might rise with size ratio is

that subjects might not be able to determine which object is the target without first performing a perceptual scale adjustment. If so, then the response time for each trial in the expected-object condition of Experiment 1 should be a combination of the time necessary for perceptual scale adjustment and the time necessary for image size adjustment. Thus the sum of the slopes from Experiments 2 and 3 (7 + 17 = 24) should be equal to the slope from the expected-object condition of Experiment 1 (29), and in fact they are not too far apart.

The second noteworthy finding is that size-scaling time increased when a distractor object was present, even when there was no target decision to make (as shown by the comparison of Experiments 2 and 3). Thus our conclusion that parsing is associated with slow size scaling still stands. A large part of the steep slope measured in Experiment 1 was apparently due to the visual separation of the target from the distractor object.

General Discussion

A number of conclusions emerge from these experiments. Among the most important is that two size-scaling processes are used in human vision. One operates by adjusting the represented size of a shape-specific memory representation. The other operates by adjusting a perceptual mechanism that selects the incoming visual information by scale.

Image Size Adjustment

We have hypothesized that subjects adjusted the represented size of visual image representations in the expected-object condition of Experiment 1 and in Experiment 3. This conclusion is consistent with a number of earlier studies in which researchers used image size transformations that produced similar response time patterns (Bundesen & Larsen, 1975; Bundesen et al., 1981; Kubovy & Podgorny, 1981; Larsen, 1985; Larsen & Bundesen, 1978; Sekuler & Nash, 1972). However, Sekuler and Nash and Bundesen et al. also found that the effects of size transformation and rotation were approximately additive. In our experiments the unexpected object is the same shape as the expected object, but rotated 45°. Therefore, we might expect the size and object expectation effects to be additive. Instead, we found a very large size effect when the target object was the expected one and a very small effect when it was unexpected. One reason why our subjects might have treated expected and unexpected objects so differently is that our instructions described the stimuli as different objects and not two rotations of the same object. Another possible reason is that some subjects judged whether diagonal stimuli were square by comparing the alignment of the left and right vertices, as mentioned in Experiment 1. This strategy could not have been used in Sekuler and Nash's successive rectangle judgment task.

Sekuler and Nash (1972) also found a large size effect, which suggests that subjects used images in performing the task. However, their stimuli were very similar to those in Experiment 2, in which our subjects used only perceptual

scale adjustment and not image size adjustment. Although their task of comparing the length/width ratio of two successive rectangles seems very similar to our task of comparing length and width of a single rectangle, our task required discrimination between a square and a rectangle, whereas theirs required discrimination between two rectangles. Perhaps squares are classified differently from other rectangles early in perceptual processing. (This could be why we have distinct names for them.) This categorization would be useful in our task, but it would not help Sekuler and Nash's subjects to compare rectangles, and thus they would be more likely to rely on visual images.

Perceptual Scale Adjustment

The results from Experiment 2 and the unexpected-object condition of Experiment 1 demonstrated a small but highly significant effect of size ratio that did not depend on knowledge of the target shape. This effect apparently reflects the presence of a selection mechanism that chooses visual inputs on the basis of size. The extra time for larger size ratios presumably is necessary to adjust the size that is to be selected. Such a mechanism would be necessary if higher level visual processes were limited to processing at a single scale. In order for inputs of different sizes to be processed, the input might be adjusted by a scale factor to match this standard before being passed on to the higher level processes. At any given time, a single scale factor must be set, and thus only one particular size of stimulus will be mapped to the standard size. When an object's size is known before its appearance, the scale factor can be chosen in advance. When an object appears at some other size, the scale factor must be adjusted until the object's size maps to the standard size. This adjustment must be fast, because higher level processing cannot

proceed until it is done. However, it is still gradual, so that larger adjustments take more time.

This selection by size is in many ways analogous to the selection by location that has been demonstrated in numerous attention experiments (Eriksen & Hoffman, 1972a, 1972b; Posner, Nissen, & Ogden, 1978; Posner, Snyder, & Davidson, 1980). Instead of expecting a stimulus at a particular size, subjects in these experiments were told to expect a stimulus at a particular location. They responded faster when the stimulus actually appeared at the expected location, which suggests that this location was chosen at the expense of other locations for special processing. Thus observers apparently set themselves to process an object at a particular size and a particular location, even when these spatial properties are not relevant to the task.

As noted earlier, Larsen and Bundesen (1978), in their Experiment 2, measured a similar perceptual scale adjustment. In Figure 6 we present their data along with data from Experiments 1, 2, and 3. To allow comparison, we organized our data as they had organized theirs. They tried to control for overall effects of stimulus size, regardless of size expectation. To do so, they adjusted the mean response time for each of the unexpected-size conditions by subtracting from it the response time for the expected-size condition with the same actual size.

Larsen and Bundesen's (1978) data are intertwined with the data from both unexpected-object conditions (see Figure 6). This is just as we would predict because these conditions should all require the same perceptual scale adjustment. The slope in the expected-object condition of our Experiment 2 is similar to the others, although it is somewhat higher. As mentioned earlier, subjects might have occasionally used imagery in this condition. All of these slopes are far below those from Experiment 3 and the expected-object condition of
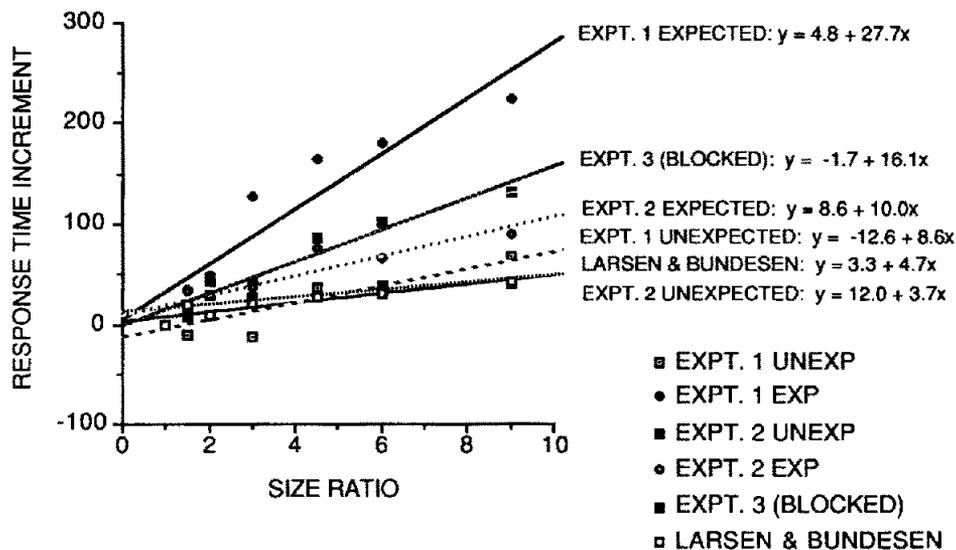

EXPT. 1 EXPECTED: y = 4.8 + 27.7x

EXPT. 3 (BLOCKED): y = -1.7 + 16.1x

EXPT. 2 EXPECTED: y = 8.6 + 10.0x
EXPT. 1 UNEXPECTED: y = -12.6 + 8.6x
LARSEN & BUNDESEN: y = 3.3 + 4.7x
EXPT. 2 UNEXPECTED: y = 12.0 + 3.7x

▫ EXPT. 1 UNEXP
● EXPT. 1 EXP
■ EXPT. 2 UNEXP
◇ EXPT. 2 EXP
■ EXPT. 3 (BLOCKED)
▢ LARSEN & BUNDESEN

Figure 6. The amount that response time increases with each size ratio for Experiments 1, 2, and 3. (Data from Larsen and Bundesen's 1978 scale transformation experiment are included for comparison. EXPT = Experiment; UNEXP = Unexpected; EXP = expected.)

Experiment 1, which presumably reflect image size adjustment.

Finding the Best Size Scaling Functions

If two distinct size-scaling processes are to be differentiated and characterized, there must be some way to tell them apart. Larsen and Bundesen (1978) argued that the two types of size transformations can be distinguished by the type of function relating response time to size ratio. As noted earlier, they claimed that image transformations produce response times that are linear with size ratio and that scale transformations produce response times that rise with the log of the ratio.

Upon analysis, we found that our data correspond about as well to the size ratio as to the log of the size ratio. Furthermore, we were unable to convince ourselves that either function accounted for Larsen and Bundesen's (1978) data better than the other. In the Appendix, we examine five different plausible size scaling functions and find that the data from our experiment and from Larsen and Bundesen's experiment are generally inadequate for distinguishing among them. Although we used the size ratio to predict the results of all three experiments, the only prudent conclusion is that response time generally increases with the disparity between expected and actual sizes. Exactly how disparity should be measured, both for image size adjustment and for perceptual scale adjustment, remains to be determined.

Whatever conclusion one might draw about the shape of the functions for image transformations and scale transformations, there is another factor that separates Larsen and Bundesen's (1978) image transformation data from their scale transformation data. In their study and in ours, the range in response times from slowest to fastest was much smaller for scale transformations than for image transformations, which indicates that scale transformations are performed more quickly. In cases in which imagery was presumably used, in the expected-object condition of Experiment 1 and Experiment 3, the slopes were 29.3 and 17.3, respectively. In contrast, when imagery was presumably not used, in the unexpected-object condition of Experiment 1 and in the two conditions of Experiment 2, the slopes were only 8.6, 10.1, and 4.2. (As mentioned earlier, the slope from the expected-object condition of Experiment 2 may be somewhat elevated because subjects may have occasionally used an image.) Similarly, a linear fit to Larsen and Bundesen's (1978) second experiment, in which imagery presumably was not used, had a slope of 4.7 ms per unit increase in size ratio, whereas slopes from their imagery conditions in Experiments 1 and 3 were around 12, as estimated from their graphs. In other experiments that were designed to elicit imagery, Bundesen et al. (1981) found a slope of 12 when the second stimulus was larger than the first and a slope of 20 when the second was smaller. The stimuli in this experiment were letters and digits. In another study designed to require imagery, Bundesen and Larsen (1975) displayed in their graphs slopes of around 50 for open line drawings and around 30 for filled random polygons.

This wide variation in slopes for image size scaling is not surprising. In these different experiments, the researchers used a variety of stimuli, some of which were more complex than others. Varying the complexity of the image could vary the time necessary to adjust its size, and this would result in the slope differences across imagery experiments.

Even though image size adjustment speed varies considerably, every case of image size adjustment is slower than the perceptual scale adjustment in our Experiment 2, the unexpected-object condition of our Experiment 1, and Larsen and Bundesen's (1978) Experiment 2 (which was designed to prevent the use of images). Given that Larsen and Bundesen's scale transformation slope is so near the slope in our two unexpected-object conditions, it is tempting to speculate that scale transformations occur at a constant rate, as might be expected from a low-level perceptual process that is operating independently of what is appearing in the visual field.

Use of Imagery in Parsing

We infer that image representations are involved when spatial transformations are relatively slow, when these transformations are only specific to a particular shape, and when the time necessary for these adjustments depends on the amount of adjustment required. A number of previous experiments have demonstrated slow, gradual image size adjustment. Subjects in these experiments usually had to distinguish a complex target shape from distractor shapes that shared component features with the target. Thus we know that images can represent at least some complex spatial configurations.

The experiments reported here are the first to provide evidence for the use of imagery in parsing figure from ground, and by revealing imagery's role in parsing, they shed new light on the nature and uses of visual images. Our results suggest that images can interact with incoming perceptual information to select some parts of the input and filter out others. This finding adds to the evidence that imagery is intertwined with perception and clarifies the nature of that relation.

Finding a link between visual parsing and imagery is not a surprise. A number of researchers have suggested that perception and imagery share mechanisms (see, e.g., Brooks, 1967; Farah, 1988; Finke & Shepard, 1986; Kosslyn, 1980; Segal & Fusella, 1970; Shepard & Cooper, 1982). Perhaps visual parsing is accomplished, at least in part, by top-down activation within the visual system. If so, then this same sort of activation might be used in the absence of perceptual input to generate images, which would then be processed like incoming visual information. This explanation is supported by Farah's (1985) evidence that imagery produces facilitation for certain visual tasks.

An alternative to the image-as-parser explanation is that the target object is compared against an image only when the task is difficult. Experiments 1 and 3 might be more difficult because of the presence of the distractor object. If images were being used as templates for comparison, we would expect our subjects to report the use of the image, because in other tasks in which images are compared against perceived shapes, subjects are usually very aware of the use of images. When subjects are looking for a difference between the image and the stimulus, they must regard the image as something sepa-

rate from the stimulus. However, when our subjects were asked after the experiment whether they had performed the equal/unequal discrimination by comparing the stimulus against an image, most responded "No." These reports are consistent with the image-as-parser explanation. In this case, the image corresponds to the target object, and forming the image is part of perceiving the object. Subjects would not perceive the image as something separate from the stimulus and would not report it.

Thus we infer that imagery can help to select a visual object from overlapping distractors. Because this selection depends on the size of the object, it is probably occurring at a relatively early, spatial-processing level. However, if this selection is actually driven by the same mechanisms used to generate mental image representations, then it involves conceptual representations and higher level visual processes. It seems that although this selection is done relatively early in the processing of a stimulus, it might be directed from higher processing levels.

Levels of Selection

Both of the selection mechanisms studied here illustrate that information at a particular scale is chosen for special processing. This selection is clearly occurring at a relatively early processing level in which information is still represented spatially, before spatial properties such as size are abstracted out.

Duncan (1984) offered evidence that selection occurs at a relatively late level of processing in which spatial properties have been removed and a separate representation has been constructed for each object. His subjects saw two superimposed objects on each trial and made two visual discriminations. (In fact, our superimposed stimuli were modeled after Duncan's.) Sometimes both discriminations involved the same object, and so only one object required attention and the other could be ignored. In one condition, however, the first discrimination involved the first object and the second discrimination involved the second object so that subjects were required to attend to both objects. Because performance was worse in this last condition, Duncan concluded that only one object at a time can be processed and that switching to another object requires effort. The interference between two superimposed objects cannot be attributed to selection of a particular region, because they both occupied approximately the same region. Duncan argued that selection must occur late in processing, when separate representations for the two objects have been constructed.

Duncan (1984) asserted not only that selection occurs late in visual processing but also that it occurs only at that time. He argued that there is no selection in the early stages of processing and that the position effects found by Posner and his colleagues (Posner et al., 1978; Posner et al., 1980) and Eriksen and Hoffman (1972a, 1972b) reflect not the spatial organization of the visual information but the perceptual grouping that results from it.

From our data, little can be said about interference between visual objects at an abstract representational level because all of the effects seem to occur at lower levels of processing.

However, the importance of parsing suggests a reinterpretation of Duncan's (1984) findings. Duncan found a decrease in performance when subjects were forced to examine both of two overlapping objects. This decrease might not have been the result of interference at some high level of processing in which object representations are manipulated, as Duncan implied. Rather, it may reflect the added difficulty of parsing a second target object. When both discriminations require attention to the same object, subjects can prepare to select that object before the trial begins. But when attention to both objects is required, subjects must first activate the representation for one object and then switch and activate the other. Thus Duncan's performance difference could be due to relatively low-level visual parsing.

In any event, the demonstration of perceptual scale adjustment shows that some part of the visual system optimally processes information at a particular size and that the chosen size can be gradually altered. The demonstration of image size adjustment with superimposed objects shows that subjects can prepare to select one object over another, but only if the object appears at a particular size. Because both of these methods of selection depend on a spatial property, they both probably occur at a level in which information is still organized spatially, rather than at a higher level in which information about visual objects is coded abstractly. Thus the existence of these two size-scaling processes constitutes evidence for selection at a relatively early level of visual processing.

References

Brooks, L. (1967). The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology, 19*, 289–299.

Bundesen, C., & Larsen, A. (1975). Visual transformation of size. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 214–220.

Bundesen, C., Larsen, A., & Farrell, J. E. (1981). Mental transformations of size and orientation. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 279–294). Hillsdale, NJ: Erlbaum.

Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing* (pp. 75–176). New York: Academic Press.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*, 501–517.

Eriksen, C. W., & Hoffman, J. E. (1972a). Some characteristics of selective attention in visual perception determined by vocal reaction time. *Perception & Psychophysics, 11*, 169–171.

Eriksen, C. W., & Hoffman, J. E. (1972b). Temporal and spatial characteristics of selective encoding from visual displays. *Perception & Psychophysics, 12*, 201–204.

Farah, M. J. (1985). Psychophysical evidence for a shared representational medium for mental images and percepts. *Journal of Experimental Psychology: General, 114*, 91–103.

Farah, M. J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review, 95*, 307–317.

Finke, R., & Shepard, R. N. (1986). Visual functions of mental imagery. In K. R. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance, Vol. II: Cognitive processes*

and performance (chap. 37). New York: Wiley.

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory & Cognition, 13*, 289–303.

Kosslyn, S. M. (1980). *Image and mind.* Cambridge, MA: Harvard University Press.

Kubovy, M., & Podgorny, P. (1981). Does pattern matching require the normalization of size and orientation? *Perception & Psychophysics, 30*, 24–28.

Larsen, A. (1985). Pattern matching: Effects of size ratio, angular difference in orientation, and familiarity. *Perception & Psychophysics, 38*, 63–68.

Larsen, A., & Bundesen, C. (1978). Size scaling in visual pattern recognition. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 1–20.

Lowe, D. G. (1987a). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence, 31*, 355–395.

Lowe, D. G. (1987b). The viewpoint consistency constraint. *International Journal of Computer Vision, 1*, 57–72.

Mach, E. (1914). *The analysis of sensations.* Chicago: Open Court Publishing.

Parasuraman, R., & Davies, D. R. (Eds.) (1984). *Varieties of attention.* Orlando, FL: Academic Press.

Pinker, S. (1984). Visual cognition: An introduction. *Cognition, 18*, 1–63.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. L. Pick & I. J. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137–157). Hillsdale, NJ: Erlbaum.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*, 160–174.

Segal, S. J., & Fusella, V. (1970). Influence of imaged pictures and sounds on detection of visual and auditory signals. *Journal of Experimental Psychology, 83*, 458–464.

Sekuler, R., & Nash, D. (1972). Speed of size scaling in human vision. *Psychonomic Science, 27*, 93–94.

Shepard, R. N., & Cooper, L. A. (1982). *Mental images and their transformations.* Cambridge, MA: MIT Press.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315.

Appendix

Almost any reasonable algorithm for gradually adjusting orientation will change it by a fixed number of degrees in each time unit. Similarly, most algorithms for gradually adjusting position will change it by a fixed distance in each time unit. Size is different, however, in that different conceivable algorithms for gradually adjusting size predict different functions for change in size over time.

The differences between size-scaling functions could be important in distinguishing between perceptual scale adjustment and image size adjustment. Larsen and Bundesen (1978) claimed that they could distinguish between the two because the time necessary for image adjustments increased linearly with size ratio, whereas the time necessary for perceptual scale adjustments increased with the log of the size ratio. In this Appendix we compare our data and Larsen and Bundesen's against the size ratio, the log of the size ratio, and other functions.

Perhaps the most straightforward assumption is that response time should increase linearly with the size ratio, but there is a good theoretical reason to expect some other function instead. Cooper and Shepard (1973) argued that while an image is being transformed from one orientation to another, it passes through states that correspond to intervening orientations. If the same is true for size transformations, and if $a < b < c$, then an image scaled from size a to size c at some point corresponds to an image at size b. If so, then size scaling should be sequential-additive. The time it takes to adjust from a to c ought to be the sum of the times that it takes to go from a to b and from b to c. If response time increases linearly with size ratio, then size scaling will not be sequential additive. For instance, a size change from 1 directly to 3 should require $\frac{3}{1}$, or 3, time units. A change from 1 to 2 followed immediately by a change from 2 to 3 should require $\frac{2}{1} + \frac{3}{2}$, or $3\frac{1}{2}$ time units.

However, size scaling will be sequential-additive if the time necessary depends on the log of the size ratio. There are other size scaling methods that are sequential-additive as well. Bundesen et al. (1981) presented an algorithm in which the viewer adjusts the size of an image by "moving the object" toward or away from the viewer at a constant rate until its apparent size is the desired size. Size scaling in this algorithm is still sequential-additive in terms of the distance from the viewer, but response time rises linearly with size ratio. This is just one of many possible algorithms, and many of them predict different sorts of response time functions.

Listed as follows are a number of plausible ways in which visual representations might be adjusted for size. Along with each is the function that it predicts between the two sizes and response time. All are built on the assumptions that scaling takes more time when more size adjustment is required and that scaling is sequential-additive. The purpose of this list is to arrive at a set of different functions that might be found in the size-scaling data. Thus none of these scaling methods is investigated in any detail, and many variations on each method that produce the same function are ignored. If image size scaling and perceptual scale adjustment are different processes, then they could easily involve different methods and produce different functions.

In the following formulas, S_1 is the beginning size, S_2 is the ending size, T is the time required for the transformation, and k is a constant (not necessarily the same from example to example). Size is proportional to diameter. In the examples it is assumed that a smaller size is being adjusted to a larger size, although all these methods can be modified to perform in the other direction scaling.

1. Perhaps the most straightforward possibility is that over each unit of time, a constant k is added to the size (see top of Figure A1):

$$S_2 = S_1 + kT;$$

$$T = \frac{1}{k}(S_2 - S_1).$$

Response time should increase linearly with the difference between the two sizes.

2. Another possibility is that over each unit of time, the size is multiplied by a constant k. This sort of pattern would appear if an image were divided into concentric rings and, in each time step, the contents of each ring were copied into the next ring out (see bottom of Figure A1):

$$S_2 = S_1 \times k^T;$$

$$k^T = \frac{S_2}{S_1};$$

$$T = \log_k \frac{S_2}{S_1}.$$

Response time should increase linearly with the log of the ratio between the two sizes.

3. Over each unit of time, a constant is added to the area.

$$S_2{}^2 = S_1{}^2 + kT;$$

$$T = \frac{1}{k}(S_2{}^2 - S_1{}^2).$$

Response time should increase linearly with the difference between the squares of the two sizes.

4. Over each unit of time, the area is multiplied by a constant:

$$S_2{}^2 = S_1{}^2 \times k^T;$$

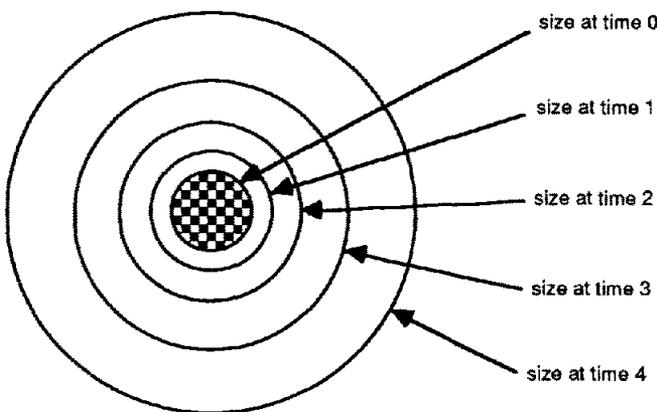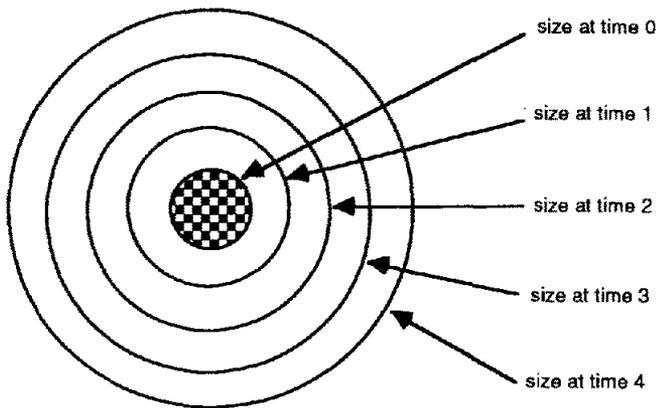$$k^T = \left(\frac{S_2}{S_1}\right)^2;$$

$$T = 2\log_k \frac{S_2}{S_1}.$$

size at time 0
size at time 1
size at time 2
size at time 3
size at time 4

size at time 0
size at time 1
size at time 2
size at time 3
size at time 4

Figure A1. Two types of size scaling. (In the upper diagram, a constant amount is added to the size in each time unit. In the lower diagram, the size is multiplied by a constant amount in each time unit. These two methods of size scaling result in different response time functions.)

As in Example 2, response time should increase with the log of the size ratio.

5. Bundesen et al. (1981) suggested a very different method of size scaling. They claimed that when their subjects saw two shapes of different sizes, they treated them as two shapes with the same real size but at different distances. Subjects made the apparent sizes equal by "moving" the shapes until they were both at the same distance in a process that would be similar to making an image of an object moving in three-dimensional space.

The two overlapping rectangles at the top of Figure A2 represent the two stimuli as they appear in the original stimulus. The angles enclosing these two shapes are marked S_1 and S_2 because shapes at those apparent sizes fit within them. Both shapes are at distance D from the viewer. In the first step, the subject adjusts the represented actual sizes of the objects by changing their represented distances while holding their apparent sizes constant. The two rectangles in the second half of Figure A2 represent the two objects after their represented sizes are made equal. This step presumably always requires a fixed amount of time.

In the second step, the distance for one object is adjusted while the actual size is held constant, so that the apparent size changes. Eventually the distance is made equal to the other object's distance, and their apparent sizes are the same. Moving the object over a larger distance requires more time. After the first step, the two shapes are represented at distances RD_1 and RD_2, with both at size RS:

$$\frac{S_1}{D} = \frac{RS}{RD_1} \text{ and } \frac{S_2}{D} = \frac{RS}{RD_2};$$

$$RD_1 = \frac{D \times RS}{S_1} \text{ and } RD_2 = \frac{D \times RS}{S_2};$$

$$T = RD_1 - RD_2 = D \times RS \left(\frac{1}{S_1} - \frac{1}{S_2}\right).$$

If we assume that a standard real size is chosen, so that RS is constant, then RT should increase with the difference of the reciprocals of the two sizes.

6. Bundesen et al. (1981) did not assume that the represented real size was constant. Instead, they made another reasonable assumption: that one of the distances, RD_2 was always the same. Because $RS \times D = RD_2 \times S_2$,

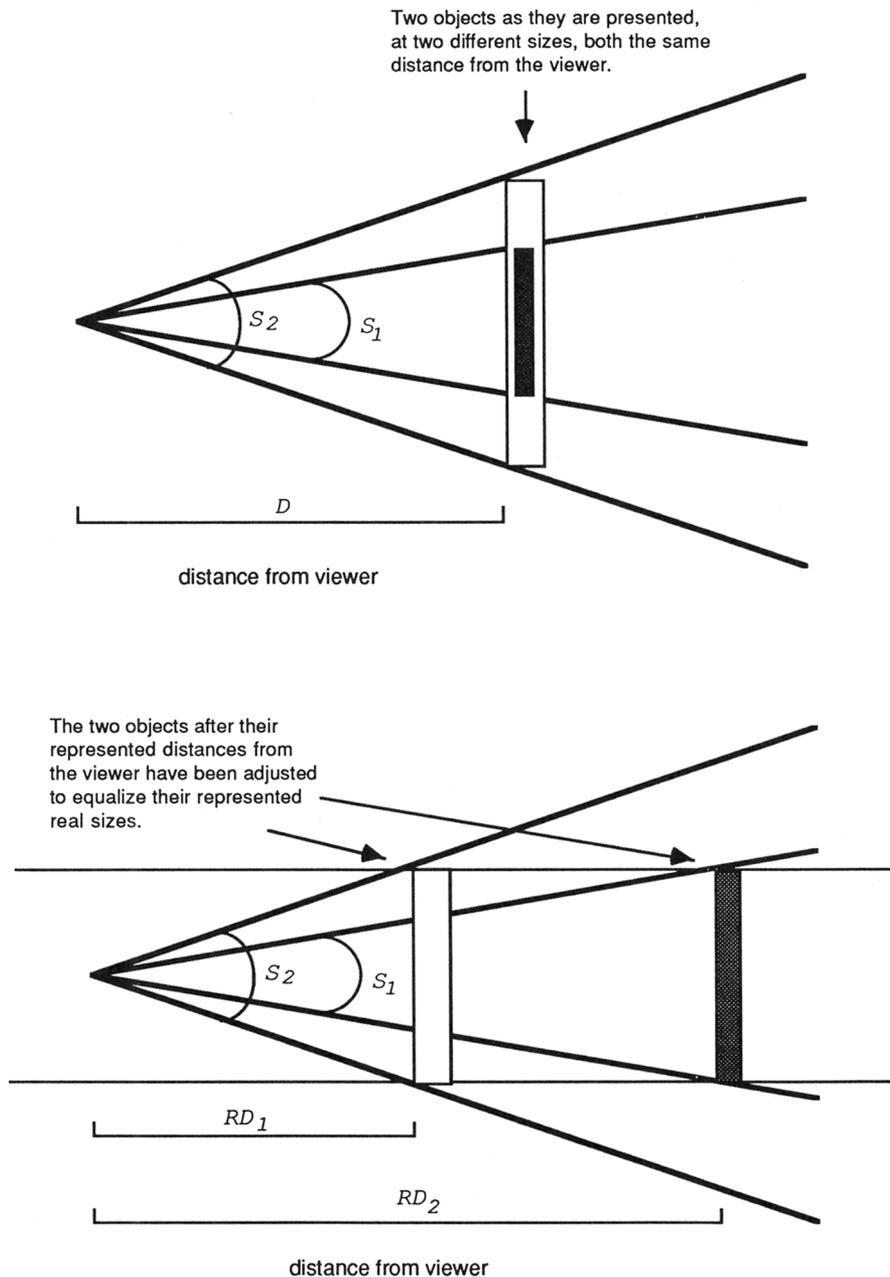$$T = RD_2 \times S_2 \left(\frac{1}{S_1} - \frac{1}{S_2}\right);$$

$$T = RD_2 \left(\frac{S_2}{S_1} - 1\right).$$

Under this assumption, RT is a linear function of size ratio.

7. Another possibility is that the two different-sized objects are seen as such, and one of them is imagined moving in depth until its apparent size matches the apparent size of the other. This problem reduces to the same problem as in the previous method, and once again RT increases linearly with the ratio between the two sizes.

Each of these functions was used to generate a set of contrast weights. These weights were then tested against the data from the expected- and unexpected-object conditions of Experiments 1 and 2 and from Experiment 3 (see Table A1).

As is evident in Table A1, none of the functions provides a notably superior fit to the data. To compare the fits from the different functions, we correlated the predictions of each function with each individual subject's means for the different size ratios. This produced a set of correlation coefficients for each function, one correlation (r)

Two objects as they are presented,
at two different sizes, both the same
distance from the viewer.

S_2 S_1

D

distance from viewer

The two objects after their
represented distances from
the viewer have been adjusted
to equalize their represented
real sizes.

S_2 S_1

RD_1

RD_2

distance from viewer

Figure A2. One can do size scaling by treating two visual objects as if they are different distances from the viewer (Bundesen, Larsen, & Farrell, 1981). (In the upper diagram, two objects of different sizes are represented at the same distance. In the lower diagram, the apparent sizes have been held constant, but the distances have been changed so that the real sizes of the two objects are equal. Once such a representation is constructed, the two objects are "brought together," and their apparent sizes automatically become equal. See text for further explanation. Adapted from Figure 16.5 of C. Bundesen, A. Larsen, & J. E. Farrell, 1981, "Mental Transformations of Size and Orientation," in J. Long & A. Baddeley [Eds.], *Attention and Performance IX*, pp. 297-294, copyright © 1981 by Lawrence Erlbaum Associates. Adapted by permission.)

for each subject. These *r*s were converted to *z*s with a Fisher transform, and every possible pairing of functions was compared with a matched-pairs *t* test. This entire operation was done separately for the expected- and unexpected-object conditions of Experiments 1 and 2 and for Experiment 3.

In none of the five sets of data tested was there ever a significant difference between the fits of the linear and log size-ratio functions. These two functions both provided a better fit than did the size difference in the blocked data from Experiment 3; in the other data sets, there were no significant differences. The remaining two func-

Table A1
Effect Sizes From Contrasts of Five Functions With the Data From Experiments 1, 2, and 3

Condition	Ratio	Log ratio	Difference	Difference of reciprocals	Difference of squares
Experiment 1					
Expected	.802	.811	.811	.747	.773
Unexpected	.382	.343	.415	.247	.447
Experiment 2					
Expected	.711	.725	.736	.653	.711
Unexpected	.366	.387	.322	.411	.263
Experiment 3 (blocked)	.840	.840	.806	.830	.744

tions, difference of reciprocals and difference of squares, both tended to do worse than the first three functions. Differences of reciprocals fared better than differences of squares in the blocked data; neither did significantly better than the other in the other experiments. Overall, there was no noticeable pattern in these results that distinguished the expected-object conditions from the unexpected-object conditions.

These comparisons illustrate the fact that there currently are no grounds for choosing one function as being the best account for our data. The results are inconclusive mainly because the values for expected and actual sizes used in these experiments were not chosen with the goal of maximizing the differences between these functions' predictions. No firm conclusions can be made without further experimentation.

We used the same combinations of expected and actual sizes as Larsen and Bundesen (1978) did in their Experiment 2 and so their data probably did an equally poor job of distinguishing between these two functions. Because we do not have individual subject data from their experiments, we cannot perform the same sort of comparison on their results. We can, however, correlate the means for each size

ratio from their scale transformation experiment (their Experiment 2) with the predictions of each of the five functions (see Table A2).

The surprising result is that the highest correlation occurs when one size is subtracted from the other and not from either of the two functions that Larsen and Bundesen (1978) mentioned. The 10 possible pairings of correlations with Larsen and Bundesen's data were compared. The only difference to approach significance was between the log of the size ratio and the difference of reciprocals, and the log of the size ratio fit better, $t(4) = 2.51$, $p < .07$, two-tailed. On the basis of the data available to us, it seems unwise to draw any conclusions about which functions best account for Larsen and Bundesen's scale transformation data.

The available data do not allow us to determine whether image transformations and scale transformations differ in the shape of the response time functions that they produce. Although a number of studies in image size scaling have been done under the assumption that response times are related to the ratio between the two sizes, experimenters would be wise to consider other functions, such as the difference between the two sizes, that could account for the data at least as well.

Table A2
Correlations of Predictions From Five Functions With Data From Experiments 1, 2, and 3 and From Larsen and Bundesen (1978)

Condition	Ratio	Log ratio	Difference	Difference of Reciprocals	Difference of Squares
Experiment 1					
Expected	.937	.990	.931	.847	.770
Unexpected	.838	.772	.648	.821	.545
Experiment 2					
Expected	.906	.953	.965	.790	.881
Unexpected	.603	.710	.442	.862	.229
Experiment 3 (blocked)	.967	.985	.885	.920	.730
Larsen & Bundesen	.925	.936	.951	.769	.867