scientific reports

OPEN

Multiple item representations in visual working memory simultaneously guide attention

Caibin Duan[1,2], Lihua Zhang[2]✉, Dequn Song[1]✉ & Bing Zhang[1]

Visual working memory (VWM) is a subject of ongoing debate regarding whether multiple item representations can simultaneously guide attention. The Single Item Template hypothesis (SIT) posits that VWM representations only allow a single item to guide attention, while the Multiple Item Template hypothesis (MIT) suggests that multiple items in VWM representations can guide attention simultaneously. This study further investigates this through a dual-task paradigm. Participants were required to complete memory-search tasks under different memory and match types, with memory items and distractors being of the same category (Experiment 1) and different categories (Experiment 2). Results show: (1) In Experiment 1, when the memory type was color, multiple item representations in VWM simultaneously guided attention, providing support for the MIT hypothesis; however, when the memory type was graphic, multiple item representations in VWM did not simultaneously guide attention, providing support for the SIT hypothesis. (2) In Experiment 2, whether the memory type was color or graphic, multiple item representations in VWM did not simultaneously guide attention, providing support for the SIT hypothesis. Whether multiple item representations in VWM can simultaneously guide attention is significantly influenced by memory types and the relationship between memory items and categories of distractors.

Keywords Visual working memory, Attention guidance, Visual search

Visual working memory (VWM) is an important component of individual cognitive processing, which not only has the function of temporary storage and processing of information, but also has a certain guiding effect on attention[1–5]. People face a large amount of complex information every day, but ultimately only some of the information is perceived and processed. Among them, VWM, as a top-down cognitive processing process, plays a certain role in guiding attention. The guidance of attention by VWM refers to the fact that when faced with multiple choices, VWM representation will prioritize guiding individuals to pay attention to the same or similar content. Specifically, VWM representation can serve as an attention template, guiding people to prioritize attention to objects that match the template in a top-down manner during visual search[6–8]. Research on VWM guided attention often adopts a dual task processing paradigm, which requires participants to complete a visual search task during memory tasks such as memorizing graphics or colors. It was found that during the visual search process, when memory content reappears, regardless of whether it is searching for the target object or interfering objects, individuals will receive priority attention[9,10]. The study was supported by both eye tracking data and ERP data[11,12]. The biased competition model explains this phenomenon: during visual searches, objects in the field of view compete with each other for limited attention resources. If the VWM representation is in a pre-activated state, the perceptual representations that match it in the visual cortex are activated, making objects in the field of view that match the VWM representation the "winner" and receiving priority attention, that is, the VWM representation guides attention[13]. Downing was the first to validate this model through dual task machining experiments[14]. In the experiment, participants were first presented with a facial image and asked to memorize it; Then two facial images are presented simultaneously as clues, one is a previously memorized image, which is the memory matching item, and the other is a new facial image, which is the non memory matching item; Finally, present the target for detection reaction. The results showed that compared to non memory matching items, participants responded faster to targets that were in the same position as the memory matching items. Bahle et al. conducted a series of studies using real-life scenes as experimental materials and eye tracking technology[15]. The results showed that distractors matched with memory objects attracted the subjects' first eye twitching in visual search.

[1]Students' Affairs Division, Shenyang Agricultural University, Shenyang 110866, China. [2]College of Psychology, Liaoning Normal University, Dalian 116029, China. ✉email: 2014500017@syau.edu.cn; dcbgreat@126.com

However, studies have also found that VWM representations do not always guide attention[16]. Olivers attributed these differences to different types of visual search, which are related to whether the search target randomly changes[17]. Specifically, in Fan et al.'s study, search targets are fixed and can be stored in long-term memory to complete search tasks[7]. Therefore, working memory has sufficient resources to represent memory items and form VWM representations. When an object appears in a search task that matches the VWM representation, it guides attention. In Downing and Dodds' study, the search target is randomly changing and needs to be stored in working memory and in a priority processing state in order to complete the search task[16]. Therefore, a large amount of working memory resources are used to maintain search targets, preventing memory items from guiding attention.

Oliver et al. proposed the Single Item Template Hypothesis (SIT) to explain the differences in VWM representation[18]. It is believed that there is only one representation in VWM that guides attention at a time. VWM representation can be divided into two different types: activation type and accessory type. The VWM representation in an activated state is the focus of the individual's central executive process and the core of cognitive processing, directly guiding attention; The VWM representation in an attached state is on the periphery of individual cognitive processing and does not guide attention. Although individual working memory capacity is generally maintained at 3–4 items, only one item can be activated at a time to guide selective attention. If multiple identical items are retained in working memory, the mutual interference between items and competition for resources can easily lead to each item not becoming an active state, but rather being referred to as an attached state[19–21]. In previous studies involving fixed search targets, there was only one memory item and no other competitors. Therefore, it belongs to the activation type and has a guiding effect on attention[9]. In the study of random changes in search targets, there are two items included: memory content and targets. SIT posits that only one item can become an activation type, and the search target has a competitive advantage, so the target belongs to the activation type, while the memory item belongs to the subsidiary type and does not guide attention[16]. Further support for SIT comes from van Moorselaa et al.'s study[22]. The study presented memory items with different loads and controlled the matching between distractors and memory items in visual search. It was found that only when the memory load was 1, distractors that matched the VWM representations would guide attention; Under other loads, distractors that match the VWM representations have no effect on attention. This indicates that only one project can become an activation type at a time to guide attention.

The Multiple Item Template Hypothesis (MIT) presents an opposing view, suggesting that the representation of multiple items in VWM can simultaneously guide attention[23]. Hollingworth and Beck conducted a study in which participants were presented with two memory items while controlling for different matching conditions[24]. In the search task, there were either 1 or 2 distractors that matched the memory items. The results indicated that under different matching conditions, interferents that matched VWM representations all had a guiding effect on attention. Similarly, Fan et al. found in visual search tasks that two disruptors matching the VWM representation simultaneously guided attention[7]. These results all support the MIT hypothesis.

The contradiction between the two hypotheses may be due to two reasons. Firstly, it may be related to whether the memory items and distractors belong to the same category. VWM represents the guidance of attention based on objects, and the similarity between memory items and distractors determines the size of the guidance effect[25]. This means that attention is more easily directed towards external stimuli that are similar to the objects stored in memory. When the memory item and the distractor belong to the same category, their similarity is higher, which may lead to attention being more easily directed towards the distractor. This similarity matching enhances the association between memory items and distractors, making attention more easily attracted to these similar stimuli during the search process. At this point, distractors may become a "stumbling block" in the search process, as they are highly similar in features to memory terms, making it difficult for searchers to distinguish between targets and interferences, resulting in a decrease in search efficiency. When memory items and distractors do not belong to the same category, their similarity is lower, which helps searchers focus more on features related to the target. At this point, the guiding role of memory items in attention is more clear and specific, as it can guide attention away from distractors that are dissimilar to memory items, thereby promoting the improvement of search efficiency. Studies supporting SIT hypothesis often involve memory items and distractors belonging to different categories; for example, van Moorselaa et al.'s study used colors as memory items and graphics as distractors resulting in a smaller guidance effect when multiple items are remembered[22]. Conversely, studies supporting MIT hypothesis typically involve both memory items and distractors belonging to the same category; for instance, Fan et al.'s study used colors for both memory items and distractors leading to a larger guidance effect when multiple items are remembered[7]. Secondly, it may be related to the type of memory being tested. Working memory is a complex cognitive system that contains different types of memory components, such as visual spatial memory (used to store location, graphic, and other information), phonological loops (responsible for processing sound and language information), episodic buffers (integrating information from different memory subsystems), and central executive systems (dividing and scheduling various memory subsystems). Different types of memory information may follow different mechanisms when guiding attention, thereby affecting the applicability of SIT and MIT. Previous studies have explored this issue using a variety of memory materials, including color, graphic, size, and other attributes, which correspond to different memory subsystems. For example, color may rely more on visual spatial memory, while size may involve more abstract numerical or spatial relationship processing. Due to the inconsistency in memory types among these studies, inconsistencies have arisen between the experimental results. Some studies have found that compared with graphics, there is a greater guidance effect on attention when remembering colors[26]. Therefore, it is necessary to select materials again under conditions where memory items and belong or do not belong to the same category; different types of memories should be selected for re-exploring SIT and MIT hypotheses.

Regarding the representation of VWM in guiding attention, it is consistent with SIT hypothesis or MIT hypothesis. Hollingworth and Beck provided a good test method: In dual task processing, there are three types of matching between memory and search terms: M2-0, memorizing 2 items, and 0 stimuli in the search material were the same as those in the memory material; M2-1 memorizing 2 items, and 1 stimulus in the search material was the same as that in the memory material; M2-2, memorizing 2 items, 2 stimuli in the search material were the same as the memorized material. If SIT hypothesis is true, when remembering 2 items, only one item can become the activated type to guide attention[24]. In the M2-1 condition, if the memory item that becomes the activated type matches the distractor, it produces a guidance effect (reaction time: M2-1 > M2-0); if the non-activated type memory item matches the distractor, it does not produce a guidance effect (reaction time: M2-1 = M2-0). In the M2-2 condition, it definitely includes a memory item that becomes an activated type matching with the distractor, producing a guidance effect (reaction time: M2-2 > M2-0). If MIT hypothesis is true, when remembering 2 items, both items can simultaneously guide attention. In both M2-1 and M2-2 conditions, guidance effects will occur (reaction time: M2-1 > M2-0; M2-2 > M2-0), and M2-2's guidance effect is greater than that of M2-1(reaction time:M2-2 > M2-1). Therefore, SIT holds whenM2-2 > M2-0, M2-1 > M2-0 and M2-2 = M2-1, or when M2-2 > M2-0, M2-1 = M2-0 and M2-2 > M2-1; MIT holds when M2-2 > M2-0, M2-1 > M2-0, and M2-2 > M2-1.

In summary, VWM representation has a certain guiding effect on attention. However, it is unclear whether VWM only allows single item guided attention (SIT) or multiple items can guide attention simultaneously (MIT). Based on the causes of SIT and MIT, this study aims to further explore the guidance of VWM representation on attention by controlling different types of memory, whether memory items and distractors belong to the same category.

Experiment 1: Memory items and distractors are of the same category

Method

Subject

32 college students participated in this experiment, with 16 males and 16 females, with an average age of 19.45 years old. The subjects were all right-handed and had no color blindness or color weakness.

Instruments and experimental materials

Presenting stimuli on a 15 inch color screen with an E-prime2.0 resolution of 1920 * 1080. The subjects are approximately 65 cm away from the screen.

The memory task material consists of six different graphics and six different colors. The graphics are squares, hexagons, circles, crosses, triangles, and diamonds (2.43° × 2.43°, with a line thickness of 0.18°); The colors are red, yellow, blue, green, purple, and orange, with RGB values of 197, 149,162; 255, 245, 79; 20, 112, 192; 64, 176, 80; 112, 48, 160; 237, 125, 49[26].

The visual search task material consists of a combination of the aforementioned memory material and three different directional line segments. The size of the line segment is 0.31° × 0.14°, including three directions: tilt 30° to the left, tilt 30° to the right, and in the vertical direction.

Experimental design

Adopting a 2 (memory type: color, graphic) × 3 (matching type: M2-0, M2-1, M2-2) within subject design. The dependent variables are reaction time for visual search tasks and accuracy for memory tasks.

Experimental process

In order to minimize the influence of verbal working memory on attentional guidance, a pronunciation inhibition was employed to prevent speech processing[27]. Prior to the experiment, participants were presented with four random numbers and instructed to verbally repeat them at a frequency of two times per second during the experiment. The participants initiated the experiment by both repeating the numbers and pressing the spacebar. The experimental procedure began with a fixation point displayed at the center of the screen for 500–1000 ms. This was followed by the presentation of a memory material lasting 1000 ms (consisting of two different graphics or colors, displayed on either side of the fixation point). Subsequently, an empty screen was shown for 800 ms before presenting four visual search materials for 3000 ms (3 combination graphics containing vertical line segments and 1 combination graphic containing left or right oblique line segments), randomly distributed at positions 1, 4, 7, and 10 o'clock on an imaginary disk with a viewing angle of 6° from the fixation point. Among them, the graphics of the left and right slanted line segments appear randomly, each accounting for 50%. Subjects were required to search for tilted line segments within these materials and respond by pressing either "F" when encountering a left slanted line segment or "J" when encountering a right angled line segment as quickly and accurately as possible. Following their response, there was an additional blank screen presented for 500 ms before displaying a memory detection interface where participants judged whether it matched with the initial memory material (pressing "F" if it matched and "J" if it did not match, each accounting for 50%). See Fig. 1.

The experiment consisted of three matching types: M2-0, memorizing 2 items, and 0 stimuli in the search material were the same as those in the memory material; M2-1 memorizing 2 items, and 1 stimulus in the search material was the same as that in the memory material; M2-2, memorizing 2 items, 2 stimuli in the search material were the same as the memorized material. The experiment was divided into 2 blocks, including graphics and colors respectively, and the blocks were presented randomly among the subjects. Each block contains 90 trials, and the three matching types appear randomly in each block, each accounting for 1/3 of the trials.

Results

One participant with an error rate of more than 20% was excluded from the analysis. The accuracy rate of memory task was analyzed by repeated measurement variance (Table 1). The results revealed a significant main effect

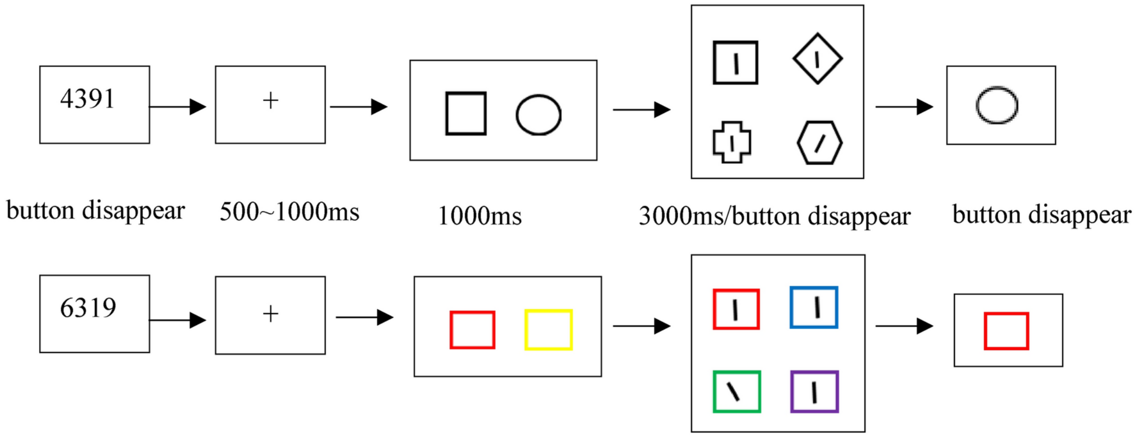

Fig. 1. Flow diagram under M2-1 condition when the memory type and the distracts are of the same category.

	M2-0	M2-1	M2-2
Colour	.94 ± .07 (718.08 ± 35.17)	.96 ± .04 (750.48 ± 35.97)	.96 ± .03 (766.97 ± 37.18)
Graphics	.93 ± .07 (892.50 ± 41.21)	.92 ± .07 (929.45 ± 45.91)	.92 ± .08 (915.30 ± 47.68)

Table 1. Memory task accuracy rate and search task response time ($\overline{x} \pm s$, $n = 31$). Accuracy outside parentheses and reaction time inside parentheses.

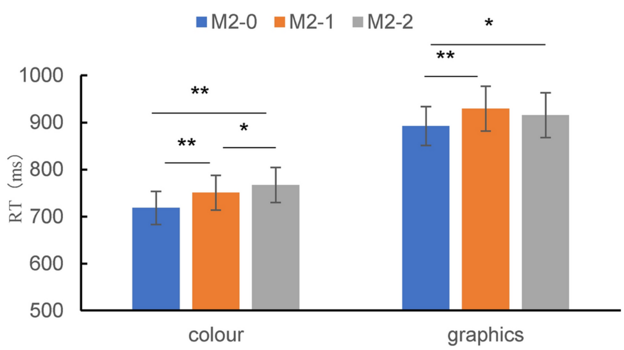

Fig. 2. Search reaction time under various conditions (the vertical line represents the standard error, *$p < .05$, **$p < .01$, same below).

of memory type, F (1,30) = 11.06, $p = .002$, $\eta_p^2 = .269$, $BF_{10} = 14.062$, with higher memory accuracy in the color condition compared to the graph condition. The main effect of matching type was not significant, F (2,60) = .417, $p > .05$, $BF_{10} = .377$. However, there was a significant interaction between memory type and matching type, F (2,60) = 4.05, $p = .025$, $\eta_p^2 = .119$, $BF_{10} = 1.728$. Simple effect analysis showed that under the color condition, the memory accuracy of M2-2 and M2-1 was significantly higher than that of M2-0, but there was no significant difference between M2-2 and M2-1. There was no significant difference in the memory accuracy of M2-2, M2-1 and M2-0 under the graph condition.

As this study focuses on the influence of VWM representation on attentional guidance, the experiment only included trials where both visual search task and memory detection task were performed accurately, and data outside these criteria were excluded[12]. To test the reliability of response time, we performed an odd–even split-half analysis for different memory types[28,29]. The results revealed a high internal reliability, $r_{\text{colour}} = .96$, $p < .001$; $r_{\text{graphics}} = .95$, $p < .001$. Repeated measures of ANOVA were conducted for response to the search task, with multiple comparisons performed using the Bonferroni correction method (Table 1 and Fig. 2). The results revealed a significant main effect of memory type, F (1,30) = 46.09, $p < .001$, $\eta_p^2 = .606$, $BF_{10} > 1000$, with longer search response time in the graphic condition compared to the color condition. The main effect of matching type was significant, F (2,60) = 17.91, $p < .001$, $\eta_p^2 = .374$, $BF_{10} > 1000$. The interaction margin between memory type and matching type was significant, F (2,60) = 2.82, $p = .068$, $\eta_p^2 = .086$, $BF_{10} = 5.144$. Simple effect analysis showed that under color conditions, the search response time of M2-2 was significantly longer than that of M2-0, $t(30) = 5.14$, $p < .001$, *Cohen's* $d = 0.24$; the search response time of M2-1 was significantly longer than that of

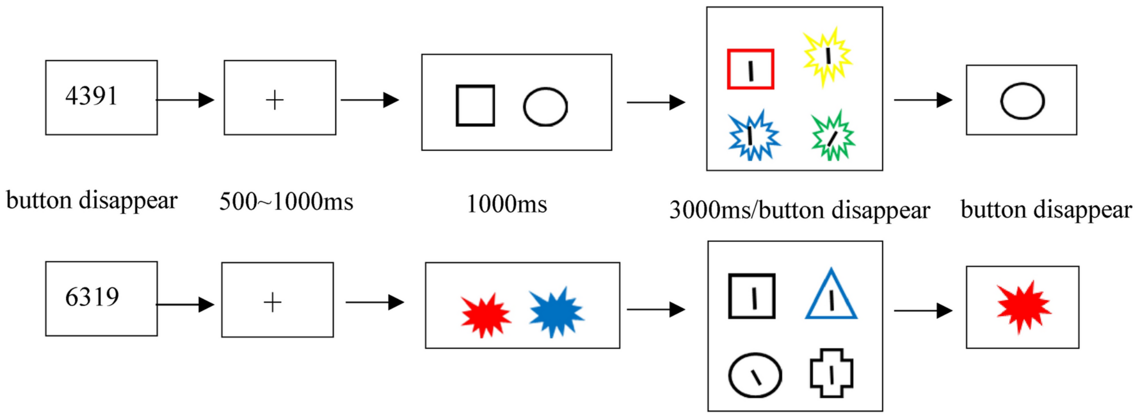

Fig. 3. Flow diagram under M2-1 condition when the memory type and the distracts are of different categories.

	M2-0	M2-1	M2-2
Colour	.91 ± .07 (942.90 ± 39.61)	.90 ± .10 (1038.3 ± 44.19)	.92 ± .09 (1046 ± 41.74)
Graphics	.92 ± .09 (811.50 ± 29.65)	.92 ± .08 (874.36 ± 32.93)	.92 ± .08 (874.72 ± 29.95)

Table 2. Memory task accuracy rate and search task response time ($\overline{x} \pm s$, $n = 30$). Accuracy outside parentheses and reaction time inside parentheses.

M2-0, $t(30) = 4.34$, $p < .001$, *Cohen's d* = 0.26; the search response time of M2-2 was significantly longer than that of M2-1, $t(30) = 5.68$, $p < .001$, *Cohen's d* = 0.70. Under the graph condition, the search response time of M2-2 is significantly longer than that of M2-0, $t(30) = 1.93$, $p < .05$, *Cohen's d* = 0.19; the search response time of M2-1 is significantly longer than that of M2-0, $t(30) = 3.304$, $p < .05$, *Cohen's d* = 0.15; there is no significant difference between M2-2 and M2-1. The results show that under the color condition, multiple item representations in VWM can simultaneously guide attention, supporting the MIT hypothesis; Under the graphic condition, Although VWM representation also has a guiding effect on attention, multiple item representations in VWM cannot simultaneously guide attention, supporting the SIT hypothesis.

Experiment 2: Memory items and distracts are of different categories
Method
Subject
32 college students participated in this experiment, with 8 males and 24females, with an average age of 20.20 years old. The subjects were all right-handed and had no color blindness or color weakness.

Instruments and experimental materials
The experiment is essentially the same as Experiment 1, with the visual search task materials containing not only the existing materials from Experiment 1, but also a combination of graphics featuring different colors and vertical line segments.

Experimental design
Same as Experiment 1.

Experimental process
It is basically the same as Experiment 1. The differences are as follows: in Experiment 1, the distracts in memory materials and visual search materials are of the same category; In Experiment 2, there were different types of distracts in memory materials and visual search materials (see Fig. 3).

Results
Two participants with an error rate of more than 20% were excluded from the analysis. The accuracy rate of memory task was analyzed by repeated measurement variance (Table 2). The results indicated that neither the main effect nor interaction effect were significant, $Fs < .78$, $ps > .384$, $BF_{10} < .259$.

As this study focuses on the influence of VWM representation on attentional guidance, the experiment only included trials where both visual search task and memory detection task were performed accurately, and data outside these criteria were excluded[12]. To test the reliability of response time, we performed an odd–even split-half analysis for different memory types[28,29]. The results revealed a high internal reliability, $r_{colour} = .97$, $p < .001$; $r_{graphics} = .94$, $p < .001$. Repeated measures of ANOVA were performed for response to the search task,

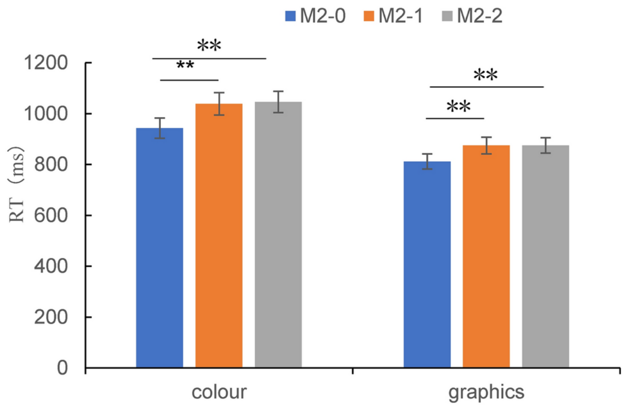

Fig. 4. Search response time for different categories of memory types and distracts.

with multiple comparisons performed using the Bonferroni correction method (Table 2 and Fig. 4). The results revealed a significant main effect of memory type, $F\ (1,29) = 40.88$, $p < .001$, $\eta_p^2 = .585$, $BF_{10} > 1000$, with longer search response time in the color condition compared to the graph condition. The main effect of matching type was significant, $F\ (2,58) = 29.511$, $p < .001$, $\eta_p^2 = .504$, $BF_{10} > 1000$. When the memory type was color or graph, the search response time of M2-2 was significantly longer than that of M2-0, $t(59) = 7.38$, $p < .001$, *Cohen's* $d = 0.40$; the search response time of M2-1 was significantly longer than that of M2-0, $t(59) = 6.16$, $p < .001$, *Cohen's* $d = 0.37$; there was no significant difference between M2-2 and M1-1. The interaction between these factors was not significant, $F\ (2,58) = 1.880$, $p > .05$, $BF_{10} < 0.613$. The results indicate that the VWM representation has a guiding effect on attention, but multiple item representations in VWM cannot simultaneously guide attention, supporting the SIT hypothesis.

Discussion

In this study, two experiments were conducted to investigate the influence of multiple item representations on attention in VWM. Experiment 1 demonstrated that when memory items and distractors belong to the same category, a color-based multiple item representations in VWM guide attention and support MIT hypothesis. However, when the memory type is graphic, the multiple item representations in VWM do not lead to attention, suggesting that MIT is not fully established. Building upon these findings, Experiment 2 manipulated the relationship between memory items and distractors. The results indicate that under different categories of memory items and distractors, when the memory type is color or graphic, VWM representation also guides attention, but multiple item representations in VWM cannot simultaneously guide attention, supporting the SIT hypothesis. The ability of multiple item representations in VWM to simultaneously guide attention is significantly influenced by memory types and the relationship between memory items and categories of distractors. Specifically, when the memory type is color and the memory item is in the same category as the distractor, multiple item representations in VWM can simultaneously guide attention, supporting the MIT hypothesis; However, when the memory type is graphic or the memory item is of a different category from the distractor, multiple item representations in VWM cannot simultaneously guide attention, supporting the SIT hypothesis.

In Experiment 1, when two colors are used as memory items, both M2-2 and M2-1 matching types have a guiding effect on attention during visual search tasks with response times showing M2-2 > M2-1. This suggests that multiple item representations in VWM simultaneously guide attention and support MIT hypothesis. This is consistent with the findings of Beck et al.[23]. They used eye tracking methods and asked participants to search for two colors. Under sequential search conditions, where participants searched for one color before searching for another, time conversion costs were observed. However, under simultaneous search conditions, participants were able to switch between the two search colors without incurring any time conversion cost. This indicates that two colors as VWM representations can simultaneously guide attention, the results support the MIT hypothesis. However, Beck's study was conducted under conditions where memory items were associated with search tasks[23]. In contrast, Experiment 1 expanded the scope by demonstrating that the guidance effect still occurs even when memory items are independent of search tasks, providing more comprehensive support for MIT hypothesis. Previous research has shown that individuals can maintain 3 to 4 items in VWM[26]. When the number of memories exceeds 4, i.e. exceeds the memory capacity of VWM, individuals need to consume a large amount of cognitive resources for maintenance. Due to the limited cognitive resources, they are less susceptible to interference in simultaneous visual search tasks, and the guidance effect will completely disappear. When the number of memories is within the capacity range, individuals do not need too many resources to maintain, and are susceptible to interference in simultaneous visual search tasks, resulting in a guiding effect.

However, in Experiment 1, when the memory item consisted of two graphics, both M2-2 and M2-1 matching types had a guiding effect on attention in visual search. However, the response time was found to be equal for M2-2 and M2-1. It was indicated that multiple item representations in VWM cannot simultaneously guide attention, the results support the SIT hypothesis. However, there are significant differences in the results under different types of memory. Whether multiple item representations in VWM can simultaneously guide attention is significantly influenced by memory types. This finding is consistent with Zhang et al.'s research, which suggested that color attribute in VWM has a stronger guiding effect in visual search compared to other stimulus attributes such as orientation and graphic[30]. The dimension weight theory suggests that individuals assign different attention weights to each dimension of visual stimuli during visual processing[31]. Attention weights are influenced both from top to bottom (such as stimuli related to the target receiving more weight and attention) and from bottom to top (such as stimuli with higher significance in search contexts receiving more weight and attention). Compared to graphics with low saliency, colors with high saliency gain greater attention weights, and stimuli with high saliency have an advantage in capturing attention from bottom to top, even eliminating top-down attention capture[32]. Therefore, in Experiment 1, compared to the memory type of graphics, it is easier to capture attention when the memory type is color, demonstrating the guidance of multiple item representations on attention. The accuracy of memory also indirectly provides evidence for this observation.. Under color conditions, there was no significant difference in memory accuracy between M2-2 and M2-1, but both were higher than M2-0; Under graphic conditions, there was no significant difference in memory accuracy between M2-2, M2-1, and M2-0. From a side perspective, compared to graphics, individuals are more sensitive to colors and have a deeper memory.

In summary, SIT and MIT are two important hypotheses regarding the role of multiple VWM representations in attention guidance, both of which have their own validity, but are regulated by memory types. SIT hypothesis believes that at any given time, VWM can only process or store one item as the focus of attention. Although the total capacity of VWM may support storing multiple items simultaneously (usually considered 3–4), only one item can be active, guiding attention. MIT hypothesis believes that VWM can simultaneously process or store multiple projects, all of which can guide attention to a certain extent. However, multiple VWM representations are regulated by memory types (color and graphic) for attention guidance.

When the memory type is color, the multiple item representations in VWM guide attention and support the MIT hypothesis; When the memory type is graphic, the multiple item representation in VWM cannot guide attention, supporting the SIT hypothesis. On the one hand, color is a highly salient visual feature that is easy to highlight and capture attention in visual scenes. Therefore, when the memory type is color, color representation is more likely to become the focus of attention. Compared to color memory, graphic memory is less likely to capture attention. Although graphic features are important, they are usually not as prominent or distinguishable as colors. On the other hand, the guiding role of VWM in attention involves top-down cognitive processing. The representation of objects in working memory will guide attention to prioritize the selection of objects that match or are similar to them in the visual context in a top-down manner. Color and graphics, as different types of memory, have different ways of representation and processing mechanisms in VWM. Color representation relies more on low-level visual feature processing, making the task relatively easy and consuming less cognitive resources; However, graphic representation involves more complex analysis of graphics and structures, making the task relatively difficult and consuming more cognitive resources.

Compared to Experiment 1, where the memory item and the distractors are of the same category, Experiment 2 was conducted under conditions where the memory item and the distracts are of the different categories. The results showed that M2-2 and M2-1 matching types had guiding effects on attention when the memory items were 2 graphics or 2 colors, but the response time is M2-2 = M2-1. This indicates that multiple item representations in VWM cannot guide attention and support SIT. This is different from Experiment 1, when the memory type is color, the multiple item representation in VWM can guide attention simultaneously. It shows that whether the stimulus can capture attention is not only determined by the characteristic attributes but also by whether the memory item and the distracter are in the same category, which reduces the guidance of color attributes to attention. This finding is consistent with previous studies. For example, van Moorselaa et al.'s study found that when memory items are colors and distractors are graphics (different categories), results support SIT[20]. In Fan et al.'s study, both memory items and interference items are colors belonging to the same category, supporting MIT[7]. Perceptual load theory can explain this phenomenon. According to this theory, the perceived task load determines the allocation of attention resources. When the perceived task load is low, only a part of attention resources are required for cognitive processing while other resources automatically process distractors causing interference effects. When the perceived task load is high, a large amount or even all attention resources are required for cognitive processing, and distractions unrelated to the task are automatically filtered and will not be processed[33]. In this study, Experiment 2 is characterized by a higher perceptual load compared to Experiment 1. This is due to the fact that memory items and distractors belong to different categories, thus requiring more cognitive resources for processing. Consequently, when attention resources are limited, the impact of memory-matched distractors on attentional allocation diminishes, while the influence of task-related targets increases. Indirect evidence supporting this notion comes from memory accuracy data. Specifically, Experiment 2 demonstrated no significant difference in the memory accuracy of M2-2, M2-1 and M2-0 under color and figure conditions when compared to Experiment 1. This suggests that the heightened perceptual load in Experiment 2 leads individuals to be less sensitive to color itself and instead prioritize attention towards the specific task at hand.

In summary, whether multiple item representations in VWM can simultaneously guide attention is influenced by whether the memory term and interference term belong to the same category. Working memory is a complex cognitive system that contains multiple memory components. When memory items and distractors belong to the same category, such as both colors, these similar memory information may rely on the same memory subsystem for processing. Due to the similarity in features, working memory needs to distinguish and select targets more finely. Therefore, multiple item representations in VWM can simultaneously guide attention, supporting the MIT hypothesis. When the memory item and distractor are not of the same category, such as remembering color and distractor being graphic, these different memory information may rely on different memory subsystems for processing. At this point, due to the mismatch of features caused by category differences, working memory is more likely to exclude distractors that match the memory item. Therefore, multiple item representations in VWM can not simultaneously guide attention, supporting the SIT hypothesis.

It's noteworthy that Experiment 2 revealed that even when the memory item and the search item belong to different categories, individual working memory still directs attention. A possible explanation is that Woodman et al. found that people can use sequential encoding to sequentially encode items into working memory, and the memory retention effect is the same as simultaneous encoding[34]. This result is also supported by Chung et al.[35]. Zhao and Vogel further revealed that sequential encoding and simultaneous encoding have the same individual differences[36]. This indicates that VWM resources have flexibility, as they can not only handle multiple projects simultaneously, but also store and extract information in sequence. Whether the memory items are encoded sequentially before or simultaneously with the search task, they can effectively guide attention during the search process. When the memory item is prioritized for encoding, its representation occupies the dominant weight in the VWM resource pool; Subsequently, when performing search tasks, although the categories of interference and memory items are different, the sustained activity of VWM resources still maintains a high level of activation of memory representations, helping individuals quickly identify and prioritize information related to memory items.

This study has certain limitations: the search target only appears in new items, and the memory items are only matched with interference items that do not include the search target, which may affect the generalizability of the research results. In the future, further exploration can be conducted on whether the search task response time is different when the search target appears in a memory item or a new item under M2-1 or M2-2 conditions, in order to further expand this study.

Conclusions

Whether multiple item representations in VWM can simultaneously guide attention is significantly influenced by memory types and the relationship between memory items and categories of distractors. Specifically, when the memory type is color and the memory item is in the same category as the distractor, multiple item representations in VWM can simultaneously guide attention, supporting the MIT hypothesis; However, when the memory type is graphic or the memory item is of a different category from the distractor, multiple item representations in VWM cannot simultaneously guide attention, supporting the SIT hypothesis.

Data availability

The data and materials used in this study are available upon request from the corresponding author.

Received: 16 November 2024; Accepted: 3 March 2025
Published online: 11 March 2025

References

1. Bartsch, L. M. & Shepherdson, P. Freeing capacity in working memory (WM) through the use of long-term memory (LTM) representations. *J. Exp. Psychol. Learn. Memory Cognit.* **48**(4), 465–482 (2022).
2. Duan, B. J., Wang, D. T., Li, W. X., Ma, X. F. & Zhou, A. B. The impairment of the visual spatial attention in Chinese children with dyslexia: A cognitive deficit or a developmental delay?. *Curr. Psychol.* **42**(6), 4341–4349 (2023).
3. Forsberg, A., Guitard, D. & Cowan, N. Working memory limits severely constrain long-term retention. *Psychon. Bull. Rev.* **28**(2), 537–547 (2021).
4. Kirmsse, A., Zimmer, H. D. & Ecker, U. K. H. Task demands differentially affect processing of intrinsic and extrinsic object features in working memory. *Exp. Psychol.* **69**(6), 320–334 (2023).
5. Rhodes, S., Buchsbaum, B. R. & Hasher, L. The influence of long-term memory on working memory: Age-differences in proactive facilitation and interference. *Psychonomic Bull. Rev.* **29**(1), 191–202 (2022).
6. Cao, Y., Li, X. Y. & Wang, B. C. The interaction between working memory and long-term memory in visual memory. *Chin. J. Appl. Psychol.* **29**(6), 1–19 (2023).
7. Fan, L. et al. Visual working memory representations guide the detection of emotional faces: An ERP study. *Vis. Res.* **119**, 1–8 (2016).
8. Fu, X., Ye, C., Liang, T., Hu, Z. & Liu, Q. The impact of retro-cue validity on working memory representation: Evidence from electroencephalograms. *Biol. Psychol.* **170**, 108320 (2022).
9. Folk, C. L., Leber, A. B. & Egeth, H. E. Made you blink! Contingent attentional capture produces a spatial blink. *Percept. Psychophys.* **64**(5), 741–753 (2002).
10. Pang, C. et al. Different attentional selection modes of object information in the encoding and maintenance stages of visual working memory. *Acta Psychol. Sin.* **55**(9), 1397–1410 (2023).
11. Che, X. W., Xu, H. Y., Wang, K. X., Zhang, Q. & Li, S. X. Precision requirement of working memory representations influences attentional guidance. *Acta Psychol. Sin.* **53**(7), 694–713 (2021).
12. Zhang, B., Hu, C. L., Chen, Y. Z., Miao, S. M. & Huang, S. The modulation of working memory load and perceptual load on attentional guidance from representations of working memory. *Acta Psychol. Sin.* **49**(8), 1009–1021 (2017).
13. Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**(3), 193–222 (1995).
14. Downing, P. E. Interactions between visual working memory and selective attention. *Psychol. Sci.* **11**(6), 467–473 (2000).
15. Bahle, B., Matsukura, M. & Hollingworth, A. Contrasting gist-based and template-based guidance during real-world visual search. *J. Exp. Psychol. Human Percept. Perform.* **44**(3), 367–386 (2018).
16. Downing, P. E. & Dodds, C. M. Competition in visual working memory for control of search. *Visual Cognit.* **11**, 689–703 (2004).
17. Olivers, C. N. What drives memory-driven attentional capture? The effects of memory type, display type, and search type. *J. Exp. Psychol. Human Percept. Perform.* **35**(5), 1275–1291 (2009).
18. Olivers, C. N., Peters, J., Houtkamp, R. & Roelfsema, P. R. Different states in visual working memory: When it guides attention and when it does not. *Trends Cognit. Sci.* **15**(7), 327–334 (2011).
19. Hu, Z. et al. Verbal learning, working memory, and attention/vigilance may be candidate phenotypes of bipolar ii depression in chinese han nationality. *Acta Psychol.* **226**, 103563 (2022).
20. Kocaarslan, M. The relationships between oral reading fluency, sustained attention, working memory, and text comprehension in the third-grade students. *Psychol. Schools* **59**(4), 744–764 (2022).

21. Song, J., Chang, L. & Zhou, R. Effect of test anxiety on visual working memory capacity using evidence from event-related potentials. *Psychophysiology* **59**(2), 1 (2022).
22. van Moorselaar, D., Theeuwes, J. & Olivers, C. N. In competition for the attentional, template: Can multiple items within visual working memory guide attention?. *J. Exp. Psychol. Human Percept. Perform.* **40**(4), 1450–1464 (2014).
23. Beck, V. M., Hollingworth, A. & Luck, S. J. Simultaneous control of attention by multiple working memory representations. *Psychol. Sci.* **23**(8), 887–898 (2012).
24. Hollingworth, A. & Beck, V. M. Memory-based attention capture when multiple items are maintained in visual working memory. *J. Exp. Psychol. Human Percept. Perform.* **42**(7), 911–917 (2016).
25. Foerster, R. M. & Schneider, W. X. Involuntary top-down control by search-irrelevant features: Visual working memory biases attention in an object-based manner. *Cognition* **172**, 37–45 (2018).
26. Zhang, B., Zhang, J. X., Huang, S., Kong, L. & Wang, S. Effects of load on the guidance of visual attention from working memory. *Vis. Res.* **51**(23), 2356–2361 (2011).
27. Fan, L. et al. Multiple representations in visual working memory simultaneously guide attention: the type of memory-matching representation matters. *Acta Psychol.* **192**, 126–137 (2018).
28. Xu, Z., Adam, K. C. S., Fang, X. & Vogel, E. K. The reliability and stability of visual working memory capacity. *Behav. Res. Methods* **50**(2), 576–588 (2018).
29. Zhao, C., Vogel, E. & Awh, E. Change localization: A highly reliable and sensitive measure of capacity in visual working memory. *Attention Percept. Psychophys.* **85**(5), 1681–1694 (2023).
30. Zhang, B., Huang, S. & Hou, Q. X. The priority of color in working-memory-driven ocular capture. *Acta Psychol. Sin.* **46**(1), 17–26 (2014).
31. Krummenacher, J., Müller, H. J. & Heller, D. Visual search for dimensionally redundant pop-out targets: Parallel-coactive processing of dimensions is location specific. *J. Exp. Psychol. Human Percept. Perform.* **28**(6), 1303–1322 (2002).
32. Wang, L. H., Yu, H. B. & Zhou, X. L. Interaction between value and perceptual salience in value-driven attentional capture. *J. Vis.* **13**(3), 1–13 (2013).
33. Lavie, N., Hirst, A., De Fockert, J. W. & Viding, E. Load theory of selective attention and cognitive control. *J. Exp. Psychol. Gen.* **133**(3), 339–354 (2004).
34. Woodman, G. F., Vogel, E. K. & Luck, S. J. Flexibility in visual working memory: Accurate change detection in the face of irrelevant variations in position. *Visual Cognit.* **20**(1), 1–28 (2012).
35. Chung, Y. H., Brady, T. F. & Störmer, V. S. Sequential encoding aids working memory for meaningful objects' identities but not for their colors. *Memory Cognit.* **52**(8), 2119–2131 (2024).
36. Zhao, C. & Vogel, E. K. Sequential encoding paradigm reliably captures the individual differences from a simultaneous visual working memory task. *Atten. Percept. Psychophys.* **85**(2), 366–376 (2023).

Acknowledgements

The work was supported by the Social Science Foundation of Liaoning Province (Grant Nos L20CSZ006).

Author contributions

C.D., D.S., and L.Z. designed and performed research; C.D. and B. Z analyzed the data and wrote the paper.

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval

This study were approved by Shenyang Agricultural University's Ethics Committee. All experiments were performed in accordance with relevant named guidelines and regulations. Prior to participation, informed consent was acquired from all study participants.

Additional information

Correspondence and requests for materials should be addressed to L.Z. or D.S.