Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

Neural evidence for an object-based pointer system underlying working memory

Halely Balaban[a,b,*], Trafton Drew[c] and Roy Luria[a,b]

[a] Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 6997801, Israel
[b] The School of Psychological Sciences, Tel Aviv University, Tel Aviv 6997801, Israel
[c] Psychology Department, University of Utah, Salt Lake City 84112, UT, USA

ARTICLE INFO

Article history:
Received 20 December 2018
Reviewed 26 February 2019
Revised 13 March 2019
Accepted 6 May 2019
Action editor Veena Kumari
Published online 18 May 2019

Keywords:
Object-based attention
Pointer system
Resetting
Updating
Visual working memory

ABSTRACT

To accomplish even rudimentary tasks, our cognitive system must update its representation of the changing environment. This process relies on visual working memory (VWM), which can actively modify its representations. We argue that this ability depends on a pointer system, such that each representation is stably and uniquely mapped to a specific stimulus. Without these pointers, VWM representations are inaccessible and therefore unusable. In three Electroencephalogram (EEG) experiments, we examined whether the pointers are allocated in an object-based, featural, or spatial manner: three factors that were confounded in previous studies. We used a feature change-detection task, in which objects moved and could separate into independently-moving parts. Despite the movement and separation being completely task-irrelevant, we found that the separation invalidated the pointers. This happened in a shape task, where the separation changed both the objects and the task-relevant features, but importantly, also in a color task, where the separation destroyed the objects while leaving the task-relevant features intact. Furthermore, even in a color task where all items had identical shapes, object-separation invalidated the pointers. This suggests that objects and not task-relevant features underlie the pointer system. Finally, when each object-part could be individuated already before the separation, the pointers were maintained, suggesting that the pointers are specifically tied to objects rather than locations. These results shed new light on the pointers which underlie VWM performance, demonstrating that the pointer system is object-based regardless of the task requirements.

* *Corresponding author*. Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 6997801, Israel.
E-mail address: halelyba@mail.tau.ac.il (H. Balaban).

https://doi.org/10.1016/j.cortex.2019.05.008

1. Introduction

Our environment is extremely dynamic, with objects constantly changing. For example, when we drive, everything around us moves, such that the incoming visual input alters from one moment to the next. Moreover, eye movements and microsaccades cause even static visual input to change its position across the retina (McConkie & Currie, 1996). Keeping track with this bustle poses a major challenge for our cognitive system. Visual working memory (VWM) holds a limited set of online representations that can be modified when the represented items change, leaving other items intact (Balaban & Luria, 2016a; Blaser, Pylyshyn, & Holcombe, 2000; Drew, Horowitz, Wolfe, & Vogel, 2012; Drew & Vogel, 2008). The nature of VWM representations is heavily investigated (e.g., whether representations constitute independent features or integrated objects, i.e., whether all features of an object are represented as an integrated unit; Brady, Konkle, & Alvarez, 2011; Luck & Vogel, 2013; Ma, Husain, & Bays, 2014), but we argue that one of VWM's defining characteristics remains under-studied: how can VWM continuously access and modify its representations? Examining the underlying mechanism responsible for this ability was the goal of the current research.

Previous work postulated a 'pointer system' that allows selecting and monitoring items' locations (e.g., object-files, Kahneman, Treisman, & Gibbs, 1992; fingers of instantiation (FINSTs), Pylyshyn, 2000). While the theoretical notion of pointers is well established, empirical support for it is limited, and comes mostly from spatial-tracking tasks (Pylyshyn & Storm, 1988). When individuals follow the movement of identical objects, they appear to select objects via a spatial code, a finding that closely corresponds to the spatial nature of this tracking task. Importantly, if a pointer system exists, it should be used not only for spatial tracking but whenever our online representations are updated, because VWM must rely on the pointers to access and modify the appropriate representation. For example, in driving, it is important to update our online representations not only following changes in other vehicles' locations, but also according to featural changes such as break lights' state (on/off).

Accordingly, we recently found evidence that VWM relies on pointers even in a feature-task, such that invalidating the pointers prevented VWM from updating its representations (Balaban & Luria, 2017). We used a shape change-detection task with moving items. The movement was task-irrelevant (because participants had to monitor for a shape change, which could only occur after a retention interval), but 'encouraged' VWM to create one representation for each shape, with one pointer supporting it. When these shapes separated into independently-moving parts, we hypothesized the pointers would be invalidated because following the separation neither part corresponded to the original item. To examine this, we relied on the electrophysiological marker of VWM, the contralateral delay activity (CDA; Vogel & Machizawa, 2004), an Event-related potential (ERP) component whose amplitude is higher when more information is held in VWM. The CDA sharply dropped after separation, and then the amplitude gradually rose. This indicates that the original single-object representations were no longer available in VWM, followed by the encoding of each shape-half as an individuated object, in a process we termed 'resetting'. We suggested that resetting involves removing the original representation from VWM, but it might also be that the representation is not removed but only inaccessible. Regardless of the exact fate of the representational content in resetting, for the present purposes what is important is that the mapping between the representation and the item in the world is lost.

We additionally found that resetting is accompanied by a temporary blindness to salient changes in the items' shape (Balaban & Luria, 2017). To demonstrate this, we used a novel "online change detection" task, in which subjects monitored the shapes of items that moved on the screen and could change during this movement. When two halves moved as a coherent object, a change in one of them was easily detected regardless of the time it occurred. Conversely, if the two halves moved together and then separated, changes were easily detected before and after the separation, but were largely missed if they coincided with the separation, i.e., with the resetting process. This is presumably because without a valid stimulus-to-representation mapping there is no way to access the representation and compare it to the ongoing input.

Numerous control experiments demonstrated that the separation is neither necessary nor sufficient for triggering resetting. First, the pointers could be invalidated by object-replacement, without separation or movement. Second, no resetting occurred when each shape-half could be individuated before the separation, allowing the pointers to be maintained throughout the separation. For example, when the shape-halves moved independently prior to combining, thereby allowing VWM to create two independent pointers, VWM did not reset when the newly-combined object split again (Balaban & Luria, 2017). These and additional control experiments (Balaban, Drew, & Luria, 2018a) demonstrated that the electrophysiological signature of resetting is specific and stable enough to serve as a novel marker for the loss of object-to-representation mapping.

Importantly, we argue that to fully account for past and present results, the additional "layer" of pointers must act as a bridge between the representations and the stimuli (Pylyshyn, 2000). The concept of pointers allows us to precisely predict when VWM updates its representations and when it resets, and to explain why extremely similar situations lead to qualitatively different processes in VWM (Balaban et al., 2018a). We return to this point in the Discussion.

Here, we examined whether pointers are based on features (in a feature-task), objecthood (even when objects were task-irrelevant), or locations (as suggested by previous studies; Pylyshyn, 2000), three factors that were all suggested to play a critical role in online processing (e.g., the multicomponent model of working memory; Baddeley & Logie, 1999). Previous spatial-tracking studies argued that the pointer system is location-based, but location-information was confounded with objects and also served as the task-relevant feature (Pylyshyn & Storm, 1988). Similarly, in our prior studies (Balaban & Luria, 2017) object-separation invalidated the pointers, but it simultaneously changed the object, the location-information, and the task-relevant feature (shape). Thus, it is impossible to conclude what information is necessary for the pointer system. The present study disentangled these three factors, examining whether pointers are allocated based on spatial, featural, or object information. Specifically, we tested whether destroying an object invalidates the pointers even when the task-relevant features remain the same (Experiments 1 and 2), and whether the pointers follow objecthood or location cues (Experiment 3).

2. Materials and methods

Data and code are available at the Open Science Framework: http://osf.io/ft462/. No part of the study procedures was pre-registered prior to the research being conducted.

2.1. Participants

Participants were Tel Aviv University students (age range 18–30 years), with normal or corrected-to-normal visual acuity and normal color-vision, who gave informed consent following the procedures of a protocol approved by the local ethics committee. Each experiment included 16 naïve participants (14 females, mean age 21.9 in Experiment 1, 12 females, mean age 22.5 in Experiment 2, and 8 females, mean age 24.9 in Experiment 3). None of the participants reached our pre-defined criterion of a >25% rejection rate due to blinks or eye movements, and hence none were replaced. Sample size was determined based on our previous work (Balaban & Luria, 2017), where we found an effect size of $d = 1.07$ for the relevant effect in a similar experiment. We found that for a power of 90% in a within-subjects design, 8 participants are required. We doubled this sample size, to make sure we have a high probability of revealing any effects of interest.

2.2. Stimuli and procedure

We used a bilateral change-detection task programmed using the Presentation software (Neurobehavioral Systems, Inc.). The bi-lateral task allowed us to isolate the CDA (Luria, Balaban, Awh, & Vogel, 2016; McCollough, Machizawa, & Vogel, 2007; Vogel & Machizawa, 2004), an ERP marker of VWM. CDA amplitude tracks the number of online representations, increasing or decreasing when the number of attended items increases or decreases online (Balaban & Luria, 2017; Drew, Horowitz, Wolfe, & Vogel, 2011; Vogel, McCollough, & Machizawa, 2005). Many studies have shown that the CDA specifically reflects VWM, and not low-level features such as the brightness of the items (Ikkai, McCollough, & Vogel, 2010; Luria, Sessa, Gotler, Jolicoeur, & Dell'Acqua, 2010; Ye, Zhang, Liu, Li, & Liu, 2014), the number of locations (Balaban & Luria, 2016a, 2016b; Ikkai et al., 2010; Luria & Vogel, 2014), or eye-movements (Kang & Woodman, 2014). The CDA is especially suited for examining online processes in VWM, due to its precise temporal resolution and the fact it can be measured not only during memory retention but also when the items are visible on the screen (Drew & Vogel, 2008; Emrich, Al-Aidroos, Pratt, & Ferber, 2009; Tsubomi, Fukuda, Watanabe, & Vogel, 2013).

Based on prior work, we analyzed CDA amplitude to specifically isolate pointer-related processes from general VWM processes (Balaban & Luria, 2017). We recently found that situations which invalidate the stimuli-to-representation mapping, such as object-separation and object-replacement, cause a transient decrease in the amplitude of the CDA. This suggests that the original representations were inaccessible and later, new representations of the items in their novel form were encoded, in line with the loss-of-mapping notion. Several control experiments showed that the CDA-drop reflects the cognitive process of resetting following the invalidation of a pointer, rather than reflecting the perceptual signal of separation. Complete details can be found in the original studies (Balaban, Drew, & Luria, 2018b; Balaban et al., 2018a; Balaban & Luria, 2017), but the key conclusions are that separation is neither necessary nor sufficient for a drop (meaning the drop is not about separation, movement, increasing the set-size, or surprise), and that very similar situations produce a drop when the pointers are invalidated, but not when they can be maintained. We therefore use the CDA-drop as a neural marker for the loss of a pointer in VWM.

Each trial started with a 750 msec fixation display of a black cross (.4° × .4° of visual angle from a viewing distance of approximately 60 cm) in the center of a gray screen. Then, two white arrows (1.9° × .4°) appeared for 200 msec, pointing either left or right (randomly determined with an equal probability), indicating the to-be-attended side for the upcoming trial. When describing the number of items throughout the paper, we always refer only to the relevant side. After a 300–500 msec (randomly jittered) fixation display, the memory array appeared, with items that were randomly chosen without replacement (independently for each side). Each side always included the same number and type of stimuli. Items appeared at random locations inside an invisible 4.5° × 3.5° rectangle (one in each side of the screen), with a minimum distance of 2° between items. The items moved for some period of time that varied between experiments, and then disappeared for 900 msec. The items then reappeared, and participants indicated, in an unspeeded manner, via button press whether one item changed relative to the memory array (using the “z” and “/” keys on a standard keyboard for "same" and "different", respectively). Changes occurred with 50% probability.

Experiment 1. Stimuli were 4 top-half and 4 bottom-half shapes (1.6° × .8°), which could form 16 different shapes. Both halves of each shape were presented in one of six shades of blue, Red-Green-Blue (RGB) values: 0,0,255; 0,0,192; 0,128,255; 128,128,255; 64,192, 255; 0,128,192 (we used these similar colors to minimize verbal encoding, but using similar as opposed to distinct colors has been shown not to affect the CDA; Ikkai et al., 2010). Each trial included 2 colored shapes. Items in the memory array moved for 1000 msec in straight trajectories, whether up, down, or horizontally towards the fixation (randomly determined), but never crossing the center of the screen. Then the items remained stationary for 300 msec. There were 2 movement conditions: the two halves in each shape either moved as a coherent object throughout, or separated after 400 msec and then moved independently. To ensure participants paid attention to the initial 400 msec of the movement, 10% of the trials (and 25% in the first block) were catch trials, in which the memory array ended after the initial 400 msec. These trials were not further analyzed. There were two tasks (blocked with order counterbalanced across participants): color, in which one color (i.e., both halves of one shape) could change to a new color and the items' shapes were irrelevant, and shape, in which one shape-half could change to a new half of the same side (i.e., a top-half could only change to a new top-half and a bottom-half to a bottom-half) and the items' colors were irrelevant. After 12 practice trials, participants completed 14 experimental blocks (7 blocks per task) with 60 trials in each.

Experiment 2. Stimuli were squares ($1.6^\circ \times 1.6^\circ$), presented in the same colors used in Experiment 1. Items moved for 800 msec, then disappeared. There were 3 conditions: 2 or 4 squares moving as coherent units throughout, and 2 squares separating into two identical upright rectangles that moved independently after 400 msec of movement. Regardless of the condition, only a single color (i.e., one square or the two rectangles that comprised it) could change to a new color. After 12 practice trials, participants completed 11 experimental blocks with 60 trials in each.

Experiment 3. The experiment was the same as Experiment 2, except for the following. Only 2 colored squares appeared on each trial (either moving coherently or separating after 400 msec). Half the blocks included the stimuli from Experiment 2, and the other half (order counterbalanced across participants) included the same stimuli with a 1-pixel black line around each rectangle. After 12 practice trials, participants completed 14 experimental blocks (7 blocks with frames and 7 without) with 60 trials in each.

2.3. *EEG recording and analysis*

EEG was recorded inside a shielded Faraday cage, using a BioSemi ActiveTwo system, from 32 scalp electrodes at a subset of locations from the extended 10–20 system, and from two electrodes placed on the mastoids. Electrooculogram (EOG) was recorded from two electrodes placed 1 cm from the external canthi, and from an electrode beneath the left eye. Data was digitized at 256 Hz.

Offline signal processing was performed using EEGLAB Toolbox (Delorme & Makeig, 2004), ERPLAB Toolbox (Lopez-Calderon & Luck, 2014), and custom Matlab (The Mathworks, Inc.) scripts. All electrodes were referenced to the average of the mastoids. The continuous data was segmented into epochs from −200 from memory array onset to the end of the retention interval (+2200 in Experiment 1; +1700 in Experiments 2 and 3). Artifact detection was performed using a sliding window peak-to-peak analysis, with a threshold of 80 μV for the EOG electrodes, and 100 μV for the analyzed electrodes (P7, P8, PO3, PO4, PO7, and PO8). This procedure resulted in a mean rejection rate of 6.5% in Experiment 1, 3% in Experiment 2, and 4% in Experiment 3 (for evidence that eye movements are not responsible for the CDA-drop, see Balaban et al., 2018a; Balaban & Luria, 2017). For plotting purposes, the epoched data were low-pass filtered using a noncausal Butterworth filter (12 dB/oct) with a half-amplitude cutoff point at 30 Hz. Only trials with a correct response were included in the analysis. Statistical analyses were performed on the unfiltered data, to avoid potential effects of filtering on the observed results.

Epoched data were averaged separately for each condition, and the CDA difference wave was calculated by subtracting the average activity at electrodes ipsilateral to the memorized side from the average activity at electrodes contralateral to the memorized side.

2.4. *Experimental design and statistical analysis*

Our main dependent measure was mean amplitude 200–300 msec after the separation (the "Drop" time-window, see Balaban et al., 2018a; Balaban & Luria, 2017). In Experiment 3, we also tested an earlier time-window of 100–200 msec after the separation (the "Pre-Drop" time-window) and the amplitude during the retention interval (the "Post-Drop" time-window; 800–1700 msec). We present only the results from the average of 3 electrode pairs (P7/8, PO3/4, and PO7/8), but we found the same patterns of activity in each pair separately.

For each experiment, we conducted two separate Analyses of Variance (ANOVAs), one on mean amplitude and one on accuracy. In Experiment 1 the independent factors were Task (shape *vs* color) and Movement condition (integrated movement *vs* separation). In Experiment 2 the independent factor was Movement condition (2 integrated, 4 integrated, and separation). In Experiment 3 the independent factors were Stimuli-Type (frames *vs* no-frames) and Movement condition (integrated movement *vs* separation). We focus on the results of planned comparisons (contrasts) between the different conditions. Specifically, to determine whether there was a drop in amplitude, we compared the Separation condition to the Integrated movement condition. We report effect sizes for all statistical tests (partial η^2 for interactions and Cohen's d for pairwise comparisons).

For completeness, we also report the split-half Spearman–Brown corrected reliability for each task. In Experiment 1, we found a reliability of $r = .92$ in the shape task and $r = .79$ in the color task. In the color task of Experiment 2, we found a reliability of $r = .86$. In Experiment 2, we found a reliability of $r = .93$ in the No-Frames blocks, and $r = .88$ in the Frames blocks. These high reliability scores are in line with previous findings in change detection paradigms (e.g., Xu, Adam, Fang, & Vogel, 2018).

3. Results

3.1. *Experiment 1: pointers are allocated to objects and not to relevant features*

In Experiment 1, we presented two colored shapes in a change-detection task (Fig. 1; throughout the paper, the number of items refer only to the relevant hemifield, see Materials and methods). The items in the memory array moved, but movement was task-irrelevant: Participants only had to remember their object-features (color or shape, see below). Even so, coherent movement is a very strong Gestalt cue (Luria & Vogel, 2014), and we expected VWM to represent each colored shape as an integrated object, creating a single mapping between the item and its representation.

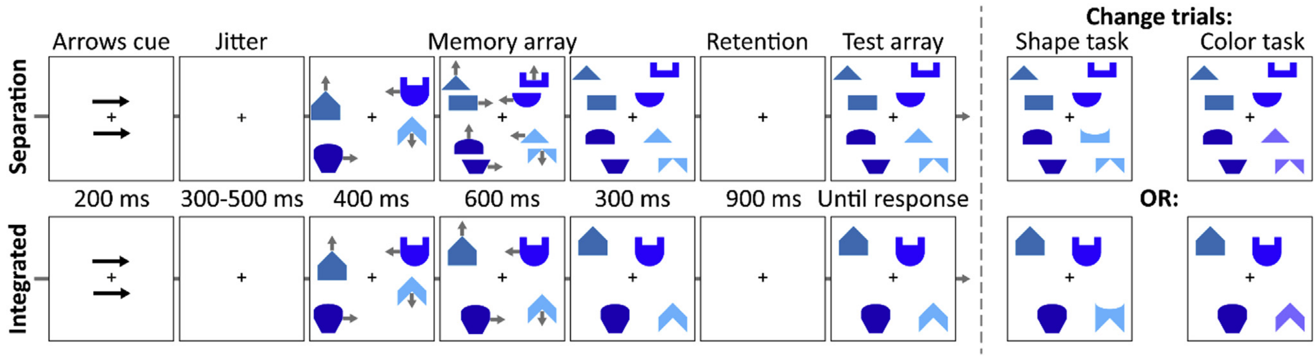

Fig. 1 – Examples of trial sequences in the different conditions in Experiment 1. On half of the trials, the colored shapes separated into independently moving halves after 400 msec (top row). On the other half, the items continued to move without separating (bottom row). The right side of the figure depicts possible change-trials in the two types of tasks: shape (left), with a change in a single shape-half, and color (right), with a change in a single color. Note that changes are the same regardless of the separation, making the separation task-irrelevant.

On half of the trials, each colored shape then separated into two independently-moving halves. In half of the blocks, participants' task was to encode only the items' shape, and their colors were irrelevant. Critically, if a change occurred in the test-array, only a single shape-half changed, regardless of the separation. Thus, even without the separation, subjects only had to attend to the separate halves of the shapes, making the separation task-irrelevant (because whether or not the halves were adjacent did not affect the task). Notably, we nevertheless reasoned that the separation will destroy the stimulus-to-representation mapping: following separation, each of the original items was replaced by two shape-halves, neither of which corresponded to the whole object. Thus, we predicted a resetting process due to the loss of a pointer, reflected by a CDA drop (Balaban et al., 2018a; Balaban & Luria, 2017). Importantly, in the current shape task, even if the separation invalidates the mapping, there are two possible explanations for this, corresponding to two theoretical structures for the pointer system. It is possible that the pointers were allocated according to the task-relevant feature (shape), such that resetting reflected the dramatic (although task-irrelevant) change in shape. Alternatively, the pointers could have been allocated to each integrated object, which the separation then destroyed, reflecting an object-based pointer system. Notably, both these hypotheses are in line with findings from spatial tracking and from our previous work, because these experiments confounded objects with the task-relevant feature.

To disentangle the two hypotheses, in the remaining blocks we used identical stimuli and design, but participants performed a color task instead of a shape task (note that the only thing that differs between the tasks is the nature of potential changes in the test array, which should not affect the CDA, because the CDA time-window ends at the end of the retention interval and before the test array appears). This means the shapes were now irrelevant and a change could only occur in a single color (i.e., both halves in separation trials). In the color task, the separation only affected the objecthood cues, and not the task-relevant features, because the color remained the same even if the object separated: the separation only "duplicated" the color. Thus, in terms of task-relevant information subjects can ignore the separation, because the same colors are present on the screen (note that holding several copies of the same color in VWM doesn't affect accuracy; Gao, Yin, Xu, Shui, & Shen, 2011). Importantly, if the pointers are allocated to integrated objects regardless of the task, the separation should still invalidate them in the color task, because the old pointer doesn't correspond to any of the new objects' parts. Conversely, if the pointers track the task-relevant features (colors), there should not be a CDA-drop, because the separation did not affect the colors. Critically, this manipulation allows us to examine whether the pointers are task-oriented or object-oriented, regardless of the task-relevant feature.

Thus, if the pointer system is feature-based, a drop should be found only in the shape task (where the separation changes the task-relevant feature) but not in the color task (where the task-relevant information remains the same). Conversely, if the pointer system is object-based, a drop should appear in both tasks, because the separation abruptly changes the objects. Since the CDA takes about 200 msec to respond (e.g., to initially rise; Vogel et al., 2005), the critical time-window we used is 200–300 msec after the separation. This time-window is identical to all our previous studies (Balaban et al., 2018a; Balaban & Luria, 2017).

For the shape task, there was a significant drop in CDA amplitude, $F(1,15) = 5.21$, $p < .05$, $d = .43$ (Fig. 2a), reflecting VWM resetting, in line with a loss of the pointer responsible for the stimuli-to-representations mapping (Balaban & Luria, 2017). The fact that resetting followed the separation still leaves open two possibilities for pointer-allocation: it can either be object-based (with the drop being due to the destruction of an object), or feature-based (with the drop being due to the task-irrelevant change in shape). If the latter is true, no resetting is expected in the color task, where the separation only duplicated the task-relevant feature of color. However, we found a significant drop in CDA amplitude for the color task, $F(1,15) = 6.39$, $p < .05$, $d = .41$ (Fig. 2b), and no significant Task by Movement condition interaction, $F < 1$, $\eta^2 = .04$, demonstrating that the effect was similar in both tasks. This

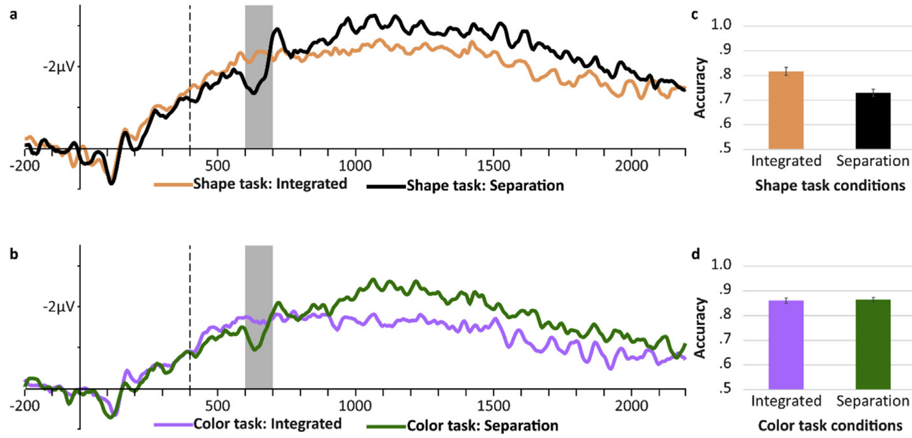

Fig. 2 – The results of Experiment 1. (a) The CDA results of the shape task. The dashed black line indicates the time of separation. The gray rectangle marks the analyzed time-window, in which the amplitude significantly dropped. (b) The CDA results of the color task, in which there was also a drop in CDA amplitude in the same time-window. (c) The accuracy results of the shape task (error bars show standard error of the mean), showing reduced accuracy following separation. (d) The accuracy results of the shape task, with the same accuracy regardless of the separation.

suggests that the pointers of VWM are allocated to objects, regardless of the task.

It is important to note that we have previously shown that the CDA-drop does not index object-separation per se, but rather the cognitive process of resetting following the loss of a pointer (Balaban et al., 2018a; Balaban & Luria, 2017). Briefly, the drop was found when the pointers were invalidated without separation, and was absent when there was separation that allowed the pointers to be maintained (e.g., when each shape-half moved independently before the joint movement, or when each shape-half had a distinct color). Thus, if the pointers can be maintained the CDA steadily rises, indicating VWM updating, and if the pointers are invalidated the CDA drops, indicating resetting. We use the CDA-drop as an indication that VWM had to reset, i.e., that the pointers were invalidated. Accordingly, the present results demonstrate that the pointers were invalidated by the destruction of an object, regardless of whether or not it changed the task-relevant features.

Interestingly, behavioral performance in the task followed a different pattern. The separation decreased change-detection accuracy in the shapes task, $F(1,15) = 28.16$, $p < .0001$, $d = 1.36$ (Fig. 2c), but not for colors, $F < 1$, $d = .10$, leading to a Task by Movement condition interaction, $F(1,15) = 39.15$, $p < .00005$, $\eta^2 = .72$ (Fig. 2d). In the present task, change detection performance was measured well after VWM recovered from the separation and created new representations. Therefore, accuracy here is not indicative of the pointer process (for a behavioral investigation of the pointer processes using another task, see Balaban et al., 2018b; Balaban & Luria, 2017). Nevertheless, accuracy can still shed light on other aspects of VWM.

For the colors, the fact that accuracy was equivalent for the Integrated and Separation conditions shows that in this task the separation did not change the task-relevant information at all, because it only duplicated the colors, and multiple copies of the same color are cost-free in terms of VWM accuracy (Gao et al., 2011). For the shape task, the lower accuracy in the Separation condition compared with the Integrated condition is further evidence for the influence of objects on VWM representations (e.g., Gao, Gao, Tang, Shui, & Shen, 2016). Namely, both conditions included the same four shape-halves, but their objecthood cues differed, and the VWM load they produced reflected this difference. When the shapes didn't separate, the four shape-halves were perceived as two objects (because of the coherent movement), while after separation, each half moved independently and was hence perceived as an independent object. The fact that four objects produce lower accuracy than two objects is not surprising, but the fact that this pattern was observed despite both conditions including the same amount of task-relevant information (i.e., four shape-halves) suggests that objects played a key role in VWM capacity-allocation (see also Balaban & Luria, 2015).

The dissociation between the two tasks indicates that accuracy reflected the task-relevant features (in line with the task instructions), contrary to the CDA which appears to track the online status of VWM, thereby revealing the dynamics of the underlying pointers. We further pursue this notion in the Discussion.

3.2. *Experiment 2: object-based pointers even in extreme conditions*

Although shapes were irrelevant in the color blocks of Experiment 1, each color had a unique form, and items separated asymmetrically. Perhaps this is why we observed evidence favoring an object-based mapping. Traditionally, in VWM color tasks, all items have identical shapes and only differ in color. In Experiment 2, we used only a color task, and the items were always squares (Fig. 3). In the Separation condition, the items separated into two rectangles, which were identical to each other not only in color but also in shape. If the pointers are indeed allocated to objects regardless of the

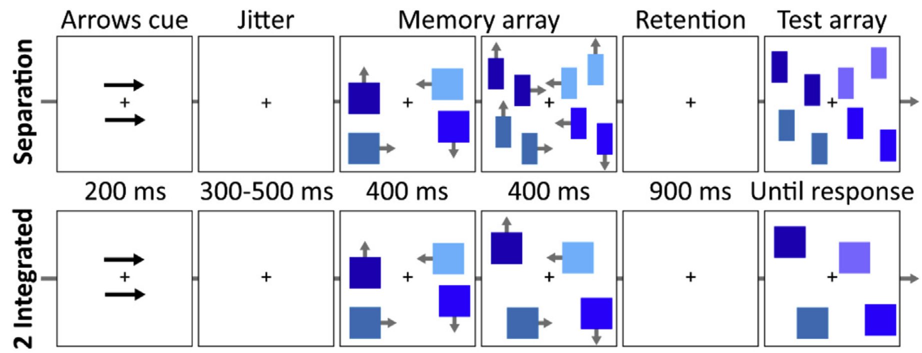

Fig. 3 – Examples of trial sequences in the different conditions in Experiment 2. Colored squares could either separate during their movement (top row), or move without separating (bottom row; there was also a condition with 4 colors, not shown here). A change could always occur in a single color (i.e., both halves in the Separation condition).

task-relevant feature, as suggested by the results of Experiment 1, the extreme prediction is that even in Experiment 2 VWM should reset following separation, because an object was destroyed and thus new pointers should be reinstated.

Indeed, following the separation the CDA amplitude significantly dropped relative to the 2 integrated colors condition, $F(1,15) = 10.13$, $p < .01$, $d = .66$ (Fig. 4a). This suggests that the destruction of an object invalidated the VWM mapping, even though the separation left the task-relevant features (i.e., colors) unchanged. This strongly supports the claim that the pointers are object-based, because here the post-separation items were identical not only in their color but also in their (task-irrelevant) shapes.

Once again, accuracy only reflected the task-relevant information, which didn't change following the separation. Therefore, accuracy following the separation was the same as when the two items didn't separate, $F < 1$, $d = .06$, and was higher than when four colors were presented, $F(1,15) = 247.73$,

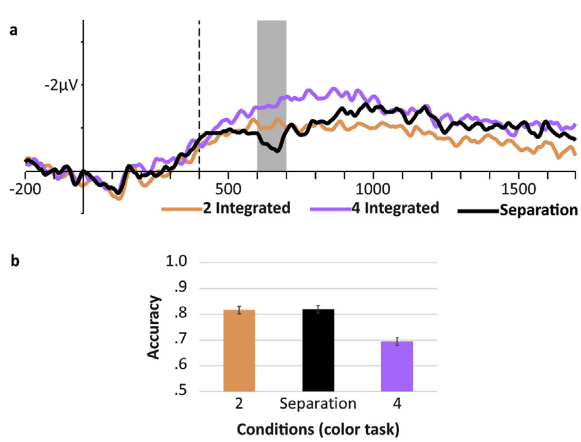

Fig. 4 – The results of Experiment 2. (a) The CDA results. The dashed black line indicates the time of separation. The gray rectangle marks the analyzed time-window, in which the amplitude significantly dropped. (b) The accuracy results of the shape task (error bars show standard error of the mean), with accuracy for the Separation condition being the same as for 2 colors and higher than for 4 colors.

$p < .000001$, $d = 2.15$ (Fig. 4b). This is another indication that separation in the color task did not affect VWM accuracy, only the underlying pointer system.

3.3. *Experiment 3: pointers are allocated based on objecthood cues, not location cues*

In Experiments 1 and 2, we found that the task-relevant features did not determine the allocation of pointers, and argued that pointers are object-based. Yet, another possibility is that the pointers are allocated to locations. Usually, each object occupies a single location in space, and each location contains one object. Thus, objects and locations are typically confounded. This was also the case in our Experiments 1 and 2. Our goal in Experiment 3 was to disentangle objects and locations, providing additional support for the object-based pointer system argument. We accomplished this by holding the location cues identical to Experiment 2, but changing the objecthood cues such that each half of the object can be perceived as an independent item even during their joint movement, by marking each half with a thin frame. If the results of Experiments 1 and 2 indeed reflected an object-based pointer system, the prediction is that marking each half as an independent object prior to the separation will eliminate the necessity to reset. Based on this logic, the frames should make it possible to create two independent pointers (one for each half) at the onset, and hence the separation should not invalidate these pointers. Therefore, in this situation we expect no CDA-drop. This diverges from Experiments 1 and 2, which only involved situations where a coherent object separates, thus triggering resetting.

Experiment 3 included Experiment 2's conditions of two colored squares that either separate (where we predicted to replicate the resetting effect) or did not separate, and in different blocks included the same conditions with a thin black frame around each rectangle (Fig. 5). These frames allowed individuating the rectangles even during their joint movement, thus encouraging participants to perceive each rectangle as an independent object. Critically, if the pointers are allocated to integrated objects, with the frames there should be two pointers already during the joint movement. Therefore, the separation should not change the pointers, and

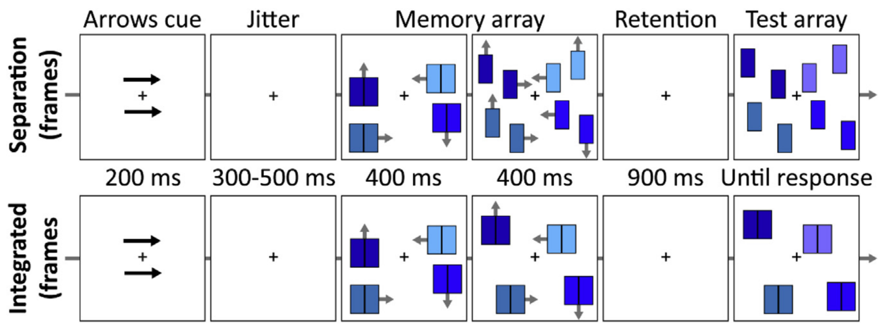

Fig. 5 – Examples of trial sequences in the different conditions of the Frames blocks of Experiment 3 (for the No-Frames blocks, see Fig. 3). On half of the trials, the colored shapes separated into independently moving halves after 400 msec (top row). On the other half, the items continued to move without separating (bottom row). A change could always occur in a single color (i.e., both halves in the Separation condition).

we expect no resetting and no CDA-drop (see also Balaban et al., 2018a; Balaban & Luria, 2017). Conversely, if the pointers are allocated solely according to locations, we should find a CDA-drop following the separation even with the added frames, because the spatial configuration of both Separation conditions is identical.

The results supported the notion that the pointer system is object-based. A CDA-drop followed separation only when the entire square was perceived as one object. Replicating the results of Experiment 2, the CDA amplitude dropped following separation in the No-Frames blocks, and this effect just missed significance level, $F(1,15) = 4.02$, $p = .06$, $d = .37$ (Fig. 6a; note that this effect is a replication of Experiment 2). The fact that the resetting effect was slightly smaller than with the same stimuli in Experiment 2 might be due to the Frames blocks, which exposed participants to the halves in a more separate form. This could potentially make the halves easier to individuate even when they appear without the frames, allowing for two separated pointers to be created in at least some of the No-Frames blocks and reducing the CDA-drop. In line with this notion, the size of the drop was numerically larger for participants who performed the No-Frames blocks first than for participants who performed the Frames blocks first (a difference of .3 μV *vs* .2 μV respectively), although our sample size does not allow us to compare them statistically. In any case, our main focus is the Frames blocks, in which there was no drop in amplitude for the Frames blocks, $F(1,15) = 1.38$, $p = .26$, $d = .19$ (Fig. 6b), producing a significant Stimuli-Type by Movement condition interaction, $F(1,15) = 13.97$, $p < .005$, $\eta^2 = .48$. This suggests that when each part could be perceived as an independent object, the pointers were not invalidated by separation. This is presumably because each rectangle was given a different pointer, allowing VWM to continuously access the representations and update them (see Balaban et al., 2018a; Balaban & Luria, 2017).

Notably, the two Separation conditions (with and without the frames), were identical in terms of spatial cues, and even produced equivalent CDA amplitudes before the separation

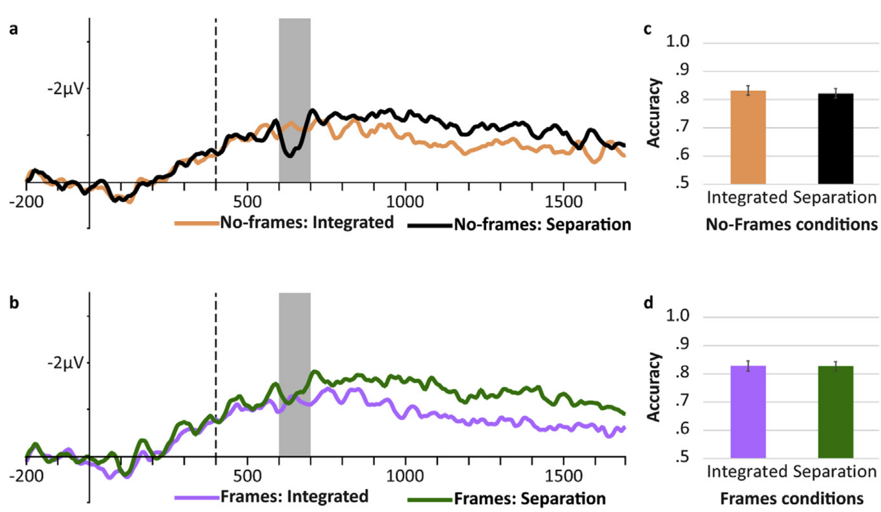

Fig. 6 – The results of Experiment 3. (a) The CDA results of the No-Frames blocks. The dashed black line indicates the time of separation. The gray rectangle marks the analyzed time-window, in which the amplitude significantly dropped. (b) The CDA results of the Frames blocks, in which there was no drop in CDA amplitude in the same time-window. (c) The accuracy results in the No-Frames blocks (error bars show standard error of the mean), with the same accuracy regardless of the separation. (d) The accuracy results in the Frames blocks, with the same accuracy regardless of the separation.

and during the retention interval (there was no interaction of Stimuli-Type and Movement condition in either time-window: $F(1,15) = 2.13$, $p = .17$, $\eta^2 = .12$ and $F(1,15) = 2.27$, $p = .15$, $\eta^2 = .13$ respectively), a point we return to in the Discussion. Nevertheless, these conditions produced different mappings, presumably because they differed in the objecthood cues they afforded. Impressively, very subtle changes in the appearance of the stimuli (a 1-pixel frame) appear to change the mapping of the pointer system, because they provide different objecthood cues. This suggests that the pointers are indeed object-based, rather than location-based.

As in the previous experiments, accuracy only reflected the task-relevant information, which was the same in all conditions, all $Fs < 1$, all $\eta^2 s < .04$ (Fig. 6c and d).

4. Discussion

The goal of this study was to examine the nature of the pointer system that connects active representations in VWM with the real-world stimuli, allowing VWM to keep track of our dynamic environment. While the nature of VWM representations has been heavily investigated (e.g., whether representations constitute integrated objects or independent features; Luck & Vogel, 2013; Ma et al., 2014), not much is known about the underlying processes that allow updating these representations when the relevant input changes. We recently established a novel tool for studying the pointer system, by demonstrating that invalidating this stimulus-to-representation mapping triggers a unique CDA-drop (Balaban & Luria, 2017). Here, the CDA-drop was used to examine whether the pointer system relies on a spatial, featural, or object-based code. We tested whether destroying an object necessarily renders the mapping associated with it unusable, even when controlling for changes in features and locations.

In Experiment 1, we found that the separation of a colored shape invalidated the mapping when shape task-relevant. Critically, object-separation also invalidated the mapping in a color task, even when shape was irrelevant. Thus, the loss of pointers was not due to changing the task-relevant features, but rather to destroying the object. In Experiment 2, we pushed this finding further, demonstrating that even in a color task where the items' shapes were identical, and the post-separation parts were equivalent (squares separating symmetrically into rectangles), the separation still made the pointers unusable, thereby necessitating resetting. Finally, in Experiment 3 we found that this pattern was indeed object-based: when each part could be easily individuated prior to the separation (by marking each half with a thin frame), the mapping was maintained throughout the separation. This demonstrates that when objecthood cues and locations are contrasted, the pointers are allocated to integrated objects.

Taken together, our results strongly suggest that a mapping is formed between each VWM representation and a single coherent object. Importantly, this was found despite the tasks always involving features (shape or color) and not objects. Of course, featural and spatial information are likely involved in establishing what is considered an independent object to begin with, but once the objects are individuated, they form the basic units on which the pointer system operates. We do not claim that features and locations are not important for the pointer system as well as other aspects of VWM, simply that it is integrated objects, and not these factors, to which pointers are allocated. Destroying an object invalidates the mapping, even if the task-relevant features of this object remain unchanged. Without a stable mapping the VWM representation is inaccessible, meaning that if the object changes, the representation cannot be accessed and updated accordingly. Therefore, the loss-of-mapping causes VWM to reset, i.e., replace the old representation with a new representation of the objects in their new form (Balaban et al., 2018a, 2018b; Balaban & Luria, 2017). Thus, the resetting process (marked by the CDA-drop) can uncover the hidden connections between VWM and the world our VWM system is representing.

Previously, studying the pointer system was mostly done in a spatial context, specifically using multiple object tracking (e.g., Pylyshyn & Storm, 1988), where participants track a subset of identical moving objects. These findings suggested that participants were able to individuate and track a handful of objects, based on their spatial locations (all items in this task were visually identical). However, the paradigm itself required individuating objects and tracking them through space, so it was not clear whether similar pointers arise naturally when spatial positions do not explicitly play a part. Here, the paradigm was based on features (shape or color), and locations were irrelevant. Furthermore, the association of features to objects was irrelevant. Therefore, the present evidence that pointers are bound to integrated objects suggest that the centrality of objects is a key characteristic of the pointer system, regardless of the task.

Notably, usually spatial location and objecthood are confounded, because a single object occupies a single place. Indeed, pointers have been shown to be used when updating the positions of moving objects (Kahneman et al., 1992; Pylyshyn, 2000; Pylyshyn & Storm, 1988). However, the present results cannot be explained by a purely location-based pointer system. In Experiment 3, the two Separation conditions (with and without the frames) were identical in terms of spatial locations, but differed in their objecthood cues, and consequently produced different pointer allocations. Unlike previous claims that regarded only locations as determining the mappings, here we found that the joint contribution of several objecthood cues, including, but not limited to, unique locations (or common motion), determines the pointers' allocation.

An object-based pointer system is in line with other object-based dynamics in the visual system generally (Chen, 2012), and in VWM specifically (Balaban & Luria, 2015; Gao et al., 2016; Luria & Vogel, 2011; Zhang & Luck, 2008). However, there are also many claims that the different features of an object are held independently in VWM (Bays & Husain, 2008; Fougnie & Alvarez, 2011; Ma et al., 2014), suggesting the organization of VWM might be feature-based. We found that even if features affect VWM representations (i.e., the contents of VWM), they are not the basic unit of the mappings between these representations and the actual stimuli that produced them. Rather, integrated objects form the building blocks of the mappings that are at the heart of VWM's ability to faithfully represent an ever-changing environment.

This points to a dissociation between the representational content of VWM, and the pointers that underlie it. In other words, we argue for a conceptual distinction between what VWM represents (e.g., whether all the features of a given object are bound in VWM) on the one hand, and the ongoing process that allows the representation to be dynamically adjusted as the environment changes on the other hand. The notion that these two levels are not identical was supported by additional dissociations revealed in the present results. First, there was a double dissociation between behavioral accuracy and the pointers. In Experiment 1, separation triggered resetting (as marked by the CDA-drop) in both tasks, but led to lower accuracy only in the shape task. In Experiment 3, separation triggered resetting only in the No-Frames blocks and not in the Frames blocks, but accuracy was the same for both block types. This is presumably because accuracy reflects the task demands and not the underlying pointers. Note, however, that we have observed massive behavioral costs associated with the loss of correspondence, but these effects are only observed when VWM is probed during the resetting process, allowing VWM no time to recover. In these situations, salient changes that coincide with the loss of pointers go undetected (Balaban et al., 2018a, 2018b; Balaban & Luria, 2017).

Second, the CDA as an electrophysiological marker of VWM is also dissociated from the pointer-specific CDA-drop, as can be seen in the pre-separation CDA of Experiment 3. While a drop was found only without frames, the amplitude of the Separation conditions was similar to the Integrated items condition, even with the added frames. We argue that VWM grouped the two framed rectangles (due to the strong common-fate cue), while still holding on to their separate pointers, allowing the squares to be later individuated without resetting. We have replicated these results with other easy-to-individuate groups (Balaban et al., 2018a; Balaban & Luria, 2017), but further work is required to establish this hypothesis. Taken together, our results suggest a dissociation between VWM representations and the pointer system.

Thus, we argue that the present results, along with previous findings regarding the resetting process (Balaban et al., 2018a, 2018b; Balaban & Luria, 2017) cannot be accounted for only by the removal of representations from VWM following dramatic changes in the environment. Instead, it appears that a bridging layer of pointers is needed to explain the full scope of evidence. For example, pointers explain why certain separation situations will result in resetting and a CDA-drop while other will lead to updating and no drop. The idea of an ongoing mapping between certain aspects of the stimuli and the representations in VWM provides a unifying framework that can accommodate all our present findings as well as findings from other tasks of spatial or featural tracking (Blaser et al., 2000; Kahneman et al., 1992; Pylyshyn & Storm, 1988), and provide testable predictions for future experiments. We conclude that VWM representations crucially rely on pointers to successfully update according to the perceptual input's dynamics, but these two systems are functionally distinct and dissociable.

Several interesting issues remain open and could be the target for future studies. First, our findings were based on simple stimuli, and it would be important to establish whether resetting occurs also for more complex items, as the ones we encounter in daily life. Previously, we found even larger effects of resetting when we tested the separation of slightly more complex items (i.e., random polygons; Balaban et al., 2018a; Balaban et al., 2018b; Balaban & Luria, 2017). This suggests that resetting might even be stronger for complex real-world stimuli, perhaps due to the stronger objecthood cues they convey. Second, while the present and previous results shed light on some aspects of the pointer system (e.g., the fact that it is object-based as revealed in the present work, and the fact it operates locally, on each item in VWM separately, see Balaban et al., 2018b), many characteristics of this system are still unclear. Specifically, one interesting question is whether the pointer system is a multidimensional or unitary construct, which we still cannot say at this point. We hope that the novel tools offered by the present line of work, namely the neural and behavioral markers of the resetting process, which indicate the invalidation of the object-to-representation mapping in VWM, can help reveal some of the answers to these and following questions.

Conflict of interest

The authors declare no competing financial interests.

CRediT authorship contribution statement

Halely Balaban: Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **Trafton Drew:** Writing - review & editing. **Roy Luria:** Funding acquisition, Resources, Supervision, Writing - original draft, Writing - review & editing.

Open practices

The study in this article earned Open Materials and Open Data badges for transparent practices. Materials and data for the study are available at http://osf.io/ft462/.

Acknowledgments

This research was supported by a Binational Science Foundation (grant number 2015301) to R.L. and T.D., an Israel Science Foundation (grant number 862/17) to R.L., and an Azrieli Fellowship to H.B.

REFERENCES

Baddeley, A. D., & Logie, R. H. (1999). *Working memory: The multi-component model*.

Balaban, H., Drew, T., & Luria, R. (2018a). Delineating resetting and updating in visual working memory based on the object-to-representation correspondence. *Neuropsychologia, 113*, 85–94.

Balaban, H., Drew, T., & Luria, R. (2018b). Visual working memory can selectively reset a subset of its representations. *Psychonomic Bulletin & Review, 25*(5), 1877–1883.

Balaban, H., & Luria, R. (2015). The number of objects determines visual working memory capacity allocation for complex items. *NeuroImage, 119*, 54–62.

Balaban, H., & Luria, R. (2016a). Object representations in visual working memory change according to the task context. *Cortex, 81*, 1–13.

Balaban, H., & Luria, R. (2016b). Integration of distinct objects in visual working memory depends on strong objecthood cues even for different-dimension conjunctions. *Cerebral Cortex, 26*(5), 2093–2104. https://doi.org/10.1093/cercor/bhv038.

Balaban, H., & Luria, R. (2017). Neural and behavioral evidence for an online resetting process in visual working memory. *The Journal of Neuroscience, 37*(5), 1225–1239. https://doi.org/10.1523/JNEUROSCI.2789-16.2016.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321*(5890), 851–854. https://doi.org/10.1126/science.1158023.

Blaser, E., Pylyshyn, Z. W., & Holcombe, A. O. (2000). Tracking an object through feature space. *Nature, 408*(6809), 196–199.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision, 11*(5), 4. https://doi.org/10.1167/11.5.4.

Chen, Z. (2012). Object-based attention: A tutorial review. Attention, perception. *& Psychophysics, 74*(5), 784–802.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

Drew, T., Horowitz, T. S., Wolfe, J. M., & Vogel, E. K. (2011). Delineating the neural signatures of tracking spatial position and working memory during attentive tracking. *The Journal of Neuroscience, 31*(2), 659–668. https://doi.org/10.1523/JNEUROSCI.1339-10.2011.

Drew, T., Horowitz, T. S., Wolfe, J. M., & Vogel, E. K. (2012). Neural measures of dynamic changes in attentive tracking load. *Journal of Cognitive Neuroscience, 24*(2), 440–450. https://doi.org/10.1162/jocn_a_00107.

Drew, T., & Vogel, E. K. (2008). Neural measures of individual differences in selecting and tracking multiple moving objects. *The Journal of Neuroscience, 28*(16), 4183–4191. https://doi.org/10.1523/JNEUROSCI.0556-08.2008.

Emrich, S. M., Al-Aidroos, N., Pratt, J., & Ferber, S. (2009). Visual search elicits the electrophysiological marker of visual working memory. *PloS One, 4*(11), e8042.

Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision, 11*(12). https://doi.org/10.1167/11.12.3.

Gao, Z., Gao, Q., Tang, N., Shui, R., & Shen, M. (2016). Organization principles in visual working memory: Evidence from sequential stimulus display. *Cognition, 146*, 277–288.

Gao, Z., Yin, J., Xu, H., Shui, R., & Shen, M. (2011). Tracking object number or information load in visual working memory: Revisiting the cognitive implication of contralateral delay activity. *Biological Psychology, 87*(2), 296–302. https://doi.org/10.1016/j.biopsycho.2011.03.013.

Ikkai, A., McCollough, A. W., & Vogel, E. K. (2010). Contralateral delay activity provides a neural measure of the number of representations in visual working memory. *Journal of Neurophysiology, 103*(4), 1963–1968. https://doi.org/10.1152/jn.00978.2009.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24(2), 175–219.

Kang, M. S., & Woodman, G. F. (2014). The neurophysiological index of visual working memory maintenance is not due to load dependent eye movements. *Neuropsychologia, 56*, 63–72. https://doi.org/10.1016/j.neuropsychologia.2013.12.028.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience, 8*, 213. https://doi.org/10.3389/fnhum.2014.00213.

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences, 17*(8), 391–400. https://doi.org/10.1016/j.tics.2013.06.006.

Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience and Biobehavioral Reviews, 62*, 100–108.

Luria, R., Sessa, P., Gotler, A., Jolicoeur, P., & Dell'Acqua, R. (2010). Visual short-term memory capacity for simple and complex objects. *Journal of Cognitive Neuroscience, 22*(3), 496–512. https://doi.org/10.1162/jocn.2009.21214.

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia, 49*(6), 1632–1639. https://doi.org/10.1016/j.neuropsychologia.2010.11.031.

Luria, R., & Vogel, E. K. (2014). Come together, right now: Dynamic overwriting of an object's history through common fate. *Journal of Cognitive Neuroscience, 26*(8), 1819–1828. https://doi.org/10.1162/jocn_a_00584.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience, 17*(3), 347–356. https://doi.org/10.1038/nn.3655.

McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex, 43*(1), 77–94.

McConkie, G. W., & Currie, C. B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception and Performance, 22*(3), 563.

Pylyshyn, Z. W. (2000). Situating vision in the world. *Trends in Cognitive Sciences, 4*(5), 197–207.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*(3), 179–197.

Tsubomi, H., Fukuda, K., Watanabe, K., & Vogel, E. K. (2013). Neural limits to representing objects still within view. *The Journal of Neuroscience, 33*(19), 8257–8263. https://doi.org/10.1523/JNEUROSCI.5348-12.2013.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature, 428*(6984), 748–751. https://doi.org/10.1038/nature02447.

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438*(7067), 500–503. doi:nature04171 [pii].

Xu, Z., Adam, K., Fang, X., & Vogel, E. (2018). The reliability and stability of visual working memory capacity. *Behavior Research Methods, 50*(2), 576–588.

Ye, C., Zhang, L., Liu, T., Li, H., & Liu, Q. (2014). Visual working memory capacity for color is independent of representation resolution. *PloS One, 9*(3), e91681.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*(7192), 233–235. https://doi.org/10.1038/nature06860.