



Nonlinear effects of spatial connectedness implicate hierarchically structured representations in visual working memory

Błażej Skrzypulec*, Adam Chuderski

Institute of Philosophy, Jagiellonian University in Krakow, Poland



ARTICLE INFO

Keywords:

Visual working memory
Visual features
Spatial relations
Change detection

ABSTRACT

Five experiments investigated the role of spatial connectedness between a pair of objects presented in the change detection task for the actual capacity of visual working memory (VWM) in healthy young adults (total $N = 405$). Three experiments yielded a surprising nonlinear relationship between the proportion of pair-wise connected objects and capacity, with the highest capacity observed for homogenous displays, when either all objects were connected or disjointed. A drop in capacity, ranging from an average of a quarter of an object out of three objects maintained in VWM, was noted when only some objects were connected while others were disjointed. As indicated by another two experiments, this effect was specific to double-feature encoding, but disappeared when single visual features had to be memorized. No existing theoretical model of VWM can directly explain this novel effect. Overall, the nonlinear effect of spatial connectedness implies that representations in VWM possess hierarchical structure defined by wholes, parts, features, and their relations, and the heterogeneity of such a structure hinders VWM performance, while homogeneity facilitates it.

Introduction

Effective visual perception of the environment is a fundamental cognitive faculty, providing input for numerous other mental functions such as categorization, imagery, learning, memory, decision making, and action. Although several theories of visual perception assume direct access to the environment (Gibson, 2015), the dominant view (Fodor & Pylyshyn, 2002; Rensink, 2000) holds that before visual information can be processed further, it must first be encoded for a short period of time via the sequence of memory buffers (including iconic, fragile, and, finally, robust memory; Sligte, Scholte & Lamme, 2008, 2009). The final (robust) buffer in this sequence, Visual Working Memory (VWM), most likely stores the most stable and most structured forms of visual information, as compared to the iconic and fragile buffer. The key property of VWM is its highly limited capacity, commonly comprising just a few objects or object features (Cowan, 2000), as shown by the hallmark studies investigating one key aspect of perception – detectability of changes in the environment (Rensink, O'Regan, & Clark, 1997).

A crucial finding along this vein, suggesting severe constraints on VWM capacity, is change blindness (Simons & Levin, 1997), that is, a severe difficulty in detecting the change to/of one given object in a scene, when the scene contains more than three to five objects and a

person does not know in advance which object is going to change. In a psychological laboratory, the standard experimental paradigm used to study the limits of VWM, called change detection task (Phillips, 1974), consists of successively presenting two patterns of objects (separated by a mask) such that either the second pattern is identical to the first or both patterns differ with respect to just a single object. The task is to report whether any difference between the patterns have occurred or not. Early studies (Luck & Vogel, 1997) suggested that the sheer number of objects presented does determine the change detection performance, with errors occurring only for displays with more than four objects. Later studies have complicated our understanding of the mechanisms and representations underlying VWM capacity, more comprehensively investigating the various characteristics of to-be-detected objects beyond their sheer number. That included the objects' features (Bays, Wu, & Husain, 2011), spatial relations (Woodman, Vecera, & Luck, 2003), and context (Clevenger & Hummel, 2013), as well as the reproduction of certain information (Wilken & Ma, 2004). These studies' findings resulted in several competing models of VWM functioning (for reviews, see Brady, Konkle, & Alvarez, 2011; Luck & Vogel, 2013).

The leading models can be categorized by referring to the role of objects' features in the change detection performance. First, according to some models VWM capacity is related only to the number of objects

* Corresponding author at: Institute of Philosophy, Jagiellonian University in Krakow, Grodzka 52, 31-044 Krakow, Poland.
E-mail address: blazej.skrzypulec@uj.edu.pl (B. Skrzypulec).

(Awh, Barton, & Vogel, 2007; Luria & Vogel, 2011), whereas other models assume that it is affected by both the number of objects and the number of features (Cowan, Blume, & Saults, 2012). Second, among the models postulating a significant role of the number of features, some emphasize the absolute number of features (Fougnie & Alvarez, 2011), whereas others claim that the number of features per object has a greater relevance (Oberauer & Eichenberger, 2013). Another question is whether the feature-related errors arise due to either the misrepresentation of some of the features or the misrepresentation of the relations between features and their objects (Bays, Wu, & Husain, 2011). The VWM models can be further subdivided according to the role of feature dimensions. Some adopt a single general limit for the number of stored features (Hardman & Cowan, 2015), while others accept separate limits for each feature dimension such as color and orientation (Olson & Jiang, 2002). Finally, both object- and feature-based capacity limits can be defined in terms of either discrete slots (Anderson & Awh, 2012; Anderson, Vogel, & Awh, 2011) or continuous resource (Bays, 2015; Fukuda, Awh, & Vogel, 2010). The slot models predict that performance drops sharply after exceeding the number of elements that occupy all the available slots. On the other hand, the resource models entail that performance decreases gradually with the rising number of objects and features, which means a lesser portion of the available resource is allocated to a single element in VWM.

However, many authors believe that models focusing solely on the role of objects and features cannot comprehensively characterize how information is stored in VWM, because such models neglect a potentially crucial role of relational and contextual factors (see Clewenger & Hummel, 2013; Johnson, Spencer, Luck, & Schoner, 2013). Specifically, it has been demonstrated that the change-detection task performance worsens when successively presented patterns differ in the spatial configuration of their objects, even if the number of objects and features remains unchanged (see Brady et al., 2011). Furthermore, such a performance was shown to improve when the presented objects become organized according to Gestalt principles, such as item-item similarity (e.g., Lin & Luck, 2009; Peterson & Berryhill, 2013), item-whole similarity (Kałamała, Sadowska, Ordziniak, & Chuderski, 2017), and layout closure (Li, Qian, & Liang, 2018); better performance is also seen after increasing the spatial proximity between objects (Woodman et al., 2003). In particular, change detection becomes easier when presented objects are organized into pairs composed of proximal or even spatially connected elements, as compared to when the objects are distributed randomly (Xu, 2002, 2006). The benefit from pair-wise presentation is especially profound when multiple visual features (e.g., both color and orientation) need to be tracked.

Spatial connectedness seems to be particularly interesting theoretically as it may decrease the number of represented objects but simultaneously increase each object's complexity. Specifically, spatially connecting two objects provides a strong cue that these objects constitute a single whole, which can reduce the number of units to be stored in VWM, making the task easier. However, a whole constituted by the connected objects will have a more complex shape and more features than a single object, so storing more complex objects in VWM can make the task more demanding. Thus, a systematic investigation of the potential VWM performance benefits resulting from spatially connecting the visual objects, both in the single and multiple feature tracking, can potentially extend our understanding of the neurocognitive mechanisms underlying momentary encoding and the storage of visual information in human memory. To date, however, all of the experiments that have investigated the relation between visual complexity and spatial connectedness compared the performance of spatially connected pairs for objects that were either all disjointed or all composed (e.g., 10 disjointed vs. 5 two-part objects in Xu, 2006).

The primary goal of the present studies was to improve our understanding of the role of spatial connectedness in VWM by investigating the effects of a gradual increase in the number of connected pairs, starting from all disjointed objects through to cases in which some

objects comprised two-element wholes, up to cases in which all objects were pair-wise connected. Moreover, the role of the amount of spatial connectedness was examined for both single and multiple feature tracking, in the search for any possible interaction. Experiments 1 and 2 identified a surprising data pattern, with performance benefiting from connecting all the objects into pairs in the same way as disjointing all the objects. Performance was hindered only when some objects were paired, while others were kept unconnected. Furthermore, Experiments 3–5¹ revealed that the considered nonlinear relationship is likely to occur only when more than one feature has to be tracked. Overall, these findings support the view that actual VWM capacity stems from hierarchically structured representations of wholes, parts, and their features (see Brady et al., 2011; Clewenger & Hummel, 2013), and suggests that fully homogeneous representations (of whatever form: either singletons or complexes, but not both at the same time) yield the best performance. At present, no existing VWM model can directly explain this novel effect.

Experiment 1

Participants

A total of 61 healthy people with normal colour vision (20 men and 41 women) aged 18–40 were recruited by internet advertisements. They were gratified with an equivalent of 15 euros in local currency. They were informed in a general manner that the study investigated memory and that their data would be anonymous. Participation could be ended at will at any moment. All other aspects of the study were carried out in accordance with the ethical principles of the 1964 Declaration of Helsinki (World Medical Association, 2001). Two participants were excluded from the analyses for random-level performance (defined thereof as the change discrimination index falling below .05). Because a gradual manipulation of the number of pairs has not previously been applied, it was difficult to predict the respective effect size a priori. Consequently, a respective sample size was adopted that allowed researchers to detect the $f^2 = .20$ effects for repeated measurements with the power of .95, and the $f^2 = .15$ effects with the power of .80 (G*Power; Faul, Erdfelder, Buchner, & Lang, 2009).

Materials and design

The stimuli were either colored or black circles with a white Greek letter placed inside. Each stimulus had an approx. 0.6 degree visual angle in diameter. Eight easily discriminable colors (blue, green, red, yellow, fuchsia, orange, brown, and cyan) and eight distinctive letters were used (α , Ω , Δ , Π , Ψ , Φ , ω , Θ).² Each stimulus could be presented either as separated spatially from all the remaining stimuli, with the minimal visual distance of 1.24° between the stimulus centers, or as spatially connected to another stimulus. The connected pairs were oriented either vertically or horizontally. All stimuli were randomly positioned within a grey rectangle placed at the center of the screen, sized 5.3° × 3.4° (8.8% × 10.2% of the screen). The number of connected pairs, manipulated randomly, was the first independent variable (*pairs*), having four levels: there could be no pairs (6 spatially separated stimuli), 1 pair (2 connected and 4 separated stimuli), 2 pairs (4 pair-wise connected and 2 separated stimuli), or 3 pairs (6 pair-wise connected stimuli). In each trial, the positions of stimuli were randomly selected from a constant set of eight locations; the same set of locations was used

¹ Another two unreported experiments applied a larger memory load (8 instead of 6 objects), but such a big load led to a close-to-floor performance (mean discrimination index of 0.249 and 0.272) and so to inconclusive results.

² Participants' knowledge of Greek has not been tested but such knowledge is very rare within the local population from which the participants were recruited.

in all of the experiments. In the 0-pair condition, six locations out of eight were selected. In the 1-pair condition, five locations out of eight were selected and an additional sixth stimulus was added to create a single, connected pair. The additional stimulus was randomly positioned to the left or right, or below or above one of the five selected locations. Analogously, in the 2-pair condition as many as four locations were selected and two stimuli were added to create two pairs, and in the 3-pair condition three locations out of eight were selected and three stimuli were added.

There were three types of trials regarding the stimulus feature (the circle color or the letter identity) that was relevant in a given trial. In the *color* trials, six circles were displayed each having a unique color. Analogously, in the *letter* trials, six black circles, each containing a different white letter were presented. Finally, in each double-feature trial, both features varied: there were three uniquely colored circles, as well as three unique letters, the latter displayed inside three black circles. The number of relevant features in a trial (either single or double) defined the second independent variable (*number of features*). An additional comparison between the color and letter single-feature trials facilitated the verification of the difficulty of the study materials.

Procedure

The task was run on a standard PC workstation with a 23-inch LCD monitor, which had a resolution of 1920 × 1080 pixels. The experiment was administered in a psychological laboratory, in groups ranging from three to seven people, with each person occupying a designated cubicle. Participants were seated comfortably about 50 cm from the computer screen. The session was started by the screen presenting the instructions; subsequent steps were initiated by pressing the designated key. Next, a training session was composed of 8 trials. After this training, participants initiated when ready the main part of the experiment, which consisted of 10 blocks, each composed of 32 trials (320 trials in total). After the end of each block, a break could be taken, if needed. The next block was initiated by pressing the key.

Each trial started with the presentation of stimuli pattern for 3.5 s. Next, a colorful mask was displayed for 1 s. The mask comprised a 16 × 16 matrix of colorful squares that covered the square-shaped fragment of the screen in which stimuli could be located. Colored squares were randomly generated each time the mask was shown. This same masking was used in all the experiments. Subsequently, another stimuli pattern was presented which could be either identical to the first pattern or different from it, either (a) in the color of one colored circle (in the cases of the color trials and half of the double-feature trials) or (b) in the letter presented inside one of the black circles (in the case of the letter trials and half of the double-feature trials). Participants were allowed 4 s to decide whether the two patterns were the same or different, by pressing the left or right arrow key (assigned randomly for each participant). The possible stimuli patterns and the course of the trial are illustrated in Fig. 1A.

Each of the ten blocks of 32 trials contained: 8 color single-feature trials (2 for each possible number of pairs), 8 letter single-feature trials (2 per number of pairs), and 16 double-feature trials (4 per number of pairs). In the change trials, exactly one feature of a single stimulus changed. These changes were distributed in such a way that in the color and letter trials there was exactly one trial with the change for each number of pairs, whereas in the double-feature trials there were two trials with the change for each number of pairs, one concerning the change in color and the other concerning the change in letter. The remaining trials involved no change. The order of trials in each block was random.

The dependent variable reflecting accuracy in each condition was the proportion of correctly detected changes in the change trials minus the proportion of incorrectly signaled changes in the no-change trials (discrimination index; Snodgrass & Corwin, 1988). Such a measure allowed correcting for random guessing. Its unity indicated perfect

change detection performance and zero reflected random performance. Moreover, the actual value multiplied by 6 (the number of stimuli in the task) could be interpreted as the stimuli load that was effectively maintained in the robust VWM buffer (*k* value; Cowan, 2000; Rouder, Morey, Morey, & Cowan, 2011).

Analogously to studies on the beneficiary effects of grouping, it was expected that an increasing number of connections would yield increased performance, at least when two features were tracked (e.g., that three two-part complexes would load VWM to a lesser extent than six singletons). Obviously, it was also predicted that tracking two features would result in lower accuracy than tracking one feature.

Results

The detection of a change in colors ($M = .458$, 95% confidence interval [.408, .508]) significantly surpassed detection for letters ($M = .229$, [.186, .272]), $t(60) = 9.90$, $p < .001$, Cohen's $d = 1.24$. This data translated into the *k* value indicating that the participants maintained an average of only 2.05 objects in VWM, which was quite a low score.

The 4 (pairs) × 2 (features) repeated measured analysis of variance (rmANOVA) was run on data presented in Fig. 2A. Both factors yielded significant effects, $F(1, 60) = 30.98$, $p < .001$, $MSE = .028$, $\eta_p^2 = .34$, and $F(3, 180) = 8.01$, $p < .001$, $MSE = .016$, $\eta_p^2 = .12$, for the number of features and the number of pairs, respectively. Surprisingly, performance was better when two features were varied ($M = .428$, [.383, .473]), as compared to a single feature ($M = .343$, [.302, .384]). No significant interaction of factors was observed, $F < 1$.

A nonlinear relationship between the number of pairs and performance (for both single and double features) was suggested by Fig. 2A. Indeed, the two extreme levels of the pairs factor (mean value on the 0 and 3 pairs) yielded a larger discrimination index than the two middle levels of this factor (mean value on the 1 and 2 pairs) was shown by a statistically significant contrast between the extreme and middle levels (a linear combination of levels with the 1, -1, -1, and 1 weights, for the 0, 1, 2, and 3 pairs, respectively), $F(1, 60) = 18.45$, $p < .001$. The contrast was significant for both single and double features, $p = .001$ and $p = .005$, respectively, with no interaction between the contrasts related to the number of features, $F < 1$.

Finally, for the 1- and 2-pair conditions, whether the change pertained either to an object connected to another object ($M = .553$, [.523, .583]) or to an unconnected object ($M = .531$, [.499, .563]) did not significantly affect accuracy during the change trials, $t(60) = 1.64$, $p = .105$ (the 0- and 3-pair conditions were excluded from this test as they contained only either unconnected or connected objects, respectively).

Discussion

The present experiment showed a nonlinear relationship between the number of pairs and performance, for both single and double feature encoding. Errors were more frequent when only some of the objects composed pairs (1 or 2 pairs), as compared to when all objects were either paired or unpaired. However, relatively low letter-tracking accuracy resulted in a worse performance in the single-feature trials than in the double-feature trials, when the participants could compensate for their performance by an increased focus on colors. Because the Greek letters are highly specific stimuli and are more difficult to encode than colors, in order to boost change detection performance the next experiment used familiar shapes instead of the Greek letters. Moreover, the unobserved nonlinear effect of pairs definitely required replication.

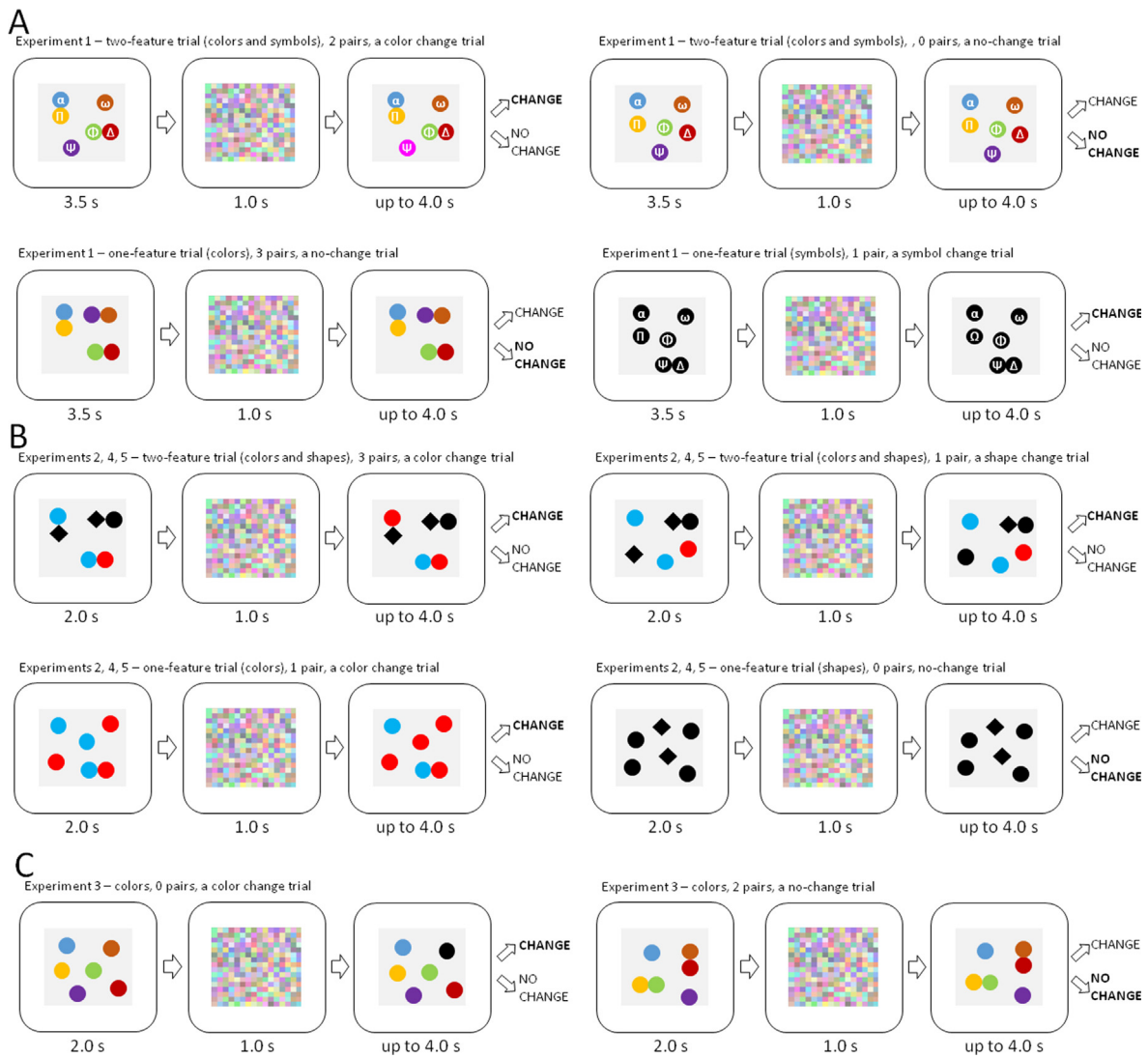


Fig. 1. A schematic illustration of stimuli and the course of events in the change and no-change trials of the change detection task. Panel A: The task variant with colors and Greek letters applied in Experiment 1. Panel B: The task variant with colors and shapes applied in Experiments 2, 4, 5. Panel C: The task variant with colors applied in Experiment 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Experiment 2

Participants

A total of 66 people with normal color vision (21 men and 45 women) aged 18–40 were recruited by internet advertisements, and subjected to an analogous procedure as in Experiment 1. One additional participant was excluded due to random performance.

Materials and design

The stimuli were circles (red, blue, and black) and squares (black); each stimulus was approx. 0.6° in diameter. Again, stimuli were presented either as separated spatially, with a minimum distance of 1.24° between their centers, or as spatially connected. Connected stimuli could form either vertically or horizontally oriented pairs. As in Experiment 1, each stimulus was randomly positioned on a 5.3° × 3.4° gray centered rectangle. Again, the number of connected pairs could be 0, 1, 2, or 3 (i.e., 0, 2, 4, or 6 pair-wise connected stimuli).

In the *color* trials, six red and blue circles were displayed. The colors were distributed randomly using the following three patterns: (a) three circles were red and three were blue, (b) two were red and four were

blue, or (c) two were blue and four were red. In the *shape* trials, black circles and black squares were presented. The shapes were distributed in three possible ways, analogously to the colors: (a) three squares and three circles, (b) two squares and four circles, or (c) two circles and four squares. In the *double-feature* trials, three red or blue circles and three black circles or squares were displayed. The proportion of colors in the colored circles as well as the proportion of squares and circles in the black figures was always 2–1. The possible stimuli patterns and the course of the trials are illustrated in Fig. 1B.

Procedure

The same training and number of experimental trials was applied as in Experiment 1. Because the stimuli patterns were now simpler, the presentation of stimuli patterns was decreased to 2 s. The second stimuli pattern could be identical to the first pattern or could differ either (a) in one colored circle (in the case of the color single-feature trials and the double-feature trials) or (b) in one black figure shape (in the case of the shape single-feature trials and the double-feature trials). Thus, the structure of the trials was identical to Experiment 1, except for the fact that shapes replaced letters. Again, 4 s were allowed to decide whether the two patterns were the same or different by the relevant key press.

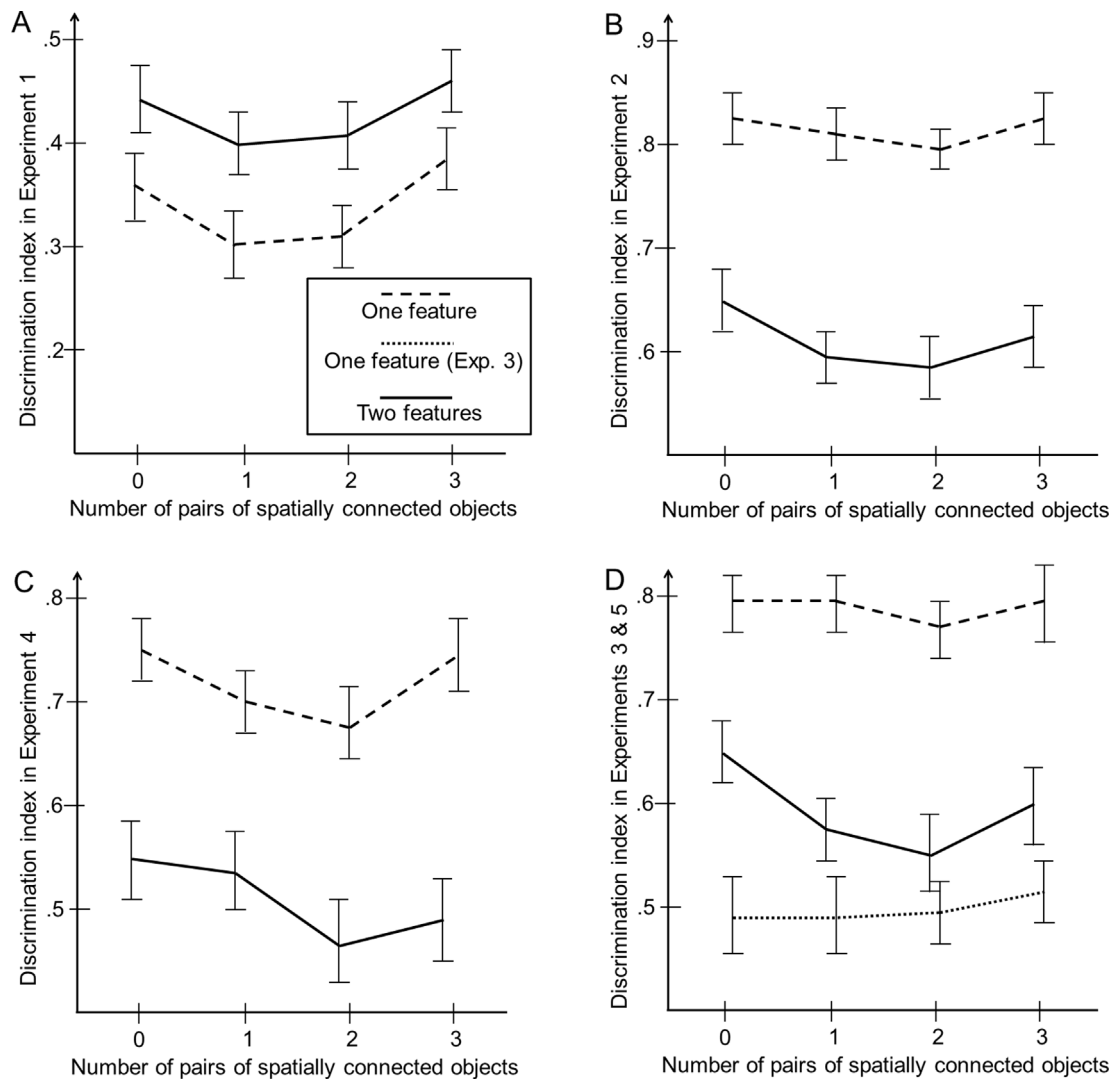


Fig. 2. Discrimination index (hit rate minus false alarm rate) in the change detection task across Experiments 1–5 as a function of the number of features to be tracked (either one or two; note that only one feature was tracked in Experiment 3) and the number of pairs of spatially connected objects (0, 1, 2, or 3). Vertical bars indicate 95% confidence intervals.

The same type of PC set was used.

Results

Discrimination for colors ($M = .857$, [.824, .890]) was significantly better than discrimination for shapes ($M = .770$, [.722, .820]), $t(65) = 6.87$, $p < .001$, Cohen's $d = 0.51$. They both surpassed performance levels in Experiment 1, this time yielding $k = 4.88$ objects on average.

Again, the 4 (pairs) \times 2 (features) rmANOVA was run on data presented in Fig. 2B. This time, the responses when two features were varied were less correct ($M = .607$, [.558, .655]), as compared to single features ($M = .814$, [.774, .853]), $F(1, 65) = 221.36$, $p < .001$, $MSE = .026$, $\eta_p^2 = .77$. The number of pairs factor again yielded a significant effect, $F(3, 195) = 4.21$, $p = .007$, $MSE = .011$, $\eta_p^2 = .06$, implying a significant contrast between the two extreme (data for the 0 and 3 pairs averaged) and the two middle levels of this factor (data for the 1 and 2 pairs averaged), $F(1, 65) = 10.39$, $p = .002$. For single features, this contrast did not reach statistical significance, $p = .054$, while for double features it was significant, $p = .007$; however, there was no interaction between the two contrasts, $F < 1$.

In the 1- and 2-pair conditions, this time accuracy was affected significantly by whether the change pertained to a connected

($M = .798$, [.763, .833]) or an unconnected object ($M = .764$, [.727, .800]), $t(65) = 2.95$, $p = .004$, Cohen's $d = 0.23$.

Discussion

To summarize the results, even with a simplified stimuli presentation leading to a good change detection performance, the findings of Experiment 1 on the nonlinear relationship pertaining to the number of connected objects were replicated.

The design of the experiment allowed, in the single-feature trials, for cases in which elements of a pair had the same feature (e.g., two connected red circles) as well as in which the features were distinct (e.g., connected red and blue circle). It seems to be plausible that remembering pairs composed of same-feature elements is easier, and so the similarity of features can introduce another factor affecting performance. However, it is unlikely that the observed nonlinear effect could result from such a factor, as in all the conditions there was exactly the same distribution of features among the objects (e.g., three red and three blue circles or two red and four blue circles), and the harder-to-encode pairs composed of elements with different features were equally probable as the easier-to-encode pairs encompassing the same feature.

Furthermore, the similarity of the features could affect performance by leading to the perceptual grouping of elements sharing colors or

shapes. Such groupings could (a) make remembering certain features easier independently from the effect of spatial connectedness, or (b) interfere with the effect of spatial connectedness by introducing an alternative way of combining the presented elements into higher-order units. However, once again, these potential influences are not likely to cause the observed nonlinear effect. If hypothesis (a) is true, then the effect of grouping should be the same no matter the number of spatially connected objects as the effect of grouping is independent from the effect of connectedness. Alternatively, if hypothesis (b) is true, then the effect of grouping should be strongest in the 0-pair condition, as there is no interfering organization determined by spatial connectedness, and weakest in the 3-pair condition.

However, as the encoding of two features was difficult, and the single-feature trials were mixed with the double-feature trials, it was possible that even in the former trials the participants to some extent encoded both features of objects, even if one of them was irrelevant. To test whether the nonlinear effect of pairs can be generalized onto encoding and storage in VWM of *actual* single features, in the next experiment only single-feature trials were applied. Moreover, to further decrease the processing load yielded by the task, the original number of single-feature trials (160) was reduced to 72, to decrease the overall length of the experiment (boredom/fatigue).

Experiment 3

Participants

Another 101 people with normal color vision (41 men and 60 women) aged 18–36 were recruited and examined in an analogous way as in Experiments 1 and 2. Five additional participants were excluded due to random performance. The sample was enlarged, in comparison to Experiment 2, in order to compensate power decreases related to a smaller total number of trials in the present experiment.

Materials and design

The stimuli were six colorful circles with the same diameter and spatial layout as in the previous experiments. Each circle was presented as having a different color chosen randomly from the following: black, white, blue, green, red, yellow, and fuchsia. Again, connected stimuli could form either vertically or horizontally oriented pairs and the number of connected pairs could total 0, 1, 2, or 3. Due to the nature of the stimuli, there were only color trials in which the second stimuli pattern could differ from the first pattern in the color of one of the circles. The colors were distributed randomly such that no two circles were presented as having the same color.

Procedure

Each participant was presented with 3 blocks consisting of 24 trials, after a training session as in Experiments 1 and 2. Again, the presentation of stimuli was 2 s, and 4 s were allowed to decide whether the two patterns were the same or different by the relevant key press. The same type of PC set was used. A fixation cross, lasting for 2 s, was presented before each new trial in order to ensure that the participants had no difficulty in discerning between subsequent trials and focusing on the center of the screen. Because all trials were color-trials, the second stimuli pattern could be identical to the first pattern or could differ in exactly one colored circle. The possible stimuli patterns and the course of the trial are illustrated in Fig. 1C.

Results and discussion

Surprisingly, in Experiment 3 the effect of pairs on performance disappeared, $F(3, 300) = 0.81, p = .488, MSE = .037, \eta_p^2 < .01$. The discrimination index values (shown in Fig. 2D) were comparable

regardless of the number of objects connected.

In the 1- and 2-pair conditions, changes to connected objects yielded comparable accuracy ($M = .682, [.617, .698]$), as did changes to unconnected objects ($M = .657, [.638, .726]$), $t(100) = 1.14, p = .257$.

Because in Experiment 3 the nonlinear effect was not observed, we formulated two hypotheses in order to explain this result. First, the lack of a fixation cross may somehow distort the results of Experiments 1 and 2. Second, it is possible that the difference resulted from the fact that in Experiment 3 all trials required tracking only one feature, while in Experiments 1 and 2 one-feature trials were mixed with two-feature trials. Explanations referring to other factors were unlikely as the stimuli presentation time in Experiment 3 was the same as in Experiment 2, and trials in which each circle had a different color were used in Experiment 1. Nevertheless, the nonlinear effect was present both in Experiment 1 and 2. To address the first hypothesis, we decided to replicate Experiment 2, in which the nonlinear effect was clearly visible, with the fixation cross added. We aimed to verify whether this simple modification of the experimental procedure could determine the presence and absence of the effect considered. Ruling this case out would suggest the other option, namely that the non-linear effect of the number of pairs is specific to multiple-feature change detection but disappears when only one feature needs to be tracked. This alternative option would be addressed directly in Experiment 5.

Experiment 4

Participants

A total of 84 people with normal color vision (36 men and 48 women) aged 18–38 were recruited and examined in an analogous way as in Experiments 1–3. Four additional participants were excluded due to random performance. The sample was increased, as compared to Experiments 1 and 2, to compensate to some extent for the power decrease related with a decreased number of trials (128 instead of 320).

Materials and design

The experimental design was the same as in Experiment 2 (see Fig. 1B for an example stimuli). The only difference was the addition of a fixation cross presented for 2 s between the consecutive trials.

Procedure

Apart from the different number of trials (4 blocks with 32 trials in each), the procedure was the same as in Experiment 2.

Results and discussion

The discrimination index for colors equaled $M = .758 [.710, .806]$, while for shapes it amounted to $M = .676 [.616, .737]$, with a significant difference for the material type, $t(83) = 4.25, p < .001$, Cohen's $d = 0.33$. Mean k value equaled 4.3 objects on average. The 2×4 rmANOVA yielded a highly significant effect of features, $F(1, 83) = 156.02, p < .001, MSE = .045, \eta_p^2 = .65$, indicating a better performance for single features ($M = .717, [.667, .767]$) than for double features ($M = .512, [.464, .560]$), and a significant effect of pairs, $F(3, 249) = 6.47, p < .001, MSE = .031, \eta_p^2 = .07$, qualifying the nonlinear relationship supported by the extreme (average on 0 and 3 pairs) vs. middle contrast (average on 1 and 2 pairs), $F(1, 83) = 9.84, p = .003$, as illustrated in Fig. 2C. However, this time performance for the three pairs did not differ from the performance for one pair, $F < 1$ (the remaining differences between the conditions were significant at the $p = .05$ level). No significant interaction of the contrasts was observed between single and double features, $F(1, 83) = 1.95, p = .165$, even though the contrast for double features did not reach the aforementioned significance level, $F < 1$.

As in Experiment 2, changes to connected objects in the 1- and 2-pair conditions yielded significantly higher accuracy ($M = .772$, [.736, .808]) than changes to unconnected objects ($M = .701$, [.667, .734]), $t(83) = 3.82$, $p < .001$, Cohen's $d = 0.31$.

In conclusion, Experiment 4 replicated the nonlinear relationship between the number of connected pairs and the change detection accuracy, ruling out any role of the fixation cross, although this time the data pattern was not as symmetrical as in Experiments 1 and 2 (the nonlinear pattern was primarily detected for the single-feature trials).

Overall, the results of Experiments 1–4 suggested that the homogeneous structure of objects (i.e., either all composed in pairs or all separate) benefits VWM performance only when more complex encoding of visual features is required; specifically, when two distinct features of a single object have to be integrated. Processing just a single feature in Experiment 3 might not have yielded enough hierarchical/structured aspects of the resulting mental representations for the level of connectedness to matter. To support such a conjecture, the final dedicated study focused on whether the same participants would deliberately encode single visual features in one condition, while in the other condition they would be forced to integrate two features.

Experiment 5

Participants

Another 93 people with normal color vision (42 men and 51 women) aged 18–30 were recruited and examined in an analogous way as in the preceding Experiments. Four additional participants were excluded due to random performance.

Materials and design

The experimental design was the same as in Experiments 4 and 2 (see Fig. 1B for an example stimuli) with one key exception; the only difference was that, within a particular block of trials, the trials from the color condition, shape condition, and mixed condition were not presented randomly. At first, all the 8 color condition trials were presented, then all the 8 shape condition trials followed, and finally all the 16 mixed condition trials were shown. The random order was applied only within each of the conditions. Thus, it was expected that after starting a new block of trials, the participants would be tracking just a single feature (as only one varied), and only in the 17th trial would they notice that as many as two features need to be tracked.

Procedure

The procedure was the same as in Experiment 4, albeit this time with 5 blocks of 32 trials. The only difference was that trials belonging to different conditions were presented in blocks and not randomly mixed, as described previously. The instruction (in the local language) stated that a participant would be presented with (a) blue and red circles, (b) black circles and squares, or (c) a mix of the figures specified in (a) and (b). It was stated that in each trial one feature might change, namely the color of red/blue circles or the shape of black circles/squares, and so to succeed in the mixed trials both features had to be memorized, while in the color-only and shape-only trials only one feature had to be tracked.

Results and discussion

Change detection performance was virtually identical for colors and shapes ($M = .782$ and $.787$, respectively), $t(92) = 0.33$. The lack of the difference in performance between colors and shapes, which was present in Experiment 2, here could result from practice. Specifically, in the present experiment the shape-only trials were always presented after the color-only trials, and so when the shape-only trials were

appearing the participants were better accustomed to the demands of the experimental task. The 2×4 rmANOVA yielded a highly significant effect of features, $F(1, 92) = 163.46$, $p < .001$, $MSE = .042$, $\eta_p^2 = .64$, indicating a better performance for single features ($M = .784$, [.741, .828]) than for double features ($M = .593$, [.546, .640]). It also yielded a significant effect of pairs, $F(3, 276) = 5.98$, $p < .001$, $MSE = .021$, $\eta_p^2 = .06$. The two-way interaction was not significant, $F(3, 276) = 2.44$, $p = .065$. Additionally, in the single-feature condition, no significant contrast pertained to any pairs number or for the extreme (0 and 3) pairs vs. middle (1 and 2) pairs, all $ps > .078$, whereas in the double-feature condition the extreme vs. middle contrast was highly significant, $F(1, 92) = 10.35$, $p = .001$. When these two contrasts were compared, such a two-way contrast indicated a significant interaction with the number of features, $F(1, 92) = 4.31$, $p = .041$.

As in Experiment 3, there was no significant difference in accuracy in the 1- and 2-pair conditions between changes to connected objects ($M = .803$, [.776, .831]) and changes to unconnected objects ($M = .782$, [.751, .813]), $t(92) = 1.69$, $p = .094$.

Overall, the results of Experiment 5 further confirmed the presence of the nonlinear effect for pairs, but also suggested that this effect is specific only to multi-feature objects that supposedly require more complex VWM representations than single features, whereas the effect is absent during single-feature tracking.

General discussion

To summarize, in four out of the five conducted experiments, a novel and surprising effect was revealed, consisting of the nonlinear “U-shaped” relationship between the proportion of spatially connected objects within the stimuli pattern and the actual VWM capacity. Capacity was highest when all the objects were either spatially pairwise connected or spatially separated. When only the subset of objects was connected, while the remaining objects were presented as spatially separate, the actual capacity dropped.

In order to comprehensively assess the overall size of this effect in the case of the double-feature trials, the data from Experiments 1, 2, 4, and 5 were transformed into k values, combined together and submitted to one rmANOVA ($N = 304$), with the four experiments acting as one factor and the two pair conditions (averaged 0 and 3 pairs vs. averaged 1 and 2 pairs) being the other factor. The actual VWM capacity dropped significantly by $\Delta k = -0.23$ objects when only some were connected ($k = 3.09$ objects, [2.93, 3.25]), as compared to when none and all objects were connected ($k = 3.32$ objects, [3.17, 3.48]), $F(1, 300) = 18.99$, $p < .001$, $\eta_p^2 = .06$. The drop was notable, as it equaled almost a quarter of the objects out of the average VWM capacity of around 3 objects. There was no interaction of the number of pairs factor with the experiment factor, $F < 1$, suggesting that the non-linear effect of the pairs number was robust across the four studies.

As the results on the difference in change detection accuracy between connected and unconnected objects were inconclusive, with significant differences in Experiments 2 and 4, but not in Experiments 1 and 5, the data from the 304 participants in these four experiments were combined into one analysis. The advantage in change detection accuracy for connected objects ($M = .741$, [.721, .760]) over unconnected objects ($M = .703$, [.684, .723]) was significant, $t(303) = 4.99$, $p < .001$, but the overall effect size was relatively weak, Cohen's $d = 0.28$. This advantage could result from the fact that in one-feature conditions both elements of a pair may have the same feature and remembering same-feature pairs is easier (see discussion on pp. 11–12). Alternatively, better performance in cases of connected objects may be present simply because spatially connected objects are processed as a single perceptual unit and so storing information about their features is less demanding than in the case of spatially disjointed elements.

The role of homogeneity of mental structures on VWM performance

The present results cannot effortlessly be explained by any existing model of VWM performance. Manipulating the number of paired items was likely to influence either the number of actually stored objects (a pair can be treated as a single object with two parts) or the featural complexity of such objects (a pair has more relevant features), or both. While each factor is commonly believed to affect performance (e.g., Awh et al., 2007; Cowan et al., 2012; Hardman & Cowan, 2015), each effect should have resulted in the linear relationship between the number of connected objects and performance. If the reduction in the number of stored objects was the primary factor influencing performance, then the trials in which only some objects were pair-wise connected would be easier than trials in which all the objects were disjointed, but more difficult than the trials in which all the objects were paired. If the featural complexity were crucial, then the trials in which some objects were connected would be more difficult than the trials in which all the objects were disjointed, but easier than the trials in which all the objects were paired. In contrast, in the present experiments the trials in which some of the objects were connected and some of the objects disjointed yielded the lowest performance.

The observed U-shaped relationship cannot be fully explained by referring to relational and contextual factors (Brady et al., 2011; Clewenger & Hummel, 2013; Johnson et al., 2013; Kałamała et al., 2017; Li et al., 2018; Peterson & Berryhill, 2013; Woodman et al., 2003), because in the trials in which all the objects were connected and in the trials in which all the objects were disjointed the overall pattern of elements did not satisfy any Gestalt principles to a larger extent, as compared to the trials with both paired and disjointed objects, and so could contribute to a better performance. Additionally, the number of features and feature dimensions was identical across all types of trials. Furthermore, the observed nonlinear effect is unlikely to result from visual crowding as in all the conditions the minimal distance between objects which did not constitute a single pair was the same. Consequently, objects in 1- and 2-pair arrays were not significantly closer to each other than objects in 0- and 3-pair arrays. Of course, it may be claimed that pairing objects by itself leads to crowding, as paired objects are obviously very close to each other. However, such crowding would result in a linear effect as the more pairs the stronger the crowding.

We propose that the existing models of VWM should be supplemented with another factor potentially influencing performance in the change detection task, that is, with a new dimension of stimulus pattern complexity, understood as the level of the heterogeneity of object structures present within the overall display. This new factor comprises something beyond the internal complexity of objects that is determined by the number of their parts and features, as advocated by the existing models of VWM. This factor cannot be reduced to the characteristics of the overall display (e.g., its compatibility with Gestalt principles) nor to contextual effects, because two displays, one containing the objects of a uniform structure and the other including two or more heterogeneous structures, can still conform to the same principles and context.

Specifically, in our study the trials with no connected objects as well as the trials with all the objects connected yielded only one type of object-structure: either simple figures or complex figures composed of two parts. Only the trials with some objects connected required the encoding of both types of structures. This additional variation of object structures might have imposed an additional load on the VWM mechanisms, because a different procedure might have been required to access the features of each type of object in order to detect the changes in an effective way; the two-part figures could require a detailed representation of each part, while the single figures could be represented in a more rudimentary fashion. In other words, a potential change to the two-part whole concerns only features of its part, but not of the whole, whereas an analogous change to the single object concerns it as a whole, while its local details can be safely neglected. Thus, in

structurally heterogeneous displays the VWM mechanism may be forced to switch back and forth between contrasting dimensions of the display elements, and that fact may yield certain processing costs that can be detrimental to overall performance.

The results of Experiments 3 and 5 show that the heterogeneity of objects' structures is more likely to affect performance when participants have to track more than one type of feature (like color and shape, in the case of our study). This fact suggests that there is some interference between processing a higher number of structures and a higher number of features, such that when the memory task is easy, as in the case of tracking one feature, the detrimental effect of heterogeneity is not observed. The presence of such interference is likely because differentiating between distinct structures introduces additional demands on the mechanisms responsible for processing features. In particular, to establish whether an item is a part of a larger whole or is a separate whole itself, it needs to recognize if its contours are connected with the contours of a different item. Furthermore, if the structures are heterogeneous, processing features becomes more complex as in addition to feature recognition it has to also recognize whether a given feature characterizes a whole or part of a complex whole.

Comparison with existing studies on spatial connectedness in VWM

The proposed hypothesis is similar to that tested by Barton, Ester, and Awh (2009). In their change-detection study, the participants were presented with displays containing Chinese characters, 3-dimensional cubes with different orientations, or both types of elements. The performance was not significantly different between uniform and mixed trials; this suggests that the heterogeneity of objects' structures did not make the task harder. However, the possible changes of objects, i.e. replacing a Chinese sign with another, or replacing a cube with one with a distinct orientation, always concerned the properties of a whole object and not only its part. As such, the heterogeneity of part-structures did not have to be encoded by VWM in order to succeed in the task. This was not the case in the present design, where the changes concerning paired objects always involved the modifying properties of one of two pairs, whereas the changes regarding simple objects always concerned a property of the whole object.

Similarly, in the change-detection experiments conducted by Peterson, Gözenman, Arciniega, and Berryhill (2015), colorful objects were presented either as disjointed or in such a way that some of them constituted spatially connected pairs. In this case, a significant difference between performance in the trials with only disjointed objects and the trials in which some objects were connected was not observed. However, the colorful objects, whose properties have to be tracked to succeed in the task, were not directly spatially connected but were separated by an additional black part. This could interfere with integrating objects into a single unit and decrease the effects of spatial connectedness.

Furthermore, in contrast to the present experiments, a linear effect of connectedness on VWM capacity was observed by Gao, Gao, Tang, Shui, and Shen (2016). The stimuli used were incomplete circles with square gaps. They were presented in three configurations: (a) all circles were organized in pairs such that their gaps created a rectangle by virtue of a modal completion, (b) some of the presented circles were pair-wise organized, or (c) none of the presented circles were organized in such pairs. The task was to remember the orientation of incomplete circles. A linear pattern was observed, and the more circles organized in pairs the better the performance. However, Gao et al. (2016) had a significantly different design from the present experiments, as it investigated the diachronic functioning of VWM (the presented circles were not displayed at once but rather in succession), which could likely affect the results.

The aforementioned discussion shows that the existing change-detection experiments which investigated structurally heterogeneous stimuli used designs which made it difficult to observe the non-linear

effect described in the present study. In particular, the former studies presented objects that were not directly connected, were not presented synchronously, or the relevant changes always concerned whole objects and not only some of their parts.

Further experiments will have to identify structural parameters, specifically those most important for the change-detection performance, and working memory functioning in general. For instance, if a display including one three-part object, one two-part object, and one simple object was more difficult to memorize than a display with two two-part objects and two simple objects, then differences in the number of parts might be of high relevance, because, even though the number of parts were equal between the displays, the former display would contain three different object structures, while the latter display would contain only two such structures. Alternatively, if a display including one three-part object, one two-part object, and one simple object was similarly difficult as a display with two two-part objects and two simple objects, then the number of structural levels, which need to be “traversed” to reach a level at which relevant changes happen, may be crucial, because in both displays the features of object parts would be at one level “below” the whole object level. Other structural parameters affecting performance are also conceivable. Overall, the present work achieves a novel perspective on the studying of working memory capacity limits in the human mind, switching the focus from the sheer numbers of objects, features, and their relations such as factors determining actual capacity, towards the intrinsic complexity of mental structures required to encode these objects, features, and relations.

Conclusion

Concluding the present study, it yielded a highly original observation: the actual capacity of VWM is neither a linear function of the number of to-be-memorized objects (assuming that two spatially connected figures form one complex object) nor a linear function of the number of relevant object features. Thus, it contradicts the predictions of the dominant object-based and feature-based models of VWM. This observation can neither be easily explained by the slot-based nor the resource-based models. It also seems to lay beyond the relational and contextual models of VWM. As a result, actual VWM capacity (at least in the change detection task studied here) is unlikely to depend on a single type of factor. In contrast, the present results suggest that VWM representations possess some hierarchical structure defined by wholes, parts, features, and their relations, and the heterogeneity of such a structure hinders the effectiveness of VWM encoding, maintenance, and retrieval, while homogeneity facilitates processing in VWM.

CRedit authorship contribution statement

Błażej Skrzypulec: Conceptualization, Methodology, Software, Validation, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Adam Chuderski:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization, Funding acquisition.

Acknowledgements

This work was sponsored by the National Science Centre of Poland (grant no. 2015/17/B/HS6/04152 awarded to A. Chuderski). The PsychoPy code for experimental procedures and the data can be obtained from an open repository (osf.io/gjwyb) or from the corresponding author upon request (via email to: blazej.skrzypulec@gmail.com).

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jml.2020.104124>.

doi.org/10.1016/j.jml.2020.104124.

References

- Anderson, D. E., & Awh, E. (2012). The plateau in mnemonic resolution across large set sizes indicates discrete resource limits in visual working memory. *Attention, Perception & Psychophysics*, 74, 891–910.
- Anderson, D. E., Vogel, E. K., & Awh, E. (2011). Precision in visual working memory reaches a stable plateau when individual item limits are exceeded. *The Journal of Neuroscience*, 31(3), 1128–1138.
- Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, 18(7), 622–628.
- Barton, B., Ester, E. F., & Awh, E. (2009). Discrete Resource Allocation in Visual Working Memory. *Journal of Experimental Psychology: Human Perception and Performance*, 35(5), 1359–1367.
- Bays, P. M. (2015). Spikes not slots: noise in neural populations limits working memory. *Trends in Cognitive Sciences*, 19(8), 431–438.
- Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia*, 49, 1622–1631.
- Brady, F. T., 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.
- Clevenger, P. E., & Hummel, J. E. (2013). Working memory for relations among objects. *Attention, Perception & Psychophysics*, 76(7), 1933–1953. <https://doi.org/10.3758/s13414-013-0601-3>.
- Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–114.
- Cowan, N., Blume, Ch. L., & Saults, J. S. (2012). Attention to attributes and objects in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 731–747.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160.
- Fodor, J. A., & Pylyshyn, Z. (2002). How direct is visual perception?: Some reflections on Gibson's “Ecological Approach”. In A. Noe, & E. Thomson (Eds.). *Vision and mind* (pp. 167–227). Cambridge: The MIT Press.
- Fougnie, D., & Alvarez, A. A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision*, 11(12), 3.
- Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology*, 20, 177–182.
- 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.
- Gibson, J. J. (2015). *The ecological approach to visual perception* (Classic ed.). New York: Taylor & Francis.
- Hardman, K. O., & Cowan, N. (2015). Remembering complex objects in visual working memory: Do capacity limits restrict objects or features? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(2), 325–347.
- Johnson, J. S., Spencer, J. P., Luck, S. J., & Schoner, G. (2013). A dynamic neural field model of visual working memory and change detection. *Psychological Science*, 20(5), 568–577.
- Kalamala, P., Sadowska, A., Ordziński, W., & Chuderski, A. (2017). Gestalt effects in visual working memory: Whole-part similarity works, symmetry does not. *Experimental Psychology*, 64, 5–13.
- Li, J., Qian, J., & Liang, F. (2018). Evidence for the beneficial effect of perceptual grouping on visual working memory: An empirical study on illusory contour and a meta-analytic study. *Scientific Reports*, 8, 13864.
- Lin, P. H., & Luck, S. J. (2009). The influence of similarity on visual working memory representations. *Visual Cognition*, 17, 356–372.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281.
- Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Science*, 17(8), 391–400.
- 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.
- Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory & Cognition*, 41(8), 1212–1227.
- Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the “strong-object” hypothesis. *Perception & Psychophysics*, 64, 1055–1067.
- Peterson, D. J., & Berryhill, M. E. (2013). The Gestalt principle of similarity benefits visual working memory. *Psychonomic Bulletin & Review*, 20(6), 1282–1289.
- Peterson, D. J., Gözenman, F., Arciniega, H., & Berryhill, M. E. (2015). Contralateral delay activity tracks the influence of Gestalt grouping principles on active visual working memory representations. *Attention, Perception & Psychophysics*, 77(7), 2270–2283.
- Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics*, 6, 283–290.
- Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Research*, 40(10–12), 1469–1487.
- Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8(5), 368–373. <https://doi.org/10.1111/j.1467-9280.1997.tb00427.x>.
- Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review*,

- 18, 324–330.
- Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, 1(7), 261–267. [https://doi.org/10.1016/s1364-6613\(97\)01080-2](https://doi.org/10.1016/s1364-6613(97)01080-2).
- Sligte, I. G., Scholte, H. S., & Lamme, V. A. (2008). Are there multiple visual short-term memory stores? *PLoS One*, 3(2), e1699. <https://doi.org/10.1371/journal.pone.0001699>.
- Sligte, I. G., Scholte, H. S., & Lamme, V. A. (2009). V4 activity predicts the strength of visual short-term memory representations. *Journal of Neuroscience*, 29(23), 7432–7438. <https://doi.org/10.1523/JNEUROSCI.0784-09.2009>.
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117, 34–50.
- Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, 4, 1120–1135.
- Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, 10(1), 80–87.
- World Medical Association (2001). World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *Bulletin of the World Health Organization*, 79(4), 373–374.
- Xu, Y. (2002). Encoding color and shape from different parts of an object in visual short-term memory. *Perception & Psychophysics*, 64(8), 1260–1280.
- Xu, Y. (2006). Understanding the object benefit in visual short-term memory: The roles of feature proximity and connectedness. *Perception & Psychophysics*, 68(5), 815–828.