Come Together, Right Now: Dynamic Overwriting of an Object's History through Common Fate

Roy Luria[1] and **Edward K. Vogel**[2]

[1]Tel-Aviv University

[2]University of Oregon

Abstract

The objects around us constantly move and interact, and the perceptual system needs to monitor on-line these interactions and to update the object's status accordingly. Gestalt grouping principles, such as proximity and common fate, play a fundamental role in how we perceive and group these objects. Here, we investigated situations in which the initial object representation as a separate item was updated by a subsequent Gestalt grouping cue (i.e., proximity or common fate). We used a version of the color change detection paradigm, in which the objects started to move separately, then met and stayed stationary, or moved separately, met, and then continued to move together. We monitored the object representations on-line using the contralateral delay activity (CDA; an ERP component indicative of the number of maintained objects), during their movement, and after the objects disappeared and became working memory representations. The results demonstrated that the objects' representations (as indicated by the CDA amplitude) persisted as being separate, even after a Gestalt proximity cue (when the objects "met" and remained stationary on the same position). Only a strong common fate Gestalt cue (when the objects not just met but also moved together) was able to override the objects' initial separate status, creating an integrated representation. These results challenge the view that Gestalt principles cause reflexive grouping. Instead, the object initial representation plays an important role that can override even powerful grouping cues.

INTRODUCTION

Gestalt grouping principles, such as proximity and common fate, play a crucial role in how we interpret visual inputs and in how we perceive, group, and integrate visual objects. The fundamental role Gestalt cues play in object grouping was demonstrated by numerous studies (for recent reviews, see Wagemans, Elder, et al., 2012; Wagemans, Feldman, et al., 2012), indicating that our perceptual and cognitive systems use these Gestalt principles to piece together the "object chaos" around us and to provide a perception of a stable and continuous world.

Notably, the objects around us constantly change and interact with each other: They move, merge, and separate. Often, these changes produce conflicts between the object's initial

representation (the object "history") and the subsequent grouped status. For example, a man and a car constitute separate objects; however, a man driving a car might be interpreted either as two separate objects or as one integrated object. Yet, these processes that deal with the dynamic nature of object grouping, when the initial separate object representation is updated by recent grouping cues, are still poorly understood. The current study investigated under which circumstances dynamic changes caused by Gestalt grouping cues would override the initial separate object representation.

Because we were interested in how Gestalt principles update the object separate status, we first established an object "history"[1] by letting the objects move independently. In the critical condition, after the independent movement, we introduced a Gestalt grouping cue, such as proximity or common fate. Thus, the initial object separate representation was updated using a Gestalt grouping cue. This allowed us to investigate under which circumstances the objects will be integrated into one unit (following the more recent grouping cue) overriding the initial separate representation (the object "history"). For example, separate objects might move toward each other and then "meet" and proceed together, as in the example of a man entering a car and then driving it. Numerous previous studies (Kerzel, Born, & Schonhammer, 2012; Gallace & Spence, 2011; Woodman, Vecera, & Luck, 2003) have shown the power of Gestalt grouping, suggesting that Gestalt grouping occurs preattentively (Moore & Egeth, 1997; Duncan & Humphreys, 1989) and leading us to expect that perception would always follow salient Gestalt cues. However, the results from this study indicate that grouping according to Gestalt principles is not always reflexive and "automatic." Instead, the initial objects' representations can remain separate, overriding a Gestalt grouping cue.

To track the objects as they evolve, we monitored their representations in working memory (WM), an on-line limited capacity storage buffer that stores the active representations to protect them from various perceptual disruptions (e.g., eye saccades, blinks, and movements; see Hollingworth, Richard, & Luck, 2008). Previous research highlighted several important interactions between Gestalt grouping principles and visual WM performance (Hollingworth & Rasmussen, 2010; Flombaum & Scholl, 2006; Wheeler & Treisman, 2002). For example, a study that measured activity in the inferior intraparietal sulcus (IPS), an area sensitive to the number of represented objects in visual WM found that stationary objects that were grouped by Gestalt cues (as compared with ungrouped objects) elicit lower activation in the inferior IPS (Xu & Chun, 2007). This relative ease of representing grouped objects then allowed for more object information to be encoded and stored in the superior IPS (an area sensitive to the object complexity). The present research enabled us to characterize the exact interplay between Gestalt grouping cues, the objects' history, and the objects' storage buffer (WM), by tracking the objects' representations on-line, when the objects were visible and interacted with each other and after they disappeared and became WM representations.

---

[1]By the term "object history," we mean the object initial representation. The goal of current study was to investigate when separate objects' representations will be integrated following a Gestalt grouping cue. Thus, we only focused on situations in which the Gestalt cue updated an initial separate object representation. The opposite case, in which a composed object separates to its parts, is out of the scope of the current study.

In the experiments reported below, we used a variant of the color change detection paradigm. We manipulated the object information by moving the objects (colored squares) before the retention interval (Hollingworth & Rasmussen, 2010). This movement was task irrelevant, as participants were asked only to remember the colors. On some trials, the colors (two or four colors) moved in separate directions (i.e., two and four separate objects conditions; see Figure 1), giving the visual system a strong cue that they are indeed separate ("independent fate"). In the color–color conjunction condition, the colors moved together, one small colored square on top of a big colored square ("common fate"). This salient gestalt grouping cues of both common fate and proximity offers strong evidence to the visual system that the two colors are actually one integrated object. In another condition (i.e., four-to-two condition), the colors moved independently but then "met" as they landed one on top of each other (again, one small colored square on top of a big colored square), which provided a new proximity cue that conflicted with the initial independent fate cue. Namely, similar to the example of man entering a car, this condition updated the history of the object as a separate item. Note that, once the objects meet, this condition is visually identical to the color–color conjunction condition (providing identical retinal stimulation and retaining similar color information), so that by comparing these two conditions we can isolate the importance of the object's history (indicating separate objects) versus the object's last perceptual input (indicating grouped objects). Across experiments we increased the saliency of these grouping cues by manipulating the degree to which the meeting objects were perceived as discrete and just "happened to meet" toward the end of their trajectory or were two objects that transform into being one integrated object.

To monitor the on-line object information during both the movement phase and as a WM representation, we relied on a neural measure named the contralateral delay activity (CDA). The CDA amplitude reflects the number of objects that are encoded at any given moment and can be measured during both visual tracking and WM retention interval (Drew, Horowitz, Wolfe, & Vogel, 2011; Drew & Vogel, 2008; Vogel & Machizawa, 2004). Thus, by monitoring the CDA amplitude, we can infer if two objects were being integrated into a single representation or if they were treated as separate objects. Importantly, the CDA amplitude is sensitive to the number of maintained objects and is not influenced by the number of distinct features that compose the object (Luria & Vogel, 2011). Moreover, because of the time precision of the EEG signal, monitoring the CDA allowed us to measure the time course of the integration process both during the movement period as well as during the WM maintenance period.

The current setup allowed us to address another related question: To what extent object grouping depends on the individual WM capacity? Note that that this integration process relies on visual WM as its workspace. However, it is not clear whether the individual capacity plays any significant role in the ability to group objects. To investigate the interaction between object grouping and WM capacity, we analyze a measure of the grouping efficiency and correlate it with the individual WM capacity.

We first analyzed how objects are integrated in the color–color conjunction (common fate) condition during their movement and then during WM retention interval. Because Gestalt principles in general and common fate in particular provide a powerful perceptual cue that

the two colors are one integrated object, we hypothesized that the items (four colors arranged in two pairs) will be integrated and represented as integrated objects. Thus, we expected the CDA amplitude for the color–color conjunction condition to be smaller than the CDA amplitude in the four separate colors condition (although these conditions retain the same amount of color information) indicating an object benefit, because identical color information is represented using fewer objects. Moreover, if the colors were fully integrated into two objects, this should be evident in a similar CDA amplitude comparing the color–color conjunction condition and the two separate colors condition (indicating that WM is sensitive only to the number of represented objects). Assuming that common fate is a powerful grouping cue, this integration should happen already during the movement phase and continue throughout the retention interval.

We then analyzed whether interacting objects were grouped when their history as separate and independent objects was updated by a grouping cue presented toward the end of their trajectory. Namely, the colors in the four-to-two condition always started their motion as independent objects, but then met (Experiments 1 and 2) or even moved together (Experiment 3). By monitoring the CDA amplitude over the course of the trial, we examined how these representations evolved as new cues were introduced. Note that the four-to-two condition and the color–color conjunction condition always become perceptually identical at some point during the trial (after 1 sec in Experiments 1 and 2 and after 600 msec in Experiment 3). Consequently, once they become perceptually identical, any differences observed between these conditions would be because of the objects' history, as integrated or independent representations. Note that the overall trial length changed between the experiments. In Experiment 1, it was 2000 msec (1000 msec movement time, followed by 100 msec stationary time and then 900 msec retention interval). In Experiment 2, the stationary time was increased to 600 msec (to increase the saliency of the meeting), resulting in a total trial length of 2600 msec. In Experiment 3, there was no stationary time (to isolate the effect of the object moving together) resulting in a total trial length of 1900 msec.

METHODS

Participants

All participants gave informed consent following the procedures of a protocol approved by the Human Subjects Committee at the University of Oregon. All volunteers were members of the University of Oregon community and were paid $10 per hour for participation. Each experiment included 16 participants.

Stimuli and Procedure

Each trial started with the onset of a white fixation cross ($0.5° \times 0.5°$) on the center of a gray screen for 750 msec. Participants were instructed to hold fixation throughout the trial. Then two arrows ($1.9° \times 0.3°$), one above and one below fixation, were presented for 200 msec, leaving only the fixation cross visible for additional 400 msec. The arrows indicated which side of the screen is relevant for the upcoming trial, and participants were instructed to pay attention only to that side. Then, the color stimuli were presented on both sides of the fixation. Note that the two sides were always balanced in terms of the visual information

that was displayed in each hemifield and that the colors stayed on the same side throughout their trajectory.

In the two separate colors condition, two colored squares (on each side), a big and a small one ($1.2° \times 1.2°$ and $0.8° \times 0.8°$), were presented, one in the upper quadrant and the other in the lower quadrant. The colors moved for 1 sec either horizontally or vertically (randomly determined), with the restrictions that movement away from fixation was not allowed, so that colors presented on the left side of fixation could not move toward the left and that the entire movement trajectory never crossed the fixation. Then only the fixation point was presented for 900 msec (the retention interval), followed by the test array in which the colors were presented at their last spatial position, but sometimes (50% of trials) one of the colors was different relative to the movement phase. Participants had to indicate if the test array is the same or different (i.e., decide if one of the colors is different) relative to movement phase. In the four separate colors condition, four colors (on each side of fixation) were presented, two at the upper quadrant and two at the lower quadrant (each quadrant had one small and one big square). The color–color conjunction condition was identical to the two separate colors condition, but each item was composed of small color on top of a big color. The four-to-two condition started with the presentation of four separate colors, two in each quadrant (similar to the four separate colors conditions); however, in each quadrant, the colors moved toward each other and ended up one of top of each other (similar to the color–color conjunction condition). In Experiment 1, after the 1-sec movement, the colors stayed stationary for 100 msec and then disappeared, and the retention interval started. In Experiment 2, the colors stayed stationary for 600 msec. In Experiment 3, the stationary phase lasted for only 17 msec (one refresh rate). In addition, in the four-to-two condition of Experiment 3, the colors met after 600 msec (instead of 1 sec) and then moved together for 400 msec, mimicking the color–color conjunction condition. The rationale for using different trial lengths across experiments was that in Experiment 2 we wanted to increase the likelihood that the meeting between the objects in the four-to-two condition will not be judged as a "coincidence," so that we let the objects stay stationary for a longer time relative to Experiment 1 (but kept the movement and retention intervals identical so that we could directly compare the experiments). In Experiment 3, we wanted to isolate the effect of the common movement, so we eliminated the stationary interval. The exact condition was random at each trial. Participants performed 20 blocks, 60 trials each.

As stimuli, we used seven potential colors (red, green, blue, cyan, purple, yellow, and white). Colors were selected randomly, with no repetition, and independently for each side.

Measuring Visual WM Capacity

For each experiment, participants first completed a behavior-only visual WM task before starting the ERP experiment. The WM task consisted of a change detection task with arrays of four and eight colored squares with a 1-sec retention interval. We computed each individual's visual memory capacity with a standard formula. The formula is $K = S(H − F)$, where K is the memory capacity, S is the size of the array, H is the observed hit rate, and F is the false alarm rate. Participants were divided into high-capacity and low-capacity groups using a median split of their memory capacity estimates.

EEG Recording

ERPs were recorded in each experiment using our standard recording and analysis procedures, including rejection of trials contaminated by blinks or large (>1°) eye movements. We recorded from 22 standard electrode sites spanning the scalp, including International 10/20 sites F3, F4, C3, C4, P3, P4, O1, O2, PO3, PO4, P7, P8, as well as PO7 and PO8 (midway between O1/2 and P7/8). The horizontal EOG was recorded from electrodes placed 1 cm to the left and right of the external canthi to measure horizontal eye movement, and the vertical EOG was recorded from an electrode beneath the right eye referenced to the left mastoid to detect blinks and vertical eye movements. Trials containing ocular artifacts, movement artifacts, or amplifier saturation were excluded from the averaged ERP waveforms. Furthermore, participants who had more than 20% of trial rejections in any condition were excluded from the analysis. The EEG and EOG were amplified by an SA Instrumentation amplifier with a bandpass of 0.01–80 Hz (half-power cutoff, Butterworth filters) and were digitized at 250 Hz by a PC-compatible microcomputer.

ERP Analyses

The CDA was measured as the difference in mean amplitude between ipsilateral and contralateral waveforms recorded at posterior parietal, lateral occipital, and posterior temporal electrode sites (PO3, PO4, P7, P8, PO7, and PO8). We used 300–1000 msec following the onset of the stimuli for tracking CDA and 1300–2000 msec, 1600–2500 msec, and 1200–1900 msec in Experiments 1, 2, and 3, respectively, for the WM CDA. Following previous studies (since Vogel & Machizawa, 2004), we will present the results from the PO7/PO8 electrodes because that is where the CDA amplitude is most evident; however, the exact same pattern of results was observed over the P7/P8 and PO3/PO4 pairs of electrodes.

RESULTS

All three experiments had the same four conditions: two colors that moved separately, four colors that moved separately, two color–color conjunction objects, and four colors that moved toward each other and became two color–color conjunction objects (four-to-two condition; Figure 1). Participants were told to encode the colors and that the movement itself was task irrelevant (note that the movement was only used to imply objecthood). Across the three experiments, we modified the four-to-two condition to gradually increase the likelihood that the meeting objects will be perceived as one integrated object and to further contrast the objects' history with their final state. Specifically, in the first experiment, after meeting each other, the object stayed stationary one top of each other for 100 msec. In the second experiment, this duration was prolonged to 600 msec, and in the third experiment, the objects met after 600 msec and moved together (similar to the color–color conjunction condition) for 400 msec. Because the trial length differed between experiments, for statistical purposes, the time window used to calculate mean amplitude was 300–1000 msec following the onset of the stimuli for tracking CDA and 1300–2000 msec, 1600–2500 msec, and 1200–1900 msec in Experiments 1, 2, and 3, respectively, for the WM CDA.

Consistent Common Fate and Proximity Grouping Cues

Movement Period—In all three experiments, color–color conjunction (common fate) objects were represented as two items, although each of these two objects had two distinct colors. First, the tracking CDA amplitude in the color–color conjunction was lower than the amplitude in the four separate objects condition [$F(1, 15) = 9.95$, $p < .05$; $F(1, 15) = 6.86$, $p < .05$; $F(1, 15) = 5.37$, $p < .05$; for Experiments 1, 2 and 3, respectively; see Figures 2, 3, and 4], indicating (some) object benefit, because the same amount of color information elicited a lower tracking CDA when it was presented with fewer objects. Moreover, common fate caused the four colors to be represented as only two items (perfect grouping), as indicated by similar tracking CDA amplitudes in the color–color conjunction condition as compared with the two separate colors condition ($F = 0.07$, $p > .79$; $F = 0.002$, $p > .92$; $F = 1.77$, $p > .20$ for Experiments 1, 2, and 3, respectively).

Memory Period—Common fate in the color–color conjunction condition resulted in perfect grouping also during WM retention interval. The WM CDA amplitude in the color–color conjunction condition was similar to the CDA amplitude in the two separate colors condition ($F = 1.11$, $p > .30$; $F = 2.66$, $p > .12$; $F = 2.95$, $p > .10$; for Experiments 1, 2 and 3, respectively) and lower relative to the four separate colors condition [$F(1, 15) = 8.51$, $p < .05$; $F(1, 15) = 45.08$, $p < .00001$; for Experiments 1 and 3, respectively]. This effect missed a significance level in Experiment 2 [$F(1, 15) = 3.61$, $p < .08$).

Conflicting Cues: Overriding the Object's History

Movement Period—As depicted in Figures 2, 3, and 4 in Experiments 1 (shared proximity for 100 msec) and 2 (shared proximity for 600 msec), we found no evidence that the meeting objects overwrote their independent history through grouping. In Experiment 3, following the common motion during the final stages of the movement, we did find evidence that the colors were dynamically grouped and that these new integrated representations were updated in WM.

The statistical analysis supported this conclusion: When the items shared a proximal location for 100 msec (Experiment 1), the tracking CDA amplitude in four-to-two condition was higher than the two color–color conjunction condition [$F(1, 15) = 29.32$, $p < .0001$], and the same pattern was observed in Experiment 2 [$F(1, 15) = 28.91$, $p < .0001$; see Figures 2 and 3]. We did find evidence for grouping in Experiment 3 (in which the object met after 600 msec and moved together for 400 msec). We separately analyzed the initial part of the movement period, before the color met (300–600 msec), and the last 100 msec of the mutual movement (900–1000 msec relative to the movement onset). The tracking CDA in the four-to-two condition was higher in the initial part of the tracking period, relative to color–color conjunction condition [$F(1, 15) = 15.98$, $p < .005$]. However, the tracking CDA amplitude from the last part of the tracking period was not different from the color–color conjunction condition ($F = 0.32$, $p > .57$; Figure 4) and was lower than the four separate colors condition [$F(1, 15) = 7.26$, $p < .05$]. Namely, toward the final phase of the movement, tracking the four-to-two conditions showed a similar amplitude as tracking two objects. Thus, an initial cue of independent fate was not overwritten by a later proximity grouping cue (in Experiments 1 and 2). However, independent fate was dynamically overwritten by a new

common fate cue (Experiment 3), and the once separate objects were quickly integrated into grouped units.

Memory Period—Mirroring the movement results, Experiment 1 showed no evidence that the meeting objects (100 msec) were integrated into a single representation: The WM CDA amplitude for the four-to-two condition was similar to the four separate objects condition ($F = 0.78$, $p > .39$) and was higher relative to the color–color conjunction condition [$F(1, 15) = 16.15$, $p < .005$; Figure 2]. Thus, a short interval proximity cue was not sufficient to cause grouping in the four-to-two condition.

Experiment 2 replicated the results from Experiment 1 at the group level (but see the Individual Differences in Grouping Objects section), so that during WM retention interval the objects in the four-to-two condition were represented separately despite the proximity cue lasting for 600 msec (instead of 100 msec in Experiment 1). The four-to-two condition had similar amplitude relative to the four separate colors condition ($F = 0.16$, $p > .26$). There was a significant difference between the four-to-two and the two color–color conjunction conditions [$F(1, 15) = 5.11$, $p < .05$; Figure 3].

In Experiment 3, following their common movement, the updated integrated representations in the four-to-two condition persisted throughout WM retention interval, so that the WM CDA amplitude in the four-to-two condition had a lower CDA amplitude relative to the four separate objects condition [$F(1, 15) = 17.04$, $p < .005$; Figure 4] and was not significantly different from two separate objects condition ($F = 3.17$, $p > .09$).

The Time Course of On-line Object Grouping

To quantify whether the objects were grouped, we computed a difference wave between the color–color conjunction condition and the four-to-two condition (conditions that become perceptually identical during the movement phase of the trial). If the colors in the four-to-two condition were integrated in the same manner as in the color–color conjunction condition, there should be no difference between the CDA in these conditions. Thus, a difference that is close to zero indicates that the items in the four-to-two conditions were grouped. Conversely, if the colors in the four-to-two condition were not integrated, then the difference wave should be positive, reflecting higher amplitudes in the four-to-two condition (indicating that more WM resources were devoted to maintain a larger number of objects). Figure 5 shows the results of this analysis: Whereas in Experiments 1 and 2 there was no indication for on-line grouping (the amplitude remained positive), in Experiment 3 participants dynamically grouped the objects just before the onset of the retention interval (900 msec), as evident in a zero amplitude of this difference wave.

To track the time course of this on-line grouping effect, we analyzed the results from Experiment 3 using a sliding window of 50 msec (starting from 300 msec post stimulus presentation) and found the first window that remained statistically not different from zero throughout the rest of the trial (using $p < .005$). This procedure indicated that the colors in the four-to-two condition were grouped 200 msec after the actual meeting[2] [namely, the tracking CDA difference was not statistically significant from zero at 800 msec poststimulus presentation, $t(15) = -2.21$, $p = .042$, and remained statistically identical to zero until the end

of the trial, i.e., $t(15) = 0.043$, $p = .96$ at 850 msec poststimulus onset]. This estimate is roughly at the same range found for updating objects in a multiple-object tracking task (i.e., 280 msec; see Drew, Horowitz, Wolfe, & Vogel, 2012).

Individual Differences in Grouping Objects

At the group level, in Experiments 1 and 2 the initial cue of independent fate was not overwritten by a later proximity cue indicating grouping. Only in Experiment 3 was independent fate dynamically overwritten by a new common fate cue.

One of the objectives of the current research was to investigate how individual differences in WM capacity interact with the gestalt grouping principles. That is, how the ability to override an existing representation because of gestalt cues differs between individuals and how these differences are connected to WM capacity. To the best of our knowledge, this question was never examined on an individual level basis. To this end, we repeated the above analyses, this time separating high from low WM capacity individuals (using a median split). To quantify whether the objects were grouped, we used the difference wave between the color–color conjunction condition and the four-to-two condition (conditions that become perceptually identical during the movement part of the trial, see the above section describing the time course of object grouping).

The only experiment in which individual differences in WM capacity interacted with the object grouping was Experiment 2. During both the movement and the memory phase, there was a significant difference between the amplitudes of low and high WM individuals [$t(14) = 2.87$, $p < .05$ and $t(14) = 2.79$, $p < .05$, for the movement phase and the memory phase, respectively]. Low WM capacity individuals did not integrate the colors during both movement [$t(7) = 6.09$, $p < .0005$] and memory phases [$t(7) = 3.52$, $p < .01$] as indicated by a positive CDA difference wave, whereas high WM individuals showed a small difference during the movement phase [$t(7) = 2.11$, $p = .07$] and interestingly showed no CDA difference during the memory phase [i.e., the CDA was not significantly different from zero, $t(7) = 0.23$, $p = .82$], indicating that high WM individuals represented the colors in the four-to-two condition as integrated items. Thus, high capacity individuals used the proximity cue and grouped the objects overriding the independent fate cue, whereas the low WM capacity group relied only on the objects' history, disregarding the proximity cue. This trend was also evident in a correlation between WM capacity and the difference score in the CDA amplitude between the four-to-two and the color–color conjunction conditions during the memory phase ($r = .56$, $p < .05$; see Figure 6), demonstrating that individuals with a high WM capacity were more likely to group the meeting objects as integrated representations. The same correlation was not significant in Experiment 1 ($r = .14$, $p > .60$), demonstrating that when the objects stayed stationary for only 100 msec, both low and high WM individuals represented the objects in the four-to-two condition as discrete, or in Experiment 3 ($r = .21$, $p > .43$), pointing out that both low and high WM capacity individuals were equally sensitive to the common fate movement and integrated the colors.

---

[2]We chose a $p < .005$ because of the multiple comparisons we performed. Using a p value of .05 with the same procedure resulted in a grouping time course estimate of 250 msec (instead of 200 msec).

Accuracy Results

In addition to the CDA, we analyzed the accuracy data. Please note that although the color–color conjunction and the two separate colors conditions are equivalent in terms of the number of the presented objects, they differ in the number of possible colors that might change. There are four colors in the color–color conjunction condition (two in each object), but only two colors in the two separate colors condition (one color per object). Thus, low accuracy in the conjunction condition (and in the four-to-two condition) may simply reflect an error-prone comparison process that needs to monitor more possible options rather than a failure in WM maintenance stage (Awh, Barton, & Vogel, 2007). The one-way ANOVA included the same conditions as the CDA analysis.

Experiment 1: Items Stay Stationary for 100 msec—The ANOVA yielded a significant main effect [$F(3, 45) = 26.03$, $p < .0005$; Table 1]. Similar to the CDA data (both movement and memory), accuracy showed a set size effect, with better performance for two separate colors relative to four separate colors [$F(1, 15) = 33.45$, $p < .0005$]. Accuracy for color–color conjunction condition was better than the four separate colors condition [$F(1, 15) = 4.75$, $p < .05$], but worse than the two separate colors condition [$F(1, 15) = 52.50$, $p < .0005$], and accuracy in the four-to-two condition was the same as in the four separate colors condition [$F(1, 15) = 2.17$, $p > .16$].

Experiment 2: Items Stay Stationary for 600 msec—The ANOVA yielded a significant main effect [$F(3, 45) = 20.82$, $p < .0005$; Table 1]. Again, accuracy showed a set size effect, with better performance for two separate colors relative to four separate colors [$F(1, 15) = 30.39$, $p < .0005$]. Accuracy for color–color conjunction condition was similar to the four separate colors condition [$F(1, 15) = 1.20$, $p > .29$], and accuracy in the four-to-two condition was marginally significant from the four separate colors condition [$F(1, 15) = 4.11$, $p > .06$].

Experiment 3: Common Motion—The ANOVA yielded a significant main effect [$F(3, 45) = 23.16$, $p < .0005$; Table 1]. Again, accuracy showed a set size effect, with better performance for two separate colors relative to four separate colors [$F(1, 15) = 27.06$, $p < .0005$]. Accuracy for color–color conjunction condition was better than the four separate colors condition [$F(1, 15) = 16.26$, $p < .005$], but worse than the two separate colors condition [$F(1, 15) = 24.99$, $p < .0005$], and accuracy in the four-to-two condition was not statistically different from the four separate colors condition [$F(1, 15) = 2.84$, $p > .11$].

DISCUSSION

Many studies have reported evidence for object-based attention (Yi, et al., 2008; Scholl, 2001; Duncan, 1984), in the sense that objects may serve as the building blocks for visual attention. There is also ample evidence that WM representations are object based (Fukuda, Awh, & Vogel, 2010; Zhang & Luck, 2008). Given the dynamic nature of objects in the real world, the current research provides important insights on how our perceptual system copes with interacting objects during their movement and after they disappear and become WM representations. Remarkably, the results demonstrate that Gestalt cues do not cause grouping

in a reflexive and automatic manner. Rather, our perceptual system can override Gestalt cues when the objects' history has strong indications for maintaining discrete representations. In general, it seems that the grouping mechanism weighs the entire object history and not just the last perceptual input when deciding to group objects.

Common fate was the most effective grouping cue we examined that quickly overwrote the object's history by integrating independent items into grouped objects, and these representations persisted both during their movement and as WM representations. This quick integration is a unique aspect for moving objects: A previous study using the exact same stimuli, with the only difference that the color–color conjunction objects were stationary, found that the objects were grouped only toward the end of the WM retention interval (Luria & Vogel, 2011). In respect to current findings, the results by Luria and Vogel further indicate that the binding process, at least when cued by proximity, has ongoing aspects that require time to develop and are less potent than a common fate grouping cue. This point is interesting in light of previous work, suggesting that Gestalt grouping occurs preattentively (Moore & Egeth, 1997; Duncan & Humphreys, 1989).

Proximity was a less effective grouping cue, especially when the objects' history as separate representations had to be overwritten. Across the three experiments, a short proximity cue was ineffective, and the objects' representations in the four-to-two condition were kept as independent. Interestingly, the end position of the colors in the four-to-two condition was identical to the color–color conjunction condition in which the items were quickly integrated. Thus, the last perceptual input of an object is not sufficient to cause grouping. Conversely, when both proximity and common fate strongly indicated that the two colors transformed to being only one object, they were rapidly grouped, updating their initial separate representation.

The results also demonstrated that high WM capacity individuals were more adaptive at overriding the objects' history and reinterpreting objecthood based on proximity cues. In Experiment 2, both high and low WM capacity individuals represented the objects when they moved separately as independent, but after the objects met(and after the objects stayed stationary for 600 msec), high capacity individuals were more prone to reinterpret the objects' status as integrated units rather than separate colors. Thus, it seems that high WM capacity individuals are more flexible at overriding previous object cues and updating the objects' representations accordingly.

The present research also shed light on the on-line grouping process when the objects are visible and how this information is being transferred to WM maintenance stage. The current experiments demonstrated that, in most cases, the object status during the movement phase (when the objects were visible) continued as WM representations. This was the case whether the objects were independent or integrated. Only high WM individuals were able to override the initial output representations during the movement phase: In Experiment 2, high WM individuals represented the meeting objects (in the four-to-two condition) as separate during their movement and when they rested one on top of each other, but then updated this representation and integrated the colors during WM maintenance period.

These conclusions are based on interpreting the CDA amplitude as reflecting the number of objects that are currently maintained in visual WM rather than the number of features that compose each object. We argue that this latter interpretation is implausible given the current set of results. Note that in Experiments 1 and 2, the exact perceptual input (comparing the color–color conjunction object and the four-to-two condition) resulted in different CDA amplitudes, depending on the object history, an outcome that is challenging to account for under the assumption that the CDA reflects the number of features instead of the number of objects. Moreover, in Experiment 3, the CDA amplitude in the four-to-two condition started as being equivalent to the four separate objects condition, but following the common movement reduced to being identical to the two separate objects condition. Again, this pattern of result cannot be explained if the CDA only reflects the number of features that compose each object. Thus, we argue that (at least under the current set of conditions) the CDA reflects the number of integrated objects maintained in visual WM.

Finally, the current results help to further develop models for object grouping. Previous research has focused on the spatial position of an object and investigated how it is being used by the grouping mechanism (Mitroff & Alvarez, 2007; Scholl, 2007; Treisman, 1998; Kahneman, Treisman, & Gibbs, 1992). For example, van Dam and Hommel (2010) demonstrated that two separate objects (e.g., an apple and a banana) were integrated into a single object file as long as they occupied the same position (i.e., appeared one on top of the other). The results from the current study demonstrate that the grouping mechanism was sensitive to the entire objects' history and not only to the object's last spatial position. Specifically, the results from Experiment 1 showed that the grouping mechanism does not reflexively group objects even when they share the same spatial location. Second, even when comparing perceptually identical objects, the grouping mechanism does not "automatically" integrate the objects. Rather, grouping depends on the entire objects' history. Third, although the grouping factors present at initial encoding carry much influence, this history can be dynamically overwritten by the introduction of a common fate grouping cue.

Acknowledgments

References

Awh E, Barton B, Vogel EK. Visual working memory represents a fixed number of items regardless of complexity. Psychological Science. 2007; 18:622–628. [PubMed: 17614871]

Drew T, Horowitz TS, Wolfe JM, Vogel EK. Delineating the neural signatures of tracking spatial position and working memory during attentive tracking. Journal of Neuroscience. 2011; 31:659–668. [PubMed: 21228175]

Drew T, Horowitz TS, Wolfe JM, Vogel EK. Neural measures of dynamic changes in attentive tracking load. Journal of Cognitive Neuroscience. 2012; 24:440–450. [PubMed: 21812558]

Drew T, Vogel EK. Neural measures of individual differences in selecting and tracking multiple moving objects. Journal of Neuroscience. 2008; 28:4183–4191. [PubMed: 18417697]

Duncan J. Selective attention and the organization of visual information. Journal of Experimental Psychology: General. 1984; 113:501–517. [PubMed: 6240521]

Duncan J, Humphreys GW. Visual search and stimulus similarity. Psychological Review. 1989; 96:433–458. [PubMed: 2756067]

Flombaum JI, Scholl BJ. A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. Journal of Experimental Psychology: Human Perception and Performance. 2006; 32:840–853. [PubMed: 16846283]

Fukuda K, Awh E, Vogel EK. Discrete capacity limits in visual working memory. Current Opinion in Neurobiology. 2010; 20:177–182. [PubMed: 20362427]

Gallace A, Spence C. To what extent do Gestalt grouping principles influence tactile perception? Psychological Bulletin. 2011; 137:538–561. [PubMed: 21574680]

Hollingworth A, Rasmussen IP. Binding objects to locations: The relationship between object files and visual working memory. Journal of Experimental Psychology: Human Perception and Performance. 2010; 36:543–564. [PubMed: 20515188]

Hollingworth A, Richard AM, Luck SJ. Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and gaze correction. Journal of Experimental Psychology: General. 2008; 137:163–181. [PubMed: 18248135]

Kahneman D, Treisman A, Gibbs BJ. The reviewing of object files: Object-specific integration of information. Cognitive Psychology. 1992; 24:175–219. [PubMed: 1582172]

Kerzel D, Born S, Schonhammer J. Perceptual grouping allows for attention to cover noncontiguous locations and suppress capture from nearby locations. Journal of Experimental Psychology: Human Perception and Performance. 2012; 38:1362–1370. [PubMed: 22428670]

Luria R, Vogel EK. Shape and color conjunction stimuli are represented as bound objects in visual working memory. Neuropsychologia. 2011; 49:1632–1639. [PubMed: 21145333]

Mitroff SR, Alvarez GA. Space and time, not surface features, guide object persistence. Psychonomic Bulletin & Review. 2007; 14:1199–1204. [PubMed: 18229497]

Moore CM, Egeth H. Perception without attention: Evidence of grouping under conditions of inattention. Journal of Experimental Psychology: Human Perception and Performance. 1997; 23:339–352. [PubMed: 9103998]

Scholl BJ. Objects and attention: The state of the art. Cognition. 2001; 80:1–46. [PubMed: 11245838]

Scholl BJ. Object persistence in philosophy and psychology. Mind & Language. 2007; 22:563–591.

Treisman A. Feature binding, attention and object perception. Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences. 1998; 353:1295–1306.

van Dam WO, Hommel B. How object-specific are object files? Evidence for integration by location. Journal of Experimental Psychology: Human Perception and Performance. 2010; 36:1184–1192. [PubMed: 20718565]

Vogel EK, Machizawa MG. Neural activity predicts individual differences in visual working memory capacity. Nature. 2004; 428:748–751. [PubMed: 15085132]

Wagemans J, Elder JH, Kubovy M, Palmer SE, Peterson MA, Singh M, et al. A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. Psychological Bulletin. 2012; 138:1172–1217. [PubMed: 22845751]

Wagemans J, Feldman J, Gepshtein S, Kimchi R, Pomerantz JR, van der Helm PA, et al. A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. Psychological Bulletin. 2012; 138:1218–1252. [PubMed: 22845750]

Wheeler ME, Treisman AM. Binding in short-term visual memory. Journal of Experimental Psychology: General. 2002; 131:48–64. [PubMed: 11900102]

Woodman GF, Vecera SP, Luck SJ. Perceptual organization influences visual working memory. Psychonomic Bulletin & Review. 2003; 10:80–87. [PubMed: 12747493]

Xu Y, Chun MM. Visual grouping in human parietal cortex. Proceedings of the National Academy of Sciences, USA. 2007; 104:18766–18771.

Yi DJ, Turk-Browne NB, Flombaum JI, Kim MS, Scholl BJ, Chun MM. Spatiotemporal object continuity in human ventral visual cortex. Proceedings of the National Academy of Sciences, USA. 2008; 105:8840–8845.

Zhang W, Luck SJ. Discrete fixed-resolution representations in visual working memory. Nature. 2008; 453:233–235. [PubMed: 18385672]
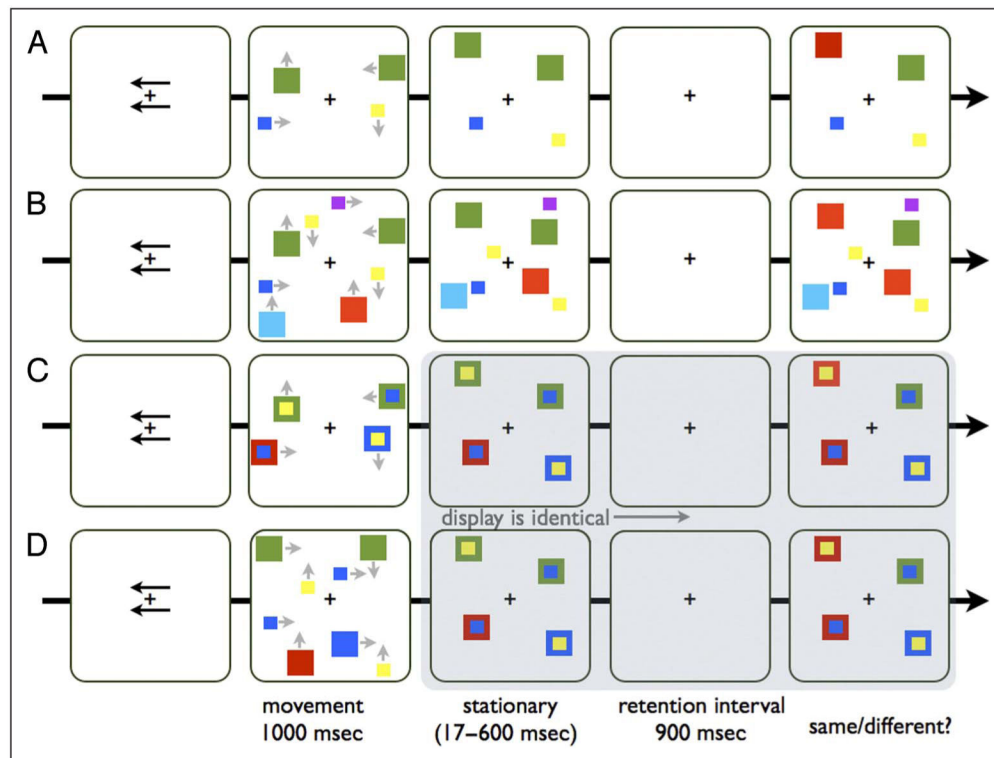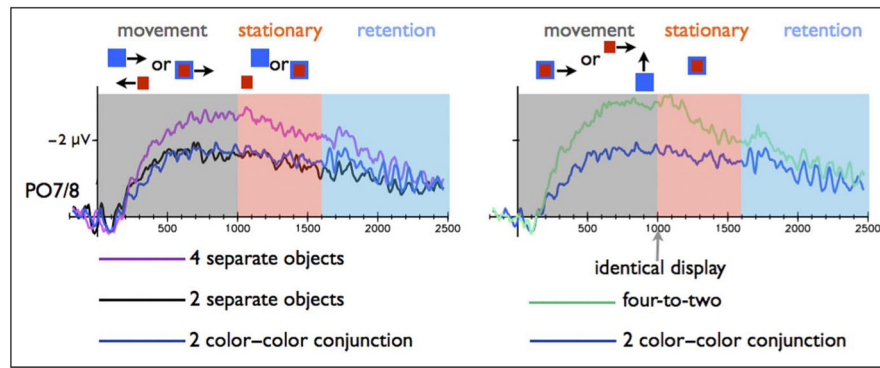
Figure 1.
A schematic example of a trial sequence in Experiment 1. Each trial started with the presentation of two arrows, one above and one below fixation, indicating the relevant side for the upcoming trial. Then the colors moved for 1 sec (the gray arrows that appear next to the colors in the movement phase only indicate their trajectory and did not appear in the actual experiment), followed by a stationary phase (100 msec in Experiment 1, 600 msec in Experiment 2, and 17 msec in Experiment 3) in which all the colors stayed in one position without moving, followed by the retention interval (900 msec), and then the test array was presented. (A) The two separate objects condition: two colors that moved separately. (B) The four separate objects condition: four colors that moved separately. (C) The color–color conjunction condition: two pairs of colors moved together, each pair is composed of one small colored square on top a big colored square. (D) The four-to-two condition: two pairs of colors that moved toward each other and ended up meeting one on top of each other. Note that the color–color conjunction condition and the four-to-two condition become perceptually identical during the movement phase.
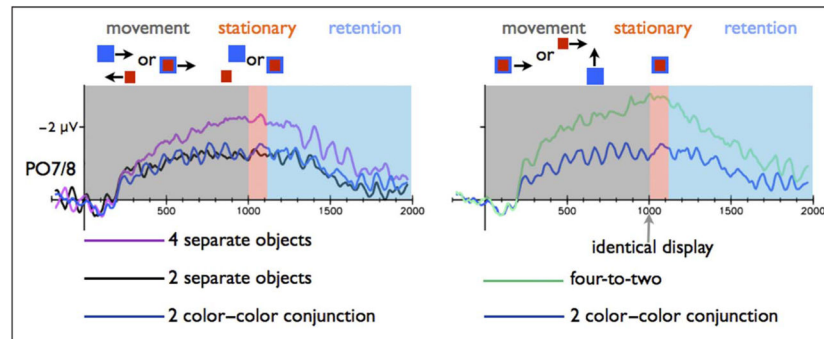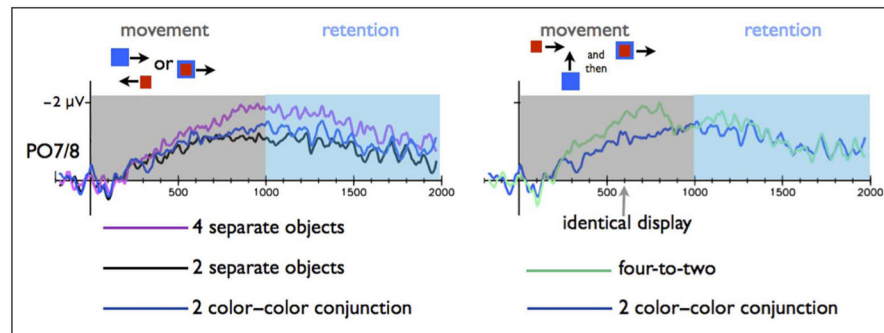
Figure 2.
CDA amplitude for the PO7/PO8 electrodes in Experiment 1 during the movement and retention intervals. Left: Results for the two separate objects and four separate objects conditions. Right: Results for the color–color conjunction stimuli and the four-to-two conditions. The CDA amplitude in the two separate objects condition was significantly lower than the four separate condition, but similar to the color–color conjunction condition (common fate), during both the movement and the memory phases. The CDA amplitude in the color–color conjunction condition was lower than the four-to-two condition during both the movement and the memory phases.
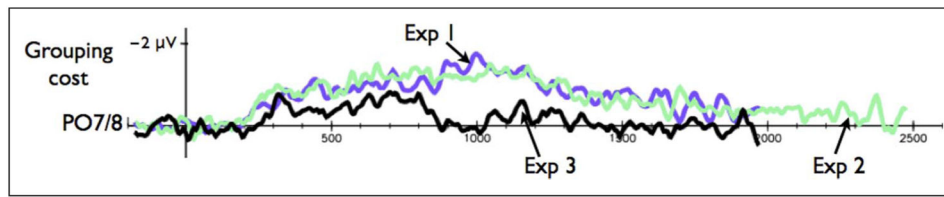
Figure 3.
CDA amplitude for the PO7/PO8 electrodes in Experiment 2 during the movement and retention intervals. Left: Results for the two separate objects and four separate objects conditions. Right: Results for the color–color conjunction stimuli and the four-to-two conditions. The CDA amplitude in the two separate objects condition was significantly different than the four separate colors condition, but similar to the color–color conjunction condition (common fate), during both the movement and the memory phases. The CDA amplitude in the color–color conjunction condition was lower than the four-to-two condition during both the movement and the memory phases.
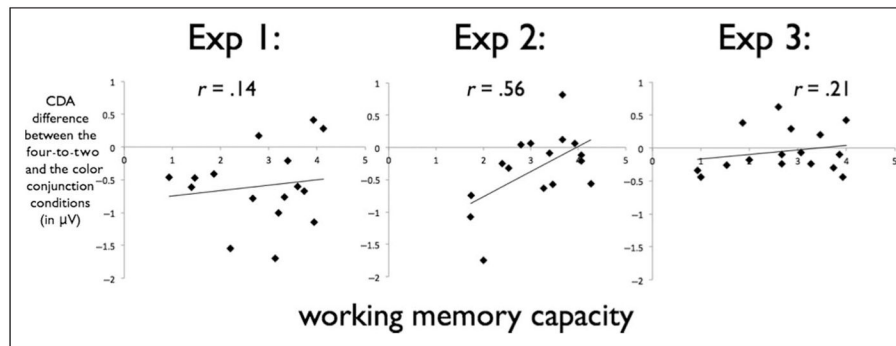
Figure 4.

CDA amplitude for the PO7/PO8 electrodes in Experiment 3 during the movement and retention intervals. Left: Results for the two separate objects and four separate objects conditions. Right: Results for the color–color conjunction stimuli and the four-to-two conditions. The CDA amplitude in the two separate objects condition was significantly different than the four separate colors condition, but similar to the color–color conjunction condition (common fate), during both the movement and the memory phases. The CDA amplitude in the four-to-two condition, started as larger relative to the color–color conjunction condition, but following the common movement dropped and became identical to the color–color conjunction condition.

Figure 5.
CDA difference wave for the PO7/PO8 electrodes between the four-to-two condition and color–color conjunction condition in Experiment 1 (purple), Experiment 2 (green), and Experiment 3 (black).

Figure 6.
Correlations between the individual WM capacity and the grouping CDA index (CDA in the four-to-two condition minus the CDA in the color–color conjunction condition) across the three experiments. Only the correlation in the Experiment 2 was significant ($p < .05$), whereas the correlations in Experiment 1 ($p > .60$) and in Experiment 3 ($p > .43$) were not.

Table 1

Accuracy Level and *SEM* across the Three Experiments for All Conditions

	2 Separate Colors	4 Separate Colors	Color–Color Conjunction	Four-to-Two
Experiment 1	.96 (.005)	.86 (.02)	.88 (.01)	.87 (.01)
Experiment 2	.96 (.007)	.87 (.02)	.89 (.01)	.90 (.01)
Experiment 3	.96 (.008)	.85 (.02)	.89 (.02)	.87 (.02)