Bindings in working memory: The role of object-based attention

Zaifeng Gao¹ · Fan Wu¹ · Fangfang Qiu¹ · Kaifeng He¹ · Yue Yang¹ · Mowei Shen¹

Published online: 6 December 2016
© The Psychonomic Society, Inc. 2016

Abstract Over the past decade, it has been debated whether retaining bindings in working memory (WM) requires more attention than retaining constituent features, focusing on domain-general attention and space-based attention. Recently, we proposed that retaining bindings in WM needs more *object-based attention* than retaining constituent features (Shen, Huang, & Gao, 2015, *Journal of Experimental Psychology: Human Perception and Performance*, doi:10.1037/xhp0000018). However, only unitized visual bindings were examined; to establish the role of object-based attention in retaining bindings in WM, more empirical evidence is required. We tested 4 new bindings that had been suggested requiring no more attention than the constituent features in the WM maintenance phase: The two constituent features of binding were stored in different WM modules (cross-module binding, Experiment 1), from auditory and visual modalities (cross-modal binding, Experiment 2), or temporally (cross-time binding, Experiments 3) or spatially (cross-space binding, Experiments 4–6) separated. In the critical condition, we added a secondary object feature-report task during the delay interval of the change-detection task, such that the secondary task competed for object-based attention with the to-be-memorized stimuli. If more object-based attention is required for retaining bindings than for retaining constituent features, the secondary task should impair the binding performance to a larger degree relative to the performance of

constituent features. Indeed, Experiments 1–6 consistently revealed a significantly larger impairment for bindings than for the constituent features, suggesting that object-based attention plays a pivotal role in retaining bindings in WM.

Keywords Attention and memory · Object-based attention · Working memory

To have a coherent experience of the external world, we have to bind different sources of information across features, modalities, time, and space. This is referred to as the “binding problem” (Treisman, 1996, 1998). Feature binding is one of the most extensively explored topics in psychological studies of perception (Wolfe & Robertson, 2012); for instance, the well-known feature-integration theory (Treisman & Gelade, 1980) addresses the binding problem at perception. Feature binding is also an integral part of working memory (WM; Zhang, Johnson, Woodman, & Luck, 2012; see Allen, 2015, for a state-of-the-art review), which is a limited-capacity system for temporarily maintaining and manipulating information used in higher-level cognitive activities (Baddeley, 2007b; Baddeley & Hitch, 1974; for reviews, see Baddeley, 2012; Cowan, 2014; Logie & Cowan, 2015; van den Berg, Awh, & Ma, 2014). Although it has been well accepted that binding features into a unified percept requires more attention (e.g., Humphreys, 2001; Treisman & Gelade, 1980; Treisman & Schmidt, 1982), the role of attention in retaining bindings in WM remains in controversy (see Allen, 2015, for a state-of-the-art review).

In this decade a debate has been raised concerning whether retaining bindings in WM is more resource demanding than retaining constituent features. Wheeler and Treisman (2002) provided the first evidence implying that retaining bindings in WM required more space-based attention than retaining

Electronic supplementary material The online version of this article (doi:10.3758/s13414-016-1227-z) contains supplementary material, which is available to authorized users.

✉ Mowei Shen
mwshen@zju.edu.cn

¹ Department of Psychology and Behavioral Sciences, Xixi Campus, Zhejiang University, Hang Zhou, People's Republic of China 310007

constituent features. They found that when the test array contained the same number of objects as the memory array (whole-probe method), accuracy for the binding condition was significantly impaired relative to the single-feature condition; however, such a binding-specific deficit relative to single-feature was absent when the test array contained only one object (partial-probe method). It was interpreted that sustained space-based attention was required to retain bindings in WM; if attention employed in a bound representation was re-allocated to other objects, which occurred in the whole-probe but not the partial-probe method, the bound representation would collapse. Although there were studies revealed indirect evidence supporting this resource-demanding view of binding (e.g., Fougner, Asplund, & Marois, 2010; Fougner & Marois, 2011), many studies did not support it. First of all, the binding-specific deficit revealed in Wheeler and Treisman (2002) seems to be unstable. For instance, using a similar design as Wheeler and Treisman (2002), the binding-specific deficit was not successfully reestablished (Johnson, Hollingworth, & Luck, 2008). Second, researchers directly examined the role of space-based attention in retaining bindings by manipulating the amount of available space-based attention during the maintenance phase of WM. In contrast to the prediction of the resource-demanding view, there was no larger influence on the binding performance relative to the constituent features (Delvenne, Cleeremans, & Laloyaux, 2010; Gajewski & Brockmole, 2006; Johnson et al., 2008; Langerock, Vergauwe, & Barrouillet, 2014; Shen, Huang, & Gao, 2015; Yeh, Yang, & Chiu, 2005; Zokaei, Heider, & Husain, 2014). For instance, Johnson et al. (2008) inserted a secondary visual search task, which requires sequential shifts of space-based attention, during the maintenance phase of a WM task. They found a similar degree of disruption between binding and constituent features. Therefore, retaining bindings in WM may be a relatively automatic process that does not require more attention than retaining constituent features, which is implicitly assumed in earlier studies of WM (e.g., Cowan, Blume, & Saults, 2013; Luck & Vogel, 1997; Luria & Vogel, 2011; Vogel, Woodman, & Luck, 2001).

Moreover, the automatic view of retaining bindings in WM received support by a series of recent studies exploring the underlying mechanisms of episodic buffer (Baddeley, 2000), which is a newly added storage buffer to the multiple-component model of WM of Baddeley (Baddeley, 1986; Baddeley & Hitch, 1974). The initial multiple-component model posits that the WM system consists of a central executive and two slave subsystems—specifically, the phonological loop and the visuospatial sketchpad. The central executive is attentionally controlling ongoing processes, while the phonological loop and the visuospatial sketchpad store acoustical and visual/spatial information, respectively. After decades of extensive studies, a few limitations, however, have emerged to the initial model (cf. Baddeley, 2000). One key limitation is

concerned with the question of how information from the phonological and visuospatial subsystems is bound given their distinct information coding mechanisms. The episodic buffer was proposed to solve these limitations (Baddeley, 2000). With the help of attention from the central executive (domain-general attention or central attention; cf. Fougner, 2008), the episodic buffer is suggested to bind information from the phonological and visuospatial subsystems and from long-term memory into a coherent episodic representation.

Baddeley and colleagues initiated the exploration of episodic buffer by focusing on the role of domain-general attention in encoding and retaining bindings in WM (Allen, Baddeley, & Hitch, 2006; for reviews, see Baddeley, Allen, & Hitch, 2010, 2011). They adopted a dual-task paradigm in which a verbal backward counting task was added to consume the domain-general attention during the memorization of shapes, colors, and shape-color bindings (Allen et al., 2006). The logic was that any condition that was particularly reliant on domain-general attention should exhibit larger disruption. A significantly substantial decrease in WM performance was observed in all three conditions, indicating a role for domain-general attention in WM (see also Allen, Hitch, Mate, & Baddeley, 2012; Brown & Brockmole, 2010; Cowan, Naveh-Benjamin, Kilb, & Saults, 2006; Morey & Bieler, 2013; Vergauwe, Langerock, & Barrouillet, 2014). However, the decrease did not differ significantly between the binding and the constituent single-feature conditions, suggesting that processing bindings in WM does not consume more domain-general resource than processing single features. This finding proved to be robust, considering that it has been replicated under several distinct binding conditions: Cross-space binding (the constituent visual features were spatially separated; Karlsen, Allen, Baddeley, & Hitch, 2010), cross-time binding (the constituent visual features were temporally separated; Karlsen et al., 2010), cross-modal binding (one feature from the verbal and one from the visual channel; Allen, Hitch, & Baddeley, 2009), verbal-verbal binding (all information is processed in the phonological loop; Baddeley, Hitch, & Allen, 2009), cross-module binding (the two constituent features were processed by two WM modules; Langerock et al., 2014),¹ and static color-dynamic motion binding (Ding et al., 2015). The absence of differential disruption was also revealed in a study directly manipulating verbal attention while retaining verbal-spatial bindings (Langerock et al., 2014). These findings prompted Baddeley to conclude that the episodic buffer was not responsible for building and maintaining bindings; instead, it passively received the bindings that had been formed elsewhere before entering into the episodic buffer (Baddeley, 2012; Baddeley et al., 2010, 2011).

¹ Other researchers termed it cross-domain binding (e.g., Langerock et al., 2014). Because readers may confuse the cross-domain binding with the cross-modal binding, we used cross-module binding instead.

Although the aforementioned studies provide important insights for the binding processing, these studies seemed to ignore a critical hallmark of binding: The binding representation is by nature a coherent and integrated unit (Treisman, 1996, 1998; Treisman & Gelade, 1980). Previous perceptual studies have suggested that *object-based attention* plays a pivotal role in forming and keeping such integrated units (e.g., Chen, 2012; Cohen & Tong, 2013; Emmanouil & Magen, 2014; Huang, 2010; Kahneman, Treisman, & Gibbs, 1992; Schoenfeld, Hopf, Merkel, Heinze, & Hillyard, 2014; Schoenfeld et al., 2003). Accumulative evidence suggests that WM and perception are very similar in several key processing aspects. For instance, spatial WM recruits space-based attention to rehearse the location information (e.g., Awh & Jonides, 2001; Oh & Kim, 2004; Woodman & Luck, 2004); the same visual cortexes are employed to process the visual features (e.g., orientation) in both perception and WM (e.g., Ester, Serences, & Awh, 2009; Harrison & Tong, 2009); and WM and perception adopt similar manners in encoding and manipulating the incoming information (e.g., T. Gao et al., 2011; Z. Gao, Bentin, & Shen, 2015; Z. Gao et al., 2010; Kiyonaga & Egner, 2013, 2014; Lu et al., 2016; Mayer et al., 2007; Shen et al., 2015). Therefore, it is possible that more resource is required in retaining bindings in WM, but is limited to *object-based attention*.

Two recent studies provided preliminary evidence supporting this object-based attention alternative. Particularly, Fougne and Marois (2009) inserted a multiple object tracking (MOT) task during the WM maintenance phase and found a significant selective binding impairment relative to single features. Although this finding was initially interpreted as that the MOT task consumed space-based attention, as aforementioned, later studies consistently revealed that the same amount of space-based attention was needed for retaining both bindings and constituent features in WM (e.g., Johnson et al., 2008; Langerock et al., 2014; Shen, Huang, & Gao, 2015). Critically, Hollingworth and Maxcey-Richard recently suggested that the findings of Fougne and Marois (2009) can be explained in terms of additional storage of object-location bindings in WM (cf. footnote 3 of Hollingworth & Maxcey-Richard, 2013). Recently, we directly examined the object-based attention hypothesis by adding a secondary perceptual task (e.g., Duncan's object feature-report task, mental rotation task, transparent motion task) into the maintenance phase of a WM task (Shen et al., 2015). In line with the prediction of the object-based attention hypothesis, we found a greater impairment for bindings than for constituent features. Moreover, this selective binding impairment was

consistently revealed for color-shape binding, color-location binding, and color-orientation binding, but vanished when the secondary task consumed space-based attention (visual search task in particular), replicating previous findings (Johnson et al., 2008; Langerock et al., 2014).

To establish the role of object-based attention in retaining bindings in WM, more empirical evidence is needed. For instance, only unitized visual bindings were tested in both Fougne and Marois (2009) and Shen et al. (2015): The constituent elements of bindings were visual features with 100 % spatial contiguity (e.g., the color red and the shape square share the same location in a colored square stimulus; cf. Karlsen et al., 2010). However, there is evidence implying that the processes involved in binding vary depending on whether features are unitized (e.g., Ecker, Maybery, & Zimmer, 2013; Wheeler & Treisman, 2002; Xu, 2002, 2006). For example, Ecker et al. (2013) revealed that automatic object-based encoding only takes place for unitized bindings. The object-based attention hypothesis of binding hence may be only limited to the unitized binding. Corroborating this possibility, most of the extant WM studies assumed that the unitized visual bindings were stored in visuospatial sketchpad (e.g., Johnson et al., 2008; Luck & Vogel, 1997; Shen, Tang, Wu, Shui, & Gao, 2013; Wheeler & Treisman, 2002; Woodman & Vecera, 2011; Xu, 2002; for a review, see Zhang et al., 2012). Therefore, to figure out whether object-based attention is crucial for retaining bindings in WM, it is of necessity to establish the role of object-based attention in retaining bindings beyond the unitized visual binding. If a positive answer is obtained for the role of object-based attention, it will imply a general rule for dynamically retaining bindings in WM in general, and sheds important light on the mechanisms of episodic buffer in particular. Otherwise, it will cast a critical constraint on the role of object-based attention in retaining bindings, which is particularly important in reframing future WM models.

To this end, we explored the role of object-based attention in retaining bindings beyond the unitized visual bindings in WM, by adopting a dual-task paradigm. We added an object feature-report task into the maintenance phase of WM. Critically, we manipulated four types of binding: The two constituent features of binding were (1) stored in phonological loop and visuospatial sketchpad (cross-module binding; Experiment 1), (2) displayed from auditory and visual modalities (cross-modal binding; Experiment 2), (3) temporarily separated (cross-time binding; Experiment 3), and (4) spatially separated (cross-space binding; Experiments 4–6). These bindings are distinct from the previously tested unitized visual bindings (Fougne & Marois, 2009; Shen et al., 2015): The first

two bindings could not be handled by visuospatial sketchpad at all, whereas the last two bindings were not unitized binding at all.² They previously had been taken as critical examples in examining the mechanisms of episodic buffer (e.g., Allen et al., 2009; Karlsen et al., 2010; Langerock et al., 2014; Wang, Allen, Lee, & Hsieh, 2015). To the best of our knowledge, no study so far has investigated the role of object-based attention under these bindings. For the object-based attention task, in line with previous studies (Matsukura & Vecera, 2009; Shen et al., 2015) we adopted the standard task to demonstrate the existence of object-based attention: Duncan's object feature-report task (Duncan, 1984). If retaining bindings in WM requires more object-based attention than the constituent features, then a selective binding impairment should be observed regardless of binding types.

Experiment 1: Object-based attention in cross-module binding

In Experiment 1, we examined the role of object-based attention in retaining cross-module bindings, focusing on the bindings between English letters and spatial locations (verbal-spatial binding), which are stored in the phonological loop and visuospatial sketchpad, respectively. Previous work revealed that human beings could retain the verbal-spatial binding in WM (e.g., Allen, Havelka, Falcon, Evans, & Darling, 2015; Morey, 2009). Moreover, Elsley and Parmentier (2009) revealed that retaining verbal-spatial bindings in WM requires domain-general attention. Langerock et al. (2014) further demonstrated that retaining verbal-spatial bindings did not require more domain-general attention than retaining the constituent features; and neither space-based attention nor verbal attention played a more important role than the domain-general attention.

² It is of note that the cross-time binding and cross-space binding used in current Experiments 3–6 were different from the bindings in the object recognition literature (e.g., Treisman, 1996). It is possible that the two types of binding are not bounded in WM but instead are about paired associate learning. Although currently we cannot refute this possibility, we had to point out that these bindings were selected from previous studies exploring the attentional mechanism of retaining bindings in WM (episodic buffer in particular; e.g., Allen et al., 2009; Karlsen et al., 2010). We consider that it is meaningful by taking them as the target bindings because this setting enables us to have direct comparisons with previous studies and to further comprehensively understand the WM mechanisms. Additionally, according to the prediction of episodic buffer (although it is in controversial), cross-space binding and cross-time binding could take place in WM. Indeed, a few WM studies provide evidence supporting this possibility (e.g., Allen et al., 2009; Karlsen et al., 2010; Vergauwe & Cowan, 2015). For instance, Allen et al. (2009) revealed a non-significant difference on accuracy between cross-space binding and unitize visual binding; Vergauwe and Cowan (2015) used the cross-space binding as the probe and found that even participants were discouraged to encode the unitized visual bindings, the participants could still form bindings and finish the WM task.

In Experiment 1, we presented the verbal-spatial binding in a way of Elsley and Parmentier (2009): Each binding comprised a letter and a location on the screen (see the memory array in Fig. 1).

Method

Participants

According to previous studies (Shen et al., 2015; Simmons, Nelson, & Simonsohn, 2011), we recruited 24 valid participants (15 males). Five participants were replaced with new participants because of chance-level performance in the secondary task. All participants were Zhejiang University undergraduates who signed consent forms, had normal color vision, and normal or corrected-to-normal visual acuity. All the participants had practiced/learned English for more than 7 years and were very familiar with these English letters. The participants were on average 20.07 ($SD = 1.75$) years old. The study conformed to Standard 8 of the American Psychological Association's Ethical Principles of Psychologist and Code of Conduct, and was approved by the Research Ethics Board of Zhejiang University.

Apparatus and stimuli

Stimuli were presented against a gray (128, 128, 128; RGB) background on a 19-inch CRT monitor with a resolution of 1024×768 pixels at 100-Hz refresh rate. Participants were seated in a dark room, approximately 60 cm from the screen.

The cross-module bindings were presented in the way of Elsley and Parmentier (2009): Each binding comprised an uppercase letter and a location (marked by a $1.87^\circ \times 1.87^\circ$ black frame) on the screen (see Fig. 1). The letters were randomly selected from a set of eight consonants (D, F, H, J, N, Q, R, T; Arial font, 48 pt.), whose upper- and lowercase physical appearances can be distinguished easily. The locations were randomly selected from a set of eight locations placed equidistantly along an imagery circle (3.64° from the center of the screen). As in Elsley and Parmentier (2009), in each trial, four distinct uppercase consonants were presented simultaneously in distinct locations.³ For the probed stimuli, a lowercase consonant at the screen center, a black frame at one of the eight possible locations, and a lower-case consonant located in one of the eight locations was used in the letter, location, and letter-location binding condition, respectively (see the probe in Fig. 1). It is of note that a lowercase English letter was adopted in the probe, so that participants

³ The memory load in Experiment 1 was four bindings, whereas in Experiments 2–5 it was three bindings. We used a higher load in Experiment 1 because our pilot study revealed that participants' performance was at ceiling when three stimuli were memorized.

Fig. 1 A schematic illustration of the experimental procedure in Experiment 1. Note in the real experiments we used Chinese characters to describe objects' attributes (e.g., “高” for high, “低” for low) for the two response windows of the object task

could not make a judgment solely based on the physical similarity of the letters between memory array and probe.

For Duncan's object feature-report task (Duncan, 1984), a box and a line were presented each time (see Fig. 1 for an illustration). Each object had two task-relevant features: The box (1.02° width) was either short (1.02°) or tall (1.43°) and had a gap (0.31° width) centered on either the left or the right side of the box. The line (3.19° long) was either dashed or dotted and tilted 8° to either the left or the right. The backward mask was 2.11° in width and 2.86° in height.

Design and procedure

The experiment adopted a 2 (task-load: with object-task vs. no object-task) $\times$ 3 (memory-condition: Feature 1 only, Feature 2 only, and binding between Features 1 and 2) within-subjects design. According to whether a secondary task was involved, the whole experiment was divided into two sessions, with their order counterbalanced between participants. Following previous studies (e.g., Allen et al., 2006; Ueno, Allen, Baddeley, Hitch, & Saito, 2011; Shen et al., 2015), the same stimuli were used in the memory array for the binding and constituent feature conditions. To ensure that the participants only memorized the task-relevant dimension, the three memorized conditions were presented in blocks (cf. Allen et al., 2006). In particular, each task-load session consisted of three blocks: letter-only, location-only, and letter-location binding, with their order fully counterbalanced using a Latin square. Each block had 32 trials, leading to a total of 192 trials. For the “with object-task” session, the two to-be-reported features could be derived equally from the same object (within-object; 16 trials) or from different objects (between-objects; 16 trials), with their order fully counterbalanced between participants. Before the formal experiment, participants first completed practice trials containing single memory tasks and single object feature-report tasks to familiarize themselves with

each task; then participants completed at least 12 practice trials of the dual tasks. The whole experiment lasted about 60 minutes.

Participants were first informed whether the secondary object-based task was conducted and what features would be reported (e.g., gap direction of box and texture of line, in Chinese) at the beginning of each block (cf. Matsukura & Vecera, 2009; Shen et al., 2015). The participants reported all possible pairings of object features. Each trial began with a 500-ms fixation (a black cross at the screen center, $0.85^\circ \times 0.85^\circ$) to inform participants of the start of a trial. Participants were required to fixate the screen center. After a 250-ms blank interval, the memory array was presented. The to-be-memorized items in Experiment 1 were presented simultaneously for 1,000 ms (see Fig. 1). After a 500-ms interval, the secondary object-task began by presenting the superimposed box-line stimulus for 200 ms, which was immediately backward masked for 200 ms (cf. Matsukura & Vecera, 2009). Then, two question screens were presented sequentially informing participants which feature dimension should be judged, each lasting a maximum of 2,000 ms. Once the participants made a response (by pressing the corresponding key on the keyboard), the procedure immediately went to the next step. In the no object-task condition, participants were instructed to ignore the questions by pressing the spacebar on the keyboard such that potential response factors were largely balanced between with-object and no-object conditions. Finally, 500 ms after the secondary task, a probe for letter was presented. The participants were explicitly instructed to determine whether the probed information had been present in the memory display, by pressing “J” for presence and “F” for absence. In 50 % of trials, the probed information was absent from the memory array, that is, a new feature or binding that had not appeared in the memory array was presented. It is of note that in the binding change condition, the new binding was formed by combining two features from two randomly

selected objects in the memory array (see probe in Fig. 1). For both the memory task and the secondary feature-report task, response accuracy was emphasized, and a response needed to be made within 2,000 ms. The intertrial blank interval was randomly selected from 1,000 to 1,500 ms.

Analysis

Only trials with correct responses on the secondary task were further analyzed. The change detection performance of the memory task was reported as corrected recognition (hits - false alarms; Allen et al., 2006; Allen et al., 2012; Snodgrass & Corwin, 1988). For the analysis of the raw data, please see Supporting Material 1. A two-way repeated-measures analysis of variance (ANOVA) was conducted on the corrected recognition, with Task-load (with object-task vs. no object-task) and Memory-condition (Feature 1 only, Feature 2 only, and binding) as the within-subjects factors.⁴ For significant Task-load $\times$ Memory-condition interactions, we first calculated the differences between “no object-task” and “with object-task” conditions. We then ran planned contrasts by separately comparing the differences under the binding condition with the differences under the two single-feature conditions.

In addition, before the performance of the secondary tasks was beyond our interest and for the sake of clarity, we provide the analyses of the secondary tasks (both accuracy and RT) as supporting material (see Supporting Material 2 for details).

Results

The overall accuracy of the secondary task was 97.58 %. Figure 2 shows the corrected recognition for the letter-location binding stimuli. The two-way ANOVA on the corrected recognition revealed a significant main effect of Task-load, $F(1, 23) = 19.69$, $p < .001$, $\eta_p^2 = .46$, suggesting that performance was significantly worse under the “with object-task” condition ($M = .76$) than under the “no object-task” condition ($M = .85$). The main effect of Memory-condition also reached significance, $F(2, 46) = 18.28$, $p < .001$, $\eta_p^2 = .44$. Further analysis (Bonferroni-corrected) revealed that the performance was significantly lower for binding ($M = .70$) than for letter ($M = .89$, $p < .001$) and location ($M = .83$, $p = .002$) conditions, while no difference was found between the

⁴ In this study we pooled the change-detection trials that corresponding to within-object (16 trials per memory condition) and between-objects trials (16 trials per memory condition) from the feature-report task. A critical consideration for doing this is that Shen et al. (2015) found that the largest disruption was observed between “no object-task” and “with object-task (the average of within-object and between-objects)” conditions, although there was significant disruption between “within-object” and “between-objects” conditions (cf. Experiment 5). Moreover, by not further splitting the “with object-task” condition into two separate conditions, we could significantly reduce the duration of experiment and make the analysis more concise.

Fig. 2 Results for Experiment 1. The cost is calculated as the difference between “no object-task” and “with object-task” conditions (no object-task minus with object-task). Error bars stand for standard errors

latter two ($p > .05$). Critically, there was a significant Task-load $\times$ Memory-condition interaction, $F(2, 46) = 12.65$, $p < .001$, $\eta_p^2 = .36$.

Planned contrasts deconstructing the interaction showed that the disruption caused by the secondary object-task was significantly larger under the letter-location binding condition ($M = .19$) than under the letter ($M = -.01$), $t(23) = 4.93$, $p < .001$, Cohen’s $d = 1.01$, and location ($M = .09$), $t(23) = 2.27$, $p = .033$, Cohen’s $d = .47$, conditions.

Discussion

Experiment 1 added a Duncan’s object feature-report task into the maintenance phase of verbal-spatial bindings. In line with Elsley and Parmentier (2009) and Langerock et al. (2014), the inserted task incurred significant disruption to the binding performance. However, different from Langerock et al. (2014), we found a selective binding impairment: Duncan’s object feature-report task exerted a larger impairment on binding than on the constituent features. This finding suggests that retaining cross-module bindings in WM requires more object-based attention than the constituent features, supporting the object-based attention hypothesis.

While Experiment 1 examined binding across different WM modules, the encoded information of bindings was still from the same visual modality. In Experiment 2, we tested the role of object-based attention when binding took place across modalities (e.g., between vision and audition).

Experiment 2: Object-based attention in cross-modal Binding

In Experiment 2, we examined the role of object-based attention in retaining cross-modal bindings, focusing on the

bindings between visual and auditory modalities (visual-auditory binding). We replaced the letter-location binding in Experiment 1 with a type of visual-auditory binding used in Allen et al. (2009): Each binding comprised an auditory color name and a visual shape (see Fig. 4). Using this setting, Allen et al. (2009) demonstrated that domain-general attention or space-based attention did not play a critical role in retaining cross-modal bindings in WM.

Method

There were 24 valid participants (17 males). Seven participants were replaced with new participants because of chance-level performance in the secondary task or extremely low-level (three standard deviations away from the mean) performance in the memory task. All participants were Zhejiang University undergraduates who signed consent forms, had normal color vision, normal or corrected-to-normal visual acuity, and normal hearing. The participants were 19.34 ($SD = 1.61$) years old on average.

The cross-modal bindings were presented in the way of Allen et al. (2009; see also Wang et al., 2015 for a similar setting): Each binding comprised an auditory color name and a visual shape presented simultaneously, with the shape always being at the screen center. The to-be-memorized color and shape values were randomly selected from a pool of six distinct values of each type (see Fig. 3). A female Chinese speaker digitally recorded the name of each tested color. In line with Allen et al. (2009), three to-be-memorized stimuli were presented sequentially; colored blobs, unfilled three-point black outlines, and colored shapes were used as the probe in the auditory color, visual shape, and color-shape binding conditions, respectively (see Fig. 4).

The procedure in Experiment 2 was identical to Experiment 1 except for the following three aspects: (1) Because the color names in Experiment 2 were delivered in a verbal format, the participants may be biased to verbally encode the visual shapes so that the bindings were eventually constructed within the phonological loop. To prevent this strategy from taking place, Experiment 2 required the participants to verbally repeat three digits “1 2 3” loudly throughout a trial.⁵ Moreover, participants were required to fixate the screen center while trying to avoid verbally encoding the visual stimuli as much as possible. (2) Three auditory-visual

bindings were presented sequentially and each lasted 1,000 ms, with an interstimulus interval of 250 ms. Three memory conditions were tested: Auditory-color, visual-shape, and shape-color binding. (3) A probe was always presented at a position 1.6° lower than the screen center, and participants made a response (cf. Allen et al., 2006).

The whole experiment lasted about 70 minutes. The other aspects were the same as Experiment 1.

Results

The overall accuracy of the secondary task was 97.67 %. Figure 5 shows the corrected recognition for the visual-auditory binding stimuli.⁶ The two-way ANOVA on the corrected recognition revealed a significant main effect of Task-load, $F(1, 23) = 27.23$, $p < .001$, $\eta_p^2 = .54$, suggesting that performance was significantly worse under the “with object-task” condition ($M = .55$) than under the “no object-task” condition ($M = .65$). The main effect of Memory-condition also reached significance, $F(2, 46) = 166.18$, $p < .001$, $\eta_p^2 = .88$. Further analysis (Bonferroni-corrected) revealed that the performance was significantly lower for binding ($M = .27$) than for auditory color ($M = .84$, $p < .001$) and visual shape ($M = .70$, $p < .001$) conditions, and the performance for auditory color was significantly higher than that for visual shape ($p < .001$). Critically, there was a significant Task-load $\times$ Memory-condition interaction, $F(2, 46) = 10.36$, $p < .001$, $\eta_p^2 = .31$.

Planned contrasts deconstructing the interaction showed that the disruption caused by the object-task was significantly larger under the color-shape binding condition ($M = .21$) than under the auditory color ($M = -.01$), $t(23) = 4.52$, $p < .001$, Cohen's $d = .92$, and visual shape ($M = .08$), $t(23) = 2.66$, $p = .014$, Cohen's $d = .54$, conditions.

Discussion

Experiment 2 added a Duncan's object feature-report task into the maintenance phase of visual-auditory bindings. In line with Experiment 1 but different from Allen et al. (2009), we found a selective binding impairment: Duncan's object feature-report task exerted a larger impairment on binding than on the constituent features. This finding suggests that

⁵ In Experiments 1 and 3–5, we did not add an articulatory suppression task. In Experiment 1, the participants had to verbally encode the English letter considering that it was stored in the phonological loop. However, it was difficult for the participants to verbally encode the location information. Our pilot study with Experiment 3 found that most of the participants did not verbally encode the memorized stimuli because we emphasized that they need to avoid verbally encoding the stimuli as much as possible. To this end, we did not add an articulatory suppression task. However, we admit that to have a stricter control, future studies may consider adding an articulatory suppression task.

⁶ In Experiment 2, we did not differentiate the accuracy at different serial positions during the data collection. Considering that the serial positions significantly affect the working memory performance (e.g., Allen et al., 2006, 2014; Hu, Hitch, Baddeley, Zhang, & Allen, 2014), it would be interesting to explore how the selective binding impairment is affected by serial positions in future studies. However, because a similar conclusion was reached in terms of the role of attention on binding regardless of the serial positions (e.g., Allen et al., 2006, 2014; Hu et al., 2014), and the results of Experiment 2 are in line with those in Experiments 1, 3–5, we currently predict that a similar conclusion may be reached even considering the serial positions.

Fig. 3 Colored blobs and unfilled shapes used in Experiments 2–5. The colors were purple (255, 0, 255; RGB), yellow (255, 255, 0), blue (0, 0, 255), green (0, 255, 0), red (255, 0, 0), and white (255, 255, 255), from

left to right. The shapes ($1.52^\circ \times 1.52^\circ$ of visual angle on average) consisted of triangle, star, chevron, circle, cross, and diamond, from left to right. (Color figure online)

retaining cross-modal bindings in WM also requires more object-based attention than the constituent features.

It is of note that the visual-auditory binding also belongs to cross-module binding considering that the visual and auditory information is handled by visuospatial sketchpad and phonological loop, respectively. Therefore, Experiment 2 adds further evidence suggesting that cross-module binding needs more object-based attention (cf. Experiment 1).

The results of Experiments 1 and 2 collectively suggest that the object-based attention hypothesis of bindings is not constrained to unitized bindings which potentially can be handled by the visuospatial sketchpad entirely. In the next three experiments, we further examined whether the object-based attention hypothesis is amenable to the visual bindings without high spatial contiguity, by temporally or spatially separated the constituent features of binding.

Experiment 3: Object-based attention in cross-time binding

In Experiment 3, we examined the role of object-based attention in retaining cross-time bindings. We achieved this by

presenting the stimuli in a way of Karlsen et al. (2010): The constituent elements of a binding were presented sequentially, with a backward mask between (see Fig. 6). Karlsen et al. (2010) suggested that domain-general attention did not play a pivotal role in retaining cross-time bindings in WM.

Method

There were 24 valid participants (12 males). Eight participants were replaced with new participants because of chance-level performance in the secondary task or extremely low-level (three standard deviations away from the mean) performance in the memory task. The participants were 21.84 ($SD = 1.89$) years old on average.

The cross-time binding in Experiment 3 were presented in a way of Karlsen et al. (2010): Each binding comprised a color and a shape, which were presented sequentially with a 500-ms backward mask between (see Fig. 6). Each trial contained three bindings. We first presented one feature dimension (e.g., shape) for 1,000 ms, followed by a 500-ms backward mask (10.21° width $\times$ 3.41° height; see Fig. 6), and then the other feature dimension (e.g., color) for 1,000 ms. Moreover, the two feature elements of a binding were presented in the same location, with the participants required to retain the

Fig. 4 An illustration of memory arrays and probes used in Experiment 2. (Color figure online)

Fig. 5 Results for Experiment 2. The cost is calculated as the difference between “no object-task” and “with object-task” conditions (no object-task minus with object-task). Error bars stand for standard errors

corresponding binding. In 50 % of trials the three colors were presented first, and in the other 50 % of trials the three shapes were presented first. The to-be-memorized colors and shapes were randomly selected from a pool of six distinct values of each type (see Fig. 3). For the probed stimuli, colored blobs, unfilled three-point black outlines, and colored shapes were used in the visual color, visual shape, and color-shape binding conditions, respectively (see Fig. 6).

The procedure in Experiment 3 was identical to Experiment 1 except for the following three aspects: (1) The memory array was replaced by three cross-time bindings (see Fig. 6). Therefore, three new memory conditions were tested: Visual-color, visual-shape, and color-shape binding. (2) Participants were required to fixate the screen center while trying to avoid verbally encoding the visual stimuli as much as possible. (3) The probe was always presented at a position 1.6° lower than the screen center and participants made a response.

Fig. 6 An illustration of memory arrays and probes used in Experiment 3. Note in the experiment both the colors and shapes can be presented first. (Color figure online)

The whole experiment lasted about 60 minutes. The other aspects were the same as Experiment 1.

Results

The overall accuracy of the secondary task was 96.92 %. Figure 7 shows the corrected recognition for the color-shape binding stimuli. The two-way ANOVA on the corrected recognition revealed a significant main effect of Task-load, $F(1, 23) = 32.27, p < .001, \eta_p^2 = .58$, suggesting that performance was significantly worse under the “with object-task” condition ($M = .63$) than under the “no object-task” condition ($M = .76$). The main effect of Memory-condition also reached significance, $F(2, 46) = 72.91, p < .001, \eta_p^2 = .76$. Further analysis (Bonferroni-corrected) revealed that the performance was significantly lower for binding ($M = .42$) than for visual color ($M = .89, p < .001$) and visual shape ($M = .78, p < .001$) conditions, and the performance for visual color was significantly higher than that for visual shape ($p = .003$). Critically, the Memory-condition factor was further modulated by the secondary task, revealing a significant Task-load $\times$ Memory-condition interaction, $F(2, 46) = 7.79, p = .001, \eta_p^2 = .25$.

Planned contrasts deconstructing this interaction revealed that the disruption caused by the secondary object-task was significantly larger under the color-shape binding condition ($M = .25$) than under the visual color ($M = .04$), $t(23) = 4.01, p = .001$, Cohen’s $d = 0.82$, and visual shape ($M = .09$), $t(23) = 2.35, p = .028$, Cohen’s $d = 0.48$, conditions.

Discussion

Consistent with Experiments 1 and 2, Experiment 3 demonstrated that Duncan’s object feature-report task incurred a larger impairment on binding than on the constituent features for

Fig. 7 Results for Experiment 3. The cost is calculated as the difference between “no object-task” and “with object-task” conditions (no object-task minus with object-task). Error bars stand for standard errors

the cross-time bindings. Therefore, the object-based attention hypothesis works for temporally separated visual bindings.

Experiment 4: Object-based attention in cross-space binding

In Experiment 4, we further examined the role of object-based attention in another type of visual-visual bindings without high spatial contiguity: Cross-space bindings. We replaced the cross-time binding in Experiment 3 with a type of cross-space binding used in Karlsen et al. (2010; see also Vergauwe & Cowan, 2015): Each binding comprised a color and a shape, with a spatial gap between (see Fig. 8). Karlsen et al. (2010) revealed that domain-general attention did not play a critical role in retaining cross-modal bindings in WM. We recently further found that by adding a visual search task to consume the space-based attention during the maintenance phase of cross-space bindings, the secondary task equally disrupted

Fig. 8 An illustration of memory arrays and probes used in Experiment 4. (Color figure online)

the performance for bindings and the constituent features, suggesting that space-based attention also did not play a critical role in retaining cross-space bindings (Gao, Wu, & Shen, unpublished data).

Method

There were 24 valid participants (14 males). Four participants were replaced with new participants because of chance-level performance in the secondary or memory task. The participants were 18.74 ($SD = 1.10$) years old on average.

Each cross-space binding comprised a color and a shape presented simultaneously, with the shape always appearing directly above the color with a center-to-center distance of 1.98° of visual angle (see Fig. 8). Each time, three bindings were displayed simultaneously in a horizontal row at the screen center, with a 3.68° center-to-center visual distance in between. The to-be-memorized colors and shapes were randomly selected from a pool of six distinct values of each type (see Fig. 3). For the probed stimuli, colored blobs, unfilled three-point black outlines, and colored shapes were used in the visual color, visual shape, and color-shape binding conditions, respectively (see Fig. 8).

The procedure in Experiment 4 was identical to Experiment 3 except for that the memory array was replaced by three cross-space bindings.

The whole experiment lasted about 50 minutes. The other aspects were the same as Experiment 3.

Results

The overall accuracy of the secondary task was 97.17 %. Figure 9 shows the corrected recognition for the color-shape binding stimuli. The two-way ANOVA on the corrected

Fig. 9 Result for Experiment 4. The cost is calculated as the difference between “no object-task” and “with object-task” conditions (no object-task minus with object-task). Error bars stand for standard errors

recognition revealed a significant main effect of Task-load, $F(1, 23) = 35.80$, $p < .001$, $\eta_p^2 = .61$, suggesting that performance was significantly worse under the “with object-task” condition ($M = .61$) than under the “no object-task” condition ($M = .77$). The main effect of Memory-condition also reached significance, $F(2, 46) = 88.17$, $p < .001$, $\eta_p^2 = .79$. Further analysis (Bonferroni-corrected) revealed that the performance was significantly lower for binding ($M = .44$) than for visual color ($M = .88$; $p < .001$) and visual shape ($M = .76$; $p < .001$) conditions, and the performance for visual color was significantly higher than that for visual shape ($p = .002$). Critically, the Memory-condition factor was further modulated by the secondary task, revealing a significant Task-load $\times$ Memory-condition interaction, $F(2, 46) = 9.82$, $p < .001$, $\eta_p^2 = .30$.

Planned contrasts deconstructing this interaction revealed that the disruption caused by the secondary object-task was significantly larger under the color-shape binding condition ($M = .29$) than under the visual color ($M = .08$), $t(23) = 3.58$, $p = .002$, Cohen's $d = .73$, and visual shape ($M = .08$), $t(23) = 3.17$, $p = .004$, Cohen's $d = .65$, conditions.

Discussion

In line with Experiment 3, Experiment 4 added further evidence that Duncan's object feature-report task incurred a larger impairment on binding than on the constituent features for the cross-space bindings, suggesting that the object-based attention hypothesis is not limited to visual bindings with high spatial contiguity.

Experiments 1–4 converged in showing that object-based attention plays a critical role in retaining bindings regardless of unitization and constituent elements of bindings. However, it could be possible that these findings were contaminated by the unbalanced task load: Participants had to retain two fold of features in retaining bindings than in retaining single feature condition, which led to worse performance for retaining bindings than for retaining single features when no secondary task was added. Critically, the worse performance in binding would further lead to larger binding impairment. We argued that these results were immune to this alternative interpretation for at least two reasons. First, our recent study revealed a nonsignificant difference in accuracy between color-location binding and color when no object-based secondary task was performed (cf. Experiment 3 in Shen et al., 2015). However, instead of revealing comparable impairments between binding and color, the added object-based task still incurred a larger impairment for binding than for color or location, implying that memory load was not the key factor affecting the degree of impairment. Second, although the binding performance was significantly worse for color-shape binding than for color or shape without the secondary task, the similar degree of impairments between binding and the constituent features were consistently found when a space-based visual search task

was inserted (cf. Johnson et al., 2008; Shen et al., 2015). Similarly, Allen et al. (2006; cf. Experiment 3), Langerock et al. (2014), and Ding et al. (2015; cf. Experiment 6) revealed worse performance for binding than for the constituent features; however, the task load in terms of domain-general attention did not result in a larger binding impairment relative to the constituent features. These results consistently suggest that task load was not the underlying factor leading to larger binding impairment.

However, to have a strict examination of the object-based attention hypothesis of binding, in Experiment 5 we further examined the influence of task load on the selective binding impairment. We presented the stimuli in a way of cross-space binding. Critically, the participants were required to remember three visual colors and three visual shapes simultaneously or the three corresponding bindings. In this vein, the same number of features was retained between retaining bindings and retaining single features. If the unbalanced task load could explain our previous findings, the selective binding impairment would vanish.

Experiment 5: Ruling out the unbalanced memory load alternative

Methods

There were 24 participants (11 males) participated in the experiment. Four participants were replaced with new participants because of chance-level performance in the secondary task or extremely low-level (three standard deviations away from the mean) performance in the memory task. The participants were 20.62 ($SD = 1.79$) years old on the average.

The procedure in Experiment 5 was identical to Experiment 4 except for the following two aspects: (1) In the single feature condition, we did not inform the participants the dimension to be tested in advance. Instead, participants were required to memorize both the three colors and the three shapes. For the probe, 50 % of trials were color-probe, and 50 % of trials were shape-probe. These two probe types were randomly presented. To have a direct comparison with Experiment 4, we also examined three memory conditions: Visual-color, visual-shape, and color-shape bindings. Therefore, similar to previous experiments, there were 32 trials under each feature-load combination. (2) Each task-load session consisted of two blocks: single feature-only and color-shape binding, with their order fully counterbalanced.

Results

The overall accuracy of the secondary task was 96.17 %. Figure 10 shows the corrected recognition for the color-shape binding stimuli.

Fig. 10 Result for Experiment 5. The cost is calculated as the difference between “no object-task” and “with object-task” conditions (no object-task minus with object-task). Error bars stand for standard errors

We first examined the effect of Memory-condition on the performance under “no object-task” condition by running a one-way ANOVA with Memory-condition as the within-subjects factor. The ANOVA revealed a significant main effect of Memory-condition, $F(2, 46) = 8.07$, $p = .001$, $\eta_p^2 = .26$. Post hoc test (Bonferroni-corrected) revealed that the performance of visual color ($M = .64$) was significantly higher than visual shape ($M = .44$, $p = .004$) and binding ($M = .51$, $p = .044$). However, there was no statistical difference between visual shape and binding ($p = .47$), even with a reversed trend. Additionally, we further pooled visual color and visual shape as one single-feature factor, and found a non-significant difference between single feature ($M = .54$) and binding ($M = .51$), $t(23) = .81$, $p = .43$, Cohen’s $d = .16$.

Then we ran a two-way ANOVA on the corrected recognition with Memory-condition and Task-load as within-subjects factors. The ANOVA revealed a significant main effect of Task-load, $F(1, 23) = 30.72$, $p < .001$, $\eta_p^2 = .57$, suggesting that performance was significantly worse under the “with object-task” condition ($M = .32$) than under the “no object-task” condition ($M = .53$). The main effect of Memory-condition also reached significance, $F(2, 46) = 31.82$, $p < .001$, $\eta_p^2 = .58$. Further analysis (Bonferroni-corrected) revealed that the performance was significantly higher for visual color ($M = .61$) than for visual shape ($M = .35$, $p < .001$) and binding ($M = .31$, $p < .001$) conditions, and the performance difference between visual shape and binding did not reach significance ($p = .79$). Moreover, the Memory-condition factor was further modulated by the secondary task, revealing a significant Task-load $\times$ Memory-condition interaction, $F(2, 46) = 13.56$, $p < .001$, $\eta_p^2 = .37$.

Planned contrasts deconstructing this interaction revealed that the disruption caused by the secondary object-task was significantly larger under the color-shape binding condition

($M = .40$) than under the visual color ($M = .06$), $t(23) = 5.07$, $p < .001$, Cohen’s $d = 1.04$, and visual shape ($M = .18$), $t(23) = 3.37$, $p = .003$, Cohen’s $d = .69$, conditions.

Discussion

By requiring the participants to memorize six features or their corresponding bindings, we found that the binding performance was comparable to visual shapes or the single-feature condition as a whole when the object-task was absent. However, we replicated the findings of Experiments 1–4, suggesting that the unbalanced memory load did not contaminate the selective binding impairment, at least for the cross-space binding in Experiment 3.

Experiment 6: Comparison between within-object and between-object conditions

It is worth noting that when revealing the existence of object-based attention, it is routinely comparing the performance between between-objects and within-object conditions (e.g., Awh, Dhaliwal, Christensen & Matsukura, 2001; Duncan, 1984), by assuming that extra object-based attention is exhausted in the between-objects condition. However, in Experiments 1–5 we examined the role of object-based attention by pooling the within-object and between-objects conditions as one condition (i.e., with object-task) and comparing the performance between with and no object-task conditions. Therefore, the selective binding impairment by the object feature-report task may be attributed to the memorization process (cf. Awh et al., 2001) in the Duncan task.⁷ Because the memorization process of WM needs more than object-based attention (see Fougine, 2008, for a review), it will make the role of object-based attention unclear.

In our previous study (cf. Experiment 5 in Shen et al., 2015) we had ruled out this possibility by splitting the object feature-report task into within-object and between-objects conditions. We found that comparing to the no-object-task condition, within-object condition led to a significantly larger impairment to unitized color-shape bindings than to the constituent single features (akin to this study); critically, a similar pattern was revealed between within-object and between-objects conditions. However, we also demonstrated that the binding disruption was significantly larger between no-object-task and within-object condition than between within-object and between-objects condition. To sensitively reveal the potential selective disruption while efficiently shortening the experiment duration (see footnote 4), we reduced the trial number from 36 in Shen et al. (2015) to 16 trials in Experiments 1–5 for each of the within-object and between-objects condition,

⁷ We thanked Dr. Geoffrey Woodman for pointing this out to us.

and pooled the two conditions as a “with object-task” condition (32 trials in total). We considered that the role of object-based attention could still be manifested in this manner. However, because distinct new bindings were tested in this study, it is of necessity to reexamine the aforementioned possibility. Looking into the data of Experiments 1–5, we found that there were not enough data (particularly under the binding condition) supporting a valid analysis between within-object and between-objects conditions. To this end, we ran a new experiment with within-object condition and between-objects condition each having 32 trials. We took cross-space binding as an example; on the base of Experiment 5, we removed the no object-task condition and focused on the comparison between within-object and between-objects conditions.

Method

There were 12 participants (three males) participated in the experiment. The participants were 20.38 ($SD = 1.01$) years old on the average. The with object-task and no object-task conditions were replaced with within-object and between-objects conditions. The to-be-reported features were from the same object in the within-object condition and were from different objects in between-objects condition (see Fig. 11). There were 32 trials under each load condition. The other aspects were the same to Experiment 5.

Results

The overall accuracy of the secondary task was 94.29 %. Figure 12 shows the corrected recognition for the color-

Fig. 12 Results for Experiment 6. The cost is calculated as the difference between “within-object task” and “between-object task” conditions (within-object task minus between-object task). Error bars stand for standard errors

shape binding stimuli. The result pattern was similar to the one in Experiment 5.

We first examined the effect of Memory-condition on the performance under “within-object” condition to ensure that any selective binding impairment was not due to harder binding condition, by running a one-way ANOVA with Memory-condition as the within-subjects factor. The ANOVA revealed a significant main effect of Memory-condition, $F(2, 22) = 13.80$, $p < .001$, $\eta_p^2 = .56$. Post hoc test (Bonferroni-corrected) revealed that the performance of visual color ($M = .64$) was significantly higher than visual shape ($M = .27$, $p < .001$) and binding ($M = .37$, $p = .027$). Moreover, the performance of binding was not significantly different from that of shape ($p = .70$), although a reversed trend was found.

Fig. 11 A schematic illustration of the secondary task in Experiment 6. The upper row illustrates the within-object condition, in which the to-be-reported features are the height of the box and the position of the gap of the box. The lower row illustrates between-objects condition, in which the to-be-reported features are the height of the box and the texture of the

tilted line. The order of the two questions were fixed in a block and was fully balanced between blocks. Note in the real experiments we used Chinese characters to describe objects’ attributes (e.g., “高” for high, “低” for low) for the two response windows of the object task

Finally, we ran a two-way ANOVA on the corrected recognition with Memory-condition and Task-load as within-subjects factors. The ANOVA revealed that the main effect of Task-load, $F(1, 11) = 1.09$, $p = .319$, $\eta_p^2 = .09$, was not significant. The main effect of Memory-condition was significant, $F(2, 22) = 22.77$, $p < .001$, $\eta_p^2 = .67$. Further analysis (Bonferroni-corrected) revealed that the performance was significantly higher for visual color ($M = .66$) than for visual shape ($M = .28$, $p < .001$) and binding ($M = .28$, $p = .001$) conditions, and the performance difference between visual shape and binding did not reach significance ($p = 1.00$). Critically, there was a significant Task-load $\times$ Memory-condition interaction, $F(2, 22) = 3.97$, $p = .034$, $\eta_p^2 = .27$.

Planned contrasts deconstructing this interaction revealed that the disruption due to processing one more object was significantly larger under the color-shape binding condition (.17) than under the visual color (-0.03), $t(11) = 2.60$, $p = .025$, Cohen's $d = .75$, and visual shape (-0.01), $t(11) = 2.41$, $p = .035$, Cohen's $d = .70$, conditions.

Discussion

By taking cross-space binding as an example, we examined whether the selective binding impairment revealed in Experiments 1–5 was due to the memorization process in the Duncan task. In line with Shen et al. (2015), we demonstrated that consuming extra object-based attention led to a larger impairment on binding than on the single features. Therefore, we considered that the currently revealed selective binding impairments were due to consuming more object-based attention instead of the memorization process (at least for the cross-space binding). Moreover, Experiment 6 provides new evidence showing that the selective binding impairment in previous Experiments 1–4 (at least for Experiment 4) was not due to unbalanced task difficulty between binding and constituent single features.

Additionally, the results of Experiments 5 and 6 also helped us to rule out an alternative explanation of the lower binding performance in Experiments 1–4. To be specific, because participants had to process more features in the binding condition than in the constituent single feature conditions, there would be stronger proactive interference in the binding condition. Moreover, there was noise in each feature decision, the decision noise for the two-featured bindings (judging whether the two features were originally bounded as one unit) may be greater than the decision noise for the single features alone (Shaw, 1982). These two aspects together would lead to higher decision noise for binding and hence poor performance. We considered that this explanation was hard to explain the findings of Experiments 5 and 6, in which the same number of features were processed between binding and single feature conditions. Moreover, recent studies had suggest that bindings and constituent single features can coexist in

WM (e.g., Brady, Konkle, & Alvarez, 2011; Cowan et al., 2013; Cowan et al., 2013; Song, Liu, Lu, & Gu, 2016; Vergauwe & Cowan, 2015), even with an independent buffer for bindings (e.g., Baddeley, 2000). Therefore, although there was noise in each feature decision, it did not necessarily imply that the decision noise of the formed binding would be greater than the feature dimension. Indeed, Experiments 5 and 6 even consistently found a trend of better performance for binding relative to visual shapes.

General discussion

This study explored an object-based attention hypothesis of retaining bindings in WM. This hypothesis claims that retaining bindings in WM requires more object-based attention than retaining the constituent features. To test this hypothesis, we added an object feature-report task, which consumed object-based attention, to the maintenance phase of a WM task. Moreover, we tested four new bindings that are distinct from unitized visual bindings and have not been tested under the framework of object-based attention hypothesis: Cross-module binding of information from the phonological loop and visuospatial sketchpad (Experiment 1), cross-modal binding of auditory and visual information (Experiment 2), cross-time (Experiment 3) and cross-space binding of visual features (Experiments 4, 5, and 6). In contrast to previous studies revealing no significant interaction between task load and memory condition with these bindings (e.g., Allen et al., 2009; Karlsen et al., 2010; Langerock et al., 2014; Vergauwe et al., 2014), Experiments 1–6 consistently demonstrated a significant interaction between the two factors: The secondary object-based task led to a significantly larger disruption for cross-module, cross-modal, cross-time, and cross-space bindings than for the constituent features. Moreover, the selective binding impairment was not due to the unbalanced memory load between binding and constituent single features (Experiment 5) or the memorization process involved in the object feature-report task (Experiment 6). This study hence provides compelling evidence in that retaining bindings in WM requires more object-based attention than the constituent features, supporting an object-based hypothesis of retaining bindings in WM.

Retaining bindings in WM needs object-based attention

Taking our study and Shen et al.'s (2015) together, we argue that the view that object-based attention plays a critical role in retaining bindings in WM is strongly supported. First of all, this study added four new distinct bindings, which enables us to understand the role of object-based attention beyond visuospatial subsystem of WM. In two studies we demonstrated that the object-based attention works not only in retaining bindings

composed of information from the visuospatial sketchpad with high spatial contiguity (Shen et al., 2015), but also in retaining bindings with the constituent features from distinct modals or modules or without spatial contiguity (this study), which theoretically could not be addressed by the visuospatial sketchpad (e.g., Baddeley et al., 2010).

Second, so far three distinct tasks had been employed in consuming the object-based attention (i.e., mental rotation task, transparent motion task, and object feature-report task) in the two studies. These tasks covered static (object feature-report task) versus dynamic (mental rotation and transparent motion task) object-based processing; one-object (mental rotation task, transparent motion task, object feature-report task) versus dual-object processing (object feature-report task); object-as-a-whole operation (mental rotation task) versus object-feature extraction (transparent motion task, object feature-report task) processing. Therefore, the revealed selective binding impairment was not limited to one specific object-based task.

Implications to episodic buffer

The tested bindings in this study offer a particularly strong test of the episodic buffer concept (Baddeley, 2000) and hence cast important implications to the mechanisms of episodic buffer. Particularly, different from Shen et al. (2015), in which the tested bindings could be built in the visuospatial sketchpad, the tested bindings in this study were initially employed in studies exploring the mechanisms of episodic buffer. According to the revised multiple-component model of WM of Baddeley (2000), the cross-module, cross-modal, cross-time, and cross-space bindings cannot be formed in the visuospatial sketchpad, but were expected to be built in the episodic buffer (Allen et al., 2009; Baddeley et al., 2010; Elsley & Parmentier, 2009; Karlsen et al., 2010; Langerock et al., 2014). However, researchers consistently revealed that consuming domain-general attention, space-based attention, or verbal attention did not lead to a larger binding impairment relative to the constituent features. These findings as well as other similar results led Baddeley to give up the active view of episodic buffer proposed in Baddeley (2000). Currently, Baddeley considered episodic buffer as a passive storage buffer, which receives bindings that had been formed elsewhere before entering into the episodic buffer (see Baddeley et al., 2010, 2011, for reviews).

This study hence provides clear evidence supporting the *active* view of episodic buffer, which is in line with the initial proposition of Baddeley (2000). On the other hand, the current object-based attention hypothesis suggests a modification of the Baddeley (2000) model, which proposed that binding was processed via the domain-general attention in the central executive. An intuitive way to resolve the discrepancy between the current findings and Baddeley's (2000) model is to assume

that there is also domain-specific attention within the central executive beyond the domain-general attention. This possibility was supported by the studies of Goldman-Rakic (1988) that investigated the neural correlates of WM via single-cell recordings in awake monkeys. She found that certain cells in prefrontal cortex, where the neural substrates of the central executive are expected to be located, were only active when the monkeys were successfully memorizing a cued location or memorizing a shape. She hence proposed that WM was represented across certain autonomous subdivisions, and each subdivision had a corresponding executive subsystem. However, this view of central executive was not supported by numerous subsequent cognitive and neuropsychological studies (for a review, see Baddeley, 2007a).

Here we suggest that the episodic buffer is not a slave buffer under the control of the central executive. Like the phonological loop or visuospatial sketchpad, the episodic buffer is an *independent* storage buffer but is fueled by *object-based attention*. Moreover, the phonological and visuospatial subsystems have *direct* access to the episodic buffer. Considering that the episodic buffer needs to bind the multicode information, there may be a hierarchical structure between the episodic buffer and the phonological and visuospatial subsystems, with the latter two subsystems serving as lower-level storage buffers (see also Brady et al., 2011). It is worth noting that a similar structure had been proposed by Baddeley et al. (2011). However, there is a key difference between Baddeley et al.'s (2011) claim and ours. In particular, Baddeley and colleagues dropped the active view of the episodic buffer, considering it as a passive recipient of the products (i.e., binding) of the lower-level subsystems (Baddeley et al., 2011). This view could well explain the binding composed of elements from the same storage subsystem; however, it cannot interpret many other bindings, such as the currently tested cross-module binding. In short, we currently suggest that episodic buffer is an active buffer fueled by object-based attention.

Meanwhile, we have to point out that more empirical evidence is needed before concluding that episodic buffer as an active buffer. For instance, according to the assumption of Baddeley (2000), the episodic buffer is in charge of both encoding and retaining bindings in WM. All of Baddeley and colleagues' studies on episodic buffer explored the role of domain-general attention in the encoding and maintenance phases. However, all the extant evidence revealing the role of the object-based attention in retaining bindings tapped the maintenance phase of WM (i.e., this study; Fournie & Marois, 2009; Shen et al., 2015). Therefore, future studies need to explore the role of object-based attention in encoding bindings into WM. In addition, because a secondary task was added to the maintenance phase, the representation was kept within WM for a relatively long period, which could make the requirement of attention particularly extensive (cf. Zhang et al., 2012). Future studies should also consider whether the

duration of maintenance phase affects the role of object-based attention. Finally, so far we have tested the role of object-based attention in five distinct types of binding (i.e., visual binding with high spatial contiguity, cross-time visual binding, cross-space visual binding, cross-modal binding, and cross-module binding) that mainly tapped the function of episodic buffer as an interface of WM subsystems. However, another core function of the episodic buffer is that it serves as an interface linking WM and long-term memory (Baddeley, 2000; see Allen et al., 2015, for an elegant example in this perspective), which needs to be examined in future studies in terms of the role of object-based attention.

Implications to the function of object-based attention

This study also contributes at least two new aspects to the object-based attention field. First, it is well accepted that there is a close relationship between attention and WM (e.g., Cowan, 2001; Fougne & Marois, 2006; Oberauer & Hein, 2012; for reviews, see Chun, 2011; Chun, Golomb, & Turk-Browne, 2011; Fougne, 2008; Kiyonaga & Egner, 2013); however, most of the WM studies focused on space-based attention, verbal attention, and domain-general attention (e.g., Allen et al., 2006, 2009; Baddeley et al., 2011; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007; Barrouillet, Portrat, & Camos, 2011; Johnson et al., 2008). Currently, only few studies directly explored the role of object-based attention in WM (Barnes, Nelson, & Reuter-Lorenz, 2001; Matsukura & Vecera, 2009, 2011; Shen et al., 2015; Woodman & Vecera, 2011). Barnes et al. (2001) and Matsukura and Vecera (2009, 2011) explored the role of visual object-based attention in retaining simple visual objects in WM; and Woodman and Vecera (2011) implied that accessing features from multifeatured visual objects in WM requires object-based attention. On the other hand, Shen et al. (2015) first revealed the role of object-based attention in retaining unitized visual bindings in WM. This study hence adds new evidence supporting the role of object-based attention in WM. Moreover, this study is the second study directly exploring the role of object-based attention in retaining bindings in WM.

Second, this study provides the first clear evidence suggesting that the role of object-based attention is beyond the visuospatial sketchpad. Particularly, all the relevant studies assumed that object-based attention plays a critical role in the visuospatial sketchpad because it is implicitly assumed that storing visual object representations needs visual object-based attention (Barnes et al., 2001; Matsukura & Vecera, 2009, 2011; Shen et al., 2015). For instance, Matsukura and Vecera (2009)

required the participants to perform an attention-consuming task (either a visual object-based feature report task or a space-based visual search task) within a concurrent object (two colors) or spatial WM task.⁸ Results revealed that object WM was disrupted to a larger degree by the object-based secondary task than by the space-based secondary task; however, a reversed profile was found for spatial WM.⁹ This study suggests that the role of object-based attention can be generalized from object WM to a type of storage buffer (e.g., episodic buffer; Baddeley, 2000; or focus of attention, Cowan, 2001) holding bindings. However, future studies are required to figure out whether there is any difference on the nature of object-based attention between the object WM and the possible buffer dedicated to storing bindings (e.g., episodic buffer).

Related to the second point, it is worth noting that recently Cowan, Saults, and Blume (2014) proposed that WM capacity could be further divided into central and peripheral components. The central component could be shared by distinct stimuli (e.g., visual and auditory stimuli), whereas the peripheral component is modality specific. Both the central and peripheral components are involved in storing a memory array. In a set of elegant experiments, they showed that the central component of WM could store around one chunk unit of information (including binding, visual color, etc.), while the capacity of peripheral component depends on the nature of stored stimuli (e.g., one binding, 2–3 visual colors). Cowan et al. suggested that the same WM capacity of central component for both binding and constituent single features could explain the comparable impairments between binding and single features when domain-general attention was exhausted (e.g., Allen et al., 2006, 2012; Cowan et al., 2006, 2013; Ding et al., 2015). By taking these results into consideration, we predicted that even the WM capacity of central component was fixed, more object-based attention should be required for binding relative to the single features. However, our current experiment settings prevented us from employing the methods in Cowan et al. (2014) to examine the role of object-based attention in the central and peripheral components of WM. We believe it is an important issue to be addressed in future studies, which will also shed light on the relationship between object WM and episodic buffer.

How does object-based attention operate?

To establish the role of object-based attention in processing bindings in WM, a key question should be addressed: How

⁸ Visuospatial sketchpad is suggested to contain at least two distinct components: Spatial and object WM (Baddeley, 1986, 2012). The former stores location information while the latter stores nonspatial object information.

⁹ Consequently, a significant disruption was also found to the visual feature conditions of the current study (e.g., visual color/shape in Experiments 3–5). Note that in Experiment 1 a significant disruption was also found for the location condition. One possible explanation is that participants had to process the place holders to obtain the detailed location information, which may require object-based attention.

does the object-based attention operate in the episodic buffer? Most studies have focused on the underlying mechanisms of object-based attention in visual perception (e.g., Cohen & Tong, 2013; Duncan, 1984, 2006; Egly, Rafal, Driver, & Starrveveld, 1994; Emmanouil & Magen, 2014; Kahneman et al., 1992; Schoenfeld et al., 2014; Schoenfeld et al., 2003; Scholl, 2001; Valdes-Sosa et al., 1998, 2000), and recently made significant progress due to sophisticated neuroimaging methods (e.g., Baldauf & Desimone, 2014; Cohen & Tong, 2013; Schoenfeld et al., 2014; Schoenfeld et al., 2003). Relevant to these perceptual findings, accumulating evidence has suggested that similar processing mechanisms are shared between perception and WM (e.g., T. Gao et al., 2011; Z. Gao et al., 2015; Z. Gao et al., 2010; Harrison & Tong, 2009; Kiyonaga & Egner, 2014; Mayer et al., 2007). Researchers even suggest that perception and WM reflect a common selective mechanism that differs only in whether attention operates over internal representations or for external features (Chun, 2011; Chun et al., 2011; Kiyonaga & Egner, 2013). Paralleling external attention, which selects information from the environment, WM is a type of internal attention that sustains information over time (Chun, 2011). From this perspective, the mechanisms for feature binding in perception can shed important light on the operating mechanisms of object-based attention in WM.

Currently, one of the most sophisticated theories for explaining the mechanisms of object-based attention is elaborated by the “integrated competition model” (Duncan, 1996; Duncan, Humphreys, & Ward, 1997). According to this model, when attention is directed to a feature of a multiple-featured object, activity in the cortical module encoding this feature will first be enhanced and subsequently spread to modules encoding other features pertaining to this object. Moreover, once attention is directed to an object’s feature, all other constituent features will become dominant in the corresponding cortical modules, resulting in a unified experience of the corresponding object. This key prediction has been verified by fMRI and MEG studies (e.g., O’Craven, Downing, & Kanwisher, 1999; Schoenfeld et al., 2014; Schoenfeld et al., 2003). Duncan (2006) further suggested that this integrated competition model fits the processing of bindings containing other nonvisual information. We thereby suggest that the process proposed by the integrated competition model may underlie the neural mechanisms for processing binding in the WM.

Moreover, this study implies that adding a secondary task consuming object-based attention impairs the normal operations suggested by the integrated competition model. According to the integrated competition model, there are two core operations involved in object-based attention: Enhancing the related feature modules in the brain and spreading the activation from one feature module to another. It is possible that one or both of them are impeded after adding a

secondary task consuming object-based attention. Relevant to this possibility, it has been suggested that there are at least two activation states for representations in WM (e.g., Cowan, 1988, 2001; Oberauer & Hein, 2012; Olivers et al., 2011): (1) focus of attention, which actively retains a limited number of representations for the ongoing task, and (2) the activated portion of long-term memory, which is in a dormant state and can become the focus of attention with top-down task demand or bottom-up salience. Baddeley and colleagues recently confirmed that there were two activation states in WM (Allen, Baddeley, & Hitch, 2014; Hu et al., 2014). Therefore, adding an object-based secondary task may convert the active representations into a dormant state, which makes the core operations in the integrated competition model into a dormant state. Future neuroimaging studies are needed to elucidate the above issues.

Conclusion

By inserting a secondary task consuming object-based attention during the maintenance phase of WM, in six experiments we consistently demonstrated that bindings were significantly disrupted compared with constituent features. This selective binding impairment was observed regardless of the tested binding types. We conclude that retaining bindings in WM requires more object-based attention than retaining constituent features.

Acknowledgments This research is supported by NSFC (No. 31271089, 31571119, and 61431015), RFDP (No. 20120101120085), and fundamental research funds for the central universities.

References

- Allen, R. J. (2015). Memory binding. In D. W. James (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., Vol. 15, pp. 140–146). Oxford, UK: Elsevier.
- Allen, R. J., Baddeley, A., & Hitch, G. J. (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General*, 135(2), 298–313. doi:10.1037/0096-3445.135.2.298
- Allen, R. J., Baddeley, A., & Hitch, G. J. (2014). Evidence for two attentional components in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(6), 1499–1509. doi:10.1037/xlm0000002
- Allen, R. J., Havelka, J., Falcon, T., Evans, S., & Darling, S. (2015). Modality specificity and integration in working memory: Insights from visuospatial bootstrapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(3), 820–830. doi:10.1037/xlm0000058
- Allen, R. J., Hitch, G. J., Mate, J., & Baddeley, A. (2012). Feature binding and attention in working memory: A resolution of previous contradictory findings. *The Quarterly Journal of Experimental Psychology*, 65(12), 2369–2383. doi:10.1080/17470218.2012.687384

- Allen, R. J., Hitch, G. J., & Baddeley, A. D. (2009). Cross-modal binding and working memory. *Visual Cognition*, 17(1/2), 83–102. doi:10.1080/13506280802281386
- Awh, E., Dhaliwal, H., Christensen, S., & Matsukura, M. (2001). Evidence for two components of object-based selection. *Psychological Science*, 12(4), 329–334.
- Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, 5(3), 119–126.
- Baddeley, A. (1986). *Working memory*. Oxford, UK: Clarendon.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417–423. doi:10.1016/S1364-6613(00)01538-2
- Baddeley, A. (2007a). Fractionating the central executive. In A. Baddeley (Ed.), *Working memory, thought, and action*. New York, NY: Oxford University Press.
- Baddeley, A. (2007b). *Working memory, thought, and action*. New York, NY: Oxford University Press.
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63(1), 1–29. doi:10.1146/annurev-psych-120710-100422
- Baddeley, A., Allen, R. J., & Hitch, G. (2010). Investigating the episodic buffer. *Psychologica Belgica*, 50(3/4), 3–4.
- Baddeley, A., Allen, R. J., & Hitch, G. J. (2011). Binding in visual working memory: The role of the episodic buffer. *Neuropsychologia*, 49(6), 1393–1400. doi:10.1016/j.neuropsychologia.2010.12.042
- Baddeley, A., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). New York, NY: Academic Press.
- Baddeley, A., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, 61(3), 438–456. doi:10.1016/j.jml.2009.05.004
- Baldauf, D., & Desimone, R. (2014). Neural mechanisms of object-based attention. *Science*, 344(6182), 424–427.
- Barnes, L. L., Nelson, J. K., & Reuter-Lorenz, P. A. (2001). Object-based attention and object-based working memory: Overlapping processes revealed by selective interference effects in humans. In C. Casanova & M. Ptito (Eds.), *Progress in brain research* (Vol. 134, pp. 471–481). Amsterdam, The Netherlands: Elsevier.
- Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 570–585.
- Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological Review*, 118(2), 175–192. doi:10.1037/a0022324
- Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, 11(5). doi:10.1167/11.5.4
- Brown, L. A., & Brockmole, J. R. (2010). The role of attention in binding visual features in working memory: Evidence from cognitive aging. *Quarterly Journal of Experimental Psychology*, 63, 2067–2079.
- Chen, Z. (2012). Object-based attention: A tutorial review. *Attention, Perception, & Psychophysics*, 74(5), 784–802. doi:10.3758/s13414-012-0322-z
- Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia*, 49(6), 1407–1409. doi:10.1016/j.neuropsychologia.2011.01.029
- Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, 62, 73–101. doi:10.1146/annurev.psych.093008.100427
- Cohen, E. H., & Tong, F. (2013). Neural mechanisms of object-based attention. *Cerebral Cortex*. doi:10.1093/cercor/bht303
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, 104(2), 163–191.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–114. discussion 114–185.
- Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, 26(2), 197–223. doi:10.1007/s10648-013-9246-y
- Cowan, N., Blume, C. L., & Saults, J. S. (2013). Attention to attributes and objects in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 731–747. doi:10.1037/a0029687
- Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult? *Developmental Psychology*, 42, 1089–1102. doi:10.1037/0012-1649.42.6.1089
- Cowan, N., Saults, J. S., & Blume, C. L. (2014). Central and peripheral components of working memory storage. *Journal of Experimental Psychology: General*, 143, 1806–1836.
- Delvenne, J. F., Cleeremans, A., & Laloyaux, C. (2010). Feature bindings are maintained in visual short-term memory without sustained focused attention. *Experimental Psychology*, 57(2), 108–116.
- Ding, X., Zhao, Y., Wu, F., Lu, X., Gao, Z., & Shen, M. (2015). Binding biological motion and visual features in working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 41(3), 850–865.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113(4), 501–517.
- Duncan, J. (1996). Cooperating brain systems in selective perception and action. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 549–578). Cambridge, MA: MIT Press.
- Duncan, J. (2006). EPS Mid-Career Award 2004: Brain mechanisms of attention. *Quarterly Journal of Experimental Psychology*, 59(1), 2–27. doi:10.1080/17470210500260674
- Duncan, J., Humphreys, G., & Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology*, 7(2), 255–261.
- Ecker, U. K., Maybery, M., & Zimmer, H. D. (2013). Binding of intrinsic and extrinsic features in working memory. *Journal of Experimental Psychology: General*, 142(1), 218–234. doi:10.1037/a0028732
- Egley, R., Rafal, R., Driver, J., & Starrveveld, Y. (1994). Covert orienting in the split brain reveals hemispheric specialization for object-based attention. *Psychological Science*, 5(6), 380–383.
- Elsley, J. V., & Parmentier, F. B. (2009). Is verbal-spatial binding in working memory impaired by a concurrent memory load? *Quarterly Journal of Experimental Psychology*, 62(9), 1696–1705. doi:10.1080/17470210902811231
- Emmanouil, T. A., & Magen, H. (2014). Neural evidence for sequential selection of object features. *Trends in Cognitive Sciences*, 18(8), 390–391. doi:10.1016/j.tics.2014.04.007
- Ester, E. F., Serences, J. T., & Awh, E. (2009). Spatially global representations in human primary visual cortex during working memory maintenance. *Journal of Neuroscience*, 29(48), 15258–15265. doi:10.1523/JNEUROSCI.4388-09.2009
- Fougnie, D. (2008). The relationship between attention and working memory. In B. J. Noah (Ed.), *New research on short-term memory* (pp. 1–45). Hauppauge, NY: Nova Science.
- Fougnie, D., Asplund, C. L., & Marois, R. (2010). What are the units of storage in visual working memory? *Journal of Vision*, 10(12), 27.
- Fougnie, D., & Marois, R. (2006). Distinct capacity limits for attention and working memory: Evidence from attentive tracking and visual working memory paradigms. *Psychological Science*, 17(6), 526–534.

- Fougnie, D., & Marois, R. (2009). Attentive tracking disrupts feature binding in visual working memory. *Visual Cognition*, 17(1/2), 48–66. doi:10.1080/13506280802281337
- Fougnie, D., & Marois, R. (2011). What limits working memory capacity? Evidence for modality-specific sources to the simultaneous storage of visual and auditory arrays. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(6), 1329–1341.
- Gajewski, D. A., & Brockmole, J. R. (2006). Feature bindings endure without attention: Evidence from an explicit recall task. *Psychonomic Bulletin & Review*, 13(4), 581–587.
- Gao, Z., Bentin, S., & Shen, M. (2015). Rehearsing biological motion in working memory: an eeg study. *Journal of Cognitive Neuroscience*, 27(1), 198–209. doi:10.1162/jocn_a.00687.
- Gao, T., Gao, Z., Li, J., Sun, Z., & Shen, M. (2011). The perceptual root of object-based storage: An interactive model of perception and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 37(6), 1803–1823. doi:10.1037/a0025637
- Gao, Z., Li, J., Yin, J., & Shen, M. (2010). Dissociated mechanisms of extracting perceptual information into visual working memory. *PLoS ONE*, 5(12), e14273. doi:10.1371/journal.pone.0014273
- Goldman-Rakic, P. S. (1988). Topography of cognition: Parallel distributed networks in primate association cortex. *Annual Review of Neuroscience*, 11, 137–156. doi:10.1146/annurev.ne.11.030188.001033
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238), 632–635. doi:10.1038/nature07832
- Hollingworth, A., & Maxcey-Richard, A. M. (2013). Selective maintenance in visual working memory does not require sustained visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 39(4), 1047–1058. doi:10.1037/a0030238
- Hu, Y., Hitch, G. J., Baddeley, A., Zhang, M., & Allen, R. J. (2014). Executive and perceptual attention play different roles in visual working memory: Evidence from suffix and strategy effects. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1665–1678.
- Huang, L. (2010). What is the unit of visual attention? Object for selection, but Boolean map for access. *Journal of Experimental Psychology: General*, 139(1), 162–179. doi:10.1037/a0018034
- Humphreys, G. W. (2001). A multi-stage account of binding in vision: Neuropsychological evidence. *Visual Cognition*, 8, 381–410.
- Johnson, J. S., Hollingworth, A., & Luck, S. J. (2008). The role of attention in the maintenance of feature bindings in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, 34(1), 41–55. doi:10.1037/0096-1523.34.1.41
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24(2), 175–219. doi:10.1016/0010-0285(92)90007-o
- Karlsen, P. J., Allen, R. J., Baddeley, A., & Hitch, G. J. (2010). Binding across space and time in visual working memory. *Memory & Cognition*, 38(3), 292–303. doi:10.3758/mc.38.3.292
- Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychonomic Bulletin & Review*, 20(2), 228–242. doi:10.3758/s13423-012-0359-y
- Kiyonaga, A., & Egner, T. (2014). The working memory stroop effect: When internal representations clash with external stimuli. *Psychological Science*, 25(8), 1619–1629. doi:10.1177/0956797614536739
- Langerock, N., Vergauwe, E., & Barrouillet, P. (2014). The maintenance of cross-domain associations in the episodic buffer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(4), 1096–1109. doi:10.1037/a0035783
- Logie, R. H., & Cowan, N. (2015). Perspectives on working memory: Introduction to the special issue. *Memory & Cognition*, 43(3), 315–324. doi:10.3758/s13421-015-0510-x
- Lu, X., Huang, J., Yi, Y., Shen, M., Weng, X., & Gao, Z. (2016). Holding biological motion in working memory: An fMRI study. *Frontiers in Human Neuroscience*, 10(251). doi:10.3389/fnhum.2016.00251
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–281. doi:10.1038/36846
- Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia*, 49(6), 1632–1639. doi:10.1016/j.neuropsychologia.2010.11.031
- Matsukura, M., & Vecera, S. P. (2009). Interference between object-based attention and object-based memory. *Psychonomic Bulletin & Review*, 16(3), 529–536. doi:10.3758/Pbr.16.3.529
- Matsukura, M., & Vecera, S. P. (2011). Object-based selection from spatially-invariant representations: Evidence from a feature-report task. *Attention, Perception & Psychophysics*, 73(2), 447–457. doi:10.3758/s13414-010-0039-9
- Mayer, J. S., Bittner, R. A., Nikolic, D., Bledowski, C., Goebel, R., & Linden, D. E. (2007). Common neural substrates for visual working memory and attention. *NeuroImage*, 36(2), 441–453. doi:10.1016/j.neuroimage.2007.03.007
- Morey, C. C. (2009). Integrated cross-domain object storage in working memory: Evidence from a verbal-spatial memory task. *Quarterly Journal of Experimental Psychology*, 62(11), 2235–2251. doi:10.1080/17470210902763382
- Morey, C. C., & Bieler, M. (2013). Visual short-term memory always requires attention. *Psychonomic Bulletin & Review*, 20, 163–170. doi:10.3758/s13423-012-0313-z
- O’Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature*, 401(6753), 584–587. doi:10.1038/44134
- Oberauer, K., & Hein, L. (2012). Attention to information in working memory. *Current Directions in Psychological Science*, 21(3), 164–169. doi:10.1177/0963721412444727
- Oh, S. H., & Kim, M. S. (2004). The role of spatial working memory in visual search efficiency. *Psychonomic Bulletin & Review*, 11(2), 275–281
- Olivers, C. N., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences*, 15(7), 327–334. doi:10.1016/j.tics.2011.05.004
- Schoenfeld, M. A., Hopf, J. M., Merkel, C., Heinze, H. J., & Hillyard, S. A. (2014). Object-based attention involves the sequential activation of feature-specific cortical modules. *Nature Neuroscience*, 17(4), 619–624. doi:10.1038/nn.3656
- Schoenfeld, M. A., Tempelmann, C., Martinez, 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*, 100(20), 11806–11811. doi:10.1073/pnas.1932820100
- Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, 80(1), 1–46.
- Shaw, M. L. (1982). Attending to multiple sources of information: I. The integration of information in decision making. *Cognitive Psychology*, 14(3), 353–409
- Shen, M., Huang, X., & Gao, Z. (2015). Object-based attention underlies the rehearsal of feature binding in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*. doi:10.1037/xhp0000018
- Shen, M., Tang, N., Wu, F., Shui, R., & Gao, Z. (2013). Robust object-based encoding in visual working memory. *Journal of Vision*, 13(2), 1. doi:10.1167/13.2.1
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant.

- Psychological Science*, 22(11), 1359–1366. doi:[10.1177/0956797611417632](https://doi.org/10.1177/0956797611417632)
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117(1), 34–50.
- Song, C., Liu, W., Lu, X., & Gu, Q. (2016). Building blocks of visual working memory: Objects, features, or Hybrid? *Chinese Journal of Applied Psychology*, 22(2), 112–126.
- Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, 6(2), 171–178.
- Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 353(1373), 1295–1306. doi:[10.1098/rstb.1998.0284](https://doi.org/10.1098/rstb.1998.0284)
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97–136.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14, 107–141.
- Ueno, T., Allen, R. J., Baddeley, A., Hitch, G. J., & Saito, S. (2011). Disruption of visual feature binding in working memory. *Memory & Cognition*, 39(1), 12–23. doi:[10.3758/s13421-010-0013-8](https://doi.org/10.3758/s13421-010-0013-8)
- Valdes-Sosa, M., Cobo, A., & Pinilla, T. (1998). Transparent motion and object-based attention. *Cognition*, 66, B13–B23.
- Valdes-Sosa, M., Cobo, A., & Pinilla, T. (2000). Attention to object files defined by transparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, 26(2), 488–505. doi:[10.1037/0096-1523.26.2.488](https://doi.org/10.1037/0096-1523.26.2.488)
- van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological Review*, 121(1), 124–149. doi:[10.1037/a0035234](https://doi.org/10.1037/a0035234)
- Vergauwe, E., & Cowan, N. (2015). Working memory units are all in your head: factors that influence whether features or objects are the favored units. *Journal of Experimental Psychology: Learning Memory & Cognition*, 41(5), 1404–1416.
- Vergauwe, E., Langerock, N., & Barrouillet, P. (2014). Maintaining bindings between features in visual working memory requires no more attention than maintaining the features themselves. *Canadian Journal of Experimental Psychology*, 68, 158–162.
- 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 and Performance*, 27(1), 92–114.
- Wang, S., Allen, R. J., Lee, J. R., & Hsieh, C. E. (2015). Evaluating the developmental trajectory of the episodic buffer component of working memory and its relation to word recognition in children. *Journal of Experimental Child Psychology*, 133, 16–28. doi:[10.1016/j.jecp.2015.01.002](https://doi.org/10.1016/j.jecp.2015.01.002)
- Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, 131(1), 48–64.
- Wolfe, J., & Robertson, L. (2012). *From perception to consciousness: Searching with Anne Treisman*. New York, NY: Oxford University Press.
- Woodman, G. F., & Luck, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review*, 11, 269–274.
- Woodman, G. F., & Vecera, S. P. (2011). The cost of accessing an object's feature stored in visual working memory. *Visual Cognition*, 19(1), 1–12. doi:[10.1080/13506285.2010.521140](https://doi.org/10.1080/13506285.2010.521140)
- Xu, Y. (2002). Limitations of object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), 458–468.
- Xu, Y. (2006). Understanding the object benefit in visual short-term memory: The roles of feature proximity and connectedness. *Attention, Perception, & Psychophysics*, 68(5), 815–828.
- Yeh, Y. Y., Yang, C. T., & Chiu, Y. C. (2005). Binding or prioritization: The role of selective attention in visual short-term memory. *Visual Cognition*, 12(5), 759–799.
- Zhang, W., Johnson, J. S., Woodman, G. F., & Luck, S. J. (2012). Features and conjunctions in visual working memory. In J. Wolfe & L. Robertson (Eds.), *From perception to consciousness: Searching with Anne Treisman* (pp. 369–378). New York, NY: Oxford University Press.
- Zokaei, N., Heider, M., & Husain, M. (2014). Attention is required for maintenance of feature binding in visual working memory. *The Quarterly Journal of Experimental Psychology*, 67, 1191–1213. doi:[10.1080/17470218.2013.852232](https://doi.org/10.1080/17470218.2013.852232)