tapraid5/zfn-xhp/zfn-xhp/zfn00611/zfn2701d11z | xppws | S=1 | 9/20/11 | 9:55 | Art: 2010-1941

Journal of Experimental Psychology:
Human Perception and Performance

0096-1523/11/$12.00 DOI: 10.1037/a0025637

The Perceptual Root of Object-Based Storage: An Interactive Model of Perception and Visual Working Memory

Tao Gao, Zaifeng Gao, Jie Li, Zhongqiang Sun, and Mowei Shen
Zhejiang University

Mainstream theories of visual perception assume that visual working memory (VWM) is critical for integrating online perceptual information and constructing coherent visual experiences in changing environments. Given the dynamic interaction between online perception and VWM, we propose that how visual information is processed during visual perception can directly determine how the information is going to be selected, consolidated, and maintained in VWM. We demonstrate the validity of this hypothesis by investigating what kinds of perceptual information can be stored as integrated objects in VWM. Three criteria for object-based storage are introduced: (a) automatic selection of task-irrelevant features, (b) synchronous consolidation of multiple features, and (c) stable maintenance of feature conjunctions. The results show that the outputs of parallel perception meet *all* three criteria, as opposed to the outputs of serial attentive processing, which fail *all* three criteria. These results indicate that (a) perception and VWM are not two sequential processes, but are dynamically intertwined; (b) there are dissociated mechanisms in VWM for storing information identified at different stages of perception; and (c) the integrated object representations in VWM originate from the "preattentive" or "proto" objects created by parallel perception. These results suggest how visual perception, attention, and VWM can be explained by a unified framework.

Keywords: visual working memory, visual search, parallel perceptual processing, object-based attention

Visual working memory (VWM) enables the online maintenance and manipulation of a limited amount of visual information (Baddeley & Hitch, 1974; Phillips, 1974). It is a critical component of both theories of memory (e.g., Baddeley, 1992) and visual cognition (e.g., Duncan & Humphreys, 1989; Kahneman, Treisman, & Gibbs, 1992). In the past two decades, there has been tremendous progress in revealing the mechanisms of VWM,[1] including the capacity of VWM (e.g., Alvarez & Cavanagh, 2004; Awh, Barton, & Vogel, 2007; Jiang, Shim, & Makovski, 2008; Luck & Vogel, 1997; Xu, 2002), the unit of the storage resources (Barton, Ester, & Awh, 2009; Bays & Husain, 2008; Zhang & Luck, 2008), and different stages of VWM, such as consolidation (e.g., Vogel, Woodman, & Luck, 2006), maintenance (Johnson, Hollingworth, & Luck, 2008; Wheeler & Treisman, 2002), retrieval (Hyun, Woodman, Vogel, Hollingworth, & Luck, 2009), and decay (Zhang & Luck, 2009). The neural signatures of VWM have also been explored by fMRI (e.g., Harrison & Tong, 2009; Todd & Marois, 2004; Xu & Chun, 2006) and ERP (e.g., Vogel, & Machizawa, 2004; Vogel, McCollough, & Machizawa, 2005; Gao et al., 2009). How VWM impacts other perceptual and cognitive processes has also been investigated, including the role of VWM in transsaccadic integration (e.g., Hollingworth, Richard, & Luck, 2008; Irwin & Andrews, 1996), mental rotation (e.g., Hyun & Luck, 2007), and emotional processing (e.g., Kensinger & Choi, 2009). VWM has also been studied with different populations, including young infants (Oakes, Ross-Sheehy, & Luck, 2006; Ross-Sheehy et al., 2003), children (Riggs, McTaggart, Simpson, & Freeman, 2006), and patients with neuropsychological disorders (e.g., Gold, Wilk, McMahon, Buchanan, & Luck, 2003).

These studies greatly enhance our understanding of the nature of VWM, and make this line of research one of the most active and successful areas of cognitive science. Nevertheless, compared with the abundant evidence of the properties of VWM, it is still far from clear (a) why those properties exist, and (b) how to interpret

Tao Gao, Zaifeng Gao, Jie Li, Zhongqiang Sun, and Mowei Shen, Department of Psychology and Behavior Sciences, Zhejiang University, Hangzhou, China.

The current article is based on a research project Tao Gao conducted when he was a master student at Zhejiang University. We thank Geoffrey Woodman, Liqiang Huang, Brandon Liverence, Tim Vickery, and Brian Scholl for insightful comments on various versions of the draft. This research is supported by the National Natural Science Foundation of China (No. 30870765), Key Project of Humanities and Social Sciences, Ministry of Education (No. 07JZD0029), the National Foundation for Fostering Talents of Basic Science (No. J0730753), and the Fundamental Research Funds for the Central Universities.

Correspondence concerning this article should be addressed to Mowei Shen, Department of Psychology and Behavior Sciences, Xixi Campus, Zhejiang University, Hang Zhou, P.R. China. 310028. E-mail: mwshen@zju.edu.cn

Fn1 [1] Visual working memory (VWM) is also referred as visual short term memory (VSTM). It occurs to us that the field as a whole does not distinguish these two terms, as they are usually used interchangeably without specific clarifications. In the current article, we hope to highlight visual memory's active interaction with on-line perception, instead of passive maintaining the outputs of perception. From this perspective, we feel VWM may be a more appropriate term. However, this does not suggest that the topic of our paper is different from studies of VSTM.

existing findings within a framework that is intrinsically coherent with theories of visual perception and attention, which are the other two major topics of visual cognition.

Interaction Between VWM and Visual Perception

Perception and VWM are two highly interactive processes. One important function of VWM is to enable coherent and stable visual experience across various changes and disruptions (e.g., saccades, occlusions, and changes of the perceiver's view, etc.). Studies of transsaccadic integration show that VWM is critical for integrating information collected across separate gaze fixations (Hollingworth et al., 2008; Irwin & Andrews, 1996). The function of VWM has also been revealed by studies of anorthoscopic perception, in which observers view a moving object through a slit. Although at each moment, only a fragment of the object is visible, the whole image of that object can nevertheless be readily perceived (Parks, 1965). In such circumstances, VWM plays an important role in constructing the whole image by integrating the fragments viewed at different times (e.g., Rock, 1981). The above studies emphasize the importance of VWM when perception is interrupted for specific reasons.

Another prominent theory, the visual routine theory, goes even further, suggesting that VWM is an intrinsic component of online perception, even when the perceptual processing is not interrupted for any reason (e.g., Roelfsema, 2005; Ullman, 1984). According to this theory, the first stage of visual perception is the "bottom-up", automatic creation of base representations by several "elemental operations" (Marr, 1982). In the second stage, the "elemental operations" are arranged into sequences to form visual routines which are applied to the base representations. In addition to the base representations to which visual routines are applied initially, representations are also being created and modified in the course of executing visual routines. These gradually modified representations are referred as "incremental representations". During online perception, it is necessary to place these intermediate representations into VWM, so that the outcomes of different visual routines can be accumulated, and the incremental representations themselves can be temporally retained for further use, until the final coherent perceptual representation is constructed. The distinction between "final output" of perception and "incremental representation" can be matched to other theories of visual perception. For instance, the "final output" of perception can be assumed as detailed information assembled by focal attention (Treisman & Gelade, 1980; Wolfe, 1994). It is robust over time, and won't be erased by subsequent stimuli (Chun & Potter, 1995). In contrast, the "intermediate representation" is created automatically and pre-attentively, which is volatile and only contains low-resolution information (Rensink, 2000).

The Perceptual Root of VWM

The relationship between VWM and online perception has great implications for investigations of VWM mechanisms. During the last 50 years, there has been a tremendous progress in understanding the mechanisms of visual perception. Many existing studies of VWM implicitly assume that VWM is a higher-level processing stage subsequent to visual perception, and store the final outputs of perception. We refer this perspective as the *Final-Output* model. According to this model, it is unclear how theories of visual perception and guide the investigation of VWM, as by the time storage in VWM starts, visual perception has already been terminated.

In contrast, here we advocate an *Interactive* model of VWM. This model assumes that VWM actively operates within different stages of perception, storing different types of intermediate perceptual representation, which are gradually assembled until the final coherent perceptual representation is constructed. This model implies that online visual perception and storage in VWM are intermixed with one another. Therefore, theories of visual perception are not just about perception itself, but also constrain the mechanisms of VWM.

The Interactive model is consistent with recent studies showing how the storage in VWM can be impacted by factors that are also important for visual perception. For instance, there are studies focused on the effects of perceptual discriminability on the capacity of VWM (e.g., Alvarez & Cavanagh, 2004; Awh et al., 2007) and the role of attention in perceiving and storing feature bindings (e.g., Johnson, Hollingworth, & Luck, 2008; Wheeler & Treisman, 2002). Recent studies have also directly explored how perceptual information is dynamically consolidated in VWM (Vogel et al., 2006; Woodman & Vogel, 2008). These studies collectively demonstrate that investigation of the storage of visual stimuli in VWM can significantly benefit from considering how those visual stimuli are perceived in the first place.

The close interaction between VWM and perception is also supported by the recent discovery that information stored in working memory can automatically and even irresistibly bias the deployment of attention in perceptual tasks (for a review, see Soto, Hodsoll, Rotshtein, & Humphreys, 2008). Neuroscientific studies also have revealed that VWM tasks not only activate brain regions for relatively high-level cognitive processes, but also recruit brain areas for sensory and perceptual coding of the visual stimuli (e.g., Awh & Jonides, 2001; Serences, Ester, Vogel, & Awh, 2009; Silvanto & Cattaneo, 2010). The sensory-recruitment theory of VWM is consistent with the hypothesis that VWM is not just a higher level process subsequent to perception, but is capable of active interaction with online perception.

Object-Based Storage in VWM

In the current study, we apply the Interactive model of VWM to studies of object-based storage, which is one of the most important aspects of the mechanism of VWM. Luck and Vogel (1997) originally found that memorizing the conjunction of multiple features (e.g., color and orientation) is as good as memorizing each individual feature. This result indicates that objects, instead of individual features, are the units of VWM. Object-based storage is provocative, not only because it provides an elegant model for VWM, but also because it fits beautifully with studies of object-based attention (e.g., Duncan, 1984; Egly, Driver, & Rafal, 1994; Watson & Kramer, 1999) and the "object file" theory in general (Carey & Xu, 2001; Gao & Scholl, 2010; Kahneman et al., 1992; Mitroff, Scholl, & Wynn, 2004). Given these object-based effects, it seems that perception, attention, and VWM can be explained by a unified framework: at the perceptual stage, a visual display is segmented into different objects, serving as the potential targets of selective attention. The coherent object representation assembled

tapraid5/zfn-xhp/zfn-xhp/zfn00611/zfn2701d11z | xppws | S=1 | 9/20/11 | 9:55 | Art: 2010-1941

by attention can then be stored into the object files in VWM. This framework fits the Final-Output model very well.

Unfortunately, this framework is not supported by a number of follow up studies. It has been shown that memory performance drops when the memory objects consist of multiple parts (Olson & Jiang, 2002; Wheeler & Treisman, 2002; also see the discussion in Vogel, Woodman, & Luck, 2001, regarding the distinction between parts and objects), or complex information (e.g., irregular polygons, Alvarez & Cavanagh, 2004). More recent studies reveal that VWM can hold a fixed number of objects, but only with limited resolution (Awh et al., 2007; Zhang & Luck, 2008).

Revealing the limitations of object-based storage is a critical step for understanding the nature of VWM. Based on these findings, the current study is aiming at addressing the following issues: (a) what is the origin of these object representations with low resolution, (b) how are these object representations are constructed in VWM, and (c) how to interpret this type of object representations with a framework that is coherent with theories of perception and attention.

Dissociated Mechanisms in VWM for Storing Objects

Here we introduce theories of visual perception to reveal how VWM stores object representations. VWM actively engages into online perception, instead of passively storing the final outputs of perception. Given their dynamic interaction, we hypothesis that the mechanism of VWM is characterized by the *processes* of constructing perceptual representation, instead of the properties of *the final outputs* of perception. This hypothesis implies that theories of different stages of online perception can serve as powerful tools for investigating the mechanism of VWM.

Visual perception has traditionally been dichotomized into two stages of processes (Neisser, 1967). The first stage consists of parallel preattentive processing, which enables a quick detection of distinctive features; the second stage is attentive processing, which binds multiple features from the same object, and recognizes its detailed information (Treisman & Gelade, 1980; Wolfe, 1994). Theories of object recognition have been developed based on this two-stage model of visual search. According to the "Preattentive Object File" theory, visual scenes are segmented into individual objects during preattentive processing (Wolfe & Bennett, 1997). These preattentive objects exist as receptacles holding local features such as color and shape. However, the conjunction of homogenous features within each "object file" (e.g., color-color conjunction) can only be identified by focal attention. Another closely related theory is the "Coherence" theory, which emphasizes the stability of object representations at different stages of perceptual processing (Rensink, 2000). According to this theory, prior to focal attention, low-level "proto objects" are continually formed rapidly in parallel across the visual field. These proto objects are volatile, being replaced when any new stimulus appears at their retinal locations. Focal attention acts as a metaphorical hand that grasps a small number of proto-objects, updating them into objects with a much higher degree of coherence over space and time.

Here we explore the nature of object representation in VWM by combining the Interactive model of VWM with the above theories of visual perception. The interaction between perception and VWM implies that there are dissociated mechanisms for representing information extracted at different stages of perceptual processing. In VWM, objects with limited resolution (e.g., Awh et al., 2007; Zhang & Luck, 2008) are not the final products of visual perception, but originate from the "preattentive objects" (Wolfe & Bennett, 1997) or "proto objects" (Rensink, 2000) created by parallel perception. Due to the limited resolution of parallel perception, these object representations only contain highly discriminable information. Once these proto objects are selected into VWM, online perception can then proceed to the next stage, during which focal attention incrementally assembles detailed information onto the proto objects, until the final coherent object representations are constructed (Ullman, 1984).[2] Fn2

The Current Study: Exploring the Perceptual Root of Object-Based Storage in VWM

The Interactive model of VWM motivates us to introduce theories from visual search into studies of VWM, and assume that the root of object-based storage in VWM is parallel perceptual processing. This hypothesis is tested empirically in the current project, by exploring how visual information is *selected*, *consolidated*, and *maintained* in VWM. Each of the above three stages of processing provides a unique criterion for determining whether or not certain types of perceptual information can be represented as part of an integrated object in VWM. The first one is *automatic selection*, which tests whether the task-irrelevant information contained in an object can be automatically selected in VWM along with the task-relevant information. (Experiments 1 to 3). The second one is *synchronous consolidation*, which tests whether certain information from an object can be consolidated synchronously with other information from the same object (Experiments 4 to 6). The third one is *stable maintenance*, which tests whether certain information can be stored stably as part of an integrated object (Experiments 7 to 9). According to the Interactive model, the outputs of parallel perceptual processing should meet all three criteria, as opposed to the outputs of focal attention, which should fail all of them.

Two types of visual stimuli are selected to represent information extracted at different stages (parallel vs. serial) of perception.[3] One Fn3 type of information is highly discriminable simple features (e.g., color and orientation) which can be identified via parallel process-

[2] Perhaps a more intuitive way to describe different types of perceptual information is to refer them as "basic features" and "complex features". However, since the term "feature" has been primarily used to describe information processed by parallel perception, we do not use this term here to avoid potential confusions. Instead, we refer different types of perceptual information as "highly discriminable information" and "detailed information".

[3] It is noteworthy that visual search is not a field without controversy. There are some enduring questions which have not been addressed definitively. For example, the parallel search for a wide range of stimuli seems hard to be explained by a set of primitive feature maps (e.g. Enns & Rensink, 1991; Nakayama & Silverman, 1986; Öhman et al., 2001). It is also controversial whether the distinction between the two stages of perception should be explained in terms of "preattention" vs. "attention", or "spread attention" vs. "focal attention" (e.g. Hochstein & Ahissar, 2002; Joseph, Chun, & Nakayama, 1997). Addressing these issues in detail is beyond the scope of the current project on VWM. Therefore, while realizing the existence of these controversial issues, we do try to avoid them in the current project. For instance, we will refer to the first stage of perception as parallel processing, without specifying whether it is achieved preattentively or by spread attention.

tapraid5/zfn-xhp/zfn-xhp/zfn00611/zfn2701d11z | xppws | S=1 | 9/20/11 | 9:55 | Art: 2010-1941

ing; the other is fine detailed information, processing of which requires focal attention (e.g., color-color conjunction, a small gap on the frame of a circular ring; for a review see Wolfe, 2003). We also manipulate perceptual discriminability by changing the perceptual context in which the memory materials are presented.

Section I: Automatic Storage in VWM

Experiments 1 to 3 explore the object-based automatic selection for simple features and detailed information. Provided an object is selected into VWM as an integrated whole, when observers are only required to store one feature, the other task irrelevant information from the object should also be automatically selected.

Pilot experiments in our lab showed that when observers were required to retain two features (colors and orientation) in VWM, changes in one feature dimension can distract from the detection of changes from the other dimension. In the present study, we took advantage of this “Irrelevant-distracting” effect to probe whether different types of perceptual information can be automatically selected into VWM. Participants were explicitly instructed to maintain only the target information and ignore any change from task-irrelevant information.

Experiment 1a: Automatic Storage of Arrows' Orientation

We first investigate whether a task-irrelevant simple feature can be automatically selected into VWM. Our hypothesis would be wrong if there is no automatic selection for basic features. Memory items consisted of two highly discriminable features (color and orientation). Participants were required to detect changes in the color dimension and ignore the orientation information.

Method

Participants. Twelve Zhejiang University undergraduates participated in this experiment. All had normal color vision and normal or corrected-to-normal visual acuity.

Stimuli and apparatus. Stimuli were presented in the center of a 17 in. monitor (100-Hz refresh rate), with a gray background (CIE: x = 0.341, y = 0.277; Y = 47.04 cd/m^2). Each memory item was an orientated and colored arrow (1.2° in length). Both the memory and test arrays contained four arrows that were randomly placed within an 8° × 8° region, with the constraint that the items were separated from each other by at least 2° (center to center). The color of each arrow was selected from a set of seven colors: red (CIE: x = 0.558, y = 0.356; Y = 10.33 cd/m^2), green (CIE: x = 0. 321, y = 0.604; Y = 52.55 cd/m^2), blue (CIE: x = 0.133., y = 0.073; Y = 11.96 cd/m^2), violet (CIE: x = 0.259, y = 0.153; Y = 28.04 cd/m^2), yellow (CIE: x = 0.433, y = 0.483; Y = 46.3 cd/m^2), black (Y = 1.047 cd/m^2), and white (Y = 76.43 cd/m^2). The orientation of each arrow was selected from a set of eight directions: from 0° to 315° (relative to the positive direction of the horizontal axis), in 45° steps.

Design and procedure. Schematic illustration of a single trial was depicted in Figure 1. After a 250 ms fixation and a 500 ms blank screen, the memory array was presented for 250 ms, followed by a 750 ms blank interval. The test array was then presented and maintained on the display until a response was initiated. Participants were instructed to focus on detecting color change, and ignore the orientations of the arrows. The colors in the memory and test arrays were identical, except that on 50% of the trials, the color of one item in the test array was different from that of the corresponding item in the memory array. Of most interest was the task-irrelevant orientation change: on 50% of the trials, the arrows' orientations were not changed in the test array (*No Irrelevant-Change* condition); on the remaining 50% of the trials, the orientation of *each* arrow in the test array was changed at least by 90° (*Irrelevant-Change* condition). Participants responded by pressing one of the two buttons (“F” and “J” on the keyboard) on each trial to indicate whether or not the colors of the memory and test arrays are identical. Both the response accuracy and reaction times were recorded. There were 80 randomly ordered trials in total, 40 for each of the Irrelevant-Change and No Irrelevant-Change conditions. F1

Results and Discussion

RT and accuracy are shown in the second row of Table 1a and 1b, respectively. In this and the following experiments, trials with inaccurate response were excluded from the RT analyses. There was a significant difference between the Irrelevant-Change (771 ms) and No Irrelevant-Change (729 ms) conditions, $t(11) = 4.131$, T1a T1b

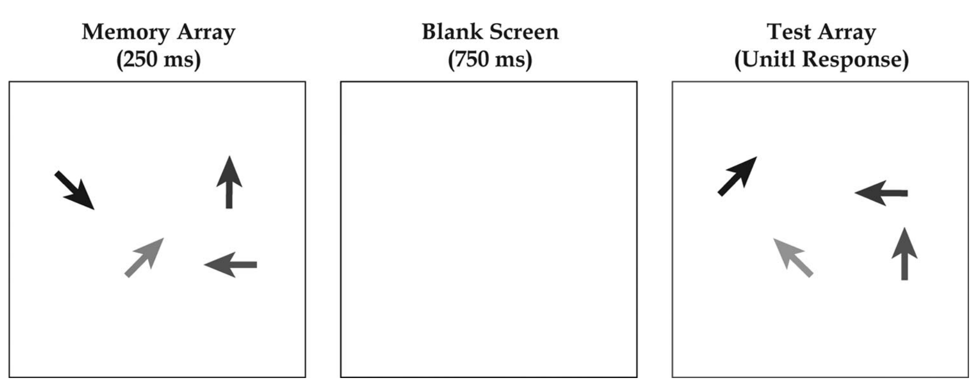

Figure 1. A schematic illustration of a single trial from Experiment 1a. The Memory array contained four oriented arrows. In this particular trial, all arrows' orientations in the test array were changed, which was irrelevant to the observers' task.

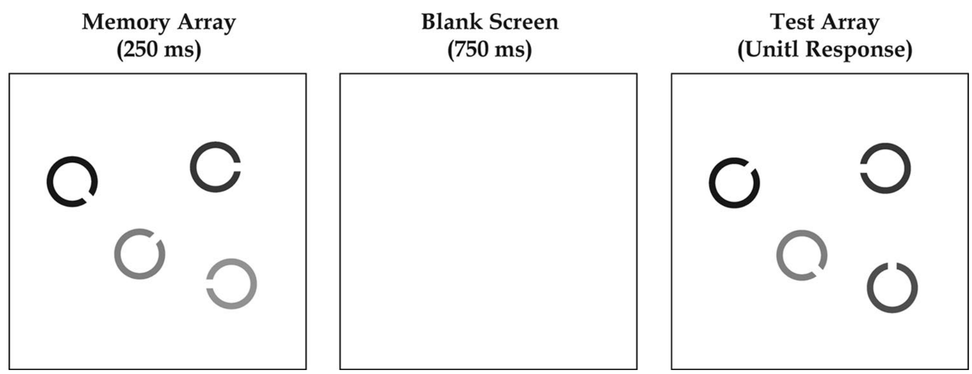

Figure 2. A schematic illustration of a single trial from Experiment 1b. Memory array contains four oriented rings. In this particular trial, the color of the yellow ring was changed to violet in the test array. All the rings' orientations in the test array were also changed, which was irrelevant to the observers' task.

p = .002, η_p^2 = .608, indicating that irrelevant orientation change impaired participants' response to color change. The accuracy of the Irrelevant Change condition (80.7%) was lower than that of the No Irrelevant-Change condition (85.7%). The difference was marginally significant, $t(11)$ = 1.974, p = .074, η_p^2 = .262. The accuracy results strongly confirmed that the significant RT difference was not due to the speed–accuracy trade-off.[4] Fn4

The current results indicate that VWM cannot just select an individual feature without processing other basic features from the same objects. Different dimensions of simple features are selected into VWM as integrated objects.

Experiment 1b: Storage of Oriented Gap

Here we investigate whether similar irrelevant-distracting effects can be obtained by changing detailed information, which requires focal attention during perception.

The task-irrelevant detailed information is produced by manipulating the orientation of a small gap on the stroke of a circular ring (see Figure 2). F2 This type of stimulus is selected for the following reasons. First, search for a ring with a particular orientation can yield a steep search slope. In fact, it has been employed as a standard material for exploring attention shifting during visual perception (e.g., Woodman & Luck, 1999; Woodman & Luck, 2003). Therefore, employing this type of stimulus in a VWM task can help to integrate findings in both visual search and VWM, a principle theoretical motivation of the current project. Second, oriented gap is similar to arrow orientation (Experiment 1a) in terms of the nature of the memory materials. In both cases, the task-irrelevant information is from several items' orientations. The only difference is whether an item's orientation is presented in a format that can be efficiently perceived or not. Therefore, comparing the storage of these two types of stimuli can highlight the impact of perceptual efficiency on VWM.

According to the Interactive model, because the oriented gap cannot be represented as part of an integrated object by the end of parallel perceptual processes, it should not be automatically encoded in VWM.

Method

This experiment was identical to Experiment 1a except as noted here. Twelve new undergraduates in Zhejiang University participated. The memory array consisted of four circular rings (1.2° in diameter). On the frame (0.1° in width) of each of the rings, there was a gap (0.2° in width). The colors of the rings and their orientations were identical to those in Experiment 1a. Participants were instructed to detect the color change and ignore the oriented gaps.

Results

The RT and accuracy were displayed in the third row of Table 1a and 1b, respectively. RT results showed that there was no difference between Irrelevant-Change (718 ms) and No Irrelevant Change (712 ms) conditions, $t(11)$ = .037, p = .722, η_p^2 = .012; there was also no difference between the accuracy of Irrelevant Change (82.8%) and No Irrelevant Change conditions (83.6%), $t(11)$ = .506, p = .623, η_p^2 = .023. The current results are consistent with the hypothesis that detailed information cannot be automatically selected into VWM.

Discussion

The limitation of the Irrelevant-distracting paradigm is that it is ultimately an indirect measurement. It can effectively demonstrate the existence of automatic storage when there is a salient Irrelevant-distracting effect. Unfortunately, it is difficult to disprove automatic storage when the distracting effect is absent. Specific to Experiment 1a and 1b, the results show that different types of task-irrelevant information are processed distinctively during the change detection task. However, it is unclear what processes in VWM are responsible for this dissociation. The change detection task involves three stages of processing: selection, maintenance, and comparison. Our hypothesis predicts that the distinction between Experiment 1a and 1b is attributable to the selection and storage of information into VWM. Alternatively, the

[4] It is interesting that irrelevant changes only influenced reaction time, but had no effect on the accuracy of the change detection task. These results suggest that reaction time is a more sensitive measurement in this particular irrelevant-distracting paradigm. It is possible that participants could ultimately distinguish irrelevant-feature change from target-feature change. However, filtering out irrelevant-feature requires an extra process, which slows down the response.

tapraid5/zfn-xhp/zfn-xhp/zfn00611/zfn2701d11z | xppws | S=1 | 9/20/11 | 9:55 | Art: 2010-1941

Table 1a
Summary of the Reaction Times Results From Experiments 1 and 2

Experiment name	Irrelevant feature	Testing condition	No-irrelevant change (ms)	Irrelevant change (ms)	T-test
Expt 1a	Arrow's Orientation	Whole Report	729	771	t = 4.131 p = .002
Expt 1b	Oriented Gap	Whole Report	712	718	t = .364 p = 0.722
Expt 2a	Arrow's Orientation	Partial Report	746	791	t = 3.518 p = .004
Expt 2b	Oriented Gap	Partial Report	796	796	t = .016 p = .988

distinction might be attributed to the comparison stage. According to this alternative hypothesis, selecting and consolidating information in VWM requires serial, attentive processing, which automatically encodes the task-irrelevant information into VWM. However, during the comparison stage, it is unnecessary for the visual system to process each individual test item with focal attention. Instead, detecting a highly discriminable feature (i.e., detecting color change) only requires a *global* comparison between the information maintained in VWM and the test array (see Jiang et al., 2000). Without focal attention, the detailed information containing the test array cannot be perceived during the global comparison. Therefore, it cannot yield any irrelevant-distracting effect. This alternative explanation is addressed in Experiment 2.

Experiment 2a: Irrelevant-Change of Basic Feature From a Single Object

Here we try to avoid the global comparison by employing a partial report method, in which only a single item is presented in the test array. Thus, the visual system has to compare the test array with individual objects maintained in VWM. According to the alternative explanation of Experiment 1b, both basic features and detailed information can cause irrelevant-distracting effects.

Note that in Experiment 1, when there is an irrelevant orientation change, *all* four items change their orientations, which may produce a global change in the test array. It is possible that without this global change, even basic features can no longer produce an irrelevant-distracting effect. Here we try to first rule out this possibility, by demonstrating that changing a basic feature of a single object can still produce a salient irrelevant-distracting effect.

Method

This experiment was identical to Experiment 1a except as noted here. Fourteen new undergraduates in Zhejiang University participated. In the test array, one colored arrow was presented at one of the four locations initially occupied by the memory items (Figure 3). Participants were instructed to focus on this single test item and detect whether its color was identical to its corresponding item in the memory array. F3

Results and Discussion

The RT and accuracy are displayed in the fourth row of Table 1a and 1b, respectively. There was a significant RT difference between Irrelevant-Change (791 ms) and No Irrelevant Change (746 ms) conditions ($t(13) = 3.518$; $p = .004$, $\eta_p^2 = .488$). The difference between the accuracy of Irrelevant-Change (83%) and No Irrelevant Change (83%) conditions was not significant, $t(13) = .090$; $p = .930$, $\eta_p^2 = .001$, showing that there was no speed–accuracy trade-off.

These results reveal that a global change is not necessary for the irrelevant-distracting effect. Changing the orientation of a single object can still disrupt the detection of color change.

These results support the hypothesis that automatic selection of basic features is based on the processing of individual objects, instead of the global structure of the memory array.

Experiment 2b: Irrelevant-Change of Detailed Information From a Single Object

Here we investigate whether the change of task-irrelevant detailed information can distract from the detection of color change with the partial report paradigm.

Method

This experiment was identical to Experiment 2a except as noted here. Thirteen new undergraduates in Zhejiang University partic-

Table 1b
Summary of the Accuracy Results From Experiments 1 and 2

Experiment Name	Irrelevant Feature	Testing Condition	No-Irrelevant Change	Irrelevant Change	T-test
Expt 1a	Arrow's Orientation	Whole Report	85.7%	80.7%	t = 1.974 p = .074
Expt 1b	Oriented Gap	Whole Report	83.6%	82.8%	t = .506 p = .623
Expt 2a	Arrow's Orientation	Partial Report	83%	83%	t = .090 p = .930
Expt 2b	Oriented Gap	Partial Report	80.7%	81.9%	t = .739 p = .474

tapraid5/zfn-xhp/zfn-xhp/zfn00611/zfn2701d11z | xppws | S=1 | 9/20/11 | 9:55 | Art: 2010-1941

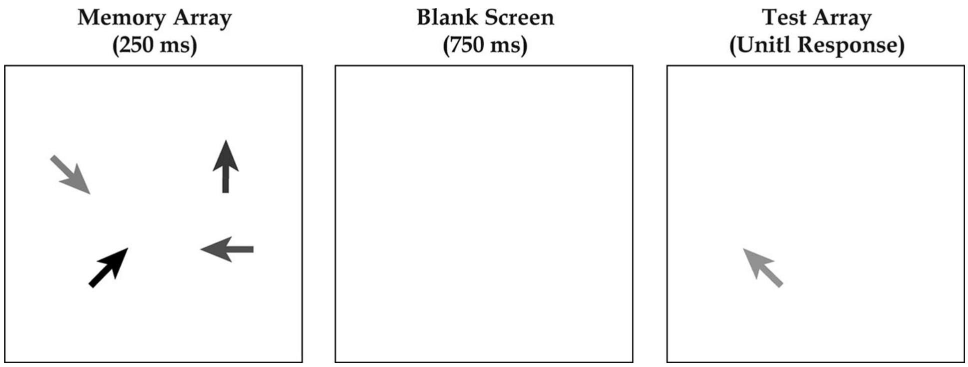

Figure 3. A schematic illustration of a single trial from Experiment 2a. There was only one item in the test array. In this particular trial, both the color and orientation of the test item were changed. The task was to detect color change and ignore the orientation change.

AQ: 2

ipated. The oriented arrows were replaced by oriented circular rings which were identical to those of Experiment 1b.

Results and Discussion

The RT and accuracy are displayed in the fifth row of Table 1a and 1b, respectively. There was no RT difference between Irrelevant-Change (796 ms) and No Irrelevant Change (796 ms) conditions, $t(12) = .017$, $p = .988$, $\eta_p^2 < .001$. The difference between the accuracy of the Irrelevant Change (81.9%) and No Irrelevant Change (80.7%) conditions was also not significant, $t(12) = .739$, $p = .474$, $\eta_p^2 = .043$.

These results do not support the alternative explanation of Experiment 1. Since the test item is the only object abruptly presented in the display, it should be able to attract focal attention (e.g., Yantis & Jonides, 1984). Moreover, the load for perceiving the test array is low, and there are abundant resources for any potential automatic processing (e.g., Lavie, 1995). However, facilitating the perception of the test item cannot produce any distracting effect from detailed information, suggesting that limitations from the comparison stage cannot explain the distinctive processes for basic features and detailed information in VWM.

Experiment 3: Color Change Versus Color-Color Conjunction Change

According to the Interactive model, there are dissociable mechanisms in VWM for retaining information extracted at different stages of perception. To explore the general validity of this model, it is important to employ memory materials other than orientation, which was focused on in Experiments 1 and 2. In Experiment 3 we investigated the automatic storage of individual colors and color-color conjunctions. Each memory object consisted of an oriented triangle containing two colors. The change of individual colors was introduced by replacing the colors of each object with new ones. The change of color-color conjunction was introduced by holding the two colors contained in each object constant, but swapping their relative positions
F4 *within* each object (see Figure 4a). Participants were required to detect changes in the orientations of the triangles, while ignoring changes in the color dimension.

The first reason for selecting color-color conjunction as the irrelevant detailed information is similar to the reason for selecting oriented gap: perception of this particular type of information has been systematically explored in visual search studies (e.g., Wolfe, 2003; Woodman & Luck, 1999, 2003). Employing it as the memory material can help to reveal the intrinsic interaction between perception and VWM. The "Preattentive Object File" theory (Wolfe & Bennett, 1997) suggests that multiple features from different dimensions are segmented into individual object files via parallel perceptual processing. However, how homogenous features conjunct (e.g., color-color conjunction) within each object file can only be resolved by focal attention. According to our hypothesis, only individual colors, instead of the color-color conjunctions, can be automatically selected into VWM.

In the second place, contrasting the storages of individual colors and color-color conjunctions can address another alternative explanation of Experiments 1 and 2. It is possible that automatic storage in VWM has nothing to do with perceptual discriminability in general, but is limited to a set of primitive feature dimensions. According to this hypothesis, automatic storage of oriented lines is attributed to a specialized memory resource dedicating to store the oriented lines. (For a detailed discussion of this model, see Olson & Jiang, 2002). For oriented rings, however, there is no such specialized resource, meaning that they cannot be automatically selected in VWM. This alternative hypothesis can be tested in the current experiment, since both individual colors and color-color conjunctions are from the same feature dimension. Whereas our hypothesis predicts that color-color conjunction cannot be automatically stored, the alternative hypothesis predicts that both types of information can be automatically stored, as long as there is a specialized memory resource for color.

Method

Participants. Twelve Zhejiang University undergraduates participated.

Stimuli. The memory array contained four orientated isosceles triangles, with a 24.8° apex angles. The length of the midline of each triangle was 1° of visual angle. The orientation of the triangle was defined by the angle between the midline of the triangle and the horizontal axis. There were 12 different angles,

tapraid5/zfn-xhp/zfn-xhp/zfn00611/zfn2701d11z | xppws | S=1 | 9/20/11 | 9:55 | Art: 2010-1941

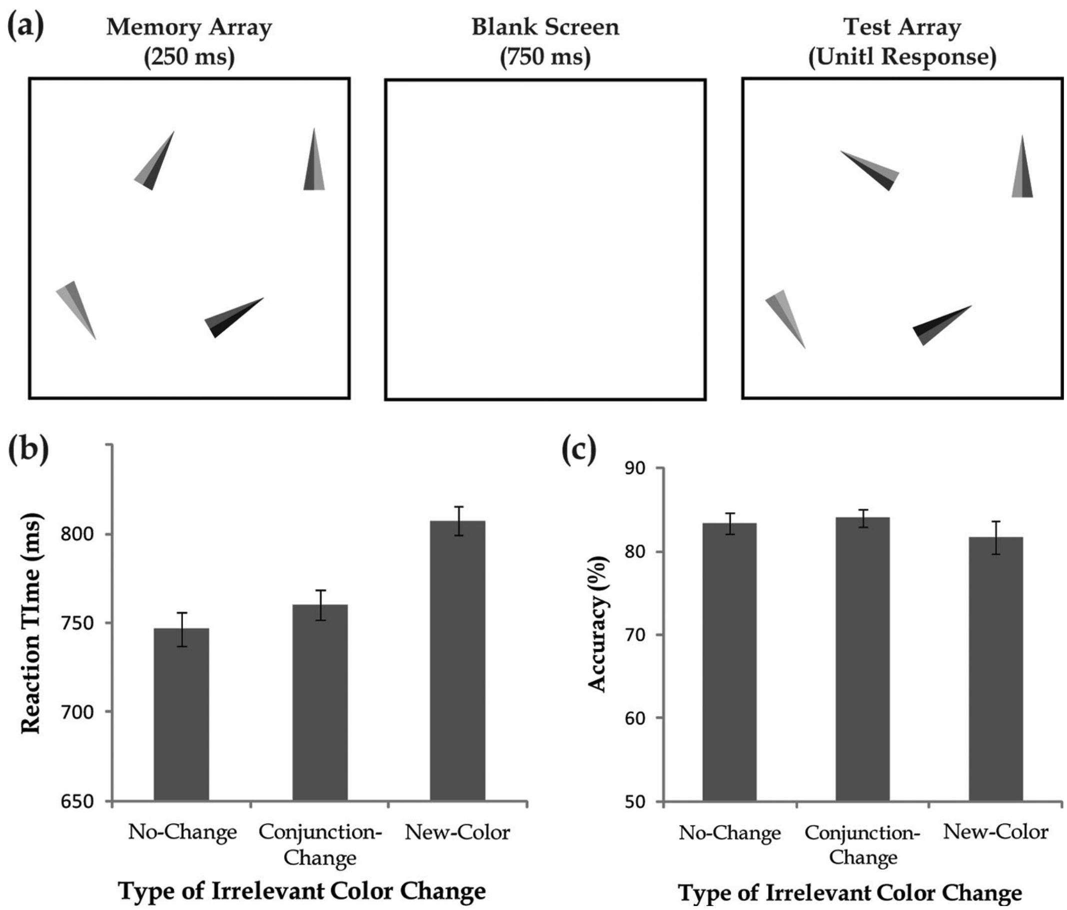

Figure 4. (a) A schematic illustration of a Conjunction-Change trial from Experiment 3. The orientation (target feature) of the upper-left triangle was changed. The colors within each triangle were swapped their positions. (b) The reaction times of detecting orientation change as a function of the type of irrelevant color change. (c) The accuracy of detecting orientation change as a function of the type of irrelevant color change.

starting from 0° with a 30° step. The triangles were divided into two parts along the midline. Each part contained a unique color, which was selected from a pool of nine colors without replacement. In addition to the seven colors used in Experiment 1, two new colors were added into the color pool, which were cyan (CIE: x = 0.181, y = 0.284; Y = 56.49 cd/m^2) and brown (CIE: x = 0.607, y = 0.351; Y = 25.43 cd/m^2).

Design and procedure. Participants were instructed to remember the orientations of the four triangles and ignore any color change. On 50% of the trials, the orientation of one triangle was changed. There were three types of irrelevant color change: the *No-Change* condition, in which colors of the triangles were kept constant; the *Conjunction-Change* condition, in which the two colors of each triangle were kept constant, but their positions within the triangle were swapped; and the *New-Color* condition, in which the two colors of each triangle were replaced by two new colors from the color pool.

Results and Discussion

The RT and accuracy are displayed in Figure 4b and Figure 4c, respectively. The most important result is that color-color conjunctions cannot be automatically selected in VWM. RT of the New-Color condition (807 ms) was significantly longer than the No-Change (747 ms, $t(11)$ = 3.971, p = .002, η_p^2 = .589) and Conjunction-Change (760 ms) conditions, $t(11)$ = 3.489, p = .005, η_p^2 = .525. RT did not differ between the No-Change and the Conjunction-Change condition, $t(11)$ = .863, p = .406, η_p^2 = .063.

Overall accuracy of the three irrelevant color change conditions were all above 80%, and the differences between them were not significant (No-Change vs. Conjunction-Change: $t(11)$ = .422, p = .681, η_p^2 = .016; No-Change vs. New-Color: $t(11)$ = .538, p = .601, η_p^2 = .026; New-Color versus Conjunction-Change: $t(11)$ = .815, p = .432, η_p^2 = .057). These results indicated that there was no speed–accuracy trade-off.

tapraid5/zfn-xhp/zfn-xhp/zfn00611/zfn2701d11z | xppws | S=1 | 9/20/11 | 9:55 | Art: 2010-1941

These results reveal that automatic storage is not determined by specific feature dimensions. There are dissociated mechanisms in VWM for perceptual information with different degrees of discriminability, even when the information is from the same feature dimension.

Intermediate Discussion of Automatic Storage

Through Experiments 1 to 3, we have obtained robust and consistent evidence showing that the outputs of parallel perception can be automatically selected and stored in VWM. Similar findings have also been reported in other recent studies (Hyun et al., 2009; Soto & Humphreys, 2009). In addition to the automatic selection of basic features, here we are particularly interested in the *dissociation* between the storage of different types of perceptual information. In contrast to basic features, there is no evidence showing that detailed information can be automatically stored as part of an integrated object. These results support the hypothesis that there are dissociated mechanisms in VWM for information extracted at different stages of perception.

However, we are indeed aware of the limitations of the Irrelevant-Distracting paradigm. Results from this paradigm are not conclusive, since one can always argue that detailed information is actually automatically selected into VWM, but for some reason, fails to impact participants' performance. We will address this issue in detail in the General Discussion. In the next two sections of the current project, we avoid this methodological limitation by employing different paradigms which can *directly* probe the object representations in VWM.

Section II: Synchronous Consolidation in VWM

The efficiencies of processing basic features and detailed information are quite different (e.g., Treisman & Gelade, 1980; Quinlan, 2003). In Experiments 4 to 6, we investigate whether there is a similar distinction during the consolidation of perceptual information in VWM. According to the Final-Output model of perception and VWM, despite the fact that an object's basic features can be perceived preattentively with high efficiency, consolidation of those features won't start until the whole object is fully selected by serial attention, which can then transfer the whole object into VWM as an integrated unit. In other words, the efficiency of consolidating basic features in VWM is constrained by the efficiency of consolidating detailed information. In contrast, according to the Interactive model, the information initially selected into VWM is the outputs of parallel perception. This model predicts that consolidating basic features should be much faster than consolidating detailed information.

To estimate the time course of consolidating different types of perceptual information, we employed a paradigm developed by Vogel, Woodman, and Luck (2006). The stimulus onset asynchrony (SOA) between the memory array and its backward mask was manipulated. The time course of consolidation was estimated by the accuracy of change detection as a function of the Memory-Mask SOA.

Experiment 4: Synchronous Consolidation of Basic Features

Here we estimated the time course of consolidating basic features in VWM. Participants were required to maintain (a) only the color feature, (b) only the orientation feature, or (c) both the color and orientation features.

Method

Participants. Twelve undergraduates from Zhejiang university participated in this experiment.

Stimuli. The memory array contained two arrows (1.5°) whose colors were identical to those of Experiment 2. Each backward mask consisted of 4 arrows whose endpoints equally spread over an imaginary circle with a 2.4° diameter. The colors of each arrow in the backward mask were randomly selected from the same color set as the memory items, with the constraint that each mask had four different colors.

Design and procedure. There were three blocks: *Single-Color*, *Single-Orientation* and *Two-Feature* groups. The order of these blocks was counterbalanced across subjects. Each block was named by the nature of the change detection task. Within each block, the duration of the memory array was randomly selected from 40, 70, 100, and 130 ms. Schematic illustration of a single trial from the Two-Feature block is depicted in Figure 5a. After F5 a 250 ms fixation and a 500 ms blank interval, a memory array was presented, immediately followed by a 200 ms backward mask. Following the mask, a word was presented for 200 ms, serving as the cue to indicate which feature (color or orientation) might be changed in the test array. After the offset of the cue, a blank interval was presented, whose duration was varied as a function of the duration of the memory array so as to keep the interval between the onsets of the memory and test arrays constant at 1,000 ms. The test array remained on the display until a response was initiated. On half of the trials, the cue was a Chinese word meaning "color"; on the remaining half of the trials, there was an "orientation" cue. On half of the trials, the test array was identical to the memory array; on the other half of the trials, one of the test items was changed in the cued feature. Since in the Two-Feature block, the cue was varied from trial to trial, and was presented after the offset of the memory array, participants were forced to consolidate both features into VWM when the memory array was visible. In the Single-Color and Single-Orientation blocks, the cue did not change across the whole block (e.g., in the Single-Color group, the cue was always "color"). Within each block, there were 32 trials for each combination of cued feature and Memory-Mask SOA.

Results and Discussion

The change detection performance as a function of feature dimension and Memory-Mask SOA is depicted in Figure 5b. The highlight of these results is the similar time courses of consolidation across different conditions. Performance of the Single-Color and Single-Orientation blocks was labeled as Color-Blocked and Orientation-Blocked, respectively. Results of detecting color and orientation changes in the Two-Feature block were labeled as Color-Mixed and Orientation-Mixed, respectively. The most robust effect came from the Memory-Mask SOA, $F(3, 33) = 65.527$, $p < .001$, $\eta_p^2 = .856$, which showed that memory performance

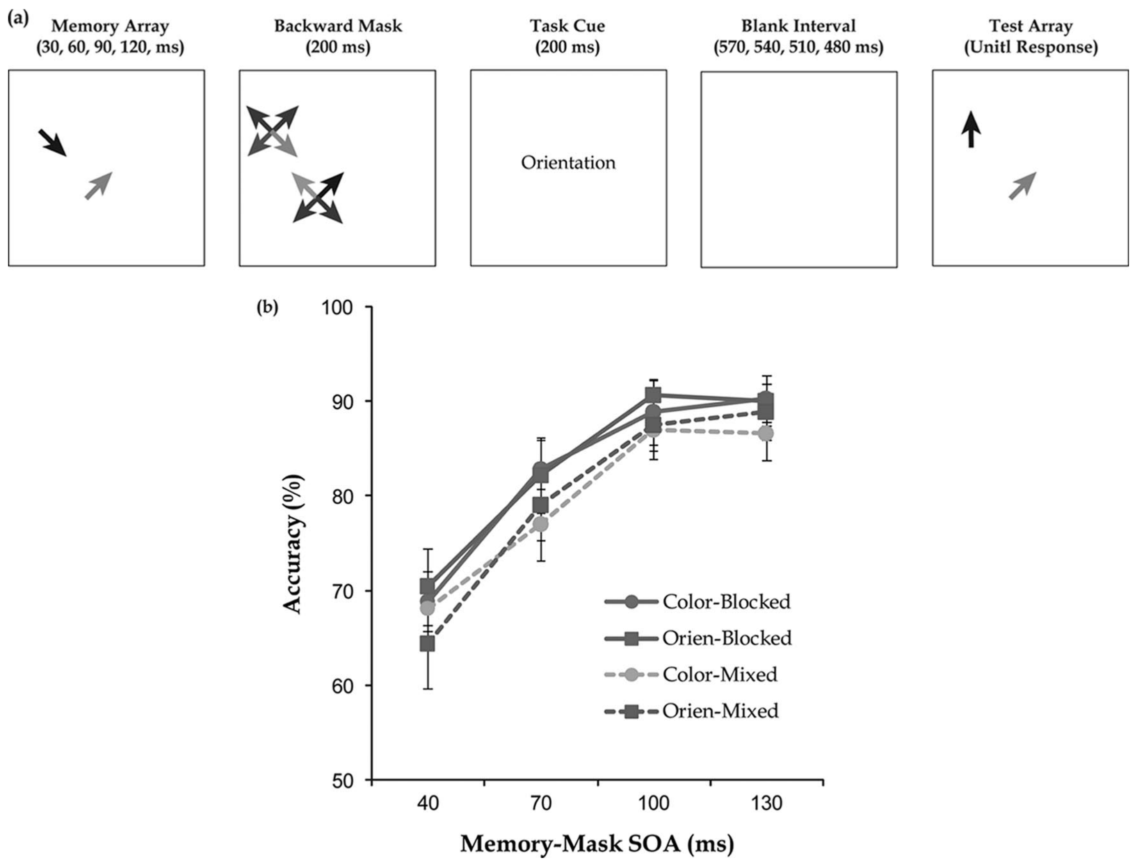

Figure 5. (a) A schematic illustration of a single trial from Experiment 4. After the offset of the backward mask, a cue (Color or Orientation) was presented to indicate which feature can be changed in the test array. In the trial depicted here, the cue was "Orientation". The orientation of the upper-left arrow was rotated by 135° in the test array. In the real experiment, the cues were Chinese words. (b) Results obtained from Experiment 4. The change detection performance is depicted as a function of Feature Dimension, Task Condition (mixed or blocked), and Memory-Mask SOA.

gradually increased as the Memory-Mask SOA was prolonged. The main effect of the Memory Tasks (Color-Blocked, Color-Mixed, Orientation-Blocked, Orientation-Mixed) was not significant, $F(3, 33) = 2.603$, $p = .116$, $\eta_p^2 = .180$. Most importantly, the interaction between Memory-Mask SOA, and Memory Tasks was not significant, $F(9, 99) = .487$, $p = .881$, $\eta_p^2 = .042$, indicating that the time course of consolidation was not impacted by the Fn5 manipulation of memory task.[5]

The above results suggest that different dimensions of simple feature can be consolidated in VWM synchronously. Consolidation of one feature won't slow down the consolidation of another feature. This pattern of results supports the hypothesis that the outputs of parallel perceptual processing can be consolidated in VWM as integrated objects.

The consolidation of multiple features is also explored in one recent study (Woodman & Vogel, 2008), which shows that the performance of consolidating color is significantly better than that of color-orientation conjunction and color-shape conjunction. However, we find it difficult to directly compare these results with ours. Perhaps the most important reason is that in Woodman and Vogel (2008), the performance of consolidating two features was not identical. The performance of maintaining color was about 10% higher than that of orientation. In addition, the performance of consolidating color-orientation conjunction

[5] Further analyses suggested a small but consistent effect of task set on the overall memory performance. Direct comparisons showed that performance of the Color-Blocked condition (82.7% on average) was higher than that of the Color-Mixed condition (79.7%), which was marginally significant, $F(1, 11) = 4.330$, $p = .062$, $\eta_p^2 = .282$. Similarly, performance of the Orientation-Blocked condition (83.3%) was significantly higher than that of the Orientation-Mixed condition (80.0%; $F(1, 11) = 8.864$, $p = .013$, $\eta_p^2 = .446$). This $\sim$3% difference suggests that top-down task set influences the storage of basic features. This effect can either be attributed to (a) top-down control modulate the salience of individual feature maps (e.g. Wolfe et al., 2003); or (b) the feature cue is constant in the blocked conditions, but varies from trial to trial in the mixed conditions, which can disrupt the maintenance of VWM. In any case, there is no hint that memory task can also influence the efficiency of consolidation, because in both feature dimensions, the interaction between memory task and SOA was not significant (Color-Mixed vs. Color-Blocked: $F(3, 33) = 1.043$, $p = .386$, $\eta_p^2 = .087$; Orientation-Mixed vs. Orientation-Blocked: $F(3, 33) = 0.557$, $p = .647$, $\eta_p^2 = .048$).

tapraid5/zfn-xhp/zfn-xhp/zfn00611/zfn2701d11z | xppws | S=1 | 9/20/11 | 9:55 | Art: 2010-1941

was even better than that of the orientation condition. Therefore, it was unclear whether consolidating two features was slower than one feature, which is the main focus of the current study.

Experiment 5: Dissociated Consolidations for Basic Features and Detailed Information

In Experiment 5, we investigated whether information extracted at different stages of perception can be consolidated in VWM synchronously. Instead of oriented arrows, we employed oriented squares with gaps.

Method

This experiment was identical to the Two-Feature condition in Experiment 4 except as noted here. Eight new Zhejiang University undergraduates participated. The memory array contained two squares (1.25°), with a gap (0.2°) on the frame. Their colors were selected from a set containing red, blue, green, and white. The orientations of their gaps were selected from four orientations: up, down, left, and right. Each backward mask was a big square (1.3° × 1.3°) containing four small square rings (0.65° × 0.65°). The color of each small square was randomly selected with the
F6 constraint that each mask had four different colors (see Figure 6a). Across trials, the Memory-Mask SOA was randomly selected from the following durations: 70, 130, 190, and 250 ms. Participants were explicitly instructed to remember both colors and oriented gaps.

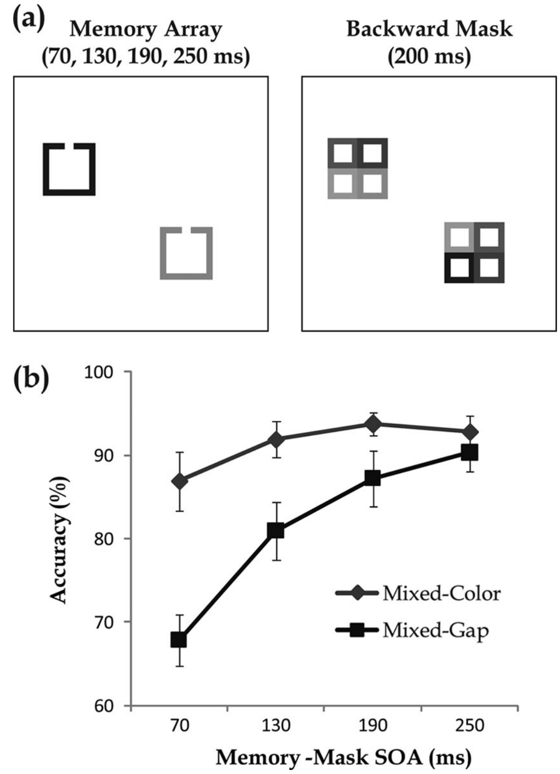

Figure 6. (a) Illustration of the square ring and the backward mask employed in Experiment 5. (b) Results obtained from Experiment 5. The change detection performance is depicted as a function of Feature Dimension, and Memory-Mask SOA.

Results and Discussion

Change detection performance as a function of Memory-Mask SOAs is depicted in Figure 6b. The most important finding is that consolidating color is much faster than orientation of the gap. A repeated ANOVA analysis showed that the main effects of target dimension: $F(1, 7) = 17.265$, $p = .004$, $\eta_p^2 = .712$, Memory-Mask SOA: $F(3, 21) = 27.090$, $p < .001$, $\eta_p^2 = .795$, and their interaction: $F(3, 21) = 4.799$, $p = .011$, $\eta_p^2 = .407$ were all significant. In the Color condition, the performance reached plateau as early as 130 ms. In the Oriented-Gap condition, the performance was high at the 250 ms condition (90.3%) and did not differ from that in the Color condition (92.8%), $t(1,7) = 1.155$, $p = .286$, $\eta_p^2 = .160$, indicating that given sufficient time, two rings' orientations can be consolidated as well as their colors. However, the performance at detecting gap orientation changes was lower than that of detecting color changes at the 70 ms, $t(1,7) = 4.303$, $p = .004$, $\eta_p^2 = .726$ and 130 ms conditions, $t(1,7) = 2.851$, $p = .025$, $\eta_p^2 = .537$.

These results reveal that consolidating basic features is much faster than consolidating detailed information in VWM.

Experiment 6: Dissociated Consolidations With Larger Set Size

In the previous two experiments, to guarantee that the dissociation in the time course of consolidation was not due to the lack of memory resource, the set size of the memory array was fixed at two. With this small set size, the performance of detecting a change in the oriented gap reached the same plateau as that of detecting a color change. However, this set size is much smaller than the maximal number of objects that can be stored in VWM (e.g., Luck & Vogel, 1997). This limitation makes it impossible to explore the consolidation process within a longer time course. We directly addressed these issues by increasing the set size to four in Experiment 6. To optimize the performance of consolidating oriented gaps, participants were *only* required to detect changes from that dimension. As a comparison, a separate group of participants were required only to detect color changes.

Method

The experiment was identical to Experiment 5 except as noted here. Nineteen new undergraduates participated, 10 for the *Oriented-Gap* group, and nine for the *Color* group. There were four colored square rings in both the memory and test arrays. In the Oriented-Gap group, one item's orientation in the test array was changed on half of the trials. The Memory-Mask SOA was randomly selected from the following durations: 84, 132, 228, 324, 420, and 516 ms. In the Color group, one item's color in the test array was changed on half of the trials. The Memory-Mask SOA was randomly selected from one of the following durations: 84, 108, 132, 156, and 192 ms. Within each group, there were 40 trials for each Memory-Mask SOA.

tapraid5/zfn-xhp/zfn-xhp/zfn00611/zfn2701d11z | xppws | S=1 | 9/20/11 | 9:55 | Art: 2010-1941

Results and Discussion

F7 Change detection performance as a function of the target dimension and Memory-Mask SOA is depicted in Figure 7. The main effects of the Memory-Mask SOA were significant in both the Oriented-Gap: $F(5, 45) = 4.525$, $p = .002$, $\eta_p^2 = .371$ and the Color: $F(4, 32) = 9.684$, $p < .001$, $\eta_p^2 = .635$ conditions. Most importantly, the curves in Figure 7 clearly reveal that the efficiency of consolidating oriented gap is much lower than that of color. The accuracy of detecting a color change stops increasing at 156 ms. In the Oriented-Gap condition, in contrast, it takes 420 ms for the performance to reach plateau.

These results indicate that the efficiency of consolidating oriented gap is much slower than that of color. These results support the hypothesis that the object representations initially consolidated in VWM are based on the outputs of parallel perception. Since detailed information is not contained in such intermediate outputs of visual perception, it cannot be consolidated synchronously with basic features.

Section III: Maintaining Basic Features and Detailed Information in VWM

Experiments 7 to 9 focus on whether different types of perceptual information can be maintained stably as part of an integrated object in VWM. When the visual display is not crowded, basic features can be segmented into individual objects via parallel perceptual processing. According to the Interactive model, these feature conjunctions are the raw materials of VWM. Therefore, they are unlikely to fall apart during the maintenance in VWM. Interestingly, this model also predicts that detailed information is incrementally added to object representations in VWM. As a result, the conjunctions of detailed information may not be stable. This prediction is tested in this section.

Previous studies have already explored the maintenance of the conjunction of basic features (e.g., Johnson et al., 2008; Wheeler & Treisman, 2002). In some critical experiments of these studies,

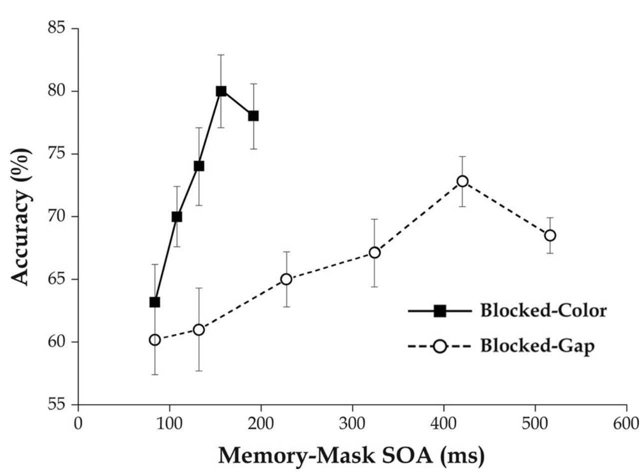

Figure 7. Results obtained from Experiment 6. The change detection performance as a function of Feature Dimension, and Memory-Mask SOA. The efficiency of consolidating oriented gap is much slower than that of color.

a single test item was presented in the center of the screen, and observers were required to identify whether the test item was identical to one of the memory items. In the *New Feature* condition, the test item contained a new feature (color or shape), which was not presented in the memory array. In the *New Conjunction* condition, the two features of the test item had been presented in the memory array, but from different objects. The results revealed that the performance of the New Conjunction condition was not significantly different from the worst performance of remembering individual features, indicating that as long as individual features are successfully stored, their conjunctions are maintained stably.

In this section, we first replicate the findings of previous studies, demonstrating that basic features can be maintained as integrated objects when presented in sparse displays. We then move forward to explore whether the conjunctions can still be maintained stably when (a) objects are composed of detailed information, and (b) objects are presented in a crowded display, in which feature conjunctions within each object cannot be resolved by parallel perception.

Experiment 7a: Maintaining Color-Shape Conjunctions in Sparse Displays

Here we attempt to replicate previous findings that the conjunction of basic features can be maintained stably in VWM. The memory materials and the logic of the current experiment are similar to those of Wheeler and Treisman (2002).

Method

Participants. Sixteen Zhejiang University undergraduates participated in this experiment.

Stimuli, design, and procedure. The details of the current experiment were identical to those of Experiment 1a, except as noted here. The memory items consisted of four colored shapes. The shapes of these items were randomly selected from one of eight distinctive shapes (see Figure 8a) without repetition. The F8 memory array was present for 200 ms, followed by a 1,000 ms blank interval. A *single* test item was then presented on the center of the screen and lasted until a response was initiated (Figure 8b). On 50% of the trials, the test item was identical to one of the memory items (*No Change* condition). On the remaining 50% of the trials, the test item was changed from one of the memory items. According to the way in which the test item was changed, these 50% of trials could fall into three categories: (a) the test items' color was changed to a new color which was not presented in the memory array (*New Color* condition, 12.5% of the trials); (b) the test item's shape was changed to a new one (*New Shape* condition, 12.5%); and (c) the test item's color and shape were all presented in the memory array, but from different items, so that the features were old but their conjunction was new (*New Conjunction*, 25%). Participants were required to make a "Change" response as long as the test item was not exactly one of the four memory items, regardless of the nature of the change. There were 160 randomly ordered trials in total.

Results and Discussion

The percentage of correct responses as a function of change type is depicted in Figure 9. In the No Change condition, the percentage F9

tapraid5/zfn-xhp/zfn-xhp/zfn00611/zfn2701d11z | xppws | S=1 | 9/20/11 | 9:55 | Art: 2010-1941

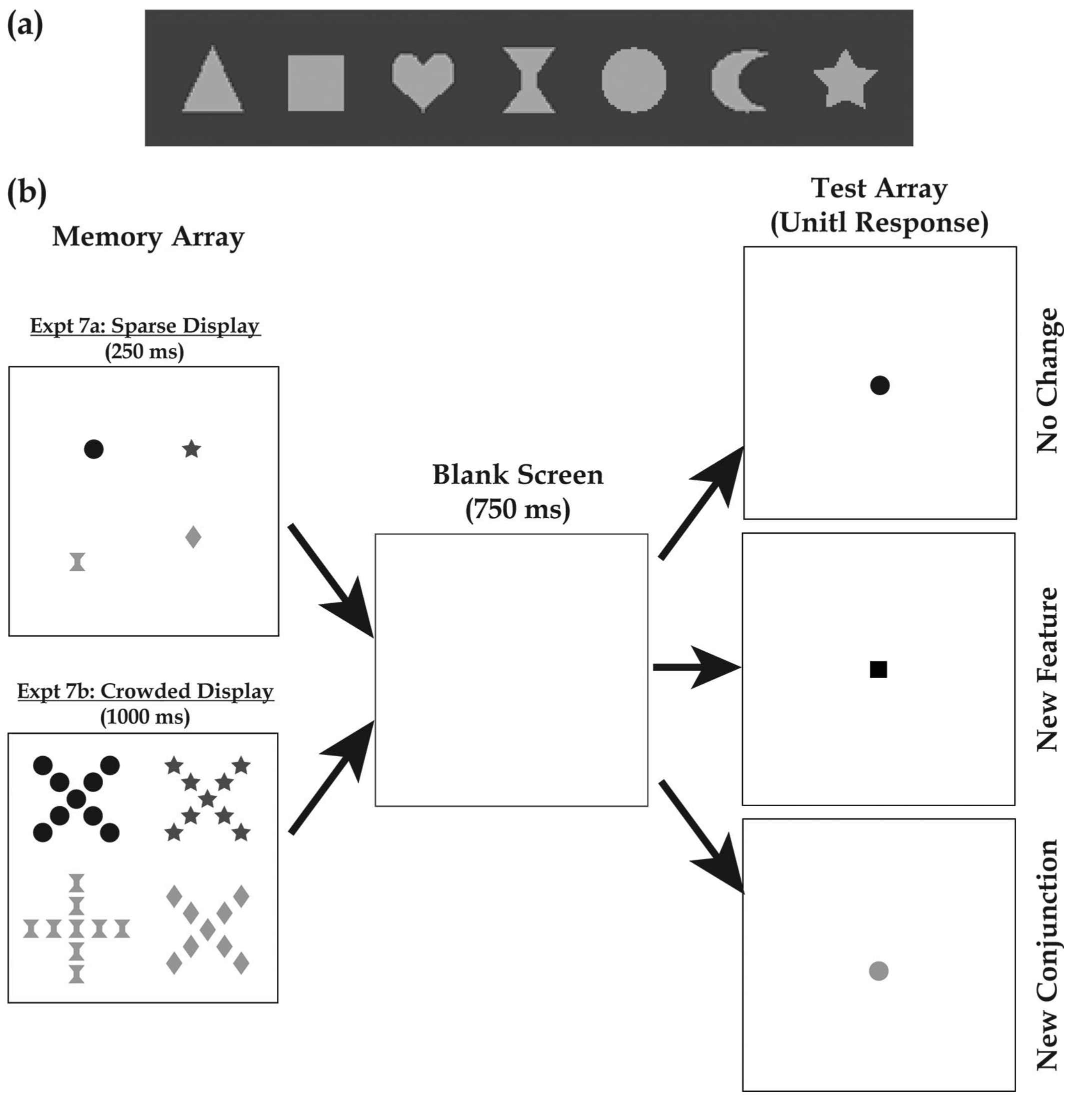

Figure 8. (a) Illustration of the shapes employed in Experiments 7a and 7b. (b) Schematic illustration of trials from Experiments 7a and 7b. In Experiment 7a, the memory items were four individual objects in a sparse display. In Experiment 7b, the memory items were presented as local components of four global objects in a crowded display.

of correct responses was 79.1%. The performance of the New Color condition (95.6%) was significantly better than that of the New Shape (87.8%) condition, $t(15) = 3.571$, $p = .003$, $\eta_p^2 = .460$, presumably because the color change in the current experiment was more salient. Of most interest, there was no significant difference between the performance in the New Shape (which was the worst feature condition) and the New Conjunction (89.8%) conditions, $t(15) = .964$, $p = .350$, $\eta_p^2 = .058$. These results replicated previous findings (Johnson et al., 2008; Wheeler & Treisman, 2002), showing that color-shape conjunction can be maintained stably in VWM.

Experiment 7b: Maintaining Color-Shape Conjunctions in Crowded Displays

The core assumption of the Interactive model is that how information is processed during perception can directly determine how such information is going to be stored in VWM. We tested this hypothesis in the previous experiments by employing various types of memory materials with different perceptual discriminability. Here we move forward and test the Interactive model in an even more extreme way: using exactly the same memory material (e.g., colored shapes), and manipulating the perceptual discriminability by changing the context in which the memory material is presented.

We achieve this goal by taking advantage of the "global precedence" effect, which is a well-known perceptual phenomenon (Navon, 1977). Specifically, we present the memory items in a crowded display, in which each memory item is a local component of a global object (see the lower-left part of Figure 8). In such a display, the proto object representations created by parallel perceptual processing should reflect the properties of those global objects. In other words, parallel processing will spontaneously

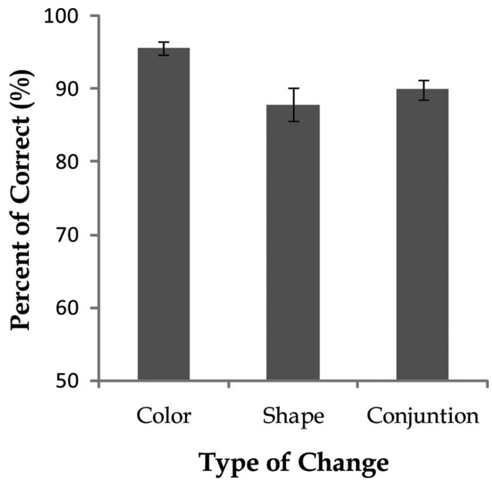

Figure 9. The results obtained from Experiment 7a. The performance of detecting change is depicted as a function of the type of change.

segment the color feature as part of the global objects, instead of binding it with each local shape. Our hypothesis predicts that when basic features are presented in such displays, the maintenance of their conjunctions will become unstable.

Method

This experiment was identical to Experiment 7a except as noted here. Sixteen Zhejiang University undergraduates participated. In each quadrant of the memory array, there were nine identical items, whose color and shape were randomly selected. These local items within each quadrant constructed a global object, whose shape was randomly selected from "+" and "×". The endpoints of the "four arms" of the global shape were equally spread on the boundary of an imaginary circle (4° in radius). The distance between the center of each global object and the center of the display was 7.8°. Participants were required to ignore the global objects and focus on remembering the four color-shape conjunctions of the local items. To guarantee that the memory array could be fully perceived, the duration of the memory array was increased
Fn6 to 1,000 ms.[6] Therefore, the results should reflect the limitations of VWM, instead of perception.

Results and Discussion

In the No Change condition, the percentage of correct responses was 74.9%. The percentage of correct response as a function of
F10 different types of change is depicted in Figure 10a. The difference between the New Color (88.9%) and the New Shape condition (83.1%) was not significant, $t(15) = .864$, $p = .401$, $\eta_p^2 = .047$. Most importantly, the performance of the New Conjunction (69.7%) condition was worse than that of the New Color, $t(15) = 5.354$, $p < .001$, $\eta_p^2 = .656$ and the New Shape, $t(15) = 4.558$, $p < .001$, $\eta_p^2 = .581$ conditions.

To make sure that the contrast between Experiment 7a and the current experiment is not due to the fact that the overall accuracy here (77.7%) was lower than that of Experiment 7a (85.0%), we further analyzed data from participants whose overall performance was higher than 80%. Data from six subjects passed this criterion. The results were depicted in Figure 10b. The performance of detecting New Conjunction (75.8%) was still significantly lower than that of New Shape (87.5%; $t(5) = 2.646$; $p = .046$, $\eta_p^2 = .583$).

The current results indicate that the conjunction of simple features can no longer be maintained stably when each memory item is presented as a local component of a global structure. These results are consistent with the hypothesis that storage in VWM is not only determined by the nature of the memory material itself, but is also directly impacted by whether the information can be represented as integrated objects at the end of parallel perceptual processing.

Experiment 8a: Maintaining Color-Orientation Conjunctions in Sparse Displays

In the next three experiments, we explored whether the conclusions of Experiment 7a and 7b can be generalized to other memory materials. We investigated whether the color-orientation conjunction can be maintained stably in sparse displays. Testing our hypothesis with different feature dimensions is important, especially because recent studies using different memory materials have reported opposite findings about the nature of VWM (Bays & Husain, 2008; Zhang & Luck, 2008).

Method

This experiment was identical to Experiment 7a except as noted here. Eleven Zhejiang University undergraduates participated. The memory array consisted of four colored arrows identical to the
stimuli used in Experiment 1b (Figure 11a). The test item could be F11
changed in three ways: New Color, New Orientation, and New Conjunction.

Results

In the No Change condition, the percentage of correct responses was 81.7%. The percentage of correct response as a function of
different types of change was depicted in Figure 12. The difference F12
between New Color (98.0%) and New Orientation conditions (88.9%) was significant, $t(10) = 2.63$, $p = .025$, $\eta_p^2 = .410$. The difference between New Conjunction (86.4%) and New Orientation conditions was not significant, $t(10) = .781$, $p = 0.453$, $\eta_p^2 = .057$.

These results are consistent with previous findings that the conjunction of basic features presented in a sparse display can be maintained stably in VWM.

Experiment 8b: Maintaining Color-Orientation Conjunction in Crowded Displays

Next we investigated whether color-orientation conjunction can be maintained stably when the memory material is perceived in a crowded display. Since the results of Experiment 7b revealed that

[6] A pilot study revealed that it was difficult to remember the shape feature when the duration of the memory array was short, possibly due to the fact that in a crowded display, local shapes could not be directly selected into VWM by the end of parallel perceptual processing.

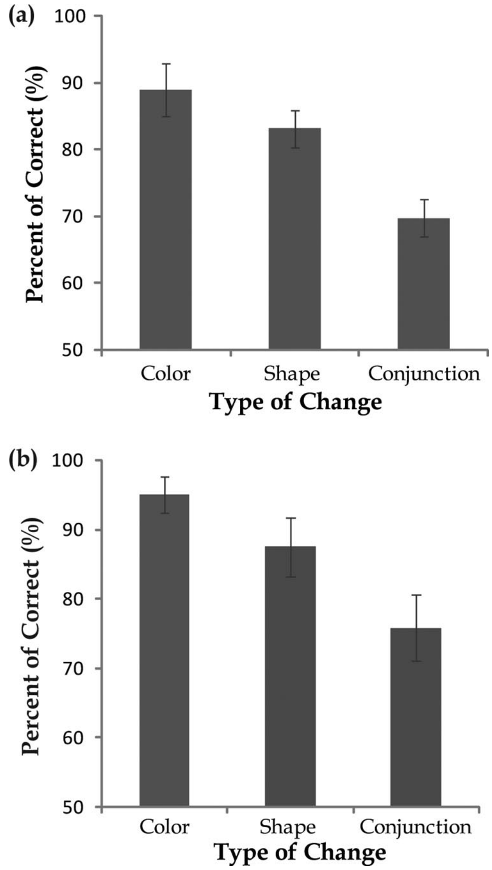

Figure 10. (a) The results obtained from all observers in Experiment 7b. The performance of detecting change is depicted as a function of the type of change. (b) The results obtained from six observers, whose performance is higher than 80%.

the overall memory accuracy dropped when the display was crowded, the set size of the memory array was reduced to three, so that individual features can be stored successfully.

Method

This experiment was identical to Experiment 7b except as noted here. Eleven Zhejiang University undergraduates participated. The shape of each global object was fixed at "O", which consisted of
Fn7 12 local arrows[7] (see Figure 11b). On each trial, the memory items were presented in three randomly selected quadrants.

Results and Discussion

In the No Change condition, the percentage of correct responses was 72.8%. The percentage of correct response as a function of
F13 different types of change is depicted in Figure 13. The difference between the performance of New Color (96.4%) and New Orientation (91.9%) was not significant, $t(10) = 1.490$, $p = .167$, $\eta_p^2 = .182$. Most importantly, the performance from New Conjunction condition (84.1%) was significantly different from that of the New Orientation condition, $t(10) = 2.771$, $p = .020$, $\eta_p^2 = .434$.

These results indicate that when presented in a crowded display, the color-orientation conjunction cannot be maintained stably in VWM. These results are consistent with the findings of Experiment 7b.

Experiment 9: Maintaining the Conjunction of Color and Oriented Gap in Sparse Displays

Here we further test our hypothesis by exploring the maintenance of objects consisting of color and oriented gaps. Since an oriented gap cannot be represented as part of the proto objects created by parallel processing, our hypothesis predicts that the conjunction of color and oriented gap will be unstable, even when the memory materials are presented in sparse displays.

Method

This experiment was identical to Experiment 8a except as noted here. Twelve Zhejiang University undergraduates participated. The memory array contained three colored circular rings identical to those used in Experiment 1b (Figure 11c). The memory array was presented for 350 ms. Participants were required to maintain the conjunction of the color and orientation of each ring.

Results and Discussion

In the No Change condition, the percentage of correct responses was 72.6%. The percentage of correct response as a function of different types of change was depicted in Figure 14. The difference F14
between the performance of the New Color (97.0%) and the New Orientation (92.7%) was not significant, $t(11) = 1.354$, $p = .202$, $\eta_p^2 = .143$. Most importantly, the performance from the New Conjunction condition (82.4%) was significantly different from that of the New Orientation condition, $t(11) = 3.514$, $p = .005$, $\eta_p^2 = .529$.

These results are consistent with the Interactive mode, which assumes that there are dissociated mechanisms in VWM for storing different types of perceptual information. These dissociated mechanisms make the representation of detailed information cannot be stored stably as part of an integrated object representations in VWM.

General Discussion

In the current project, we assume that VWM dynamically engages in different stages of perception. The information initially selected into VWM is *not* the final product of perception, but intermediate perceptual representations. This Interactive model of VWM motivates us to introduce insights from visual search studies into the exploration of VWM. We find converging evidence showing that outputs of parallel perception can be selected, consolidated and maintained as integrated objects. In contrast, there is no object-based effect for storing perceptual information requiring

[7] The reason to select this particular shape is to avoid the potential conflicts between the orientation of the global shape and the orientation of the local items. For instance, the arms of "+" and "×" employed in Experiment 7b have their own orientations, which may distract with the perception of the local arrows' orientations (see Navon, 1977).

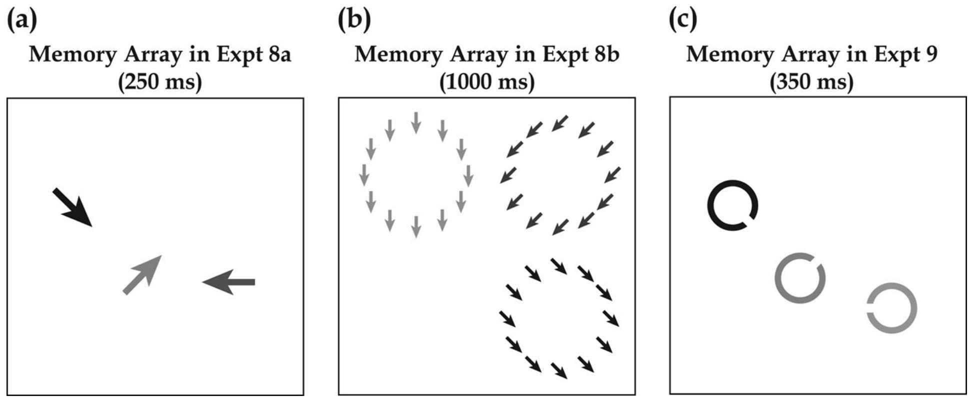

Figure 11. Three different types of memory arrays employed in Experiment 8a, 8b, and 9.

serial attentive processing. These results are consistent with the hypothesis that in VWM, the object representations with limited resolution are the proto objects created by parallel perceptual processing. To help readers to keep track of the main findings, we summarize all experiments and their results in Table 2. T2

Automatic Selection in VWM

Section I of the current project explores the selection of information in VWM. The dependent measure we used is the "Irrelevant-distracting" effect. The results suggest that while changes of task-irrelevant basic features can impair observers' performance, changes of irrelevant detailed information exhibit no such effect. These results support the hypothesis that object-based automatic selection is limited to the outputs of parallel perceptual processing. This finding is further supported by one of our recent ERP studies (Gao, Li, Yin, & Shen, 2010) showing that a task-irrelevant basic feature change in the test array can enhance the ERP component N270. However, changing task-irrelevant detailed information has no effect at all, suggesting that detailed information is not automatically encoded into VWM. Of course, given the intrinsic limitation of the task-irrelevant change paradigm, one can always argue that detailed information can be automatically selected, but somehow it just cannot impair the detection of target information or modulate an ERP component. While generally agreeing that this is a reasonable argument, we still think the current results are important, for the following reasons.

First, this paradigm has the power to *disprove* our hypothesis. Our hypothesis would be wrong if detailed information can produce an irrelevant-distracting effect, which is quite possible given previous studies on the automaticity of object-based attention (e.g., Egly et al., 1994; Watson & Kramer, 1999). Even though the null result from detailed information change is not conclusive, it at least supports the validity of our hypothesis and has greatly encouraged us to develop Sections II and III of the current project, in which our hypothesis is tested by more direct and powerful paradigms.

Second, the nonautomatic selection of detailed information is highly consistent with the results of Section II, which reveals salient dissociated time courses of consolidating different types of perceptual stimuli. In Experiment 6, the consolidation of color reaches plateau around 156 ms. In contrast, the consolidation of oriented gap continues until 420 ms, showing that this process is much slower and takes a lot of effort. It is unlikely that such a slow

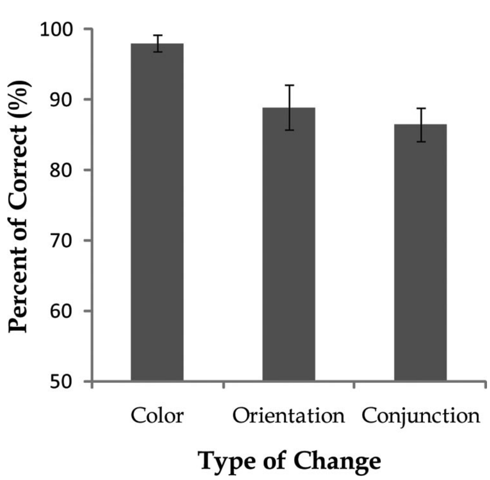

Figure 12. The results obtained from Experiment 8a. The performance of detecting change is depicted as a function of the type of change.

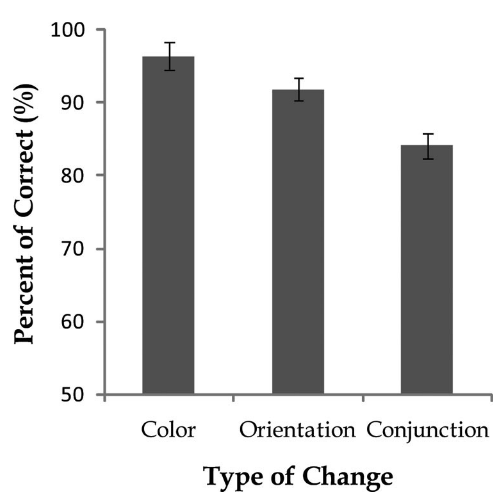

Figure 13. The results obtained from Experiment 8b. The performance of detecting change is depicted as a function of the type of change.

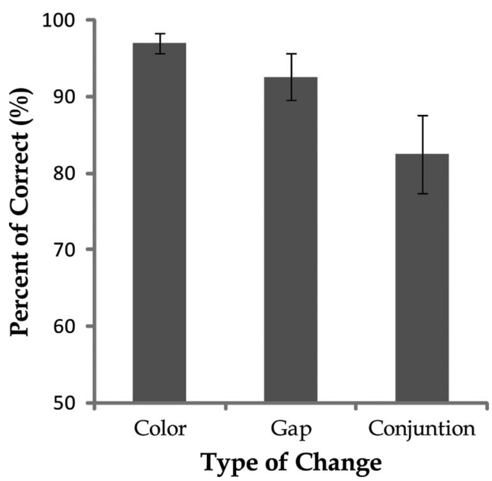

Figure 14. The results obtained from Experiment 9. The performance of detecting change is depicted as a function of the type of change.

and effortful process can be executed automatically when the oriented gap is task-irrelevant. In summary, we do not draw a strong conclusion based upon the results from Section I alone. However, we do believe that these results are important for understanding the nature of VWM, especially when they are analyzed together with the results from Sections II and III, which are discussed in detail in the next section.

We also want to note here that by suggesting that basic features are automatically stored, we are not denying the existence of top-down control on the processing of basic features. Visual search studies have clearly demonstrated that when a primitive feature is inconsistent with the top-down control, the perceptual salience of that feature will be decreased (e.g., Gao, Shen, Shui, & Gao, 2007; Müller, Reimann, & Krummenacher, 2003; Wolfe, Butcher, Lee, & Hyle, 2003). What we suggest here is that requiring participants to maintain color features will not result in a "shapeless" color representation in VWM.

Dissociated Mechanisms in VWM

The dissociated mechanisms for storing different types of perceptual information are directly explored in Sections II and III of the current project. The results show that multiple basic features can be consolidated synchronously without interference, and their conjunctions can be maintained stably. In contrast, detailed information cannot be consolidated synchronously (Section II), and cannot be maintained stably as part of an integrated object (Section III).

The results of Section III are especially important, as they show that the differences between processing of basic features and detailed information are not merely due to the fact that basic features are perceived faster than detailed information. When both types of information are sufficiently perceived and consolidated in VWM, there are still distinctive processes for maintaining them in VWM. These results indicate that detailed information is not an intrinsic component of the object representations in VWM. It can be incrementally assembled into the object representations with additional processes. However, these additional processes require time and effort, and the products of these processes will decay when maintaining them in VWM.

The existence of dissociated mechanisms in VWM is also consistent with other recent studies. One study measured change detection performance as a function of the discriminability between memory and test arrays (Awh et al., 2007). Here we are particularly interested in Experiment 2 of that study, in which

Table 2
Summary of all Experiments Reported in the Current Project

Experiment	Results	Summary
1a	Irrelevant-distracting effect from basic features with a whole report method	Object-based automatic selection for the outputs of parallel perception. No evidence of automatic selection for detailed information.
1b	No Irrelevant-distracting effect from detailed information with a whole report method	
2a	Irrelevant-distracting effect from basic features with a partial report method	
2b	No Irrelevant-distracting effect from detailed information with a partial report method	
3	No Irrelevant-distracting effect from color-color conjunction change	
4	Synchronous consolidation of basic features without mutual interference	Object-based synchronous consolidation is limited to the outputs of parallel perception.
5	Asynchronous consolidation for basic features and detailed information	
6	Dissociation in the capacity and efficiency of storing basic features and detained information	
7a	Stable maintenance of Color-Shape conjunctions presented in sparse displays	Object-based stable maintenance is limited to the outputs of parallel perception.
7b	Unstable maintenance of Color-Shape conjunction presented in crowded displays	
8a	Stable maintenance of Color-Orientation conjunctions presented in sparse displays	
8b	Unstable maintenance of Color-Orientation conjunction presented in crowded displays	
9	Unstable maintenance of Color-Gap's Orientation conjunctions in sparse displays	

tapraid5/zfn-xhp/zfn-xhp/zfn00611/zfn2701d11z | xppws | S=1 | 9/20/11 | 9:55 | Art: 2010-1941

changes with different degrees of discriminability were mixed within a block, forcing observers to encode both highly discriminable and finely detailed information. The results show that storing both types of information is as efficient as storing each type of information alone, suggesting that these two types of information can be stored in VWM without mutual interference. Consistent with this finding, one study from our lab also show that when the memory resource for storing detailed information (e.g., oriented gap) is fully occupied, task irrelevant basic features (e.g., color) can still be automatically stored in VWM and impair change detection performance (Gao et al., 2007). These results indicate that the storage of highly discriminable and finely detailed information may recruit separated resources in VWM. The Interactive model is also consistent with evidence from neuroimaging studies. For instance, one study revealed that whereas representations in the inferior intraparietal sulcus (IPS) are fixed to about four objects regardless of object complexity, those in the superior parietal and lateral occipital regions are reduced for complex objects, tracking the behavioral memory performance (e.g., Xu & Chun, 2006). In future studies, it would be interesting to investigate whether these dissociable neural circuits correspond to the outputs of different stages of perceptual processing.

The existence of dissociated mechanisms may have general implications for studies of VWM, beyond the object-based storage which is the focus of the current project. For instance, one intensely debated issue is whether the capacity of VWM consists of a fixed number of discrete slots, or a continuous pool of resource (e.g., Barton et al., 2009; Bays & Husain, 2008; Zhang & Luck, 2008). We find it intriguing that our results here are consistent with previous findings showing that VWM consists of a fixed number of slots with limited resolution (Zhang & Luck, 2008). Despite these seemingly consistent empirical results, there is an important difference in terms of the theoretical explanations of these findings. According to Zhang and Luck (2008), the low resolution of VWM is attributable to the *architecture* of the memory resources. Information with high resolution can be represented with the same architecture by deploying two or more slots to the same object. In contrast, according to our suggestion, the low resolution of VWM reflects the *contents* or *inputs* of VWM, instead of the intrinsic properties of the slots. The low resolution is determined by the products of parallel perpetual processing, which provides the initial inputs to VWM. In addition, representing information with high resolution requires processes dissociated from the storage of objects with low resolution.

Nature of the "Information Load" of VWM

Two studies have explored how the "complexity" or the "information load" of the memory array impacts the capacity of VWM (Alvarez & Cavanagh, 2004; Awh et al., 2007). However, it is not clear what determines the information load of the memory material. Here we suggest that information load is not simply characterized by the amount of information represented in *the final perceptual outputs*, but is also determined by the complexity of the computations to extract such information during *online perceptual processing*.

The first piece of evidence supporting this argument is the contrast between the storage of a bar's orientation and a ring's orientation, which are extensively explored in the current project with several different paradigms. A circular ring is apparently more complex than a straight bar. However, these two types of stimuli actually convey the same amount of orientation information. In both cases, the orientations of the stimuli are selected from the same set of eight orientations, which change in intervals of 45°. In fact, these two types of stimuli are indeed adopted by engineers to represent the same information when designing human-computer interfaces. (, e.g., in a Windows PC system, the progress of an application is represented by changing the orientation of a ring. While on a Macintosh system, the same information is represented by a rotating bar). If VWM acts at a higher level of processing, beyond perception, it should be able to extract the critical information, and get rid of the specific format in which the information is initially represented. Therefore, there should be no difference between the storage of these two types of stimuli, since they are ultimately about the same orientation information. However, our results show these two types of stimuli are processed distinctively in VWM. In contrast to the object-based storage for bar's orientation, the oriented gaps are consolidated with much lower efficiency, and cannot be maintained stably as part of an integrated object. Results of Experiment 6 and another recent ERP study (Gao et al., 2009) also show that even the capacity of storing gap orientations is much less than the capacity for oriented bars. These results indicate that storage in VWM is not only determined by the critical information contained in the final outputs of perception, but is also very sensitive to whether such information is represented in a format that can be efficiently processed during visual perception.

The second piece of evidence is the distinctive storage for memory materials represented in a sparse display or a crowded display (Experiments 7b, 8b). In both display conditions, the memory materials were selected from the same set of stimuli. Moreover, the duration of the crowded memory arrays was substantially increased, in order to guarantee that the memory items could still be fully perceived. Therefore, the results should be attributed to the limitations of VWM, instead of perception. The results show that the memory items can be stored as integrated objects only when they are presented in sparse displays. These results suggest that storage in VWM is not simply determined by the information contained in the memorized items. How the information is extracted during perception will directly impact how the information is going to be stored in VWM.

It is a general principle that information processing in the human mind is not only determined by the contents of the information, but is also greatly impacted by the format in which the information is presented (e.g., Pinker, 1999). Here is one example of this general principle: Humans can easily recognize the contents of a gray scale picture. If one simply transfers the luminance of each pixel in the picture to an array of digits from 0 to 255, the digit array will contain exactly the same information as the picture. However, it becomes almost impossible for humans to recognize the information contained in the digit array, because now the same information is represented in a format that is not supported by our visual system. With respect to the mechanism of VWM, when the information is represented in a format that can be efficiently processed by parallel perception (e.g., a bar's orientation, feature conjunctions in a sparse display), it can be then stored in VWM as integrated objects. However, if the same information is represented in a format that requires focal attention (e.g., oriented gaps, feature conjunctions in a crowded display), it will be processed by a different submechanism in VWM, which cannot exhibit significant object-based benefits.

Object-Based Attention and Object-Based Storage in VWM

Attention plays a critical role in selecting perceptual information into VWM (e.g., Chun & Potter, 1995). As attention is closely associated with VWM, the current results of object-based storage also have implications for the nature of object-based attention.

It has been shown that object-based attention and object-based VWM resemble each other in several important ways. For instance, one classic study on object-based attention reveals that the efficiency of perceptual processing is influenced by whether the target features are from the same object, instead of two spatially overlapping objects (Duncan, 1984). One study of the capacity of VWM has employed the same manipulation, showing that memory performance is determined by the number of objects in the memory array, instead of the number of spatial locations occupied by the memory array (Lee & Chun, 2001). Studies of attention have also revealed that the recognition of two features is facilitated when they are located within the boundary of a single object, as compared with when they are located within the boundaries of different objects (e.g., Egly et al., 1994; Watson & Kramer, 1999). This same-object advantage has recently been shown in the storage of VWM by both behavior and neuroimaging results (Xu & Chun, 2007).

Due to the close relationship between object-based attention and VWM, one might intuitively expect that object-based attention plays an important role in the object-based benefits in VWM. In particular, an item's individual features are integrated into coherent object representations *after* focal attention has been deployed onto that item. The coherent object representation is then transferred into the object files storing in VWM. However, the current project provides an opposing view regarding the role of attention in object-based storage: object-based storage of basic features is not attributed to the *postattentive* coherent object representations, but due to the fact that these features can be segmented into the proto objects *prior* to the focal attentive processing. For information that actually requires attentive processing, our results suggest that it is costly to consolidate and maintain it in VWM.

Our hypothesis is also supported by one recent study which explores the role of attention in maintaining integrated objects in VWM (Johnson et al., 2008). In this study, an attention-demanding visual search task is interposed into the maintenance stage of the change-detection task. The results show that the visual search task impaired overall memory performance. However, the memory for individual features and feature conjunctions is equally impaired by the search task, indicating that attention does not play a special role in maintaining the conjunction information. Note that we are not arguing against object-based attention in general. In contrast, the hypothesis that attention selects proto objects into VWM is intrinsically consistent with the nature of object-based attention, which asserts that attention is directed to individual objects, as opposed to features or spatial locations. What we argue against here is a specific strong interpretation of object-based attention, assuming that once an object is selected by attention, all of its features will be automatically processed and stored in object files. Our results indicate that such automatic processing of detailed information rarely happen. Instead, once attention is directed toward a proto object, the processing of detailed information takes effort and is under top-down control.

The limitations of object-based storage revealed by our study resemble the findings of one recent study of object-based attention (Huang, 2010). This study introduces the Boolean theory of visual awareness (Huang & Pashler, 2007; Huang, Treisman, & Pashler, 2007) into the investigation of object-based attention, which reveals some rather surprising findings. For instance, connecting two features into a single object won't necessarily facilitate the perception of these two colors. There is robust object-based advantage only when the two features are from different dimensions (e.g., color and orientation), but not when they are different values of the same dimensions (e.g., red and green). This finding is highly consistent with Section I of the current study, which shows automatic selection of the conjunction of different features (e.g., color-orientation conjunction), but not the conjunction of different values of the same feature (e.g., color-color conjunction). The resemblance between Huang (2010) and the current study highlights that online perception and storage in VWM are intrinsically related, and perhaps governed by the same rules. These results also suggest that a promising direction for future research may be to introduce the Boolean map theory into investigation of VWM.

Nevertheless, there are still some differences between Huang (2010) and the current study, which should be noted here. In Huang (2010), the motivation is to demonstrate that object-based advantage is controlled by the Boolean maps. Therefore, most stimuli employed in that study are features which can be easily matched to certain Boolean maps. In contrast, our motivation here is to dissociate the outputs of serial attentive processing from the outputs of parallel perception. As a result, we employed several types of materials which may be outside the scope of the most intuitive understanding of Boolean maps. For instance, in Section III of our study, we manipulated perceptual discriminability by employing the classic "global precedence" effect (Navon, 1977). It is not obvious to us how such an effect can be incorporated into the Boolean map framework. Apparently, further research needs to be launched to bridge these gaps.

Conclusion

Perhaps the most important finding of the current project is that how information is extracted during visual perception will directly impact how the information is going to be selected, consolidated and maintained in VWM. This finding would be odd if one assumes that VWM stores the final outputs of perception, since by the time information is transferred in VWM, the perceptual processing should have been terminated. However, this finding is exactly what one would expect if perception and VWM are two dynamically interacting processes. The current study reveals the validity of the Interaction model of perception and VWM, and shows how this model can generate novel predictions, guiding the investigation of the mechanisms of VWM.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects.

tapraid5/zfn-xhp/zfn-xhp/zfn00611/zfn2701d11z | xppws | S=1 | 9/20/11 | 9:55 | Art: 2010-1941

Psychological Science, 15, 106–111. doi:10.1111/j.0963-7214.2004.01502006.x

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18,* 622–628. doi:10.1111/j.1467-9280.2007.01949.x

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and working memory. *Trends in Cognitive Sciences, 5,* 119–126. doi:10.1016/S1364-6613(00)01593-X

Baddeley, A. (1992). Working memory. *Science, 255,* 556–559. doi:10.1126/science.1736359

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (pp. 647–667). New York, NY: Academic Press.

Barton, B., Ester, E. F., & Awh, E. (2009). Discrete resource allocation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1359–1367. doi:10.1037/a0015792

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321,* 851–854. doi:10.1126/science.1158023

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: Beyond object-files and object tracking. *Cognition, 80,* 179–213. doi:10.1016/S0010-0277(00)00154-2

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 109–127. doi:10.1037/0096-1523.21.1.109

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113,* 501–517. doi:10.1037/0096-3445.113.4.501

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458. doi:10.1037/0033-295X.96.3.433

Egly, R., Driver, J., & Rafal, R. (1994). Shifting visual attention between objects and locations: Evidence for normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123,* 161–177. doi:10.1037/0096-3445.123.2.161

Enns, J. T., & Rensink, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review, 98,* 335–351. doi:10.1037/0033-295X.98.3.335

Gao, T., & Scholl, B. J. (2010). Are objects required for object files?: Roles of segmentation and spatiotemporal continuity in computing object persistence. *Visual Cognition, 18,* 82–109. doi:10.1080/13506280802614966

Gao, T., Shen, M., Shui, R., & Gao, Z. (2007). The reconfiguration of task set has no effect on the efficiency of feature search. *Perception & Psychophysics, 69,* 345–352. doi:10.3758/BF03193755

Gao, Z., Li, J., Liang, J., Chen, H., Yin, J., & Shen, M. (2009). Storing fine detailed information in visual working memory: Evidence from event-related potentials. *Journal of Vision, 9,* 17. doi:10.1167/9.7.17

Gao, Z., Li, J., Yin, J., Gao, T., & Shen, M. (2010). Dissociated mechanisms of extracting perceptual information into visual working memory. *PLoS ONE, 5,* e14273.

Gold, J. M., Wilk, C. M., McMahon, R. P., Buchanan, R. W., & Luck, S. J. (2003). Working memory for visual features and conjunctions in schizophrenia. *Journal of Abnormal Psychology, 112,* 61–71. doi:10.1037/0021-843X.112.1.61

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature, 458,* 632–635. doi:10.1038/nature07832

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron, 36,* 791–804. doi:10.1016/S0896-6273(02)01091-7

Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual-short-term memory: Transsaccadic memory, object correspondence, and gaze correction. *Journal of Experimental Psychology: General, 137,* 163–181. doi:10.1037/0096-3445.137.1.163

Huang, L. (2010). What is the unit of visual attention? Object for selection, but Boolean map for access. *Journal of Experimental Psychology: General, 139,* 162–179. doi:10.1037/a0018034

Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review, 114,* 599–631. doi:10.1037/0033-295X.114.3.599

Huang, L., Treisman, A., & Pashler, H. (2007). Characterizing the limits of human visual awareness. *Science, 317,* 823–825. doi:10.1126/science.1143515

Hyun, J. S., & Luck, S. J. (2007). Visual working memory as the substrate for mental rotation. *Psychonomic Bulletin and Review, 14,* 154–158. doi:10.3758/BF03194043

Hyun, J. S., Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (2009). The comparison of visual working memory representations with perceptual inputs. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1140–1160. doi:10.1037/a0015019

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

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual-short term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 683–702. doi:10.1037/0278-7393.26.3.683

Jiang, Y. V., Shim, W. M., & Makovski, T. (2008). Visual working memory for line orientations and face identities. *Perception and Psychophysics, 70,* 1581–1591. doi:10.3758/PP.70.8.1581

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,* 41–55. doi:10.1037/0096-1523.34.1.41

Joseph, J. S., Chun, M. M., & Nakayama, K. (1997). Attentional requirements in a "preattentive" feature search task. *Nature, 387,* 805–807. doi:10.1038/42940

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of files: Object-specific integration of information. *Cognitive Psychology, 24,* 175–219. doi:10.1016/0010-0285(92)90007-O

Kensinger, E. A., & Choi, E. S. (2009). When side matters: Hemispheric processing and the visual specificity of emotional memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 247–253. doi:10.1037/a0013414

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 451–468. doi:10.1037/0096-1523.21.3.451

Lee, D., & Chun, M. M. (2001). What are the units of visual short-term memory, objects or spatial locations? *Perception and Psychophysics, 63,* 253–257. doi:10.3758/BF03194466

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–281. doi:10.1038/36846

Marr, D. (1982). *Vision.* New York, NY: W. H. Freeman.

Mitroff, S. R., Scholl, B. J., & Wynn, K. (2004). Divide and conquer: How object files adapt when a persisting object splits into two. *Psychological Science, 15,* 420–425. doi:10.1111/j.0956-7976.2004.00695.x

Müller, H. J., Reimann, B., & Krummenacher, J. (2003). Visual search for singleton feature targets across dimensions: Stimulus- and expectancy-driven effects in dimensional weighting. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 1021–1035. doi:10.1037/0096-1523.29.5.1021

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320,* 264–265. doi:10.1038/320264a0

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353–383. doi:10.1016/0010-0285(77)90012-3

Neisser, U. (1967). *Cognitive Psychology.* New York, NY: Appleton-Century-Crofts.

Oakes, L. M., Ross-Sheehy, S., & Luck, S. J. (2006). Rapid development of feature binding in visual short-term memory. *Psychological Science, 17,* 781–787. doi:10.1111/j.1467-9280.2006.01782.x

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A thread advantage with schematic stimuli. *Journal of Personality and Social Psychology, 80,* 381–396. doi:10.1037/0022-3514.80.3.381

Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the “strong object” hypothesis. *Perception & Psychophysics, 64,* 1055–1067. doi:10.3758/BF03194756

Parks, T. E. (1965). Post-retinal visual storage. *American Journal of Psychology, 78,* 145–147. doi:10.2307/1421101

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16,* 283–290. doi:10.3758/BF03203943

Pinker, S. (1999). *How the Mind Works.* New York, NY: WW Norton & Company.

Quinlan, P. T. (2003). Visual feature integration theory: Past, present, and future. *Psychological Bulletin, 129,* 643–673. doi:10.1037/0033-2909.129.5.643

Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition, 7,* 17–42. doi:10.1080/135062800394667

Riggs, K. J., McTaggart, J., Simpson, A., & Freeman, R. P. J. (2006). Changes in the capacity of visual working memory in 5- to 10-year-olds. *Journal of Experimental Child Psychology, 95,* 18–26. doi:10.1016/j.jecp.2006.03.009

Rock, I. (1981). Anorthoscopic perception. *Scientific American, 244,* 145–153. doi:10.1038/scientificamerican0381–145

Roelfsema, P. R. (2005). Elemental operations in vision. *Trends in Cognitive Science, 9,* 226–233. doi:10.1016/j.tics.2005.03.012

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development, 74,* 1807–1822. doi:10.1046/j.1467-8624.2003.00639.x

Serences, J., Ester, E., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science, 20,* 207–214. doi:10.1111/j.1467-9280.2009.02276.x

Silvanto, J., & Cattaneo, Z. (2010). Transcranial magnetic stimulation reveals the content of visual short-term memory in the visual cortex. *NeuroImage, 50,* 1683–1689. doi:10.1016/j.neuroimage.2010.01.021

Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences, 12,* 342–348. doi:10.1016/j.tics.2008.05.007

Soto, D., & Humphreys, G. W. (2009). Automatic selection of irrelevant object features through working memory: Evidence for top-down attentional capture. *Experimental Psychology, 56,* 165–172. doi:10.1027/1618-3169.56.3.165

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428,* 751–754. doi:10.1038/nature02466

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136. doi:10.1016/0010-0285(80)90005-5

Ullman, S. (1984). Visual routines. *Cognition, 18,* 97–159. doi:10.1016/0010-0277(84)90023-4

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature, 428,* 748–751. doi:10.1038/nature02447

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to visual working memory. *Nature, 438,* 500–503. doi:10.1038/nature04171

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,* 92–114. doi:10.1037/0096-1523.27.1.92

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1436–1451. doi:10.1037/0096-1523.32.6.1436

Watson, S., & Kramer, A. (1999). Object-based visual selective attention and perceptual organization. *Perception & Psychophysics, 61,* 31–49. doi:10.3758/BF03211947

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General, 131,* 48–64. doi:10.1037/0096-3445.131.1.48

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1,* 202–238. doi:10.3758/BF03200774

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences, 7,* 70–76. doi:10.1016/S1364-6613(02)00024-4

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research, 37,* 25–44. doi:10.1016/S0042-6989(96)00111-3

Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 483–502. doi:10.1037/0096-1523.29.2.483

Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature, 400,* 867–869. doi:10.1038/23698

Woodman, G. F., & Luck, S. J. (2003). Serial deployment of attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 121–138. doi:10.1037/0096-1523.29.1.121

Woodman, G. F., & Vogel, E. K. (2008). Selective storage and maintenance of an object’s features in visual working memory. *Psychonomic Bulletin & Review, 15,* 223–229. doi:10.3758/PBR.15.1.223

Xu, Y. (2002). Limitations in object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 458–468. doi:10.1037/0096-1523.28.2.458

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature, 440,* 91–95. doi:10.1038/nature04262

Xu, Y., & Chun, M. M. (2007). Visual grouping in human parietal cortex. *Proceedings of the National Academy of Sciences, USA, 104,* 18766–18771. doi:10.1073/pnas.0705618104

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception & Performance, 10,* 601–621. doi:10.1037/0096-1523.10.5.601

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453,* 233–235. doi:10.1038/nature06860

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science, 20,* 423–428. doi:10.1111/j.1467-9280.2009.02322.x

Received June 21, 2010
Revision received March 7, 2011
Accepted March 15, 2011 ■