Journal of Experimental Psychology:
Learning, Memory, and Cognition
2007, Vol. 33, No. 5, 959–969

0278-7393/07/$12.00 DOI: 10.1037/0278-7393.33.5.959

Binding Facilitates Attention Switching Within Working Memory

Min Bao
University of Science and Technology of China

Zhi-Hao Li
Emory University

Da-Ren Zhang
University of Science and Technology of China

The authors investigated the units of selective attention within working memory. In Experiment 1, a group of participants kept 1 count and 1 location in working memory and updated them repeatedly in random order. Another group of participants were instructed to achieve the same goal by memorizing the verbal and spatial information in an integrative way as a moving digit. The behavioral data showed that switching attention between properties of an integrated working-memory item was faster than switching between respective properties of different items. Experiment 2 demonstrated that this switching facilitation cannot be simply ascribed to the different amount of working-memory items maintained by the two groups of participants. Finally, by adopting a pure verbal task in Experiment 3, the authors observed the same binding facilitation, with the possibility of "location-based selection" excluded. They summarize the observations of all 3 experiments in the study and suggest both a location- and object-based mechanism for attention selection in working memory.

Keywords: working memory, attention, location-based, object-based, binding

Supplemental materials: http://dx.doi.org/10.1037/0278.7393.33.5.959.supp

Attention allows people to selectively process the environment that is most relevant to their goals (Desimone & Duncan, 1995). One of the central issues in the neuropsychological research of attention is the debate over the units of attentional selection, such as locations, features, and objects.

Theories of Selective Attention

In the past decades, three typical theories of visual attention, which, respectively, proposed feature-based, location-based, and object-based selection, have been intensively studied, and each has received respective evidential support. The feature-based theory considers specific features (e.g., color, orientation, or motion direction) to be the units of attentional selection (Maunsell & Treue, 2006; Saenz, Buracas, & Boynton, 2002; Treue & Martinez Trujillo, 1999). However, the theory of location-based attention involves the selection stimuli from spatial locations (Eriksen & Hoffman, 1973; Eriksen & St. James, 1986; LaBerge & Brown, 1989; Posner, 1980). For example, the classical "spotlight" model describes attention as a flashlight spot that moves between different spatial locations so that stimuli within the spot are processed more and faster. Besides these two theories, behavioral and neurophysiological evidence have also shown that selective attention can operate on an object-based representational medium, in which the boundaries of segmented objects determine what is selected and how attention is deployed (Blaser, Pylyshyn, & Holcombe, 2000; Duncan, 1984; O'Craven, Downing, & Kanwisher, 1999; Roelfsema, Lamme, & Spekreijse, 1998; Schoenfeld et al., 2003; Vecera & Farah, 1994; Yantis, 1992). One prediction of the object-based theory is that once attention is directed to a particular feature of a visual object, in actuality the whole object and all of its features (including those currently task-irrelevant) are selected. This prediction is supported by many visual-attention studies. For instance, Rodríguez, Valdés-Sosa, and Freiwald (2002) found that the participants' judgment performance was much better when attention was directed to two features within the same surface instead of two features separated into different surfaces.

Attentional Selection in Working Memory

Despite the relatively large number of studies on visual–spatial attention, few researchers have considered attentional selection in working memory (i.e., the selective access to different items held in working memory). However, as the experimental data from

---

Min Bao and Da-Ren Zhang, Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, China; Zhi-Hao Li, Department of Biomedical Engineering, Emory University.

This research was supported by National Nature Science Foundation of China Grants 30370478 and 30328017 and Ministry of Science and Technology of China Grant 2006CB500705. We thank all of the volunteers who participated in the study. Great thanks also go to Sheng He (University of Minnesota), Alex O. Holcombe (University of Sydney), Stephan Lewandowsky (University of Western Australia), and Shaun P. Vecera (University of Iowa). We also thank Robert Shannon (University of Minnesota), Priya Santhanam, and Callie Tyner (Emory University) for their help in improving the writing.

Correspondence concerning this article should be addressed to Da-Ren Zhang, Hefei National Laboratory for Physical Sciences and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230026, China. E-mail: drzhang@ustc.edu.cn

Garavan (1998) and Oberauer (2002) show, there does exist a focus of attention within working memory, which is analogous to the focus of attention in outer visual space. According to the "embedded-process" model (Cowan, 1988, 1995) of working memory, a representation in the focus of attention is what a person is aware of at any time. Any cognitive operation directed by the goal will be applied to the representation in this focus but not to the other representations, even those also currently held in working memory. Because attention in working memory is so closely related to the central executive function, studies in this field have recently become more prevalent, as researchers attempt to expand insights into the operations of this mental resource. The present research was thereby designed to study the units of this internal attentional selection.

The Object-Switching Paradigm

The object-switching paradigm is one of the commonly used approaches in studies of attention in working memory (Bao, Li, Chen, & Zhang, 2006; Garavan, 1998; Garavan, Ross, Li, & Stein, 2000; Gehring, Bryck, Jonides, Albin, & Badre, 2003; Kübler, Murphy, Kaufman, Stein, & Garavan, 2003; Li et al., 2004, 2006; Oberauer, 2002, 2003, 2005; Sylvester et al., 2003). It was originally introduced by Garavan (1998) as the "serial count" task. In this task, two kinds of geometric figures (e.g., squares and circles) are serially presented in a random order at the participant's own pace (the current displaying figure is replaced by the next one each time the response key is pressed). Participants in this task must count the number of times that each of the two figure types appear in a sequential presentation. Although the figures are presented one by one, participants have to keep two running counts for each figure type and make corresponding increments each time the figure updates. On the basis of whether the successive stimuli were of the same type, "nonswitch" (successive stimuli were the same) or "switch" (successive stimuli were different) updates could be defined. Garavan (1998) showed that the "nonswitch" reaction times (RTs) were about 500 ms shorter than those of the "switch" condition. He interpreted this difference as the cost for attention shifting from one memory count to the other and proposed that the capacity of this internal focus of attention is limited to just one item.

To further study the crossmodal characteristics of internal attention switching, Kübler et al. (2003) modified this verbal counting task by including spatial working-memory items. Their study included three successive tasks. The first was a verbal task that repeated the above-mentioned serial counting paradigm. The second was a visuospatial task with each trial session initiated by one red and one blue dot appearing in different cells of a 2 × 2 grid. After participants pressed the space bar, the initial grid and dots disappeared and never reappeared during the trial. The subsequent presentation was a list of red and blue arrows displayed in random order at participants' own paces (only one arrow was presented at a time). Participants were instructed to memorize the initial position of each dot and keep updating these two positions in their minds' eyes (i.e., to move the red/blue dot to a nearby cell in the direction of the appearing red or blue arrow). Thus, each time an arrow was presented, the participants were to update the memorized position of the dot with a color corresponding to the appearing arrow, while keeping the position of the other dot unaltered, and press the key as soon as they finished the mental update. These first and second tasks addressed attention switching within the verbal and visuospatial modalities, respectively. Furthermore, in the third (crossmodal) task, each trial session was initiated by the display of a blue dot that occupied one cell of a 2 × 2 grid. The subsequent stimuli were a series of blue arrows and red squares presented in random order at participants' own paces. The participants were to count the number of times that the square appeared in a trial and update the memorized position of the dot according to each appearing arrow. In all of these tasks, the space-bar pressings were only for recording the RTs and bringing the next stimulus. The participants only reported the final results at the end of a trial session.

Kübler et al.'s (2003) results showed that an attention-switching cost persisted in all three tasks. Moreover, they found that the switching costs in all three tasks were mutually correlated, which suggested the existence of a general, supramodal attention-switching process that transcends specific working-memory slave systems.

The Present Study

In the third task of Kübler et al. (2003), the verbal (count) and visuospatial (position) information do not have any intrinsic association, and the participants need to keep them as separate items in working memory. However, instead of updating two working-memory items, one can complete the same task in an alternative way by memorizing the count and position together as two properties of an integrated item (see Figure 1B).[1] Namely, the participant can imagine a digit in the 2 × 2 grid and, when he or she sees an arrow, move it to a nearby cell without changing its value or, when he or she sees a cross, increase the digit by one, leaving its position unaltered.

In other words, there are two distinct strategies in performing this same task: Memorize the verbal and spatial information separately (i.e., as two different working-memory items, under the paradigm shown in Figure 1A), or memorize the information jointly (i.e., as two features belonging to one integrated working-memory item, under the paradigm shown in Figure 1B). On the basis of these two strategies, we designed the present study to investigate the units of attentional selection in working memory. If attention in working memory is object based, then attending to one property of an integrated item would also enhance the attention (activation) level on the other property (O'Craven et al., 1999; Vecera, 2000). As a result, attention switching between the two properties of an integrated item would be faster than that between two separately memorized properties. In contrast, if such a binding facilitation cannot be observed, the attentional selection in working memory cannot be object based. For the purpose of simplicity, we call these two contrasting strategies *binding* and *separate*, respectively, in the rest of this article.

To examine these two possibilities, in Experiment 1, we randomly assigned participants to two groups, each adopting one of

---

[1] It is controversial and hard to define an object, even in visual-attention studies. In this article, we define a *memory object* as a working-memory item that is the target of the updating operation in the task, no matter how many features it has.

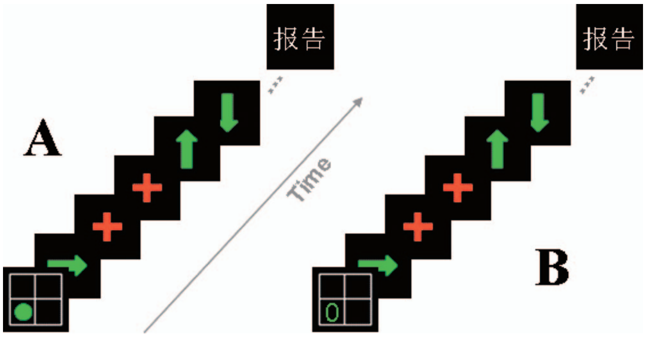

Figure 1. Schematic depiction of Experiment 1. (A) The depiction of stimuli used for the *separate* group; (B) the depiction of stimuli used for the *binding* group. Stimuli were presented at each participant's own pace. The response-stimulus interval was fixed at 100 ms.

the two strategies. We used a between-subjects design instead of a within-subject design, because once a participant becomes aware of the binding strategy, it would be difficult for him or her to separate the two properties of the object.

Experiment 1

Method

Participants. The participants in all the experiments of the present study were graduate or undergraduate students from the University of Science and Technology of China, Hefei, Anhui, China (USTC). They gave informed consent to participate in our experiments and, in return, received monetary compensation or extra course credit. Thirty participants were recruited and randomly assigned into the *separate* (5 female, 10 male; *M* age = 20.5 years, range = 19–22) and *binding* (5 female, 10 male; *M* age = 20.8 years, range = 16–25) groups.

Materials and procedure. For the *separate* group, each participant completed 40 experimental trials after 5 practice trials (a trial was defined as the task session, starting with the initial grid presentation and ending with the participant's report). If the participant correctly performed fewer than two trials in practice, he or she had to practice five more trials. This rule for practice was applied to all the experiments in the present study. As shown in Figure 1A, at the beginning of each trial, a 2 × 2 grid was centrally presented, with one cell filled by a green dot. Participants were asked to memorize the spatial location (i.e., which quadrant) of this dot. After the participant pressed the space bar, the 2 × 2 grid and the dot disappeared (never appearing again during the trial) and a sequence of centrally displayed red crosses and green arrows appeared on the screen. Each individual display contained only one cross or one arrow, and the type of stimulus varied randomly in each trial. The stimulus presentation was self-paced. By pressing the space bar, the participant erased the current display, and after a 100-ms blank screen (to make successive stimuli distinguishable), a new stimulus appeared. The number of total individual displays within a trial varied from 8 to 17.

The participants' task was to imagine the green dot moving to a nearby cell, indicated by the direction of a green arrow presented to them (e.g., if the dot occupied the upper left cell after the last updating, and the following stimulus was a right-pointing green arrow, the participant then should move the dot to the upper right cell in his or her mind's eye), and to subvocally count the number of times that the red cross appeared. Participants were instructed to complete each update as accurately and quickly as possible and to press the space bar immediately after each update was completed. The participants were told that the green dot would never go out of the 2 × 2 grid. If at any time the participant determined that appearing arrow direction would cause the green dot to move outside the grid, he or she was to report this to the experimenter so that the relevant trial could be considered unsuccessfully performed, and the corresponding data could be excluded from the subsequent analysis. The time duration from each individual display onset to the subsequent pressing of the space bar was recorded as the RT. At the end of each trial, a message would appear on the screen asking participants to report the final cross count and dot position in their memories.

To avoid two-digit number counting and to ensure sufficient times of update, we limited the occurrence of crosses to no more than nine and no fewer than three. We encouraged participants to rest about every ten trials. However, participants were free to take a brief break after completing (but never during) each trial. The same resting rule was applied in all of the other experiments in the present study. Before starting the experiment, all participants were told that the two tasks, moving dots and counting crosses, were independent of each other, so that they would handle the location and the count as two distinct items, stored separately in spatial and verbal working memory.

For the *binding* group, all the presenting stimuli were the same as those used in the *separate* group, except that the initial green dot was replaced by a green zero in the 2 × 2 grid (see Figure 1B). During the subsequent process, participants were instructed to imagine the green digit moving to a nearby cell without changing its value when they saw a green arrow or to increase the digit by one without changing its location when they saw a red cross. At the end of a trial, they were to report both the final location and the final value of the digit.

There were a total of four types of successive stimulus pairs for both groups: arrow followed by arrow, cross followed by cross, arrow followed by cross, and cross followed by arrow. The first two were in the condition of nonswitch updating, and the others were in the condition of switch updating. There were 88–90 responses for each of these four types, and they were pseudorandomly distributed in each trial. Because switch cost can depend on the number of consecutive preceding nonswitch updates, in Experiment 1, the numbers of consecutive nonswitch updates interspersed between switch updates were designed comparably for the two groups. They were also comparable for every concerned comparison in all of the other experiments in the present study.

Results and Discussion

Each participant described his or her own strategy at the end of the experiment. Two participants who were instructed to use the *separate* strategy reported that they adopted the *binding* strategy, whereas all of the others adopted the instructed strategy. These 2 participants were excluded from the subsequent data analysis (the participant information for the *separate* group in the *Participants* section does not include the 2 excluded participants). To match the participant numbers between the two groups, we recruited 2 new participants for the *separate* group. They adopted the *separate* strategy as instructed.

For both groups, most errors were of the type in which only the value of the digit was incorrect, and the incorrect value was off by only one. It is therefore reasonable to infer that the participants were diligent in updating the items in these trials. Without considering this type of error, the mean accuracies of the *separate* and *binding* groups were 95.6% and 96.3%, respectively. There was no significant between-groups difference; group t test, $t(28) = 0.455$ ($p = .652$).

RTs from the incorrectly reported trials (i.e., the final location was wrong, or the value of the digit was incorrect and off by more than one) were excluded from data analysis. In addition, RTs that exceeded 10 s or were shorter than 300 ms were excluded first, then those longer than the mean RT of each condition by three times the standard deviation were also excluded. These excluded RTs were deemed outliers (the eliminating rate was between 1.6% and 2.4% of RTs per condition).

With the attention-switching status (no shift vs. shift) and the operation type (counting vs. dot/digit moving) as the within-subject factors, and the group (separate vs. binding) as the between-subjects factor, a repeated-measures analysis of variance (ANOVA) on the RTs yielded significant main effects for switching status, $F(1, 28) = 90.89$, $p < .001$, and operation type, $F(1, 28) = 11.25$, $p = .002$, as well as significant two-way interactions of Switching Status × Group, $F(1, 28) = 12.21$, $p = .002$, and Operation Type × Group, $F(1, 28) = 7.866$, $p = .009$. The mean RTs of each condition and group are graphically shown in Figure 2 and Table 1. The main and interaction effects not mentioned here were all nonsignificant.

Besides the attention-shift cost, which was repeatedly observed in many previous studies (Bao et al., 2006; Garavan, 1998; Gehring et al., 2003; Li et al., 2004; Oberauer, 2002), the significant interaction between the switching status and participant group was one of our new findings in the present study. Our study showed smaller shift cost in the *binding* group (182 ms) than in the *separate* group (393 ms). This effect cannot be explained by merely a general slowing of responses in the *separate* group to allow for updating information, because the effect was still significant, $F(1, 28) = 10.06$, $p = .004$, when RTs in each condition for each participant were divided by the participant's mean RT.

Experiment 1 showed that the binding strategy can make attention switching faster in working memory. However, some researchers have demonstrated that the cost of attention shift can increase with memory-set size (Oberauer, 2002, 2003). The group difference we observed could therefore also be due to a memory-load difference, instead of the binding strategy, because the two pieces of information were bound as one item for the *binding* group, whereas they were still independent items for the *separate* group. One simple way to test this memory-load explanation would be to introduce an extra nonupdating memory digit for the *binding* participants so that the memory load could be balanced between the two groups (e.g., present an extra digit at the beginning of a trial and require participants to memorize but never update it throughout the trial). However, adding only a passive (nonupdating) memory digit would not be enough to balance the load. As described in Oberauer's (2002) concentric working-memory model, the passive and active (i.e., those to be updated) memory items should be of two relatively independent working-memory statuses. Items in the passive mode will have no significant influence on the attention-switching time between active working-memory items (Oberauer, 2002, 2005). To ascertain the binding effect while avoiding this "passive/active" issue at the same time, in Experiment 2, we required participants to simultaneously update two integrated items. If switching internal attention between properties of the same object is indeed faster than switching attention between properties of different objects, we can expect a faster response in an intraobject switching than in an interobject switching.

Experiment 2

Method

Participants. Another group of 10 students (6 female, 4 male; M age = 20.7 years, range = 19–23) participated in Experiment 2.

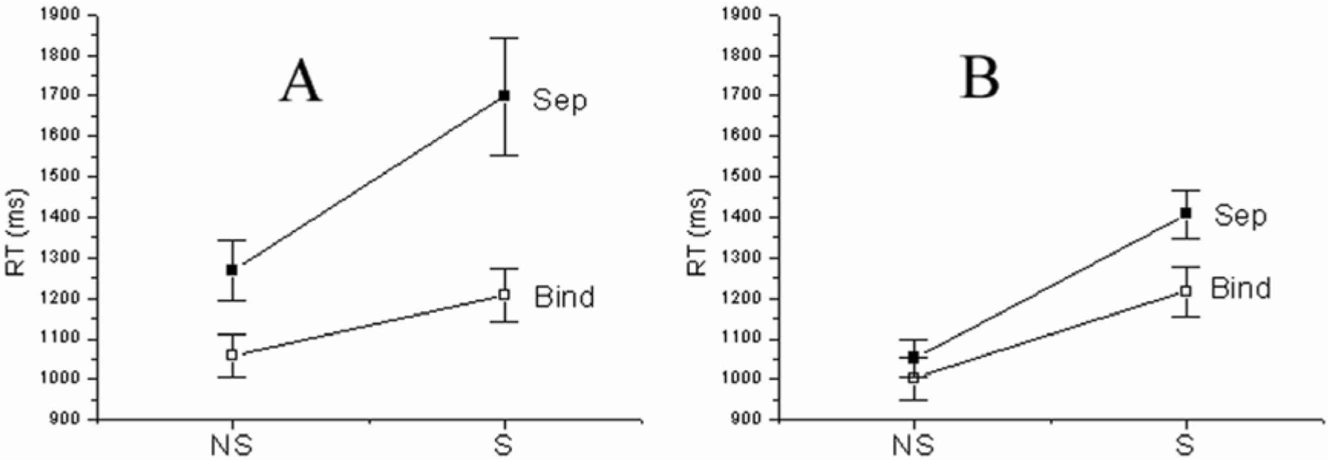

Figure 2. No-shift (NS) and shift (S) reaction times (RTs) for the *separate* (Sep) and *binding* (Bind) groups. The error bars represent standard errors. Panels A and B, respectively, show the results of dot/digit moving and counting.

Table 1

Grand Mean Reaction Time (RT) for Experiment 1

Operation type	Group	Switching status	RT (ms) *M*	RT (ms) *SD*	Shift cost (ms)	Between-group Δ shift cost (ms)
Counting	Bind	NS	1,002	207	115	240
		S	1,217	239		
	Sep	NS	1,052	174	355	
		S	1,407	229		
Dot/digit moving	Bind	NS	1,058	202	149	381
		S	1,207	260		
	Sep	NS	1,269	289	430	
		S	1,699	561		

Note. Bind = *binding* group; Sep = *separate* group; NS = nonswitch; S = switch.

Materials and procedure. Each participant completed 50 trials (excluding practice) in Experiment 2. The initial display was a 2 × 2 grid presented in the center of the screen with one cell filled by a red zero and one cell filled by a green zero. The subsequent stimuli were a series of red crosses, red arrows, green crosses, and green arrows, which were presented one by one in random order at participants' own paces. The trial length varied from 17 to 30 individual displays. Also, as in Experiment 1, the occurrence of each kind of cross was no more than 9 times and no fewer than 3. The digit manipulation (moving or adding) was the same as that for the *binding* group in Experiment 1, except that two stimulus colors (red and green) were respectively mapped to two memory digits. At the end of each trial, participants were instructed to report the final value and location of both of the two integrated digits.

In this task, participants were asked to keep two digits, each with two integrated properties, in working memory. They needed to switch attention between the two properties of the same digit (intraobject switching) or between properties of two different digits (interobject switching), according to the stimulus list. We refer to the four kinds of stimulus briefly as c1 (red cross), a1 (red arrow), c2 (green cross), and a2 (green arrow). There are a total of 16 permutations of sequential item pairs, which can be categorized into the following four conditions:

1. object nonswitch and feature nonswitch (c1c1, a1a1, c2c2, a2a2),
2. object nonswitch and feature switch[2] (c1a1, a1c1, c2a2, a2c2),
3. object switch and feature nonswitch (c1c2, c2c1, a1a2, a2a1), and
4. object switch and feature switch (c1a2, a2c1, c2a1, a1c2).

Each of these 16 conditions included 78 responses. The intraobject shift cost was defined as the RT calculation with expression $[(\mathrm{c1a1} - \mathrm{a1a1}) + (\mathrm{a1c1} - \mathrm{c1c1}) + (\mathrm{c2a2} - \mathrm{a2a2}) + (\mathrm{a2c2} - \mathrm{c2c2})]/4$, whereas the interobject shift cost was calculated as $[(\mathrm{c1a2} - \mathrm{a2a2}) + (\mathrm{a2c1} - \mathrm{c1c1}) + (\mathrm{c2a1} - \mathrm{a1a1}) + (\mathrm{a1c2} - \mathrm{c2c2})]/4$. The main aim of this experiment was to compare the intraobject shift cost with the interobject shift cost. If switching internal attention between properties of the same object is indeed faster than that between different objects, the intraobject shift cost should be smaller than the interobject shift cost.

Results and Discussion

The criteria for error-trial exclusion were similar to those in Experiment 1 (i.e., one of the final locations was wrong, or the value of one digit was incorrect and off by more than one). The participants completed their tasks with an accuracy of 88.4% ± 10.9%. For the RT data, we applied a method of outlier elimination similar to that used in Experiment 1 (eliminating 1.7%–2.4% of RTs per condition).

We observed that the intraobject shift cost (418 ms) was significantly paired *t* test, $t(9) = 5.079$, $p < .001$, smaller than the interobject shift cost (634 ms), which supports the object-based theory. Table 2 and Figure 3 show the RT comparisons in more detail.

To provide a comprehensive picture of RT data under each condition, we report the exhaustive ANOVA results. A 2 (Object-Switching Condition: switch vs. nonswitch) × 2 (Feature-Switching Condition: switch vs. nonswitch) × 2 (Operation Type: adding vs. digit moving) × 2 (Digit Being Updated: one of the digits vs. the other digit) repeated-measures ANOVA on the RTs yielded significant main effects for the Object–Switching Condition, $F(1, 9) = 49.65$, $p < .001$, and Feature-Switching Condition, $F(1, 9) = 35.75$, $p < .001$, as well as a significant interaction of these two factors, $F(1, 9) = 23.72$, $p = .001$. No significant main effect for the operation type or digit being updated was found. A significant two-way interaction was found for Feature-Switching Condition × Operation Type, $F(1, 9) = 7.883$, $p = .020$. Other interactions were all nonsignificant. Because the factorial design is fairly complicated, cell means listed in Table 2 show the detail more perspicuously.

---

[2] Switching from a numerical value to a location involves both an operation switch and a switch of the attended feature, but in the present context, the contrast between objects and features is what matters, not the contrast between object switch and operation switch. Therefore, we used the term *feature switch* in this article, although it should be noted that the feature switch here also involves an operation switch, and the two cannot be disentangled within the present paradigm. When comparing intraobject and interobject switching, the operation switch was assumed to be cancelled with the "simple insertion" premise.

Table 2
Grand Mean Reaction Times (RTs) for Experiment 2

Object-switching condition	Feature-switching condition	Item pairs	RT (ms)		Grand *M* RT (ms)	Shift cost (ms)
			M	*SD*		
		Intraobject				
NS	NS	a1a1	1,398	215	1,275	1,693 − 1,275 = 418
		c1c1	1,131	229		
		a2a2	1,407	250		
		c2c2	1,164	220		
	S	a1c1	1,707	215	1,693	
		c1a1	1,664	334		
		a2c2	1,721	256		
		c2a2	1,681	440		
		Interobject				
S	NS	a1a2	1,817	246	1,766	1,909 − 1,275 = 634
		a2a1	1,852	352		
		c1c2	1,761	258		
		c2c1	1,634	190		
	S	a1c2	1,902	227	1,909	
		a2a1	1,897	474		
		a2c1	1,917	241		
		c1a2	1,920	346		

Note. NS = nonswitch; S = switch.

In Experiment 2, attention was switched between two properties within one integrated object or between two objects while the memory load was kept constant (two objects, or four properties) throughout all trials. The faster intraobject attention switching therefore could not be interpreted simply by the memory-load effect.

However, exclusion of the memory-load account was still not enough to make object-based selection an acceptable/reasonable option. This is because in intraobject attention switching, the starting and target features are in the exact same location, whereas in the interobject switching, the two features are anchored at different locations (i.e., the two objects are in two different cells, or, even if they are in the same cell, participants could imagine the two digits being spatially offset instead of exactly in the same location). If spatial attention does work, the farther the two features are from each other in memory, the slower the attention switches will be between them. In fact, supporting this location-based account, in the explorative analysis with switch type (intraobject vs. interobject) and object distance (two objects within the same cell vs. two objects in different cells) as factors, we did find significant main effects for both factors: switch-type factor, $F(1, 9) = 19.57$, $p = .002$, responses in the intraobject condition were faster; distance factor, $F(1, 9) = 11.24$, $p = .008$, responses in the same cell condition were faster. We also found a significant two-way interaction, $F(1, 9) = 12.02$, $p = .007$, suggesting a larger RT difference between the interobject and intraobject conditions when the two objects were in different cells (235 ms) rather than in the same cell (119 ms). Therefore, the intraobject facilitation found in this experiment should at least include the contribution of shorter spatial distance.

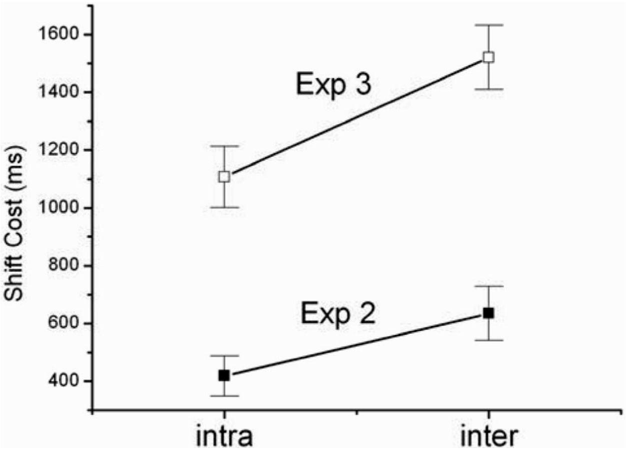

Figure 3. Mean reaction times for Experiments 2 and 3 plotted as a function of different shift-cost types. The error bars represent standard errors. Intra = intraobject shift cost; inter = interobject shift cost.

To further resolve this shifting-distance-related issue, we conducted a third experiment with a pure verbal task. This task had no spatial manipulation demand, so participants should have been unlikely to assign spatial locations to memory objects. An intraobject switching facilitation observed in such a paradigm would therefore be stronger evidence for object-based attentional selection in working memory.

Experiment 3

Method

Participants. Another 10 students (4 female, 6 male; *M* age = 20.5 years, range = 19–22) participated in Experiment 3. They were all native Mandarin Chinese speakers.

Materials and procedure. In this experiment, we designed a new object-switching paradigm that took advantage of Chinese pronunciation characteristics. Mandarin Chinese is a tonal language, in which each syllable can generally be pronounced in four different tones,[3] called "Yin-Ping" (the first tone), "Yang-Ping" (the second tone), "Sang-Sheng" (the third tone), and "Xia-Sheng" (the fourth tone). Because the Chinese pronunciation of each digit from one to nine contains only one syllable, a specific count in working memory can have two verbal properties, the syllable and the tone. By using memory objects with two verbal properties, we eliminated the "switching-distance" contaminations in object-switching RT.

At the beginning of each trial, two zeros pronounced in "Yin-Ping" were defined as the initial counts (see Table 3). A sequence of randomly selected single stimuli (red cross, green dot, yellow cross, or blue dot) was then presented centrally, one by one, and kept visible until a response was made. In Experiment 2, the updating cues of the binding features were displayed in the same color. An intraobject switch could benefit from the perceptual priming of the unchanged successive stimulus colors. Here, we used four colors, each color mapped to one of the four features. This way, the possible color-priming effect in the intraobject condition could be excluded. The trial length varied from 17 to 30 displays. When they saw a red cross, the participants were required to add one to one count without changing its rehearsing tone. When they saw a green dot, the participants were required to change the tone of the count with its numerical value kept constant. Similarly, the yellow crosses and the blue dots, respectively, corresponded to the adding and tone-changing manipulation on the other count. The tone changing followed the cycle of "Yin-Ping → Yang-Ping → Sang-Sheng → Xia-Sheng → Yin-Ping → Yang-Ping →. . . ." For instance, if the stimulus sequence was "red cross, green dot, blue dot, yellow cross, blue dot, red cross. . .," the participant would accordingly rehearse (subvocally) like this: "one (first tone), zero (first tone); one (second tone), zero (first tone); one (second tone), zero (second tone); one (second tone), one (second tone); one (second tone), one (third tone); two (second tone), one (third tone). . . ." See Table 3 for details. The sound file "sequence.wav" in the supplementary materials shows this rehearsing course in a more vivid way. Participants were told to press the space bar as soon as they completed the current operation. At the end of the trial, a message on the screen asked participants to orally report the final values of the two counts in their final tones. For simplicity, we call the adding and tone-changing operations on one count "c1" and "t1" and the adding and tone-changing operations on the other count "c2" and "t2." Exactly as in Experiment 2, there were four types of updating in this task:

Table 3
Schematic Sequence for a Typical Trial in Experiment 3

	Digit 1		Digit 2	
Stimulus	Value	Tone	Value	Tone
None (start)	0	Yin	0	Yin
1 (red cross)	1	Yin	0	Yin
2 (green dot)	1	Yang	0	Yin
3 (blue dot)	1	Yang	0	Yang
4 (yellow cross)	1	Yang	1	Yang
5 (blue dot)	1	Yang	1	Sang
6 (red cross)	2	Yang	1	Sang

Note. The first column represents the temporal sequence of an example stimulus list, in order of presentation. The Value and Tone columns denote the numerical value and tone of the corresponding digit after each update (according to the stimulus). Red crosses and green dots, respectively, correspond to the +1 operation and the tone-changing operation on Digit 1. Yellow crosses and blue dots are similarly mapped to Digit 2. Yin = Yin-Ping; Yang = Yang-Ping; Sang = Sang-Sheng.

1. object nonswitch and feature nonswitch (c1c1, t1t1, c2c2, t2t2),
2. object nonswitch and feature switch (c1t1, t1c1, c2t2, t2c2),
3. object switch and feature nonswitch (c1c2, c2c1, t1t2, t2t1), and
4. object switch and feature switch (c1t2, t2c1, c2t1, t1c2).

The method used to compute the intraobject-shift cost and the interobject-shift cost was similar to that used in Experiment 2.

Results and Discussion

This task is more difficult than the traditional serial-count task. Because the experiment time was restrained to 1 hour, only 1 participant finished the total 50 trials. All the others completed either 30 or 40 trials.

The task was completed with a mean accuracy of 80.8% ($SD = 10.8$). An outlier-eliminating method similar to that used in Experiment 1 was used (i.e., if one of the final tones was wrong or the value of one digit was incorrect and off by more than one, the trial was excluded, eliminating between 1.5% and 2.4% of RTs per condition).

The paired t test result showed that the intraobject shift cost (1,107 ms) was significantly smaller, $t(9) = 4.194$, $p = .002$ (see Figure 3), than the interobject shift cost (1,520 ms).

As in Experiment 2, we also reported the exhaustive results of the full-factors ANOVA. A 2 (Object-Switching Condition: switch vs. nonswitch) × 2 (Feature-Switching Condition: switch vs. nonswitch) × 2 (Operation Type: adding vs. tone changing) × 2 (Digit

---

[3] A specific Chinese word is commonly integrated into one syllable and only one fixed lexical tone, and the meaning of a Chinese word depends on both features. For example, /ma:/ spoken in the tone of Sang-Sheng means "horse," but the same syllable pronounced in the tone of Yin-Ping means mum. This does not mean that a Chinese word always has several tones. In fact, the two example words are two totally different Chinese words (characters) and have different written shapes (马 for /ma:/ in Sang-Sheng; 妈 for /ma:/ in Yin-Ping), and each of them only has a single tone. Likewise, each digit (0–9) corresponds to a syllable spoken in only one of the four tones (see the "digit.wav" file in the supplementary materials). However, to obtain two updateable verbal properties for a count (syllable and tone), in Experiment 3, we artifactually told the participants that the tone of a count was changeable on the basis of the tone-changing rule, even if sometimes, according to the appearing stimulus and the tone-changing rule, a count might be spoken in a different tone from the one normally spoken in Chinese.

Table 4

Grand Mean Reaction Time (RT) for Experiment 3

Object-switching condition	Feature-switching condition	Item pairs	RT (ms)		Grand *M* RT (ms)	Shift cost (ms)
			M	*SD*		
		Intraobject				
NS	NS	t1t1	1,988	663	1,847	2,954 − 1,847 = 1,107
		c1c1	1,770	572		
		t2t2	1,878	623		
		c2c2	1,751	528		
	S	t1c1	2,866	684	2,954	
		c1t1	3,060	1,110		
		t2c2	2,919	826		
		c2t2	2,971	1,044		
		Interobject				
S	NS	t1t2	2,867	805	2,960	3,367 − 1,847 = 1,520
		t2t1	3,181	812		
		c1c2	3,026	634		
		c2c1	2,764	667		
	S	t1c2	3,494	748	3,367	
		c2t1	3,452	902		
		t2c1	3,011	560		
		c1t2	3,512	1,060		

Note. NS = nonswitch; S = switch.

Being Updated: one count vs. the other count) repeated-measures ANOVA on the RTs yielded significant main effects for the object-switching condition, $F(1, 9) = 83.38$, $p < .001$, and feature-switching condition, $F(1, 9) = 118.5$, $p < .001$, as well as a significant interaction between these two factors, $F(1, 9) = 63.17$, $p < .001$. No significant main effect of operation type or digit being updated was found. A significant two-way interaction was found for Object-Switching Condition × Digit Being Updated, $F(1, 9) = 7.129$, $p = .026$, and Feature-Switching Condition × Digit Being Updated, $F(1, 9) = 6.384$, $p = .032$. Other interactions were all nonsignificant. See Table 4 for further details.

General Discussion

The aim of the present study was to investigate the units of attentional selection within working memory. If attention can select memory objects, once a particular object feature is being attended, all the other features of the same object should also be

Table 5

Concerned Comparisons, Logic, and Judgment

Hypothesis	T1	T2	T3	Prediction	Data	Outcome
1	Op	O	Op,O	T2 > NS	Exp2: T2 > NS	Correct
					Exp3: T2 > NS	Correct
				T1 < T3	Exp2: T1 < T3	Correct
					Exp3: T1 < T3	Correct
				T1 < or = or > T2 (depends on specific tasks)		
2	Op,O	O	Op,O	T1 > T2	Exp2: T1 < T2	Wrong
					Exp3: T1 = T2	No support
				T1 = T3	Exp2: T1 < T3	Wrong
					Exp3: T1 < T3	Wrong

Note. The table shows the relevant prediction and actual experiment results for each hypothesis. Hypothesis 1 is the hypothesis in the present study (i.e., the two features of a digit are selected together); Hypothesis 2 is the contrasted hypothesis (i.e., the two features of a digit are selected independently; thus, they are effectively two different objects). T1 = Type 1 switch; T2 = Type 2 switch; T3 = Type 3 switch; NS = nonswitch; Op = operation switch; O = object switch; Exp2 = Experiment 2; Exp3 = Experiment 3. Whether T1 results in faster RT than T2 does is unpredictable. The comparison result depends on specific tasks. Therefore, there should not be a clear prediction of the comparison result. The > symbol indicates that the switch type to the left of the symbol resulted in faster reaction times than did the switch type to the right of the symbol, the < symbol indicates the opposite, and the = symbol indicates that the reaction times were equivalent.

selected and should therefore all have high accessibility. As a result, faster attentional switching between features of the same object should be observed. Experiment 1 showed that attention switching was significantly faster when different memory items were bound together, rather than separated. This supports the prediction of the object-based theory. Our subsequent Experiments 2 and 3 strengthened this viewpoint (at least in verbal working memory) by excluding the possible "memory load" and "spatial distance" explanations, respectively, of the faster intraobject attention switching.

Because this finding may not be specific to visual memory, it may be difficult to argue that working memory is object based in the way that memory representations of visual objects are stored and selected. Therefore, to be more conservative, we assert that our findings indicate that the unit of attentional selection in working memory is object, either crossmodal or linguistic. The term "units" may be a more accurate description of the substrate of working memory in our experiments.

It should be noted that in the present analyses, we did not compare the RTs of the intraobject switching conditions with the RTs of the object-switch/feature-nonswitch conditions in Experiments 2 and 3. Our line of reasoning is as follows (also see Table 5). Suppose we have Object 1, which has features c1 and a1, and Object 2, which has features c2 and a2. Then, there are three types of successive operations: Type 1, updating c1 then updating a1 (i.e., the intraobject switching condition); Type 2, updating c1 then updating c2 (i.e., the object-switch/feature-nonswitch condition); Type 3, updating c1 then updating a2 (i.e., the interobject switching condition). If our hypothesis that the two features of a digit are selected together is correct, then Type 1 transitions involve only an operation switch, whereas Type 2 transitions involve only an object switch. Generally, object-switching costs and operation-switching costs could both be made greater or smaller with various manipulations of the specific experimental task, which would result in one or the other coming out greater. If our hypothesis is not correct and the two features of digit are selected independently, then the two features are effectively two different objects. In that case, Type 1 transitions involve both an object switch and an operation switch, whereas Type 2 transitions involve only an object switch. Therefore, Type 1 transitions should be slower than Type 2 ones. However, an additional analysis showed that in Experiment 2, the Type 1 switch was significantly faster than the Type 2 switch ($p = .023$), but the difference was not significant ($p = .90$) in Experiment 3. On the basis of these considerations, we did not compare Type 1 and Type 2 transitions. Instead, we focused only on the comparison of Type 1 and Type 3. Specifically, if the two features of a digit are indeed selected together, then Type 1 transitions involve only one switch (operation switch), whereas Type 3 transitions involve two switches (object switch and operation switch). This is why Type 3 was predicted to take longer. However, if the two features are selected independently, thus becoming two distinct objects, Type 1 transitions should also involve two switches (object switch and operation switch) and would therefore be predicted to take as long as Type 3 transitions. All of our observations jointly support the hypothesis that attention selects the two features of a digit together.

Limitations

One unresolved issue of the present study is the differentiation of object *selection* and *updating*. When performing the present tasks, each time a specific stimulus appeared, participants needed to first select the appropriate memory object to attend and then modify (update) the corresponding feature of the selected object according to the presented stimulus. These two stages were not separated in the present study, and we therefore will not be able to tell which stage is indeed facilitated.

In addition to the interobject attention-shift cost, we found also shift cost in intraobject switching (e.g., switching from the numerical value to the spatial position of an integrated digit). This may be due to transitions between different tasks or to the priming contribution (Gehring et al., 2003; Li et al., 2006), because the moving and adding processes are certainly different tasks, and the visual stimuli for these two tasks were also different. Similar factors (i.e., operation type and stimuli) also exist in interobject switching, but their contributions to the RTs in the inter- versus intra- comparison were assumed to cancel each other out. Besides, it is not impossible that two features of the same object have different activation levels. This may be an alternative explanation for the intraobject switching cost. Given that previous researchers (Vecera, Behrmann, & Filapek, 2001) have suggested that visual attention can also be part based (e.g., participants are more accurate in reporting two features from the same part of an object than from different parts of an object), the observed intraobject-shift cost might also reflect a part-based attentional selection in working memory, which still remains open for future study.

Reflection of the Binding Facilitation

On the basis of the results of all three experiments, we ascribe the binding facilitation found in the present study to the contributions of both location-based and object-based attention. The existing literatures on visual attention suggest that attentional selection, at least on the perceptual level, can operate on many different types of representations (e.g., locations, features, and objects), and different types of attentional selection are not mutually exclusive (Duncan, 1984; Vecera & Farah, 1994). For example, Duncan (1984) presented participants with two overlapping objects, one box and one line. The box could be either tall or short, with a gap either on the left or right. The line could be either dotted or dashed and be tilted either clockwise or counterclockwise. Participants were required to report one or two dimensions of the two overlapping objects. When two dimensions were reported, they could be either from one object or from both objects. An intraobject advantage was found: Participants were no worse at reporting the two intraobject dimensions than only one dimension, but they were more accurate at reporting two intraobject dimensions than two interobject dimensions. These results were presented as evidence for object-based attention. However, an alternative interpretation ascribes the findings to a location-based representation grouped according to whether locations are part of one object or another. To test this so-called "grouped location-based" account, Vecera and Farah (1994) modified Duncan's task by presenting the two objects either superimposed or spatially separated. The grouped location-based account predicts Duncan's object effect will be larger in the separate condition, because location-based attention must be moved much farther in the separate condition. The results of their Experiments 1 and 2 failed to verify the grouped location-based account and instead supported the object-based account. However, a different pattern was observed by including a Posner's cuing task (Posner, 1980), in which boxes or lines were cued and participants responded to a target that appeared on either the box or the line. As in their previous experiments, the box and the line could be presented either together or separately. In accordance with the prediction of the location-based account, Vecera and Farah found a larger cue-validity effect (i.e., slower response to invalid cued targets than to valid ones) in the separate condition than in the together condition.

Turning back to the present study, on the one hand, the binding facilitation observed in our three experiments (especially in the third experiment) lends support to the object-based account. On the other hand, the distance effect (i.e., longer RT for an interobject switch when the two digits were in different cells than when they were in the same cell) found in our Experiment 2 strongly confirms the contribution of location-based attention. Therefore, the present results suggest that multiple types of attentional selection may occur in attention in working memory. Following this view, the present study might be able to throw more light on the relationship between outer visual attention and internal memory attention. Recent neuroimaging and neurophysiological studies have revealed common neural substrates underlying the two kinds of attention (Griffin & Nobre, 2003; Nobre et al., 2004; Sylvester et al., 2003). Likewise, the findings in the present study provide further evidence of some similarity between external and internal attention.

One may wish to describe the status of working-memory items in terms of their activation (accessibility) levels. Garavan (1998) proposed that the item on which attention is focused presumably has the highest activation level. This high activation level is maintained until attention is taken away from that item, at which time its activation will return to a baseline level that is required to keep it within working memory. For the *separate* group in our Experiment 1, the two items were independent of each other. On the basis of Garavan's conclusion that people are limited to attending to just one item in working memory at any one time, we presume that the two items under the *separate* strategy always had different activation levels. Once the attention was switched to the unattended item, some effort to increase its activation level would have been necessary. In contrast, participants who used the *binding* strategy could switch attention faster than could those in the *separate* group. Moreover, this binding facilitation persisted even when the two switching conditions were of the same working-memory load. In the activation-level framework, our results thereby suggest that in the *binding* group, the integrated object may always be in the focus of attention, with both properties being maintained at a high activation level. This may be the reason for the observed strategy-dependent switching facilitation.

Broader Theoretical Implications

The present results may also be a part of the full picture of working memory within the framework of the concentric model (Oberauer, 2002). Building on Cowan's embedded-process model (Cowan, 1988, 1995), the concentric model assumes that there are three embedded components in working memory: the activated part of long-term memory, the region of direct access, and the focus of attention. The activated part of long-term memory consists of those memory representations that are activated above baseline and are therefore easy to move into the region of direct access. The region of direct access is a capacity-limited system that holds a small subset of activated representations and their relations. The focus of attention selects one of these elements for processing at a particular time. On the basis of this concentric model, the memory items in all tasks of the present study should be in the region of direct access. Because (a) the focus of attention only selects representations from the region of direct access, (b) our results suggest that the focus of attention actually selects the whole object (unit) integrated by the two properties rather than only the property to be updated, and (c) the two properties that compose an object (unit) can be of either the same (as in Experiment 3) or different (as in Experiments 1 and 2) modalities, it can be inferred that the region of direct access may be a general, supramodal component. This inference is consistent with Oberauer and Göthe's (2006) recent findings. In their Experiment 2, participants had to remember and update a set of digits and a set of spatial locations, both varying in set size. It was found that increasing the set size in one modality remarkably impaired performance in the other modality. The interference of the two sets was presented as evidence that the region of direct access is not modality specific. By using an alternative methodological approach, we found more evidence in support of this viewpoint. Similar issues on other levels or components of the concentric model remain unclear and should be resolved in future studies.

References

Bao, M., Li, Z., Chen, X., & Zhang, D. (2006). Backward inhibition in a task of switching attention within verbal working memory. *Brain Research Bulletin, 69,* 214–221.

Blaser, E., Pylyshyn, Z. W., & Holcombe, A. O. (2000, November 9). Tracking an object through feature space. *Nature, 408,* 196–199.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104,* 163–191.

Cowan, N. (1995). *Attention and memory: An integrated framework.* New York: Oxford University Press.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222.

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

Eriksen, C. W., & Hoffman, J. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Perception & Psychophysics, 14,* 155–160.

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics, 40,* 225–240.

Garavan, H. (1998). Serial attention within working memory. *Memory & Cognition, 26,* 263–276.

Garavan, H., Ross, T. J., Li, S. J., & Stein, E. A. (2000). A parametric manipulation of central executive functioning. *Cerebral Cortex, 10,* 585–592.

Gehring, W. J., Bryck, R. L., Jonides, J., Albin, R. L., & Badre, D. (2003). The mind's eye, looking inward? In search of executive control in internal attention shifting. *Psychophysiology, 40,* 572–585.

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience, 15,* 1176–1194.

Kübler, A., Murphy, K., Kaufman, J., Stein, E. A., & Garavan, H. (2003). Co-ordination within and between verbal and visuospatial working memory: Network modulation and anterior frontal recruitment. *NeuroImage, 20,* 1298–1308.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96,* 101–124.

Li, Z., Bao, M., Chen, X., Zhang, D., Han, S., He, S., & Hu, X. (2006). Attention shift in human verbal working memory: Priming contribution and dynamic brain activation. *Brain Research, 1078,* 131–142.

Li, Z., Sun, X., Wang, Z., Zhang, X., Zhang, D., He, S., & Hu, X. (2004). Behavioral and functional MRI study of attention shift in human verbal working memory. *NeuroImage, 21,* 181–191.

Maunsell, J. H., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences, 29,* 317–322.

Nobre, A. C., Coull, J. T., Maquet, P., Frith, C. D., Vandenberghe, R., & Mesulam, M. M. (2004). Orienting attention to locations in perceptual versus mental representations. *Journal of Cognitive Neuroscience, 16,* 363–373.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 411–421.

Oberauer, K. (2003). Selective attention to elements in working memory. *Quarterly Journal of Experimental Psychology, 50,* 257–269.

Oberauer, K. (2005). Control of the contents of working memory: A comparison of two paradigms and two age groups. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 714–728.

Oberauer, K., & Göthe, K. (2006). Dual-task effects in working memory: Interference between two processing tasks, between two memory demands, and between storage and processing. *European Journal of Cognitive Psychology, 18,* 493–519.

O'Craven, K., Downing, P., & Kanwisher, N. (1999, October 7). fMRI evidence for objects as the units of attentional selection. *Nature, 401,* 584–587.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Rodríguez, V., Valdés-Sosa, M., & Freiwald, W. (2002). Dividing attention between form and motion during transparent surface perception. *Cognitive Brain Research, 13,* 187–193.

Roelfsema, P. R., Lamme, V. A., & Spekreijse, H. (1998, September 24). Object-based attention in the primary visual cortex of the macaque monkey. *Nature, 395,* 376–381.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience, 5,* 631–632.

Schoenfeld, M. A., Tempelmann, C., Martinezm, A., Hopf, J. M., Sattler, C., Heinze, H. J., & Hillyard, S. A. (2003). Dynamics of feature binding during object-selective attention. *Proceedings of the National Academy of Sciences of the United States of America, 100,* 11806–11811.

Sylvester, C. Y., Wager, T. D., Lacey, S. C., Hernandez, L., Nichols, T. E., Smith, E. E., & Jonides, J. (2003). Switching attention and resolving interference: fMRI measures of executive functions. *Neuropsychologia, 41,* 357–370.

Treue, S., & Martinez Trujillo, J. C. (1999, June 10). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature, 399,* 575–579.

Vecera, S. P. (2000). Toward a biased competition account of object-based segregation and attention. *Brain and Mind, 1,* 353–384.

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., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General, 123,* 146–160.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24,* 295–340.

Received December 15, 2006
Revision received June 9, 2007
Accepted June 11, 2007 ■