Understanding the object benefit in visual short-term memory: The roles of feature proximity and connectedness

YAODA XU
Yale University, New Haven, Connecticut

Past research has identified visual objects as the units of information processing in visual short-term memory (VSTM) and has shown that two features from the same object can be remembered in VSTM as well (or almost as well) as one feature of that object and are much better remembered than the same two features from two spatially separated objects. It is not clear, however, what drives this object benefit in VSTM. Is it the shared spatial location (proximity), the connectedness among features of an object, or both? In six change detection experiments, both location/proximity and connectedness were found to be crucial in determining the magnitude of the object benefit in VSTM. Together, these results indicate that location/proximity and connectedness are essential elements in defining a coherent visual object representation in VSTM.

A typical visual scene we encounter usually contains a fair number of objects. Although we can extract the gist of a scene quite rapidly (Potter, 1976), as has been illustrated in numerous change detection studies, our detailed representations of the objects in a scene in visual short-term memory (VSTM) are quite poor (e.g., Rensink, O'Regan, & Clark, 1997; see Simons & Levin, 1997, for a review). With simple stimuli, such as letters and colored squares, studies have documented that we can retain a maximum of about four items at a time in our VSTM (e.g., Irwin, 1992; Luck & Vogel, 1997; Pashler, 1988; see also Vogel, Woodman, & Luck, 2001). With the aid of long-term memory, however, more items may be remembered (Hollingworth, 2004; see also Hollingworth & Henderson, 2002; Hollingworth, Williams, & Henderson, 2001). Recently, possible neural mechanisms behind VSTM have been documented (Todd & Marois, 2004; Vogel & Machizawa, 2004; Xu & Chun, 2006).

Despite the capacity limitation of VSTM, two features from the same object have been found to be remembered as well (or almost as well) as one feature of that object and are much better remembered than the same two features located on two spatially separated objects. These object effects were first discovered with the brief report paradigm, in which observers are asked to report visual features from briefly presented and then masked visual objects (Allport, 1971; Duncan, 1984, 1993; Duncan & Nimmo-Smith, 1996; Wing & Allport, 1972; see also Irwin & Andrews, 1996). More recently, these findings have been replicated with the change detection paradigm, in which observers are presented briefly with multiple visual objects and, after a short delay, detect a visual feature change to one of the objects (Luck & Vogel, 1997; Olson & Jiang, 2002; Wheeler & Treisman, 2002; Xu, 2002a, 2002b). This object benefit, or object-based feature integration, in VSTM has led to the notion that the units of information processing in VSTM are discrete objects, rather than individual features, and that an object's features may be integrated in VSTM with minimum cost (Luck & Vogel, 1997; Vogel et al., 2001). It is worth emphasizing that the object benefit, or object-based feature integration, in VSTM increases the total amount of information that can be retained in VSTM, so that with the same number of features distributed in a similar spatial envelope, more feature information can be retained in VSTM when the features are grouped into a few discrete objects than when they are distributed over many objects (Xu, 2002a).

It is not clear, however, what drives the object benefit in VSTM. Given that the presence of an object always corresponds to a spatial location and that features from an object are always located at or close to the same location, it could simply be that VSTM can retain features only from a limited number of distinct spatial locations and that all features present at the same location will be remembered together with minimum cost. This grouping of features by spatial location in VSTM corresponds to the Gestalt principle of grouping by proximity (Wertheimer, 1924/1950), which states that nearby items are more likely

to be grouped together than are distant items. We will, therefore, term this the *location/proximity account* of the object benefit in VSTM. On the other hand, features of an object share more than just a common spatial location; they are also connected to each other. The connections formed among the different features and the different parts of an object give rise to the sense that different features and different parts belong to the same object (Palmer & Rock, 1994), and the connectedness between object parts has been shown to facilitate shape perception (Saiki & Hummel, 1998). As such, connectedness between object parts may be the key in determining how object features are retained in VSTM. We will term this the *connectedness account* of the object benefit in VSTM. A third possibility is that both spatial location and connectedness play a role and, together, determine how object features are retained in VSTM. We will term this the *hybrid account* of the object benefit in VSTM.

Examining what determines the object benefit in VSTM not only is essential in understanding the nature of this effect, but also is important in understanding what defines a visual object in VSTM. Many visual processings have been shown to be object based, such as the shift of visual attention (e.g., Egly, Driver, & Rafal, 1994), the storage of visual objects in VSTM as described above, and the behavior of patients with visual neglect (e.g., Mattingley, Davis, & Driver, 1997; Tipper & Behrmann, 1996) and of patients with Balint's syndrome (e.g., Humphreys & Riddoch, 1993). Exactly what is a visual object, however, can be quite ambiguous, as Marr (1982) once commented:

> What . . . is an object, and what makes it so special that it should be recoverable as a region in an image? Is a nose an object? Is a head one? Is it still one if it is attached to a body? What about a man on horseback? These questions show that the difficulties in trying to formulate what should be recovered as a region from an image are so great as to amount almost to philosophical problems. There is really no answer to them—all these things can be an object if you want to think of them that way, or they can be part of a larger object. (p. 270)

Despite the seeming subjectiveness and ambiguity associated with defining an object, there are rules, such as the Gestalt principles (Wertheimer, 1924/1950), that have been shown to govern how we perceive and organize our perceptual world (for a review, see Palmer, 1999). Moreover, Scholl (2001) has argued that although we are free to consider almost anything as an object, given enough time and leisure, at earlier levels of visual analysis, the definition of objecthood can be well constrained: If a factor related to an object is shown to have a significant impact on a particular visual processing involving objects, this factor must be an essential part of the object representation for that visual processing. For example, Watson and Kramer (1999) showed that uniformly connected regions can facilitate attentional shift between different regions of an object. This, in turn, indicates that a uniformly connected region may define objecthood in the object-based attention shift task. So what are the essential elements for an object representation in VSTM? A shared spatial loca-

tion between parts? The connectedness between parts? Or both?

Woodman, Vecera, and Luck (2003) reported that with attention cuing, given the limited capacity of VSTM, perceptual grouping by proximity and connectedness could bias the selection of the items that would be subsequently stored in VSTM. However, it is not clear from that study whether these two grouping factors can modulate the object benefit in VSTM and are important in determining the total amount of information that can be stored in VSTM. Other studies in which perceptual grouping in VSTM has been manipulated have not addressed this question either (Jiang, Chun, & Olson, 2004; Walker & Davies, 2003).

Lee and Chun (2001) studied whether the limited VSTM capacity is determined by the total number of objects or the total number of spatial locations that objects occupy. By varying whether objects overlapped or not and, thus, varying the total number of object locations in a change detection study, Lee and Chun found no effect of object location on the total amount of information that could be held in VSTM. Because overlapping objects were both connected and closer to each other than were nonoverlapping objects, the results in Lee and Chun suggested that neither location nor connectedness plays a role in feature integration in VSTM. However, as Xu (2002a) has pointed out, the displays containing overlapping objects were visually more complex than those containing nonoverlapping objects, and as a result, it would be harder to extract features from overlapping objects. Therefore, in Lee and Chun, the difficulties associated with perceiving overlapping objects might have masked any benefit resulting from grouping by location/proximity and connectedness in VSTM.

Treating two overlapping objects as two parts of one object by perceptual grouping of proximity and connectedness, Xu (2002a, Experiment 2) reexamined the object benefit in VSTM by presenting objects consisting of an oriented bar passing through a color circle (the conjunction display). In the control display (the disjunction display), the oriented bar and the color circle of each object were detached, nonoverlapping from each other, and presented in different locations of the display, thereby forming two separate objects, instead of two parts of one object, as in the overlapping case. Following Luck and Vogel (1997), observers were asked to hold in VSTM just the colors of the circles, just the orientations of the bars, or both types of features from each type of display. When only one type of feature was remembered, performance was significantly worse in the conjunction than in the disjunction displays, showing greater difficulties in extracting and subsequently remembering the feature from an object part when it overlapped with another object part. When both types of features were monitored, performance was identical for both types of displays and replicated the findings of Lee and Chun (2001). However, there was a significant interaction between display type and feature-monitoring condition, so that the performance drop for monitoring one versus both types of features was bigger in the disjunction than in the conjunction displays. These results indicated that when higher visual complexity as-

sociated with the conjunction display was accounted for, there was indeed an object benefit in VSTM, so that more feature information could be retained in VSTM when features were grouped into smaller numbers of objects by location and connectedness.

In subsequent experiments, instead of using objects containing overlapping parts, Xu (2002a) used objects whose parts were defined by negative minima of curvature (Hoffman & Richards, 1984; Hoffman & Singh, 1997; Xu & Singh, 2002) and obtained similar results. Figure 1 provides an example of the displays used in Xu (2002a). Observers were asked to remember a color, an orientation feature, or both features of a given display and to detect a change after a brief delay. The two features (color and the bent of the stems) could be contained in the same part of an object (Figure 1A), in different parts of an object (Figure 1B), or in spatially separated objects (Figure 1C). Features were found to be best remembered when they were from the same part of an object (see also Vecera, Behrmann, & Filapek, 2001; Vecera, Behrmann, & Mc-Goldrick, 2000); although less well remembered in comparison, features from different parts of the same object were still better remembered than features from spatially separated objects.

Xu (2002a) showed that the object benefit in VSTM could be extended to retaining features from different object parts. This finding by itself, however, could be accounted for by the location/proximity account, the connectedness account, or the hybrid account of the object benefit in VSTM. The location/proximity account would argue that features from different parts of an object are located closer to each other than are those from separated objects and that spatial location is still the key in determining the object benefit in VSTM. The connectedness account would argue that features from different parts of an object are connected to each other, whereas those from different objects are not. As such, the connectedness between object parts is what drives the object benefit in VSTM. The hybrid account would simply incorporate both of these arguments and insist that both spatial location and connectedness are crucial in determining the object benefit in VSTM. It is, therefore, still unresolved which one of these three accounts captures the nature of the object benefit in VSTM.

Nonetheless, because Xu (2002a) showed that the object benefit in VSTM could be extended to features from different object parts, it allows us to carry out the following manipulations: By keeping the distance between object parts constant while manipulating how object parts are connected to each other, we can examine how grouping by connectedness independently of location affects the object benefit in VSTM; conversely, by keeping part connection constant while varying the distance between object parts, we can examine how grouping by location/proximity independently of connectedness affects the object benefit in VSTM. These two types of manipulations were, therefore, carried out in the present study to examine the roles of spatial location and connectedness in VSTM feature integration.[1]

As was described earlier with the overlapping objects used in Lee and Chun (2001) and Xu (2002a), if not accounted for, perceptual differences between displays could significantly affect VSTM performance and obscure the object-based benefit in VSTM. Therefore, instead of directly comparing performance between the different displays when two object features were integrated and retained in VSTM, as in previous studies (Xu, 2002a, 2002b), in this study the performance drop for remembering one versus two object features was compared between the different displays (i.e., the interaction between the monitoring condition and the display condition). This particular paradigm was used on the basis of Xu (2002a). In that study, in one case, features were integrated in a similar way in two displays but differed in how well they could be perceived. Although the overall performance differed between the two displays, the amount of perfor-
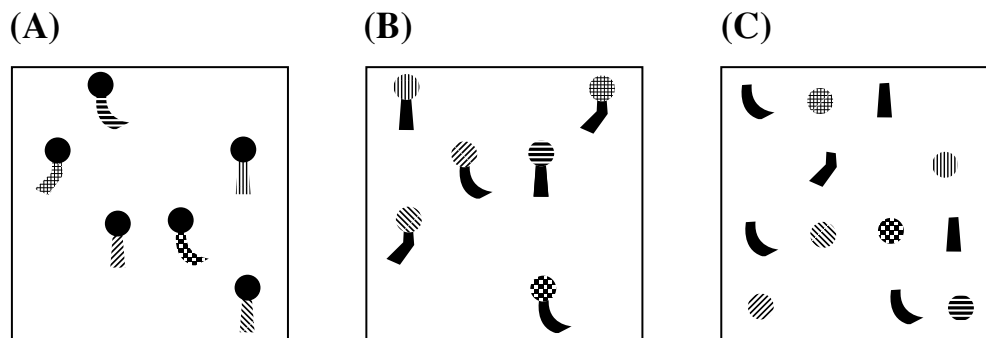

(A)   **(B)**   **(C)**

Figure 1. Stimuli adapted from the study by Xu (2002a). The texture patterns represent the different colors used in the experiment. (A) The relevant color and orientation features were contained in the same part of the object—the stem part. (B) The relevant color feature was on the circle part, the relevant orientation feature was on the stem part, and the two parts were connected. (C) The relevant features were in different parts, as in panel B, but the circle and stem parts were not connected to each other. Features were found to be best retained in visual short-term memory when they were from the same part of an object, as in panel A; they were less well retained when they were from different parts of an object, as in panel B, and were least well retained when they were from spatially separated objects, as in panel C.

mance drop for monitoring one versus two features was identical between the two displays, reflecting that the features in those two displays were equally well integrated in VSTM. In a second case, features from two displays were well matched in their perceptibility but differed in how they were integrated in VSTM. Although performance was the same for the two displays when only one features was monitored, the amount of performance drop for monitoring one versus two features differed between the two displays, again reflecting how well the features in these two displays were integrated in VSTM. Therefore, by comparing performance differences for remembering one versus two features, we could assess how well two features were integrated in VSTM independently of the perceptibility of the individual features.

EXPERIMENT 1A

In Xu (2002a), it was found that features from different parts of an object were better remembered than the same features from spatially separated objects. This effect was reexamined in the present experiment by varying how object parts were connected to each other. As in Xu (2002a),

observers were asked to remember the colors of the circles and the orientations of the bent of the stems in a display. The experiment included four types of displays, as shown in Figures 2A–2D. In Figure 2A, a circle is directly connected to a stem. In Figure 2B, a circle is connected to a stem via a third object part—a bar. In Figure 2C, there is a gap between a circle and a stem. In all three of these displays, the distance between the circle and the bent of the stem is the same. As a control condition, in Figure 2D, the color circles and the stems are completely detached from each other and are presented in different locations of the display.

According to Palmer and Rock (1994), connected parts form a stronger perceptual group than do unconnected parts. Both the connectedness and the hybrid accounts would, therefore, predict that connected parts will generate a bigger object benefit in VSTM than will unconnected parts. The location/proximity account predicts, on the contrary, that as long as the distance between parts are kept constant, the magnitude of the object benefit in VSTM will not be affected by how the parts are connected.

As was stated earlier, to account for perceptual differences associated with the different display types (e.g., the
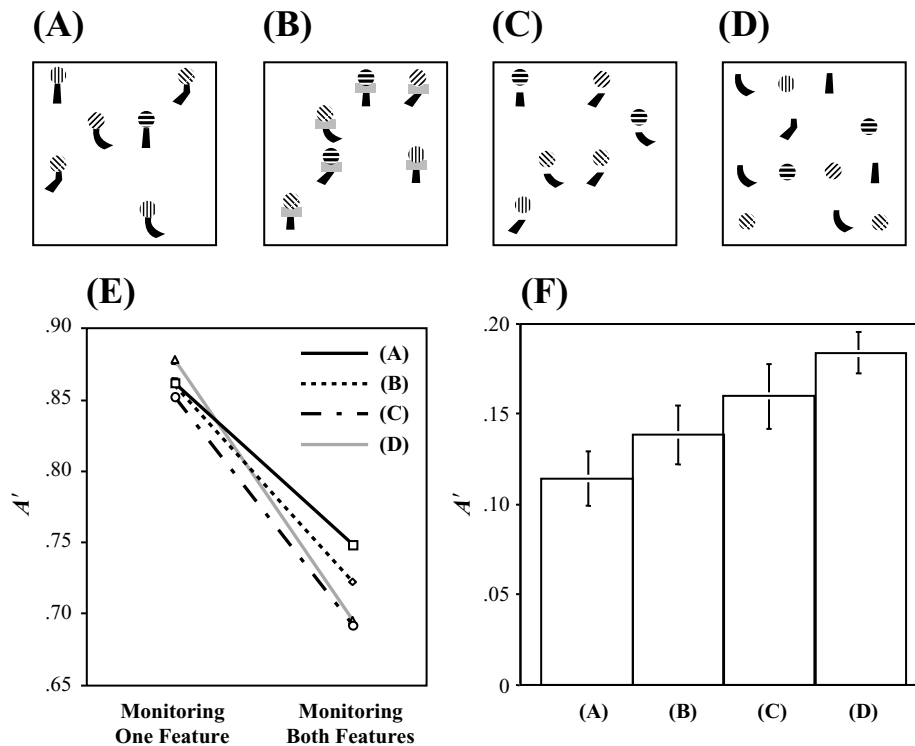

Figure 2. Panels A, B, C, and D show examples of the displays used in Experiment 1A. The texture patterns represent the different colors used in the experiment. (A) The circle and the stem parts were directly connected. (B) The two parts were connected by a third part—a horizontal bar. (C) The two parts were separated by a gap. (D) The two parts were unconnected. (E) Change detection accuracy as measured by A'. Error bars were not included here to better illustrate the differences among the conditions. (F) The results shown in panel E replotted as the differences in A' for monitoring one versus two features. It is evident that as the connection strength between parts decreases from panel A to C, the magnitude of the object-based benefit in visual short-term memory decreases accordingly. Error bars indicate standard errors.

stems are longer for the display shown in Figure 2A than for those shown in Figures 2B and 2C, which could make orientation perception easier for Figure 2A than for Figures 2B and 2C), observers were asked to remember, in different trial blocks, either just one of the two features (circle color or stem orientation) or both features. Performance differences for monitoring one versus two features among the different types of displays were then compared to examine how part connectedness affected the object benefit in VSTM.

Method

Participants. Twenty-four volunteers (14 of them female) were recruited from the Harvard University campus. They were between 17 and 40 years of age; all had normal color vision and received either payment or course credits for their participation.

Materials and Design. There were four display types in the experiment: (1) the circle and the stem parts were directly connected (Condition A; Figure 2A), (2) the two parts were connected by a third part—a horizontal gray bar (Condition B; Figure 2B), (3) the two parts were separated by a gap (Condition C; Figure 2C), and (4) the two parts were unconnected (Condition D; Figure 2D). The diameter of the color circle was 1.0º, and the maximum extent of the black shape was 1.3º × 1.6º. The size of the dark gray part in Figure 2B was 1.3º × 0.6º. The width of the gap in Figure 2C was 0.6º. The whole display extended 10.3º × 12.0º, with the objects in a given display separated by at least 2.6º (center to center). The circles were presented in one of four colors: pink, cyan, white, or yellow. The black shapes were presented in one of three orientations: right, center, or left, with the right shape slightly different from the left shape, so that an orientation change could be more easily detected. All the items were presented on a light gray background.

In different trial blocks, the observers were asked to (1) monitor and detect a color change in the circles in 50% of the trials, (2) monitor and detect an orientation change in the stems in 50% of the trials, or (3) monitor and detect either a color change in the circles or an orientation change in the stems in 50% of the trials. The participants were informed of the monitoring condition at the beginning of each block.

The computer randomly assigned the color and orientation to each object in such a way that a particular color would not appear more than twice and a particular orientation would not appear more than three times in a given display. When a change occurred, the changed color or orientation value was randomly selected. When a single feature was monitored, there were 32 change trials and 32 no-change trials, evenly distributed into two 32-trial blocks; when both features were monitored, there were 32 color change trials, 32 orientation change trials, and 64 no-change trials, evenly distributed into four 32-trial blocks. Blocks of trials from the same display type and the same monitoring condition were grouped together during testing. Ten practice trials proceeded the experimental trials at the beginning of each group of blocks.

To prevent the participants from getting exhausted at the end of a long testing session, the whole experiment consisted of two short testing sessions (about 35 min each), carried out on 2 different days for each participant. Each testing session contained all the trials for two of the four display types. The orders of presentation were balanced among the participants.

Apparatus. The MacProbe Macintosh programming software and an Apple computer with a 400-MHz Power PC G3 processor and a 17-in. monitor were used to generate the displays.

Procedure. The participants were seated in a dimly lit and quiet room, about 50 cm from the screen. They initiated each trial by pressing the space key on the computer keyboard. Each trial began with the presentation of a fixation dot at the center of the screen for 500 msec, followed by the presentation of the sample display for 500 msec. The sample display was then replaced by a blank gray background. After 1,000 msec, the test display appeared. The test display remained on the screen until the participants made a keypress. The participants were asked to press the left control key with their left index finger if they detected a change and to press the enter key on the number keypad with their right index finger if they did not detect any change. The keys were labeled with the words "Different" and "Same," respectively. As soon as the participants responded, feedback was given as either a happy face for a correct response or a sad face for an incorrect response. The feedback stayed on the screen for 250 msec. The next trial began about 1 sec after the feedback for the previous trial had disappeared. Within a block, trials came one after another. There were breaks between the blocks, during which the participants could rest for as long as they wanted.

Results

Following Xu (2002a, 2002b), the measure of A' from signal detection theory was used to measure the accuracy of the change detection performance (Grier, 1971; Pollack & Norman, 1964; see also Macmillan & Creelman, 1991).[2] Because independent guessing rates for color and orientation change detections were not available when both features were monitored, performance was averaged over color and orientation change detection before A' scores were calculated. The means of A' averaged over color and orientation change detection are plotted in Figure 2E. Overall, the differences among the different display types were marginally significant [$F(3,69) = 2.19$, $p = .097$], performance was much higher for monitoring one than for monitoring two features [$F(1,23) = 506.30, p < .001$], and there was a significant interaction between display type and monitoring condition [$F(3,69) = 3.46, p = .021$].

Change detection performance was then examined in detail for monitoring one versus two features when the connections between the circle and the stem parts varied from Condition A to Condition D. In Figure 2F, the results are replotted, showing the difference (i.e., the cost) in A's for monitoring two features versus one across the different display types. As is evident in Figure 2F, the cost for monitoring two features versus one increased, and thus, the object benefit decreased as parts became less connected from Condition A to Condition D. Indeed, the cost was higher in Condition D than in either Condition A or Condition B [$F(1,23) = 14.02$, $p = .001$, and $F(1,23) = 8.64$, $p = .007$, respectively], thus replicating the VSTM object benefit reported in Xu (2002a), but the cost did not differ between Conditions D and C [$F(1,23) = 1.01$, $p = .325$]. In addition, the difference between Conditions A and C was marginally significant [$F(1,23) = 2.97$, $p = .098$]. The differences in cost among Conditions A, B, and C, although consistent with the costs increasing from Condition A to Condition C, was not significant [$F(2,46) = 1.64, p = .206$].

Although color and orientation change detection performance could not be assessed independently when both features were monitored (because separate guessing rates could not be obtained for each), how feature perceptibility differed among the four display conditions could be examined by comparing performance when only one feature was monitored. Change detection accuracy in A' for

monitoring just orientation change was .80, .82, .82, and .85 for Conditions A–D, respectively. Although the overall difference was not significant [$F(3,69) = 1.123, p = .35$], having the stems shorter and detached from the circles in Conditions B and C actually made orientation change detection a little easier and not harder, as one would imagine. Orientation change detection was the easiest in Condition D, possibly because the orientation feature needed to be segmented from an object in Condition A, to a lesser degree in Conditions B and C, but not in Condition D. Performance for just color change detection was .90, .89, .88, and .91 for Conditions A–D, respectively. The differences among the conditions were not significant [$F(3,69) = 1.62, p = .19$]. When only one feature was monitored, the overall change detection accuracy was higher for color than for orientation [$F(1,23) = 33.00, p < .001$].

The results of this experiment will be discussed together with those of Experiment 1B.

EXPERIMENT 1B

In Experiment 1A, although the distance between object parts was kept consistent, the object benefit in VSTM decreased as the parts became less connected from Condition A to Condition C. This was illustrated as (1) a significant cost difference between Conditions A and D and between Conditions B and D, but not between Conditions C and D, and (2) a marginally significant cost difference between Conditions A and C. However, a direct comparison of the cost difference among Conditions A, B, and C did not yield a significant difference. It is possible that because two test sessions were carried out for each observer on 2 different days (possibly different times of the day) in Experiment 1A, a given observer might have been more tired and less focused in one of the test sessions than in the other. This added variance in the observers' performance might have increased the noise level in the data and made the results less significant. To confirm the finding that the connection strength between object parts play an important role in the object benefit in VSTM, Conditions A and C from Experiment 1A were retested in this experiment.

Method

Participants. Eighteen volunteers (12 of them female) were recruited from the same participant pool. Two of the 18 participants had also participated in Experiment 1A (separated by a few months).

Materials, Design, and Procedure. Conditions A and C from Experiment 1A were included here (see Figures 3A and 3B). Other aspects of the materials, design, and procedure were identical to those in Experiment 1A. The whole experiment lasted about 35 min.

Results

As in Experiment 1A, the means of A' averaged over color and orientation change detection were calculated; the results are plotted in Figure 3C. In Figure 3D, the results are replotted as the differences in A' for monitoring one versus two features. The overall difference between the two displays was not significant [$F(1,17) = 2.48, p = .13$], with performance higher for monitoring one feature than for monitoring two features [$F(1,17) = 42.55, p <$

.001]. Most important, there was a significant interaction between monitoring condition and display condition [$F(1,17) = 5.80, p = .028$], so that the drop in performance for monitoring two versus one feature was smaller for the connected parts than for the unconnected parts. When only one feature was monitored, performance did not differ between the two displays ($F < 1$); however, when both features were monitored, performance was higher for the connected parts than for the unconnected parts [$F(1,17) = 6.46, p = .021$]. These results indicate that the connection strength between object parts can modulate the magnitude of the object benefit in VSTM.

Discussion of Experiments 1A and 1B

Together, the results of Experiments 1A and 1B showed that as the connection strength weakened between two part features, the magnitude of the object benefit, or object-based feature integration, in VSTM and the total amount of feature information that could be retained in VSTM decreased accordingly.[3] Because the distance between object parts was constant, these findings could not be explained by the location/proximity account, which argues that spatial location is the only factor driving the object benefit in VSTM. On the other hand, these findings are consistent with both the connectedness and the hybrid accounts, which propose that the connection strength between object parts plays an important role in determining the object benefit in VSTM.

EXPERIMENT 2

Although Experiments 1A and 1B showed that connected part features were better integrated and resulted in a bigger object benefit in VSTM than did unconnected part features, one could argue that this effect was caused solely by the differences in stem length between the connected and the unconnected parts. Although the observers needed only to detect an orientation change at the bent part of the stem, the length of the straight part of the stem might, nonetheless, have affected how well the orientation of the bent part could be perceived. In other words, when the parts were directly connected, the stem was longer, which might have facilitated orientation perception in this condition, in comparison with when the stem was shorter in the unconnected condition. Although there was no difference in orientation change detection among the different displays when only orientation was monitored, one could argue that the task was relatively easy under this condition and that the differences in stem orientation perception might have affected performance only when the task became more demanding, such as when both features were monitored.

To address this concern, in this experiment, stem orientation change detection was carried out with only stems presented in the display. The stems were either long, as in the connected parts, or short, as in the unconnected parts in Experiment 1B. To make the task demanding, a given display contained 10 stems. If the length of the stem affected how well orientation features could be perceived in Experi-
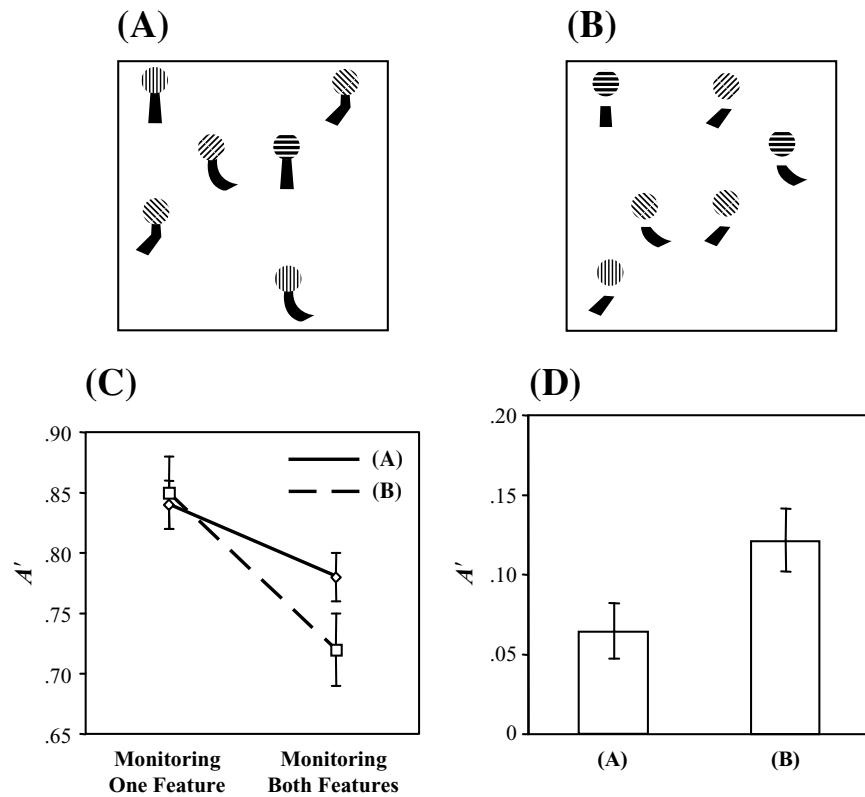
(A)

(B)

(C)

(D)

Figure 3. Panels A and B show examples of the displays used in Experiment 1B, which corresponded to Conditions A and C in Experiment 1A (see Figure 2). The texture patterns represent the different colors used in the experiment. (C) Results of Experiment 1B. (D) The results shown in panel C replotted as the differences in A' for monitoring one versus two features. The results showed that as the connection strength between the parts weakened, the object benefit in visual short-term memory (VSTM) decreased, confirming the findings from Experiment 1A that connectedness plays an important role in determining the object benefit in VSTM. Error bars indicate standard errors.

ments 1A and 1B, orientation change detection should be easier in displays containing the longer stems than in those containing the shorter stems. On the other hand, if stem length had no effect on orientation change detection, no differences in performance should be observed between displays containing the longer and the shorter stems.

Method

Participants. Six volunteers (2 of them female), fulfilling the same criteria as those in Experiments 1A and 1B, were recruited from the Yale University campus.

Materials, Design, and Procedure. The stem parts used in the two conditions in Experiment 1B were used here. The display contained 10 long stems for the *long* display condition (Figure 4A) and 10 short stems for the *short* display condition (Figure 4B). For each display condition, there were 32 change trials and 32 no-change trials, evenly distributed into two 32-trial blocks. Other aspects of the materials, design, and procedure were identical to those in Experiments 1A and 1B. The whole experiment lasted about 15 min.

Results and Discussion

As in Experiments 1A and 1B, A' scores were calculated, and the results were plotted in Figure 4C. The overall change detection accuracy was comparable to that in

Experiments 1A and 1B when both color and orientation features were monitored, indicating that remembering 10 stem orientations in VSTM was, indeed, a demanding task. However, there was no difference in orientation change detection between when the stems were long and when they were short ($F < 1$).

This finding suggests that the results of Experiments 1A and 1B could not be attributed to a difference in stem orientation perception between the connected and the unconnected part features. Rather, the results of Experiments 1A and 1B indicate that feature integration and the object benefit in VSTM is greater when part features are directly connected than when they are not.

EXPERIMENT 3

Another way of addressing the concerns raised for Experiments 1A and 1B regarding differences in stem length would be to replicate the results with a different set of stimuli in which the perception of the shape features is completely independent of the stem length. To do so, Experiment 1B was repeated with the stimuli shown in
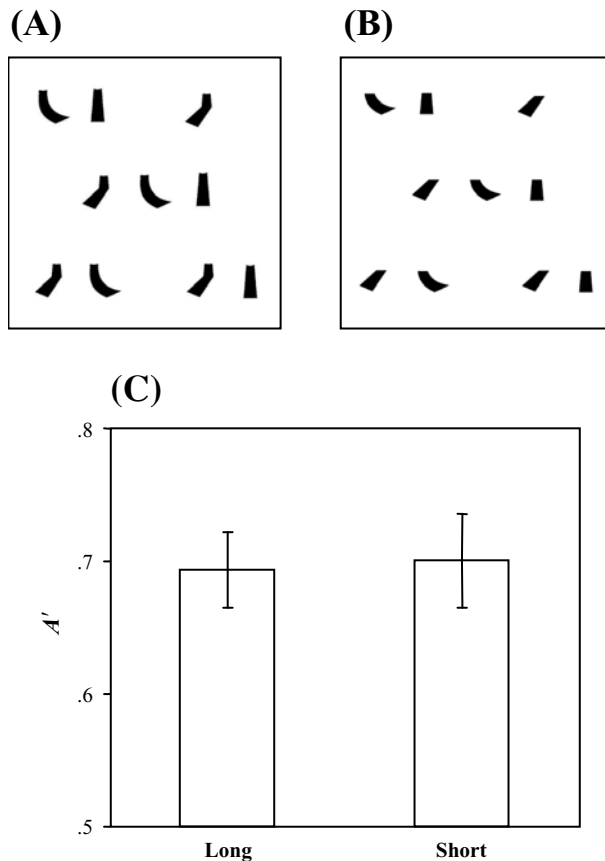
(A)

(B)

(C)

Figure 4. Panels A and B show examples of the displays used in Experiment 2. (A) The display contained the 10 long stems used in Experiments 1A and 1B. (B) The display contained the 10 short stems used in Experiments 1A and 1B. (C) Results of Experiment 2. There was no difference in stem orientation change detection between displays containing the long and the short stems. Error bars indicate standard errors.

Figures 5A and 5B. Here, each object had a solid round bottom, a curved inverted-Y bottom, or a sharp inverted-Y bottom, with the length of the stem having no effect on the ease in which these bottom shapes could be perceived. As in Experiments 1A and 1B, the observers were asked to remember the color features, the shape features, or both features of each display. A replication of Experiments 1A and 1B with these stimuli would provide stronger evidence that connected part features are better integrated and remembered in VSTM than are unconnected part features.

Method

Participants. Twelve volunteers (5 of them female), fulfilling the same criteria as those in the previous experiments, were recruited from the Yale University campus.

Materials, Design, and Procedure. The displays used are shown in Figures 5A and 5B. In different trial blocks, the observers monitored top part color features, bottom part shape features, or both types of features for a possible feature change. The two types of features could be either directed connected, as in Figure 5A, or separated by a gap, as in Figure 5B. The same colors as those in

Experiments 1A and 1B and three new bottom shape features were used here. The new bottoms had a solid shape, a curved inverted-Y shape, or a sharp inverted-Y shape. The maximum extent of the shape features was $1.94º \times 1.39º$. On the basis of the results from a pilot study, each display contained five, instead of six, objects, as in Experiments 1A and 1B, in order to avoid a floor effect. Other aspects of the materials, design, and procedure were identical to those in Experiment 1B. The whole experiment lasted about 35 min.

Results

As in Experiments 1A and 1B, the means of A', averaged over color and shape change detection, were calculated; the results are plotted in Figure 5C. In Figure 5D, the results are replotted as the differences in A' for monitoring one versus two features. The overall difference between the two display conditions was not significant $[F(1,11) = 2.56, p = .138]$, with performance higher for monitoring one feature than for monitoring two features $[F(1,11) = 197.78, p < .001]$. When only one feature was monitored, performance was significantly lower when the parts were connected than when they were unconnected $[F(1,11) = 14.12, p = .003]$; when both features were monitored, performance did not differ between the connected and the unconnected parts $(F < 1)$. Most relevant to the purpose of the present study, there was a significant interaction between monitoring condition and display condition $[F(1,11) = 6.66, p = .026]$, so that the drop in performance for monitoring two versus one feature was smaller for the connected parts than for the unconnected parts. These results indicated that features from connected parts are better remembered in VSTM than are those from unconnected parts.

Discussion

In this experiment, when either color or shape feature was retained in VSTM, performance was lower for the connected than for the unconnected parts, indicating that it was more difficult to extract features from the connected parts than from the unconnected parts (possibly due to a stronger shape-masking effect between the top and the bottom part shapes when the parts were directly connected). When both types of features were monitored, there were no longer differences between the two displays, indicating that a positive object benefit effect canceled out the negative perceptual difficulty effect associated with the connected parts (see also Xu, 2002a, Experiments 2 and 3). By measuring the drop in performance for monitoring two versus one part feature, rather than comparing performance directly for monitoring two part features, however, we still observed a significantly smaller drop in performance for the connected parts than for the unconnected parts.[4]

Therefore, with a different set of stimuli in which the perception of the shape features was completely independent of the stem length, the present results replicated those of Experiments 1A and 1B and showed that connected part features are better integrated and remembered in VSTM than are unconnected part features. Together with the results from Experiments 1A and 1B, these results showed that the formation of perceptual groupings by connected-
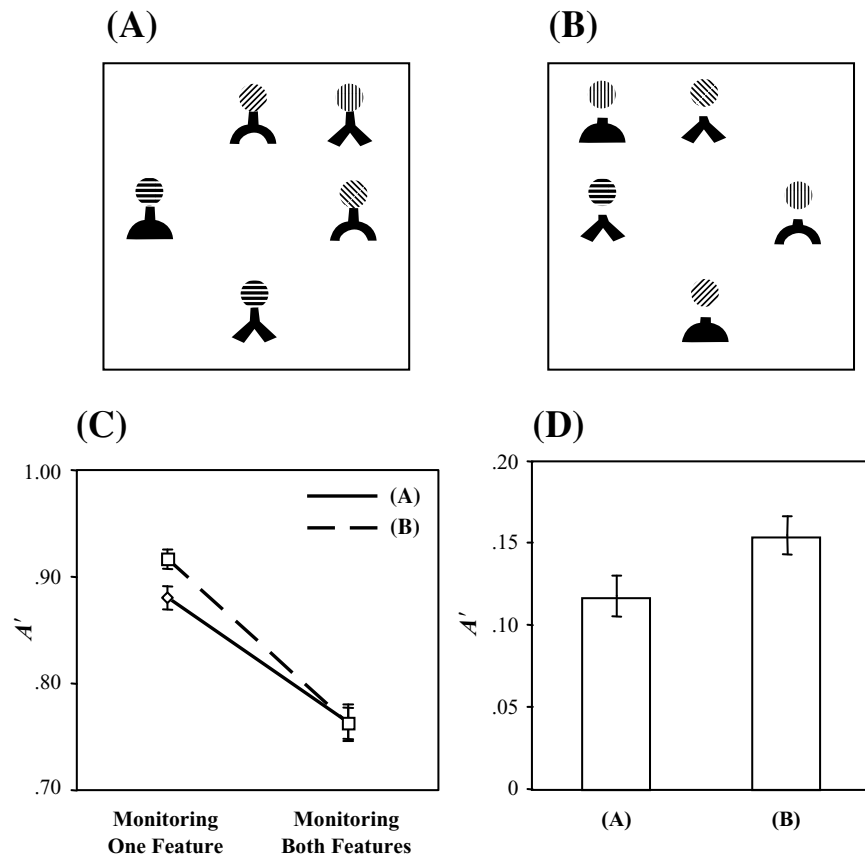
(A)

(B)

(C)

(D)

Figure 5. Panels A and B show examples of the displays used in Experiment 3, which corresponded to the two conditions used in Experiment 1B (see Figures 3A and 3B). Here, the perception of the bottom shape feature was independent of the stem length. The texture patterns represent the different colors used in the experiment. (C) Results of Experiment 3. (D) The results shown in panel C replotted as the differences in A' for monitoring one versus two features. The results showed that as the connection strength between parts weakened, the object benefit in visual short-term memory (VSTM) decreased accordingly, thus replicating the findings of Experiment 1B and showing that connectedness plays an important role in determining the object benefit in VSTM. Error bars indicate standard errors.

ness independently of location/proximity can modulate the object-based feature integration in VSTM and determine the total amount of information that can be stored into VSTM.

Although a pure location/proximity account is insufficient in accommodating the results of Experiments 1A, 1B, and 3, does location play a role at all in driving the object benefit in VSTM? A pure connectedness account would argue that as long as the grouping between the parts remains the same, varying spatial distance between the parts will not change the magnitude of the object benefit in VSTM. On the other hand, the hybrid account, which incorporates both the connectedness and the location/proximity accounts, argues that although connectedness plays an important role in the object benefit in VSTM, the distance between parts also influences how object features are stored in VSTM.

To distinguish between these two types of accounts, in Experiments 4A and 4B, the connection strength between object parts was kept constant, and the distance between parts was varied to examine how spatial distance influences the object benefit in VSTM.

EXPERIMENT 4A

To investigate the role of spatial location in VSTM information processing, this experiment used the same type of object as that in Experiments 1A and 1B. Here, the distance between the color circle and the bent of the stem was varied, and it was either short (Figure 6A) or long (Figure 6B). In both cases, the two parts were directly connected. In a third condition, the control condition, the two parts were detached from each other and were presented in different locations of the display (Figure 6C). As in Experiments 1A and 1B, the relevant features were the colors of the circles and the orientations of the bent of the stems. The participants monitored for color, orientation, or both color and orientation features and performed change detection on the relevant feature(s).
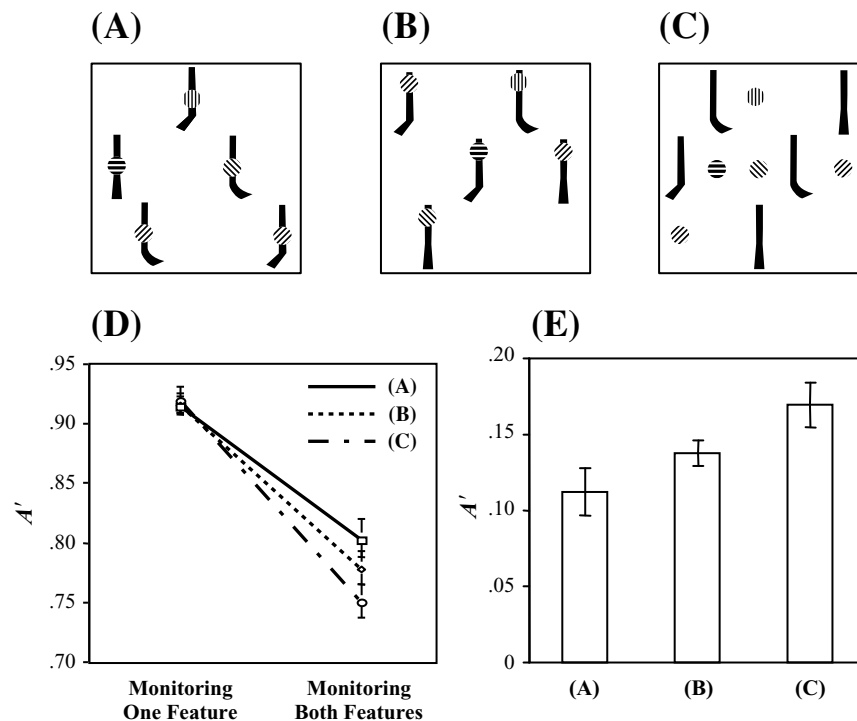
Figure 6. Panels A, B, and C show examples of the displays used in Experiment 4A. The texture patterns represent the different colors used in the experiment. (A) The circle and the stem parts were directly connected, and the distance between these two parts was short. (B) The circle and the stem parts were directly connected, as in panel A, but the distance between these two parts was long. (C) The two parts were unconnected. (D) Change detection accuracy as measured by A'. (E) The results shown in panel D replotted as the differences in A' for monitoring one versus two features. As the distance between object parts increases from panel A to B, the object benefit in visual short-term memory is attenuated. Error bars indicate standard errors.

A pure connectedness account would insist that location/proximity has no effect on the object benefit in VSTM and would, therefore, predict comparable performance for the displays shown in Figures 6A and 6B. The hybrid account, on the other hand, would argue that increasing the distance between object parts will decrease the object benefit and decrease the total amount of information that can be held in VSTM, and it would, therefore, predict worse performance for the display shown in Figure 6B than for that shown in Figure 6A.

Method

Participants. Eighteen participants were recruited from the same pool as that in Experiments 1A and 1B. Nine of the 18 participants had also participated in Experiment 1A on a different testing day.

Materials and Design. There were three display types in the experiment: (1) The circle and the black bar were attached, and the distance between the circle and the bent part of the bar was short (Condition A; Figure 6A)—the same distance as that used in Experiment 1; (2) the circle and the black bar were attached, and the distance between the circle and the bent part of the bar was long (Condition B; Figure 6B); and (3) the two parts were not attached (the control condition; Figure 6C). The extent of the black bar was $1.4° \times 3.8°$. The distance from the center of the circle to the end of the bend in Figure 6A was $1.9°$, and that in Figure 6B was $2.9°$. The extent of the whole display was $13° \times 14°$. On the basis of the results

from a pilot study, each display contained five, instead of six, objects, as in Experiments 1A and 1B, in order to avoid a floor effect.

The experiment consisted of one 50-min testing session. Other aspects of the experiment were identical to those in Experiments 1–3.

Results

As in Experiments 1–3, performance was averaged over color and orientation change detections before A' was computed for each observer in each condition. The final means of A' are plotted in Figure 6D. In Figure 6E, the results are replotted as the differences in A' for monitoring one versus two features. Overall, the differences among the different display types were significant [$F(2,34) = 3.27, p = .05$]; performance was much higher for monitoring one feature than for monitoring two features [$F(1,17) = 172.17, p < .001$], and there was a significant interaction between display type and monitoring condition [$F(2,34) = 8.75, p = .001$].

In pairwise comparisons, the drop in performance for monitoring one versus two features was much bigger in the control condition than in Condition A [$F(1,17) = 24.24, p < .001$], replicating the object benefit reported in Experiment 1A and in previous studies (Xu, 2002a, 2002b). As the distance between the two parts increased,

there was still a significant amount of object benefit present in Condition B, in comparison with the control condition [$F(1,17) = 4.67$, $p = .05$]. Although short of being significant [$F(1,17) = 3.21$, $p = .091$], the amount of object benefit present in Condition B, in comparison with Condition A, was consistent with the idea that spatial distance between parts may modulate the magnitude of the object benefit in VSTM.

The results of this experiment will be discussed together with those of Experiment 1B.

EXPERIMENT 4B

In Experiment 4A, the difference in object benefit in VSTM between Conditions A and B was short of being significant. It was possible that the distance between the circle and the bent part of the stem was not great enough to show a significant effect of spatial distance. In this experiment, the effect of location/proximity was reevaluated by repeating Conditions A and B from Experiment 4A, but with increased spatial distance between the parts.

Method

Participants. Six volunteers (5 of them female) were recruited from the same participant pool. One of the 6 participants had also participated in Experiment 4A on a different testing day.

Materials, Design, and Procedure. Conditions A and B from Experiment 4A were included here, with the spatial distance between the circle and the stem parts modified, so that as compared with Figures 6A and 6B for Experiment 4A, the distance between the parts in Figure 7A were decreased, whereas those in Figure 7B were increased. Here, the extent of the black bar was $1.4° \times 4.3°$; the distance from the center of the circle to the end of the bend in Figure 7A was $1.6°$, and that in Figure 7B was $3.3°$; the extent of the whole display was now $13° \times 19°$. Other aspects of the materials, design, and procedure were identical to those in Experiment 4A. The whole experiment lasted about 35 min.

Results

As in the previous experiments, the means of A' averaged over color and orientation change detection were calculated. The results are plotted in Figure 7C. In Figure 7D, the results are replotted as the differences in A' for monitoring one versus two features. The overall difference between Conditions A and B was not significant [$F(1,5) =$
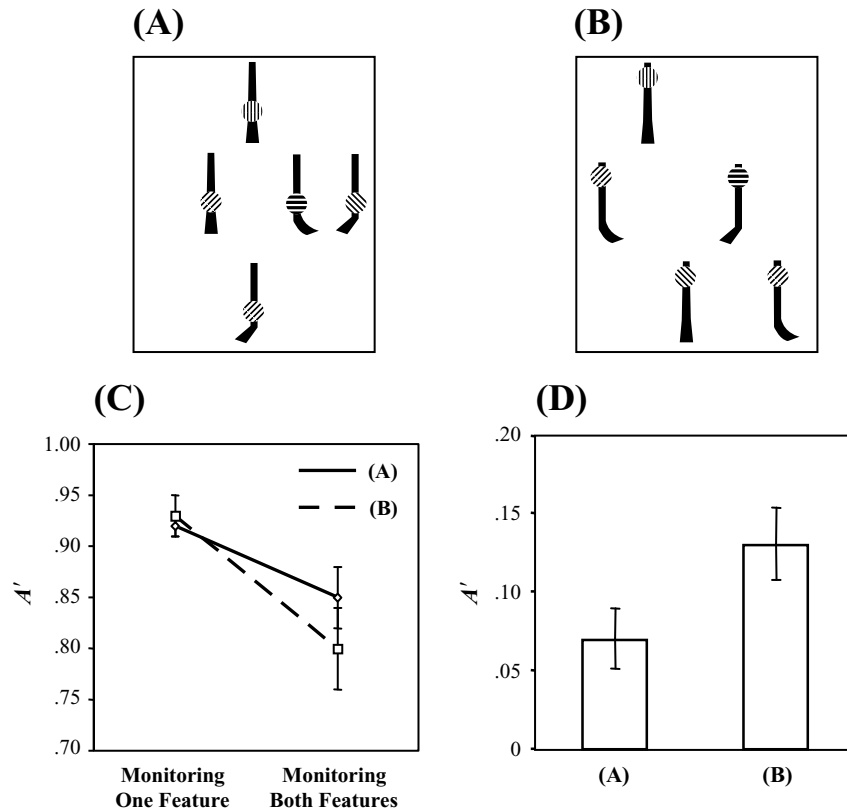

Figure 7. Panels A and B show examples of the displays used in Experiment 4B. The texture patterns represent the different colors used in the experiment. In comparison with the corresponding conditions in Experiment 4A, the distance between parts decreases in panel A and increases in panel B. (C) Results of Experiment 4B. (D) The results shown in panel C replotted as the differences in A' for monitoring one versus two features. As the distance between object parts increases from panel A to B, the object benefit in visual short-term memory (VSTM) decreases. These results confirmed the findings of Experiment 4A and showed that spatial location plays an important role in determining the object benefit in VSTM. Error bars indicate standard errors.

1.72, $p = .25$], with performance higher for monitoring for one than for two features [$F(1,6) = 31.02, p = .003$]. Most important, there was a significant interaction between monitoring condition and display type [$F(1,5) = 7.11$, $p = .045$], showing that as the distance between the parts increased, the object benefit in VSTM was modulated accordingly. None of the other effects reached significance.

Discussion of Experiments 4A and 4B

In both Experiments 4A and 4B, while keeping the connection strength between object parts fixed, increasing the spatial distance between the parts attenuated the object benefit in VSTM and resulted in a smaller total amount of information being stored in VSTM. These results argue against a pure connectedness account that denies the involvement of location/proximity in determining the object benefit in VSTM. The hybrid account, which proposes that both location/proximity and connectedness contribute to the object benefit in VSTM, therefore seems to best capture how object features are retained in VSTM.

GENERAL DISCUSSION

Previous studies have reported the object benefit, or object-based feature integration, in VSTM (e.g., Allport, 1971; Duncan, 1984; Irwin & Andrews, 1996; Luck & Vogel, 1997), which has led to the notion that the units of information processing in VSTM are objects. It is not clear, however, what drives this object benefit in VSTM. This benefit could be caused by the fact that the features of an object always share a common spatial location (the location/proximity account), the features of an object are always connected to each other (the connectedness account), or both of these factors (the hybrid account). To distinguish among these three accounts, in the present study, the connection strength and the spatial distance between object parts were varied independently in change detection experiments. In Experiments 1A, 1B, and 3, although the distance between the object parts was fixed, decreasing the connection strength between the parts decreased the object benefit in VSTM. Experiment 2 provided a control that ruled out an alternative account for the results of Experiments 1A and 1B. In Experiments 4A and 4B, although the connection strength between the object parts was fixed, increasing the spatial distance between the parts also decreased the object benefit in VSTM. Together, these results indicated that both location and connectedness are crucial in determining object-based feature integration and the total amount of information that can be stored in VSTM, so that features from connected and closer parts are better retained in VSTM than are features from less connected parts and from parts that are further apart. This is consistent with the hierarchical feature integration mechanism proposed by Xu (2002a, 2005), which argues that in VSTM, a feature is most strongly integrated with features sharing the same spatial location, moderately integrated with features from another part of the same object, and least integrated with features from other spatially separated objects.

Further studies, however, will be needed to investigate whether location and connectedness operate independently or whether there is interaction between the two in determining the object benefit in VSTM. This will, in turn, provide us with a better understanding of the detailed mechanisms our visual system uses to represent these two types of object properties.

A recent study by Davis and Holmes (2005) used stimuli similar to those in Experiments 1A, 1B, and 3 and showed no effect of part connectedness. However, those authors did not control for perceptual differences between the different displays used. A greater difficulty associated with perceiving the connected parts could easily have erased an object benefit from these objects in Davis and Holmes. In the present Experiment 1B, perceptual differences were well matched between the connected and the unconnected parts, so that change detection performance was comparable when only one object feature was retained in VSTM. As such, when two object features were retained in VSTM, performance was higher for the connected than for the unconnected parts. In Experiment 3, however, perceptibility differed between the two displays, since change detection performance for just one object feature was lower for the connected than for the unconnected parts, possibly due to a stronger shape-masking effect between the top and the bottom part shapes when the parts were directly connected. When both types of features were monitored, there were no longer differences between the two displays, indicating that a positive object benefit effect canceled out the negative perceptual difficulty effect of the connected parts. If we compare the drop in performance for monitoring two versus one part feature, there was still a significantly smaller drop for the connected than for the unconnected parts, indicating that features from the connected parts were better retained in VSTM than were those from the unconnected parts. It is, therefore, crucial to take perceptibility differences into account when comparing performance differences between different displays.

Where does the object benefit occur in VSTM? Is it at feature encoding, feature maintenance, feature retrieval, or more than one of these three stages? Visual search studies have shown that conjunction search is more efficient for two features from the same object part than for two features from different object parts (Goldsmith, 1998; Xu, 2002c). Because visual search requires online processing of the search objects, it shares only feature encoding, but not feature maintenance and retrieval, with VSTM tasks. The presence of the object benefit in visual search suggests that the object benefit in VSTM is occurring at least at the feature-encoding stage. It is likely that the integrated object representation formed at the encoding stage of VSTM will be subsequently maintained throughout the memory delay and then retrieved for comparison with the test display. More research, however, is needed to verify this hypothesis.

The findings of the present study challenge the notion that only features from a fixed number of objects can be processed in VSTM at a given time (e.g., Luck & Vogel, 1997). Depending on the connection strength and the distance between object parts, with the same number of ob-

jects and part features, the present findings showed that the total amount of information and, thus, the number of objects that can be held in VSTM may differ. This is, in general, consistent with Alvarez and Cavanagh (2004), who argued that the capacity of VSTM is not limited to a fixed number of objects but, rather, depends on the nature of the visual stimuli stored.

The present study also addressed what the essential elements are for an object representation in VSTM. As was described in the introduction of this article, although the definition of object can be ambiguous and subjective, the present study shows that such properties as location/proximity and connectedness are obligatory and essential elements in defining a coherent object representation in VSTM. These object properties, however, should not be considered as unique and essential only in retaining objects in VSTM. Connectedness between object parts has also been shown to influence the shift of vision attention (Watson & Kramer, 1999), multiple-object tracking (Scholl, Pylyshyn, & Feldman, 2001), and the performance of patients with visual neglect (Mattingley et al., 1997; Tipper & Behrmann, 1996). Although the role of location/proximity in the visual perception of objects has not been much explored, it is perceivable that location/proximity may be important in other types of object-based visual processing as well.

This brings up the question of whether we have a fixed object representation for all types of visual processing or whether the details of our object representation differ for different types of visual processing, so that certain properties matter more than others for a given visual task. In other words, is our object representation independent of the visual processing that we are engaged in, so that different types of visual processing tap into different aspects of the same underlying object representation? Or is our object representation dynamic, flexible, and influenced by the particular visual processing that we are involved in? Given that certain object properties, such as connectedness, have been shown to influence many different types of visual processing involving objects and that we do have intuitions about what counts as a visual object, it seems parsimonious to take the first view. Further research is needed to fully understand how objects are represented in the visual system.

REFERENCES

AARONSON, D., & WATTS, B. (1987). Extensions of Grier's computational formulas for *A′* and *B″* to below-chance performance. *Psychological Bulletin*, **102**, 439-442.

ALLPORT, D. A. (1971). Parallel encoding within and between elementary stimulus dimensions. *Perception & Psychophysics*, **10**, 104-108.

ALVAREZ, G. A., & CAVANAGH, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, **15**, 106-111.

DAVIS, G., & HOLMES, A. (2005). The capacity of visual short-term memory is not a fixed number of objects. *Memory & Cognition*, **33**, 185-195.

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

DUNCAN, J. (1993). Similarity between concurrent visual discrimina-

tions: Dimensions and objects. *Perception & Psychophysics*, **54**, 425-430.

DUNCAN, J., & NIMMO-SMITH, I. (1996). Objects and attributes in divided attention: Surface and boundary systems. *Perception & Psychophysics*, **58**, 1076-1084.

EGLY, R., DRIVER, J., & RAFAL, R. (1994). Shifting visual attention between objects and locations: Evidence for normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, **123**, 161-177.

GOLDSMITH, M. (1998). What's in a location? Comparing object-based and space-based models of feature integration in visual search. *Journal of Experimental Psychology: General*, **127**, 189-219.

GRIER, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, **75**, 424-429.

HOFFMAN, D. D., & RICHARDS, W. A. (1984). Parts of recognition. *Cognition*, **18**, 65-96.

HOFFMAN, D. D., & SINGH, M. (1997). Salience of visual parts. *Cognition*, **63**, 29-78.

HOLLINGWORTH, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 519-537.

HOLLINGWORTH, A., & HENDERSON, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 113-136.

HOLLINGWORTH, A., WILLIAMS, C. C., & HENDERSON, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin & Review*, **8**, 761-768.

HUMPHREYS, G. W., & RIDDOCH, M. J. (1993). Interactions between object and space systems revealed through neuropsychology. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 183-218). Cambridge, MA: MIT Press.

IRWIN, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 307-317.

IRWIN, D. E., & ANDREWS, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 125-155). Cambridge, MA: MIT Press.

JIANG, Y., CHUN, M. M., & OLSON, I. R. (2004). Perceptual grouping in change detection. *Perception & Psychophysics*, **66**, 446-453.

LEE, D., & CHUN, M. M. (2001). What are the units of visual short-term memory, objects or spatial locations? *Perception & Psychophysics*, **63**, 253-257.

LUCK, S. J., & VOGEL, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.

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

MARR, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Freeman.

MATTINGLEY, J. B., DAVIS, G., & DRIVER, J. (1997). Preattentive filling-in of visual surfaces in parietal extinction. *Science*, **275**, 671-674.

OLSON, I. R., & JIANG, Y. (2002). Is visual short-term memory object based? Rejection of the "strong-object" hypothesis. *Perception & Psychophysics*, **64**, 1055-1067.

PALMER, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

PALMER, S. [E.], & ROCK, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review*, **1**, 29-55.

PASHLER, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, **44**, 369-378.

POLLACK, I., & NORMAN, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science*, **1**, 125-126.

POTTER, M. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 509-522.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, **8**, 368-373.

Saiki, J., & Hummel, J. E. (1998). Connectedness and the integration of parts with relations in shape perception. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 227-251.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, **80**, 1-46.

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multi-element tracking. *Cognition*, **80**, 159-177.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, **7**, 261-267.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, **117**, 34-50.

Tipper, S., & Behrmann, M. (1996). Object-centered not scene-based visual neglect. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 1261-1278.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, **428**, 751-754.

Vecera, S. P., Behrmann, M., & Filapek, J. C. (2001). Attending to the parts of a single object: Part-based selection limitations. *Perception & Psychophysics*, **63**, 308-321.

Vecera, S. P., Behrmann, M., & McGoldrick, J. (2000). Selective attention to the parts of an object. *Psychonomic Bulletin & Review*, **7**, 301-308.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, **428**, 748-751.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 92-114.

Walker, P., & Davies, S. J. (2003). Perceptual completion and object-based representations in short-term visual memory. *Memory & Cognition*, **31**, 746-760.

Watson, S. E., & Kramer, A. F. (1999). Object-based visual selective attention and perceptual organization. *Perception & Psychophysics*, **61**, 31-49.

Wertheimer, M. (1950). Gestalt theory. In W. D. Ellis (Ed.), *A sourcebook of Gestalt psychology* (pp. 1-11). New York: Humanities Press. (Original work published 1924)

Wheeler, M., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, **131**, 48-64.

Wing, A., & Allport, D. A. (1972). Multidimensional encoding of visual form. *Perception & Psychophysics*, **12**, 474-476.

Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, **10**, 80-87.

Xu, Y. (2002a). Encoding color and shape from different parts of an object in visual short-term memory. *Perception & Psychophysics*, **64**, 1260-1280.

Xu, Y. (2002b). Limitations in object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 458-468.

Xu, Y. (2002c). Object-based feature integration in visual search. *Perception*, **31**, 1335-1347.

Xu, Y. (2005). *Integrated object feature representations in visual short-term memory: Evidence from encoding same-dimension features of an object*. Manuscript submitted for publication.

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, **440**, 91-95.

Xu, Y., & Singh, M. (2002). Early computation of part structure: Evidence from visual search. *Perception & Psychophysics*, **64**, 1039-1054.

NOTES

1. It is important to understand how other Gestalt grouping factors, such as common fate, convexity, etc., influence VSTM feature storage. However, given that the focus of this article is on understanding the cause of object-based feature integration in VSTM, which concerns only connectedness and proximity, a systematic study of all the Gestalt grouping factors on VSTM feature storage is beyond its scope.

2. A' was calculated for each participant in each condition, following the formula by Grier (1971):

$$A' = .5 + (H - g) * (1 + H - g) / [4 * H * (1 - g)],$$

where H is the hit rate and g is the guessing rate or false alarm rate. If the guessing rate was greater than the hit rate, the following formula was used (Aaronson & Watts, 1987; Snodgrass & Corwin, 1988):

$$A' = .5 - (g - H) * (1 + g - H) / [4 * g * (1 - H)].$$

3. One may wonder whether the decrease in performance in Experiment 1A Condition B, relative to Condition A, was indeed due to the visual system's treating objects in Condition B as having more parts with less well connected circle and stem parts or, alternatively, to amodal completion of the stem behind the gray bar, incurring a cost in Condition B. Note that in Experiment 1A, Condition B, the object was constructed in such a way that the entire circle was visible in front of the horizontal gray bar (see Figure 2B; the actual object was larger on the computer screen during testing so that this could be clearly seen). As such, the object in Condition B could not be perceived as the object in Condition A behind a gray bar. Rather, the object in Condition B would be more naturally perceived as an object having three connected parts: the circle, the gray bar, and the stem.

Whereas the stem in Condition B could be perceived as extending behind the horizontal gray bar, due to amodal completion, the stem in Condition A could be perceived as extending behind the circle as well. The presence of amodal completion was, therefore, matched between Conditions A and B. It is possible, however, that amodal completion might have been more important in Condition B than in Condition A because it would extend the stem in Condition B and facilitate orientation perception. The slightly higher performance for Condition B than for Condition A when only stem orientation was monitored indicated that even if amodal completion was more crucial in Condition B than in Condition A, it certainly did not impose extra cost in VSTM change detection. Moreover, the difference in performance between Conditions A and B was observed only when both color and orientation features had to be retained in VSTM. These results suggest that the performance difference between Conditions A and B was likely caused by how color and orientation features were connected and was not due to the influence of amodal completion.

4. One could argue that for the connected parts, when only one feature was monitored, the performance was low because observers encoded both features automatically, even though the encoding of only one feature was required. This account, however, would suggest only that features from the connected parts are better integrated than are those from the unconnected parts, which was the exact implication of the present findings. On the other hand, differences in performance between displays in the single-feature condition was not observed in Experiment 1B, which is inconsistent with this account. Moreover, if both features from the connected parts were always automatically encoded regardless of task, this account cannot explain why performance still dropped significantly for the connected parts when both features had to be remembered, in comparison with when just one feature was remembered. As such, the lower performance observed for the connected parts than for the unconnected parts when only one feature had to be monitored was likely due to perceptual differences between the two types of displays.