



A snapshot is all it takes to encode object locations into spatial memory



Harry H. Haladjian^{a,*}, Fabien Mathy^{b,1}

^a School of Social Sciences and Psychology, University of Western Sydney, Australia

^b Département de Psychologie, Université Nice Sophia Antipolis, France

ARTICLE INFO

Article history:

Received 1 May 2014

Received in revised form 17 December 2014

Available online 31 December 2014

Keywords:

Spatial attention

Visual short-term memory

Short-term memory

Subitizing

Grouping

Clustering analysis

ABSTRACT

This study examines the encoding of multiple object locations into spatial memory by comparing localization accuracy for stimuli presented at different exposure durations. Participants in the longest duration condition viewed masked displays containing 1–10 discs for 1–10 s (durations typically used in simple span tasks), and then reported the locations of these discs on a blank screen. Compared to conditions that presented the same stimuli briefly for 50 or 200 ms (exposures more typical of simultaneous spatial arrays), localization accuracy did not improve significantly under longer viewing durations. Additionally, a clustering analysis found that responses were spread among different clusters of discs and not focused on individual clusters, regardless of viewing duration. A second experiment tested this performance for displays containing two distinct clusters of discs to determine if clearly grouped subsets of objects would improve performance, but there was no substantial improvement for these two-cluster displays when compared to displays with one cluster. Overall, the results indicate that spatial information for a set of objects is extracted globally and quickly, with little benefit from extended encoding durations that should have favored some deliberative form of grouping. Such results cast doubt on the validity of Corsi blocks or equivalent common neuropsychological tests purportedly designed to evaluate specifically spatial short-term memory spans.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Visual memory is often studied to identify the stages in perception with information processing limitations. The capacity limit in memory, for example, can affect the quality of simultaneous object representations, where having to remember more objects reduces the amount of detail that can be encoded about those objects (e.g., Alvarez & Cavanagh, 2004; Ma, Husain, & Bays, 2014). Others argue that there is a limit on the total number of objects remembered regardless of the amount of information encoded per object—the often cited “four slots” limit found in working memory studies (e.g., Cowan, 2001; Luck & Vogel, 1997; Zhang & Luck, 2008). Such processing limits also relate to the fast and error-free counting of up to four items, called “subitizing”, where enumeration errors and response latencies increase substantially for sets larger than four (e.g., Burr, Turi, & Anobile, 2010; Dehaene & Cohen, 1994;

Kaufman et al., 1949; Pylyshyn, 1989; Revkin et al., 2008; Trick & Pylyshyn, 1993, 1994). Thus, it may be possible that both visual memory for objects and enumeration share a common resource with similar capacity limitations and variations that correlate between subjects (Cutini & Bonato, 2012; Piazza et al., 2011). This has been hypothesized to result from an initial competitive process of the individuation of the objects present in a visual scene (Melcher & Piazza, 2011).

The number of items that can be processed quickly (i.e., the subitizing range), however, tends to vary depending on the stimuli and reporting methods used. For example, there is an interaction between the intensity of the stimulus and the duration of exposure to it, with higher intensity stimuli requiring less time for detection (Hunter & Sigler, 1940) while also facilitating the detection of larger sets of items (Palomares & Egeth, 2010). Typical verbal reports produce a subitizing range of around four items (e.g., Revkin et al., 2008; Trick & Pylyshyn, 1993, 1994), with higher ranges when polygons forming prototypical configurations of dots are used, like the patterns found on a dice (e.g., Mandler & Shebo, 1982; Yantis, 1992). A recent localization study also identified a higher subitizing range when participants reported numerosity by marking the locations of briefly-viewed objects (Haladjian & Pylyshyn, 2011). In that study, participants were shown masked displays with randomly-placed discs at brief durations (50–350 ms), and then

* Corresponding author at: School of Social Sciences and Psychology, University of Western Sydney, Bankstown Campus, Building 24, Locked Bag 1797, Penrith, NSW 2751, Australia. Fax: +61 02 9772 6757.

E-mail addresses: haroutioun@gmail.com (H.H. Haladjian), fabien.mathy@unice.fr (F. Mathy).

¹ Département de Psychologie, Laboratoire BCL: Bases, Corpus, Langage – UMR 7320, Université Nice Sophia Antipolis, Campus SJA3, 24 avenue des diables bleus, 06357 Nice Cedex 4, France.

marked the locations of each disc on a blank computer screen. In addition to measuring spatial memory for sets of objects, this reporting method provided a numerosity estimate. Enumeration performance was high for displays with up to six items when using the localization method, but only up to four items (the “typical” subitizing limit) when using a conventional reporting method with Arabic numerals in that study.

The motivation for the current study is to better understand the factors that may enhance spatial memory for a number of simple objects (e.g., [Franconeri, Alvarez, & Enns, 2007](#); [Haladjian & Pylyshyn, 2011](#)) when using this location-based reporting method. One explanation for a higher capacity for remembering object locations may be related to the act of “pointing” to the locations of the discs, since this also engages a memory involved in motor responses (e.g., [Goodale & Milner, 2004](#)). Another possible explanation for this increased capacity is perceptual grouping, where nearby discs are grouped together for more efficient storage (e.g., [Anderson, Vogel, & Awh, 2013](#); [Brady & Tenenbaum, 2013](#); [Feldman, 1999](#); [Korjoukov et al., 2012](#)). Effectively, a grouping process involves the ability for proximal discs to form a group and produce non-independent spatial information for those discs, which could be encoded compactly into a single “slot” in memory. This would allow the encoding of information about other discs (or groups of discs) into the remaining free “slots”, and thereby increase the number of individual items that can be encoded. Such abilities for information processing systems to overcome capacity limitations whenever relational information can be computed has received particular attention recently in the visual short-term memory literature (e.g., [Alvarez & Cavanagh, 2004](#); [Bays, Catalao, & Husain, 2009](#); [Bays, Wu, & Husain, 2011](#); [Brady, Konkle, & Alvarez, 2009, 2011](#); [Fougnie & Alvarez, 2011](#); [Sargent et al., 2010](#); [Wheeler & Treisman, 2002](#); [Xu, 2002](#)). The present study tests this possible explanation within a localization task by enhancing grouping effects with longer viewing durations (Experiment 1) and by presenting displays that have clearly groupable sets of objects (Experiment 2).

The manner in which object locations are encoded into memory can be described in two different ways. One view proposes that resource allocation is continuous (e.g., [Alvarez & Cavanagh, 2004](#); [Bays & Husain, 2008](#); [Gorgoraptis et al., 2011](#); [Ma, Husain, & Bays, 2014](#); [Wilken & Ma, 2004](#)), which suggests that the key factor for memorization is the accuracy with which all the material is encoded. A contrasting view is that resource allocation is discrete, or slot-based, which proposes that the key factor for memorization is the number of objects that can be encoded (e.g., [Donkin et al., 2013](#); [Luck & Vogel, 1997](#); [Zhang & Luck, 2008](#)).

Resource allocation also can be framed in terms of information compression (e.g., [Brady, Konkle, & Alvarez, 2009](#)). One information compression method encodes information in a “lossless” manner (e.g., [Mathy & Feldman, 2012](#)), which allows the exact original data to be reconstructed from memory. In terms of spatial memory, it is possible that exact information about groups of items can be compressed in a lossless manner so that a greater number of items can be unpacked from a few groups. Similar to the ability to recall a series of 50 numbers, such as 2-4-6-8-10-, ..., 100, by retaining the shorter description “even numbers from 2 to 100”, it might be possible to retain the coarse locations of several groups of items (e.g., [Aksentijević, Elliott, & Barber, 2001](#); [De Lillo, 2004](#); [Dry, Preiss, & Wagemans, 2012](#); [Feldman, 1999](#); [Korjoukov et al., 2012](#)) without the loss of the original information regarding the number of items within each group. This process that supports the encoding of local perceptual structures, however, does not prevent any subsequent forms of distortion of the represented structures within groups. If present, this preliminary encoding of local structures can be detected using specific analyses. For example, this lossless encoding of local groups would produce a correct

report of a limited number of items, with total loss of information for items that could not be encoded due to capacity limitations (essentially indicative of a slot-based fixed resource). One example is when an observer is shown a display with seven items, and she could encode a group of three items on the top of the screen and another group of two items on the left bottom part of the screen. This observer would in this case perfectly report the presence of two groups, and would report five discs with great accuracy, but would not correctly report the presence of the two other remaining discs on the right bottom part of the screen (again, this approach does not expect the individual locations to be reported perfectly for any of the discs).

An alternative encoding method that may help increase capacity can be described as “lossy”. This may include the computing of a summary statistic, such as a global summary of spatial relationships (e.g., [Jiang, Olson, & Chun, 2000](#); [Sargent et al., 2010](#)). Reproducing this information will result in more systematic errors distributed among all objects in memory and would be indicative of a more continuous and flexible resource model (e.g., see [Franconeri, Alvarez, & Cavanagh, 2013](#)). This would suggest a non-independent encoding of spatial information. (By analogy, these two forms of compression are similar to digital file formats such as .png/gzip or .jpeg/mpeg, which are respectively lossless and lossy.)

In the current study, we examine whether or not perceptual grouping of proximal objects improves spatial memory and capacity, and also characterize how spatial information tends to be encoded (i.e., lossless or lossy). Clustering measures were used to determine if participants use grouping to remember the number of items and possibly encode more precise spatial information about multiple object locations. A first hypothesis is that displays with more groupable arrays generally will facilitate spatial memory by inducing the grouping of a limited set of proximal objects, thus increasing capacity. A second hypothesis is that longer viewing durations enhance grouping by enabling more deliberate grouping processes. In this case, more objects are encoded more precisely due to the perception of a local structure since the focus of attention can be directed on individual groups. Alternatively, spatial information may be encoded globally in a quick “snapshot” and if so, would not benefit from longer exposures. In this latter case, object locations are encoded via the perception of a global structure (i.e., from a more diffuse or global focus of attention).

To investigate these questions, we used the localization task described above that required participants to remember object locations on displays containing randomly-placed discs, and we compared performance between exposure durations typical of simple span tasks ([Shipstead, Redick, & Engle, 2012, p. 629](#)) and change detection or continuous report tasks ([Brady, Konkle, & Alvarez, 2011, p. 3](#)). This essentially corresponds to comparing performance when local encoding is encouraged, to performance when global encoding is likely to occur. Note that we use a “pseudo-opposition” between the short-term memory durations only to reflect the fact that, generally, simple span tasks use longer presentation rates (to aid neuropsychological assessment and to facilitate instructions and computation of memory span) while rapid displays prevent the use of various conscious strategies. In typical simple short-term memory span tasks, the stimulus (verbal or spatial) is usually presented at a rate of one item per second (e.g., [Gmeindl, Walsh, & Courtney, 2011](#) for the Corsi block-tapping test; see also the computerized spatial short-term memory tests of [Lewandowsky et al. \(2010\)](#)); the present study examines the gain that is expected with such longer durations. To find a common procedure for reporting the locations in the present study, however, both our conditions used a free recall procedure of the whole display in order to focus on grouping processes, rather than using either single-probe or whole-display recognition (see [Rouder](#)

et al., 2011, p. 325) for our rapid conditions, or serial report like in simple span tasks for our slow condition.

The Corsi block-tapping test, for example, requires a participant to repeat a sequence of blocks that have been identified by an experimenter (or illuminated on a computer screen; Gmeindl, Walsh, & Courtney, 2011) at a rate of one item per second until the participant makes a mistake (Berch, Krikorian, & Huha, 1998; Richardson, 2007). A more complex visuospatial span task would include a concurrent processing task between the presentation of each to-be-remembered item (e.g., judging the symmetry of a matrix), but still, both spans are defined as the amount of information one can recall in the correct order over a brief period of time (Aben, Stapert, & Blokland, 2012; Shipstead, Redick, & Engle, 2012). The slow speed at which the material is presented in a simple task allows deliberative processes to occur (e.g., since there is no concurrent task, participants can rehearse the to-be-remembered material) and is known to increase capacity up to seven items (Miller, 1956) in contrast to more recent studies using either faster procedures such as simultaneous spatial arrays or complex span tasks, which have been both devised to prevent rehearsal and grouping processes and have shown a capacity limit of four items (Cowan, 2001). Such a difference in capacity may indicate that visual memory and subitizing processes share resources (Miller, 1956), which could account for the higher capacity in some cases when presentation duration allows short-term memory encoding, but also more rarely for faster presentations.

The distinction between short-term memory and other temporary memory processes is arguably hard to distinguish (see Aben, Stapert, & Blokland, 2012), but we generally take the long exposure condition as one that allows more deliberate or serial attentional processing to better encode local information into memory while the short exposure condition limits attentional processing to a brief global “snapshot”. Although previous work has studied how memory precision declines in a sequence (e.g., Gorgoraptis et al., 2011), to the best of our knowledge, we are not aware of a study that has attempted to characterize simultaneously “short-term memory capacity” under long viewing durations (typical of a Corsi task) and under short durations (typical of simultaneous spatial array tests).

In our long exposure condition, participants viewed masked displays containing 1–10 discs using typical short-term memory exposures, that is, 1-s per item (e.g., used in the spatial span task by Lewandowsky et al. (2010), although their presentation of the stimuli is serial while ours is simultaneous). In Experiment 1, we compared these results to our previous studies that presented the same stimuli for 50 or 200 ms for all items (using data from Haladjian & Pylyshyn, 2011; Haladjian et al., 2010). We expect that with longer viewing exposures, intentional grouping strategies may be used to increase the local grouping of objects and thus improve spatial memory. In Experiment 2, we used a within-subjects design to test the differences between exposure durations as well as how displays with two clear groups of discs would affect performance. Therefore, we examined if longer and more groupable displays produce both better accuracy in reporting locations and less spatial distortions within reporting patterns.

Spatial distortion was studied through spatial compression effects that tends to result in participants remembering objects as being closer to each other than they actually were on the stimulus displays. This spatial compression is different from the information compression discussed above, in which compression was associated with many potential encoding processes that can make one spatial representation more economical. Spatial compression is only one form of distortion or bias that can be present in the report of locations (e.g., Haladjian et al., 2010; Sheth & Shimojo, 2001), which we thought could be useful to identify whether the distortions could occur locally or globally. The hypothesis was that

a better encoding of local structures with longer durations would reduce a global form of compression around the display's centroid. Although enumeration was not of primary interest (because longer durations obviously offered enough time to count the discs), we still report this measure to contrast it with another more refined measure that determined how many clusters were missed by participants under the hypothesis that a more local encoding of spatial information is associated with a greater risk of missing specific regions. The two experiments in this study aimed at characterizing the encoding of spatial information into memory and addresses questions regarding the fixed-slot versus flexible resources models of memory.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Thirty-four students and staff members from the Université de Franche-Comté were recruited for voluntary participation in this experiment and provided their informed consent; no payment was given. This research was approved by the university ethics board and carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.1.2. Apparatus

The experiment was programmed in MATLAB using Psychophysics Toolbox (Brainard, 1997). The stimuli were presented on a laptop running Windows XP, with a 38 × 22 cm LCD screen (1600 × 900 pixels; 60 Hz refresh).

2.1.3. Stimuli

The stimuli were comprised of 1–10 identical dark gray discs on a slightly lighter gray background ($\sim 1^\circ$ viewing angle in diameter). The discs were randomly placed within a $22^\circ \times 13^\circ$ (960×540 pixels) region in the center of the laptop screen with a minimum distance of $\sim 3^\circ$ between discs; this minimum distance was used to avoid crowding effects even at the periphery of our viewing displays (Bahcall & Kowler, 1999; Intriligator & Cavanagh, 2001). The discs were presented simultaneously for 1–10 s for a total duration of 1-s per disc (e.g., a display with 2 discs was presented for 2 s, and a display with 9 discs was presented for 9 s); all discs were presented simultaneously for the full duration of the trial. These low-contrast stimuli were designed to optimize the effectiveness of the subsequent random-dot texture mask that was presented for 1 s, which was used to prevent any after-images that could be used to aid localization.

2.1.4. Procedure

Except for the longer stimulus duration, the procedure is identical to that implemented in our previous studies using this localization task (Haladjian & Pylyshyn, 2011; Haladjian et al., 2010). Participants sat approximately 60 cm in front of the display; head movement was not restricted, therefore any visual degree angle computation is approximate. Each trial began with a 2-s presentation of a blank gray screen with a central fixation cross. The stimulus then appeared for the designated duration that was based on the number of discs (stimulus durations ranged from 1 to 10 s). This display was followed by a 1-s random-dot texture mask to limit viewing durations and prevent after-images. Finally, a blank gray screen appeared and participants used a computer mouse to place markers (“x”) on each of the perceived disc locations, and pressed the space bar to initiate the next trial. They were instructed to place markers for each object they saw, even if they were unsure about the exact location in order to provide us with

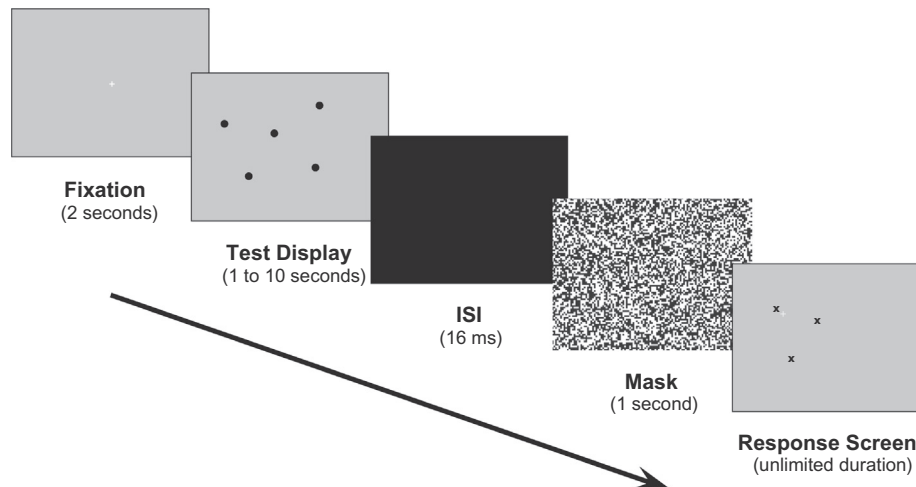


Fig. 1. Schematic of the experimental design for the localization task. This is an example of the long exposure condition in Experiment 1; this 5-disc display would be shown for 5 s (all discs shown simultaneously).

an estimate of enumeration accuracy. See Fig. 1 for a schematic of a trial in this experiment. The program presented 10 trials for each of the numerosities in a randomized order, for a total of 100 trials (with 9 separate practice trials). The experimental session lasted less than 30 min.

2.1.5. Analyses

Enumeration accuracy was determined by comparing the number of discs on a stimulus display to the number of markers placed on the response display. For localization accuracy, stimulus–response pairs were established in MATLAB (see Appendix Fig. A for examples of the stimuli and responses). For each trial, the responses were first transformed using Procrustes methods (Goodall, 1991) so that the response locations better fit the stimulus locations—these transformations included uniform expansion, translation, or rotation of the responses. Next, using nearest-neighbor methods with Delaunay triangulation, each response was individually matched to its most likely stimulus disc; any duplicate matches were corrected so that only one response was paired to a unique stimulus disc. Once these pairs were established, the Euclidian distance (from the raw data before the Procrustes transformations) between these two locations was computed to provide an estimate of the localization error. Outliers were removed from these analyses (i.e., errors >500 pixels, which accounted for 0.2% of cases). Note that trials with the incorrect number of responses were not included in some of the localization analyses in order to only report data from correctly-matched trials (although the overall results do not differ); this will be indicated in the relevant results section. Overall, in trials that had response number errors (20% of all trials), the majority of errors were undercounts (i.e., 86% of the miscounts were undercounts).

To capture a more global form of distortion, we computed a spatial compression measure by determining the distance between the centroid (center of mass of all the stimulus discs) and each stimulus disc for a trial, and compared it to the distance between the responses and their centroid in that trial. This allowed us to characterize localization errors in terms of whether spatial compression occurred, which would be indicated by shorter average distances to the centroid in the responses when compared to the stimulus distances. Such systematic compression of space is not uncommon in visual memory studies (e.g., Sheth & Shimojo, 2001) and was seen in the localization data from a previous study (Haladjian et al., 2010) where the shortest exposure duration produced the most spatial compression. (See Appendix Fig. A for an

example of this spatial compression in a participant's response.) Another way to examine systematic errors is to detect whether or not responses were biased toward the central fixation cross on the displays (instead of the center of mass of the discs), since that is where the participants were instructed to look during the experiment. When examining this measure in a preliminary analysis, we found less error in smaller numerosities under long viewing durations, but overall the trend did not indicate any substantial differences. Therefore, we use the centroid computation as it better reflects the relationships among the object locations in our stimuli.

Since the only difference between the current study and our previous studies is the longer presentation durations, we compared the results among these studies to determine if increased exposure to the stimuli improved task performance in Experiment 1. That is, we wanted to measure the benefit of increasing the presentation durations from 200 ms or less to several seconds for displays with multiple discs. In all the ANOVA models reported, subject ID was included as a random factor to control for between-subject variability. Multiple pairwise comparisons were adjusted using the Bonferroni method. Effect sizes are reported as partial eta-squared (η_p^2). All errors bars represent 95% confidence intervals.

2.2. Results

2.2.1. Enumeration accuracy

Although enumeration accuracy is not the main focus of this study, the errors do provide some insights on how these stimuli are encoded into memory. We take the number of locations that were marked on the response screen as an indirect measure of enumeration accuracy, since the observers are placing a discrete number of items to represent the items they saw on the stimulus displays. With longer viewing durations, we naturally expected better performance on reporting the correct number of items. A mixed model ANOVA on the proportion of trials that were enumerated correctly indicated significant main effects for numerosity ($F(7,1706) = 451.2$, $p < .001$, $\eta_p^2 = .65$) and duration ($F(1,151) = 261.0$, $p < .001$, $\eta_p^2 = .63$), with an interaction ($F(7,1057) = 43.4$, $p < .001$, $\eta_p^2 = .22$). See Fig. 2.

Additionally, pairwise comparisons were performed to examine duration effects within each of the numerosity conditions. The only significant differences in performance among the three duration conditions appeared for numerosities of 6 and greater (F 's > 15, p 's < .01), with performance near ceiling for displays with 1–5 discs (>95% accuracy). These results indicate, not surprisingly, that

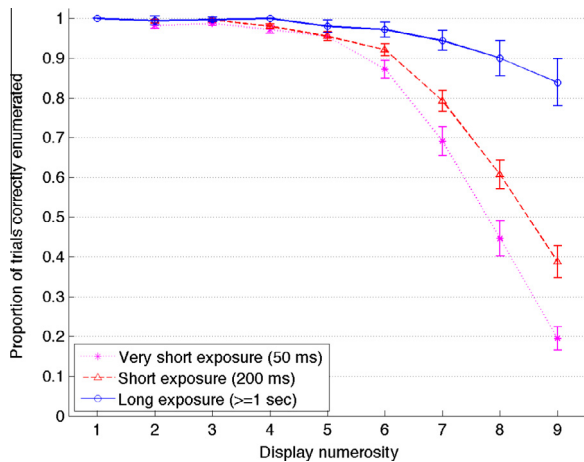


Fig. 2. Enumeration accuracy. Proportion of trials with the correct number of responses; long exposure ($N = 34$) and short/very short exposures (for both, $N = 152$; data from Haladjian and Pylyshyn (2011) and Haladjian et al. (2010)). Note: all error bars in this manuscript represent 95% confidence intervals.

longer viewing durations allow for more accurate recall of the number of items present on stimulus displays with 6 or more items, and that there was a significantly increasing rate of error for larger sets especially for short durations. Although this is not a surprising result, this indicates that there is not much loss of information during the long duration condition, which could potentially be aided by a lossless form of grouping. This process could thus potentially be involved in the almost perfect enumeration of 6 discs observed across durations. The primary interest of this study, however, is localization performance, which will be the focus of the remainder of the analyses that will target whether several groups of discs were effectively encoded to increase capacity above the four-item limit.

2.2.2. Localization accuracy

Localization accuracy was measured as the pixel distance between paired response markers and stimulus discs (only correctly enumerated trials were used in this analysis). Again, we expected that longer viewing durations would improve localization accuracy. Fig. 3a plots this experiment's results along with results from the previous studies with 50-ms and 200-ms exposures. The ANOVA examining the magnitude of localization errors indicated significant main effects for numerosity ($F(7,1864) = 195.4$, $p < .001$, $\eta_p^2 = .42$) and duration ($F(1,156) = 101.9$, $p < .001$, $\eta_p^2 = .40$), but with no interaction ($F(7,994) = 1.7$, $p = .11$, $\eta_p^2 = .01$). Pairwise comparisons indicated that errors increased significantly with each additional disc present on displays with 2–5 items in the 50-ms and 200-ms conditions (p 's $< .01$); errors did not increase significantly on displays with 5–9 items. For the long durations, errors increased significantly with each item for displays with 1–4 items (p 's $< .05$), after which the errors did not increase significantly. Comparing the effect of duration for each numerosity, the longest duration showed a decrease in errors compared to the 50-ms condition for displays with 2, 3, and 5 items (p 's $< .01$); there was no significant decrease in errors between the longest and 200-ms displays except for those with 7 items (p 's $< .01$). These results suggest that spatial information can be encoded globally in 200 ms, since there was little improvement in accuracy for longer viewing conditions where participants had time to visit all locations individually. [Note: when analyzing the precision of responses, which we calculated as the standard deviation of the localization responses (i.e., the variability in responses), we observed similar results, with more variability in responses as the display numerosity increased ($F(7,1967) = 75.9$, $p < .001$,

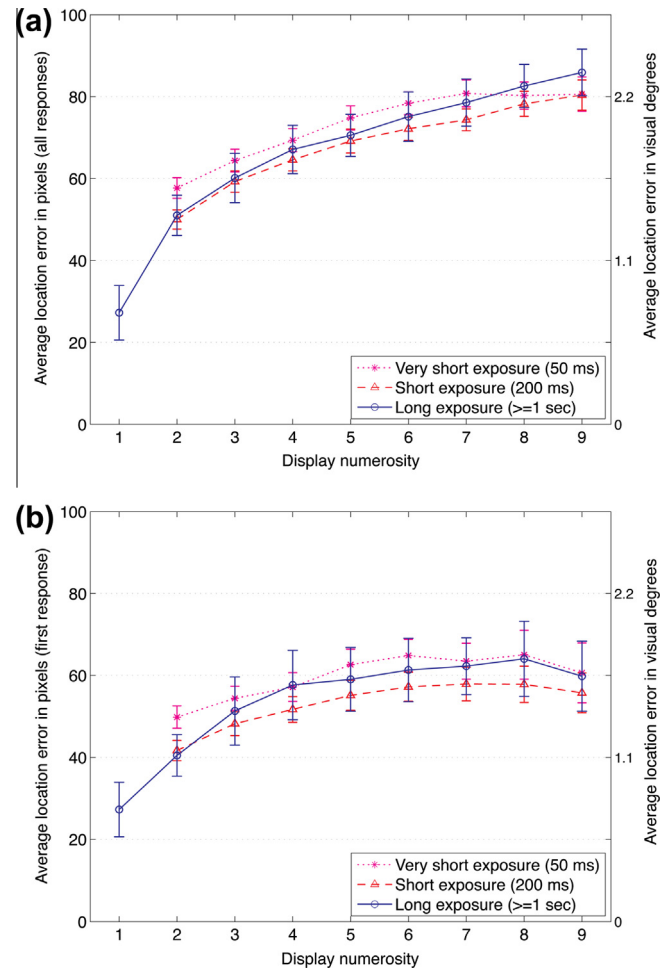


Fig. 3. Localization errors. Performance shown for (a) all responses in a trial, and also for (b) the first response made in a trial; long exposure ($N = 34$) and short/very short exposures ($N = 152$).

$\eta_p^2 = .21$), and more variability as duration decreased ($F(1,157) = 10.2$, $p = .002$, $\eta_p^2 = .06$), with no interactions.]

The magnitude of localization errors also was examined for the first response made on a display (Fig. 3b). This analysis intended to test whether similar encoding quality for all discs in a trial occurs by spreading attention among them, which would be shown by increasing errors in the first response for greater numbers because resources would be spread among all the discs within a numerosity condition. The alternate prediction is that a more local focus of attention would function serially (especially with longer durations), thus resulting in a better allocation of attention on the first object in memory and would produce flat performance for objects encoded within the proposed four-slot capacity limit. The ANOVA on the magnitude of localization errors for the first response made on the display indicated a similar trend as above, with significant main effects for numerosity ($F(7,1874) = 30.8$, $p < .001$, $\eta_p^2 = .10$) and duration ($F(1,155) = 64.5$, $p < .001$, $\eta_p^2 = .29$), with no interaction ($F(7,994) = 0.4$, $p = .92$, $\eta_p^2 = .003$). Pairwise comparisons indicated significant increases in error in the first response for the 50-ms and 200-ms conditions as numerosities increased between 2 and 3, and 4 and 5 (p 's $< .01$), after which errors leveled out; for the long durations, errors increased significantly for each display for up to 3 items (p 's $< .05$), after which the errors did not increase significantly. Additionally, pairwise comparisons among durations for each numerosity indicated less error in the longest duration than in the 50-ms displays with 2 items ($p < .001$) and 200-ms

displays with 4 items ($p < .05$), but otherwise the longest duration did not differ from the other two.

These results suggest that the longer exposure decreases localization errors minimally, and that errors tend to increase along with the number of items that needed to be encoded within the smaller numerosity conditions. Since this trend in errors was generally true even for the first response made on the screen, it suggests that the reproduction of locations for making these responses is obtained from a global representation or some type of non-independent representation of disc locations (e.g., by encoding spatial relationships instead of individual items or clusters into separate “slots”). The reduction in the magnitude of errors (or “leveling out” of errors) observed in Fig. 3b for greater numerosities may be due to some sort of density effect with larger numerosities, since the possible magnitude of localization errors would decrease when there are more stimulus discs on the display (i.e., the possible distance between a response and a stimulus disc would be reduced when there are more items on the display due to the constraints of the screen size). Nevertheless, we can still make conclusions from the increase in errors within the smaller numerosities and among the different durations.

2.2.3. Spatial compression effects

As described above, a global form of distortion was measured by determining the centroid of the discs on each of the stimulus displays and calculating the distance of each disc from that display's centroid, as well as the distances of the responses from the centroid of the response coordinates. Shorter distances from the centroid in the responses (compared to the stimulus) would indicate that participants reported disc locations as being closer to each other than they actually were. Fig. 4 plots the magnitude of these spatial compression effects as the average of the stimulus-to-centroid distances *minus* the response-to-centroid distances in a trial. A larger value on Fig. 4 indicates a greater degree of compression in participant responses. We expected that longer viewing durations would reduce spatial distortion in the responses due to better encoding into short-term memory. The hypothesis was that a better encoding of local structures with longer viewing durations would reduce the global spatial compression around the centroid.

An ANOVA on this compression measure found significant main effects for numerosity ($F(7,1718) = 43.1$, $p < .001$, $\eta_p^2 = .15$) and duration ($F(1,155) = 39.8$, $p < .001$, $\eta_p^2 = .20$), with an interaction ($F(7,998) = 4.7$, $p < .001$, $\eta_p^2 = .03$). Overall, these results suggest a

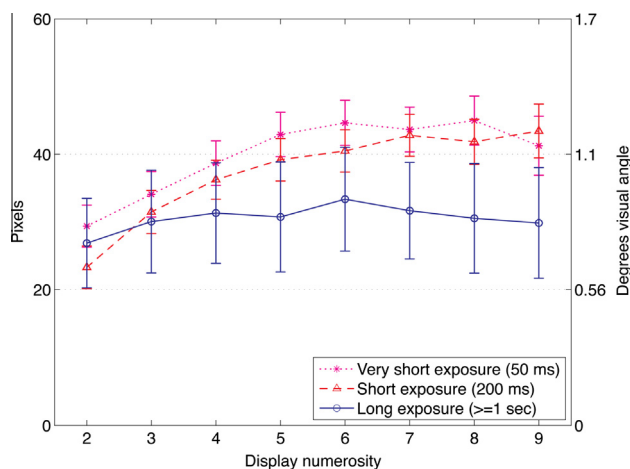


Fig. 4. Magnitude of spatial compression. Estimate of the spatial compression in responses based on the stimulus and response distances from the display and response centroids. *Note:* the values represent the average distances of the stimulus discs from the stimulus display centroid *minus* the distances of the responses from the response centroid; a larger value indicates greater compression errors.

slight reduction in compression when the stimuli were viewed for longer durations (a decrease in compression of between 5 and 15 pixels on average, or 0.1–0.3°, in the longest duration condition). Although the overall localization errors (in Fig. 3) were comparable in the short and long conditions, they were qualitatively different since performance was less compressed in the longer duration. (See Appendix Fig. B for an example of this pattern of error.) Furthermore, when taking into account the overall dispersion of the stimuli on a display (measured as the radius of the minimally enclosing circle) as a covariate in the ANOVA models, we get similar trends with significant main effects for numerosity ($F(7,1815) = 2.5$, $p = .01$, $\eta_p^2 = .01$) and duration ($F(1,156) = 34.4$, $p < .001$, $\eta_p^2 = .18$), with an interaction ($F(7,993) = 6.4$, $p < .001$, $\eta_p^2 = .04$). Although there is a reduced effect of numerosity (due to the overall spread of the stimuli), we still get an effect of duration, where longer durations reduce spatial compression regardless of the amount of dispersion of the stimuli discs. This supports the idea that items recalled from memory with shorter encoding time are more prone to a systematic global distortion (Sheth & Shimojo, 2001).

2.2.4. Grouping analysis

To examine whether or not grouping strategies were used to aid spatial memory, the discs on the stimulus displays were grouped into four regions using k -means clustering methods (for related topics, see Pothos & Chater, 2002). We used a $k = 4$ constraint based on the idea that working memory capacity is commonly said to be limited to four slots, and we hypothesized that the encoding process would likely rely on any available clusters that could be used to maximize encoding efficiency. (Furthermore, performing this analysis at $k = 3$ produced no variation in performance, with participants responding in one of the three cluster regions in 99% of trials.) For trials containing 5–9 discs, this clustering method designated a stimulus disc's membership to one of four cluster regions on a display based on the mean distances between the discs—a process that identified the discs most likely to be grouped together based on proximity. Each response was then assigned to one of these four regions by associating it with the nearest centroid of a cluster within a region. (See Appendix Fig. C for an example of this clustering procedure.) If observers encoded spatial information locally, we expect some cluster regions to be missed; if spatial information is encoded globally, then a response would be made in each cluster region even when responding with an incorrect number of items.

The ANOVA on the proportion of trials with no missed clusters (for numerosities 5–9) indicated significant main effects for numerosity ($F(4,1184) = 11.8$, $p < .001$, $\eta_p^2 = .04$) and duration ($F(1,158) = 30.0$, $p < .001$, $\eta_p^2 = .16$), with a nearly significant interaction ($F(4,632) = 2.2$, $p = .07$, $\eta_p^2 = .01$). See Fig. 5. To examine the effects of duration, pairwise comparisons were conducted by numerosity and indicated significantly more missed clusters on 50-ms displays than 200-ms displays for 5-item ($p < .05$) and 6-item displays ($p < .001$); no other differences were significant, including those for the longest viewing condition.

These results also support a global encoding of spatial locations instead of a sequential encoding of locations by cluster region: in approximately 90% of trials with 5 or more objects on the screen, there was at least one response in all four cluster regions. Responses being distributed across the display is consistent with the idea that more stimulus discs would also be distributed across space, thus more cluster regions are likely to have responses. In other words, participants noticed that something was present in each of the four cluster regions but, as we will discuss later, they did not remember the exact content within these regions. An opposite strategy favoring the deliberate grouping of objects into short-term memory for the longer duration condition would have shown a constant proportion of missed regions given that the participant would have targeted a sequential report of a constant

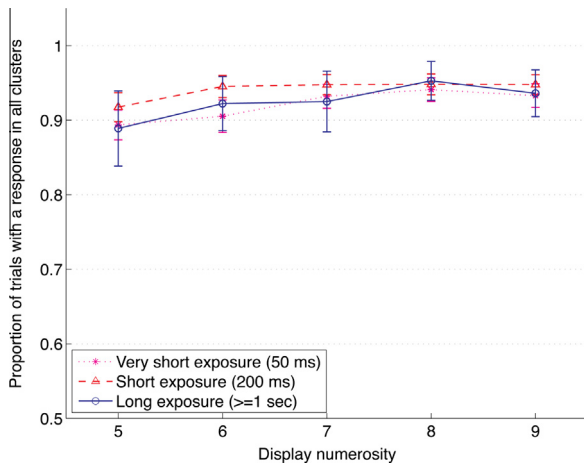


Fig. 5. Grouping analysis. Proportion of trials with a response made in each of the four cluster regions for the different exposure durations.

number of clusters based on the participant's capacity. For example, a capacity of three clusters would always miss the encoding of one cluster, no matter the numerosity (the reasoning is similar for lower capacities, where a fixed number of clusters would be left unreported no matter the numerosity). Since responses were made in all four regions in a majority of the trials (with a minor effect of duration), a local encoding of each cluster is not a likely strategy used.

3. Experiment 2

In order to confirm and extend the results from Experiment 1, we designed Experiment 2 to replicate the first experiment using a within-subjects design that directly compares performance between the short and long duration conditions. We also tested 2-cluster displays containing two clear subgroups of discs that appeared on the right and left sides of the display. Experiment 1 examined grouping effects for groups that could inadvertently appear due to the random placement of discs, and although our *k*-mean clustering method necessarily provides a separation of four clusters in such situations, it does not mean that those clusters were clearly perceived by the participants. We therefore designed more identifiable 2-cluster displays in Experiment 2 and compared performance to 1-cluster displays that minimized grouping. The 2-cluster displays were hypothesized to enhance a local encoding process that is potentially an account for the higher capacity to enumerate with the localization method. Our test conditions included two display durations that were presented in separate blocks (200-ms total or 1-s per item), with three numerosities (1, 5, and 9) and two display types (1-cluster and 2-cluster displays) that were intermixed within the duration blocks.

3.1. Methods

3.1.1. Participants

Thirty-six new participants were recruited for this experiment and provided informed consent. Twenty-nine students from the Université de Franche-Comté (UFC) participated voluntarily, and seven students from the University of Western Sydney (UWS) participated for course credit; there were no statistically significant differences in performance between these two groups of participants. The research protocol was approved by the ethics committees at both universities (in accordance with the Declaration of Helsinki).

3.1.2. Apparatus

The experiment was programmed in MATLAB using Psychophysics Toolbox (Brainard, 1997). At UFC, the stimuli were presented on a laptop running Windows XP, with a 38×22 cm LCD screen (1600×900 pixels; 60 Hz refresh). This was identical to the setup of Experiment 1. For the participants at UWS, an Apple MacBook Pro laptop was used, running OS10.7.5 (2.4 GHz Intel Core i7 with 8 GB memory), with a 33.5×21 cm LCD screen (1600×1000 pixels; 60 Hz refresh).

3.1.3. Stimuli

The test stimuli were identical to those used Experiment 1, except that only numerosities of 1, 5, and 9 were tested. These low-contrast stimuli were designed to minimize after-images and optimize the effectiveness of the subsequent random-dot texture mask that was presented for 1 s. There were two display duration conditions in Experiment 2. The discs were either presented simultaneously for a total of 200-ms (short condition) or they were presented simultaneously at a fixed duration that corresponded to the number of discs on the display at 1-s per disc, as in Experiment 1 (e.g., a display with 1 disc, which served as a baseline, was presented for 1 s, and a display with 9 discs was presented for 9 s). Again, we chose this 1-s-per-item timing since it is typically used in short-term memory experiments, and we wanted to avoid presenting low-numbered stimuli for the maximum duration to maintain participant engagement in the task.

To examine grouping effects in a more controlled manner, there were two display types in this experiment, either 1-cluster displays or 2-cluster displays (which only included the numerosities of 5 and 9 in this condition). For the 2-cluster displays, the discs appeared in two groups on each side of the display (left or right side) and were separated by a minimum of 5° (horizontally centered at the fixation cross). This created two clearly segregated sets of discs, which appeared in any number of combinations that would equal the total numerosity condition (e.g., for 5-item displays the panels were comprised of either 1 and 4, 2 and 3, 3 and 2, or 4 and 1 discs). For 1-cluster displays, the discs were scattered around the center of the screen in a manner that produced one group of discs and minimized the occurrence of subgroups. See Fig. 6 for an example of these stimulus displays. The discs were randomly placed within a $22^\circ \times 13^\circ$ (960×540 pixels) region in the center of the screen with a minimum distance of $\sim 3^\circ$ between discs to avoid crowding, as in Experiment 1. All stimuli were generated before testing and confirmed for clarity of the grouping by one of the experimenters.

3.1.4. Procedure

The procedure was identical to Experiment 1, except for the following modifications. All participants received the same two sets of stimuli that were presented in two separate blocks by stimulus presentation duration (short and long durations), which were counter-balanced for order. Within each block, the order of the numerosity and cluster conditions was randomized for each participant. The complete experimental session (consisting of 100 test

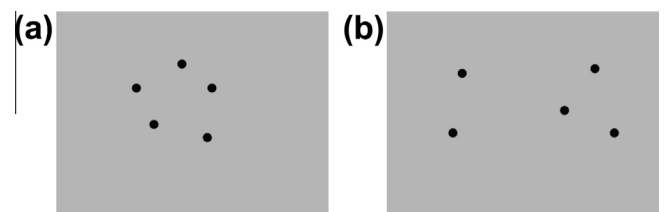


Fig. 6. Example stimuli used in Experiment 2. Panel (a) shows a 1-cluster display and panel (b) shows a 2-cluster display.

trials and 9 practice trials) lasted less than 30 min. In this experiment, 18% of trials had enumeration errors (with 86% of these errors being undercounts).

3.2. Results

3.2.1. Enumeration accuracy

We expected that longer viewing durations would allow better enumeration performance, with 2-cluster displays improving performance under longer viewing conditions if grouping strategies were used to enhance performance. A within-subjects 2 (cluster) by 3 (numerosity) by 2 (duration) ANOVA on the enumeration accuracy measure indicated no significant main effect for the number of clusters, but there were main effects for numerosity ($F(2,70) = 132.6$, $p < .001$, $\eta_p^2 = .79$), and duration ($F(1,35) = 174.4$, $p < .001$, $\eta_p^2 = .83$), with a three-way interaction ($F(1,35) = 11.3$, $p = .002$, $\eta_p^2 = .25$), indicating better performance for longer durations and smaller numerosities. See Fig. 7. We do not detail the repeated-measures ANOVA that we performed to examine differences among the different conditions (they were all significant including interactions), but we will mention the most interesting result from the pairwise comparisons (Bonferroni corrected). For 5-item displays, the short 2-cluster condition was significantly worse than the short 1-cluster condition ($p = .03$) and both the long 1-cluster and 2-cluster conditions ($p = .001$). This decrease in accuracy for short 2-cluster 5-item displays might be because the five discs were better encoded under brief presentations when they formed a single cluster due to the proximity of the discs when they formed only one cluster. This result might be due to the effect of the increased viewing eccentricity on 2-cluster displays, which might reduce the ability to detect items in the periphery under brief presentation durations (e.g., see Palomares et al., 2011).

3.2.2. Localization accuracy

Again, we expected that longer viewing durations would improve localization performance, with 2-cluster displays possibly improving performance under longer viewing conditions. The ANOVA results on average localization errors (for correctly enumerated trials) indicated a significant main effect for the number of clusters ($F(1,36) = 120.0$, $p < .001$, $\eta_p^2 = .77$), numerosity ($F(2,71) = 148.5$, $p < .001$, $\eta_p^2 = .81$), and duration ($F(1,35) = 44.6$, $p < .001$, $\eta_p^2 = .56$), with no interactions. Pairwise comparisons found significantly larger localization errors in the short compared to the long durations in both the 1-cluster and 2-cluster conditions

(p 's $< .05$). The 2-cluster displays showed significantly higher errors than the 1-cluster displays for each duration condition (p 's $< .05$). Also, there was a significant increase in error as numerosity increased (p 's $< .05$). See Fig. 8a.

The ANOVA for the first response (Fig. 8b) indicated a similar trend in localization errors, with a significant main effect for the number of clusters ($F(1,36) = 23.4$, $p < .001$, $\eta_p^2 = .39$), numerosity ($F(2,71) = 32.0$, $p < .001$, $\eta_p^2 = .47$), and duration ($F(1,35) = 15.5$, $p < .001$, $\eta_p^2 = .30$), with no interactions. The pairwise comparisons for this measure were not as clear, with the only significant trend being the effect of numerosity in the long duration condition, with errors increasing as numerosity increases (p 's $< .05$). The only significant effect of duration was in the 1-cluster 1-item displays ($p < .01$), with longer durations producing more accurate localization than the short duration. These results again suggest that spatial information for multiple objects tends to be encoded quickly and globally. First, longer presentation durations only decreased localization errors for 1-cluster displays. Second, the results do not support the idea that spatial information from more groupable sets of discs is better encoded, since there were systematically more errors in the 2-cluster displays, particularly under short viewing durations. Again, this supports the view that the encoding of spatial information into memory is more characteristic of a global and flexible resource.

3.2.3. Spatial compression effects

We examined the spatial compression effects for both 1-cluster and 2-cluster displays. This was computed using the same method

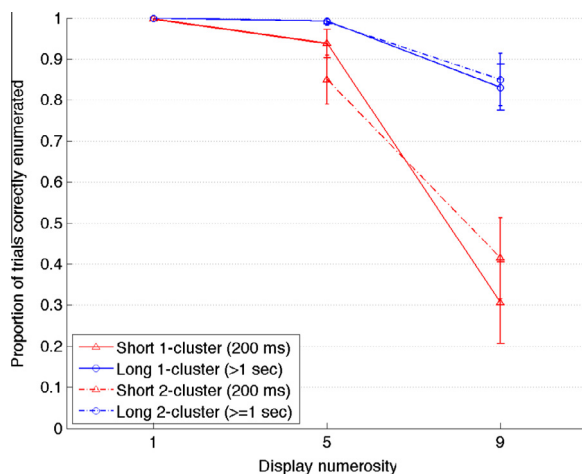


Fig. 7. Enumeration accuracy. Proportion of trials with correct number of responses in Experiment 2; short and long exposures ($N = 36$).

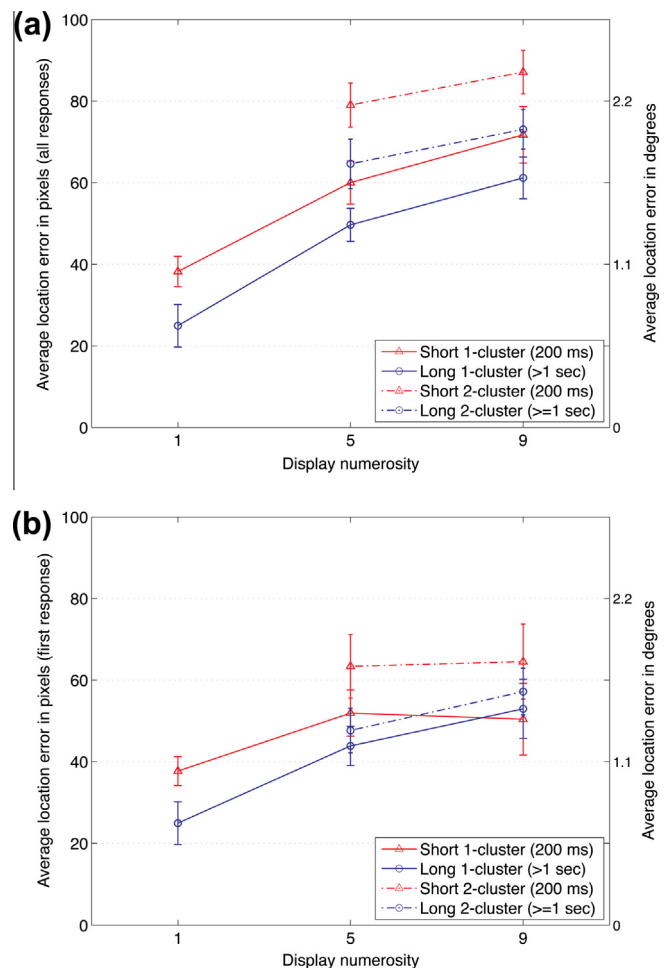


Fig. 8. Localization errors. Performance shown for (a) all responses in a trial, and for (b) the first response made in a trial.

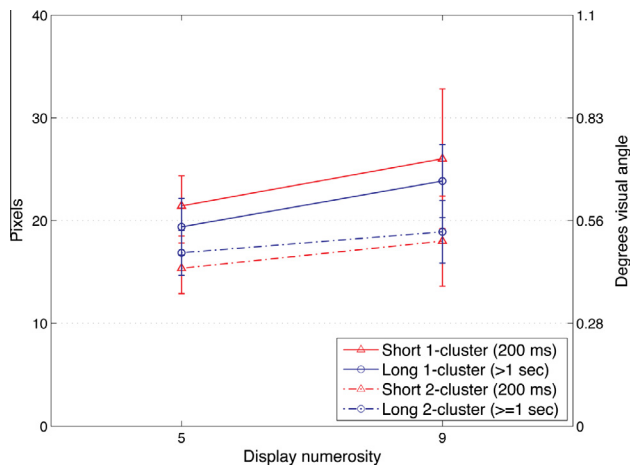


Fig. 9. Magnitude of spatial compression. Spatial compression is based on the computation of the distance of each stimulus disc from the display centroid *minus* the distance of each response from that centroid (the centroids are computed separately for each cluster of discs on 2-cluster displays); a larger value indicates greater overall spatial compression.

as in Experiment 1 except that in this analysis two centroids were computed for 2-cluster displays (one for each side of the display). A larger value in Fig. 9 indicates greater compression in participant responses. Again, we expected less compression under longer viewing durations and two-cluster displays (indicating that local encoding was prioritized). The ANOVA on this compression measure found a significant main effect for the number of clusters ($F(1,40) = 49.9$, $p < .001$, $\eta_p^2 = .56$) and numerosity ($F(1,38) = 9.8$, $p = .003$, $\eta_p^2 = .21$), but no effect of duration or interactions. Pairwise comparisons indicate slightly less compression in the 2-cluster displays compared to the 1-cluster displays (p 's $< .05$) and less compression in 5-item displays when compared to the 9-item displays of the same condition ($p < .05$), but this advantage is small (with reductions of roughly 0.2 – 0.5°).

These results indicate that extended viewing durations do not significantly reduce compression effects, but 2-cluster displays may help in producing a less spatially compressed representation overall (see Appendix Fig. B for similar errors). It may be possible that such errors are based on a summary statistic computation that tends to produce overall shifts in localization (as would be detected as translation errors), but the most relevant distortion is that of spatial compression in our data overall. No effect of duration was observed here, unlike in Experiment 1, although the trend toward higher errors in the short duration is present in the 1-cluster condition, which was significantly higher than the other 2-cluster conditions (p 's $< .01$).

4. General discussion

The present study focused on whether or not there are different processes for memorizing spatial information under short and long exposure durations—a test that we thought could explain the higher subitizing capacity when using a location-based reporting method (Haladjian & Pylyshyn, 2011). We therefore focused our analyses on the memory for spatial information from stimuli presented for long durations that allowed participants enough time to encode the number of objects into short-term memory, and we compared these data to the results from more rapid stimuli presentations. Specifically, our analyses examined how grouping can affect the number of items that can be remembered and the accuracy of spatial information, which could indicate whether encoding is local or global. The rationale was that a global form of perceptual grouping allows items to be aggregated into a larger

structure. An alternative prediction was that a more local form of perceptual grouping allows items to be grouped into smaller separate groups, which may improve localization performance. Our hypothesis was that the long exposure durations would enhance a sequential encoding of the available groups (based on a more lossless encoding process of the different groups, even if relative imprecision was expected within groups) and thus result in more accurate memorization of spatial information, especially for the first groups encoded into the capacity-limited working memory slots. Also, the faster presentation durations may encourage a global encoding of spatial information resulting in more systematic errors typical of a lossy compression process, which would suggest that spatial information is encoded in a “snapshot”. Spatial compression was used as a proxy to a global form of encoding of spatial information under the hypothesis that several local spatial compressions (for instance, one per cluster) would cancel out a global form of compression around the display's centroid.

The results from Experiment 1 indicate that the encoding of spatial information occurs globally and quickly (by 200 ms), but benefits little from the extended exposure to the stimulus. Although we found that spatial information was more prone to a global distortion with shorter durations (cf. Section 2.2.3), the longer viewing duration in this study, which provided ample opportunity to encode locations into short-term memory, did not improve localization performance when compared to the results from displays with shorter viewing durations. In Experiment 2, we only found a significant improvement of localization accuracy in the longer viewing duration for 1-item displays, with no other substantial improvement in spatial memory under longer viewing durations. Experiment 2 also found no improvement in enumeration or localization accuracy on displays with two clear groups of objects, which suggests that local grouping strategies are not systematically used to enhance spatial memory, or if so, they do not prove efficient. When looking at systematic spatial transformations (e.g., scaling, translations, or rotational shifts), the primary effect was that of spatial compression, which supports previous studies that found similar biases (e.g., Sheth & Shimojo, 2001).

These results suggest that there is no substantive optimization of spatial memory for longer durations that *a priori* offers more time for the effortful encoding of spatial information locally (that would allow a lossless memory compression process). Participants tend to encode global spatial properties and not individual items or individual groups of discs—even when the number of items to recall was below the capacity of short-term memory (around four, if we refer to the theories mentioned in the introduction). This implies that the resource for processing spatial information is distributed across all items rather than divided into slots dedicated to encoding a few items more precisely, which may be indicative of a mechanism without a fixed slot-based capacity but rather with flexible shared resources (e.g., Franconeri, Alvarez, & Cavanagh, 2013). As a result, we only observed a lossy global compression effect. If each subgroup was assigned a memory slot in a more lossless manner, there should be no performance difference (i.e., finding ceiling effects, regardless of the degree of accuracy within subgroups) within the capacity of short-term memory—contrary to our results that indicate an increase in localization errors for displays with small numerosities (1–4 items). Since this localization error is even present in the first response made on a display (similar to the errors found in the last response in the study by Gorgoraptis et al., 2011), it is likely that spatial memory relies on a flexibly allocated resource whose quality will be affected overall by the total number of items that must be remembered. This limited resource is distributed according to task demands and the type or amount of information that is being processed (see Ma, Husain, & Bays, 2014). Furthermore, our results also support the idea that the visual memory span limit observed in the Corsi

block task (Kessels et al., 2000) depends on factors other than merely encoding spatial locations serially in short-term memory (e.g., Della Sala et al., 1999; Gmeindl, Walsh, & Courtney, 2011; Page & Norris, 1998), such as remembering both the temporal order and the locations of the items. Such results question the validity of Corsi blocks or equivalent tests that purport to measure spatial short-term memory processes (see Colom et al., 2006; Unsworth & Engle, 2007). The extra time that is given in the Corsi block task in comparison to our fast presentation times might only be detecting the ability to use conscious verbal strategies that interact with the visual memorization of spatial information.

In terms of participant strategies to facilitate object encoding, we did not find evidence that the randomly-placed discs on the stimulus displays in Experiment 1 were clustered into subsets and encoded as separate groups into short-term memory. When dividing the stimulus displays into four regions using *k*-means clustering (to match the proposed four “slots”), it was evident that participants tended to report that something was present in all regions of the display, even when making errors as to the precise number of items present within each region. The results indicate that there is no grouping effect for either short or long viewing durations, but rather there is a global encoding that is influenced by the overall spread of discs on a display and susceptible to spatial compression (see Sheth & Shimojo, 2001). Since the stimuli were designed to avoid crowding by maintaining a distance of at least 3° between discs (Bahcall & Kowler, 1999; Intriligator & Cavanagh, 2001), as opposed to previous manipulations in which the items were organized in accordance with Gestalt grouping principles (e.g., Woodman, Vecera, & Luck, 2003), we cannot attribute crowding as the reason for participants missing some items within regions. Even in Experiment 2 with clear 2-cluster displays (below the fixed slot capacity of working memory), there was no benefit of grouping to the overall localization accuracy, although there was a slight reduction in spatial compression.

A possible limitation of the current stimuli is that since they were designed to prevent crowding in Experiment 1, the presence of groupable subsets of objects was limited and thus grouping strategies would not be beneficial. In Experiment 2, however, we created 2-cluster displays to test grouping strategies for instances when the groups were well under the purported capacity limit of memory slots, and we found that displays with two clear groups of discs did not improve enumeration accuracy substantially or the encoding of spatial information. Furthermore, although the detection of individual items (i.e., enumeration accuracy) might not be affected greatly when the items appeared in the periphery, the accuracy of localization might be affected on 2-cluster displays since the viewing eccentricity pushes more objects into the periphery (Palomares et al., 2011). Another possible reason why localization accuracy may have not improved in this study is due to the lack of landmarks to aid spatial memory, which have been shown to increase localization accuracy in previous studies (e.g., Lee, Shusterman, & Spelke, 2006). This may also account for the increase in localization error found with each additional response made, as there is no stable frame of reference to constrain or bias such errors, for example, as the tendency to mislocalize items away from boundaries (Huttenlocher, Hedges, & Duncan, 1991). The borders of our display screens, however, could have been used as such a frame of reference, but we placed the stimuli at least 5° away from the edges to minimize this possibility (see Huttenlocher et al., 2004). Versions of this experiment using constant landmarks to guide localization may reveal further useful information about how information is encoded into spatial memory under varied presentation durations.

Based on the fixed capacity and the flexible resource views, there were two possible predictions for the resulting localization performance in this study. Regarding the view where there is a memory limit of four “slots”, memory accuracy would have shown

a plateau before fixed-item capacity limits are exceeded. Such a result would support the hypothesis that each perceptual group is counted as a discrete item and that a lossless compression process occurs to encode these perceptual groups (e.g., Anderson, Vogel, & Awh, 2013), especially if the participant is allowed enough time to memorize the display. This would produce localization errors that are similar for each response made for stimuli within the capacity limits (again, this does not imply that the lossless encoding of the subgroups leads to a perfect report of the locations within groups, but only that constant magnitudes of errors would be expected within groups across the available span). Alternatively, the flexible resource view would predict that accuracy progressively decreases as a function of the number of stored items (including groups), since memory resources are thought to be distributed across an “unlimited” number of items. Consequently, this encoding would suffer from a more lossy form of compression process and would especially be evident in the shorter exposure durations due to the lack of time to encode the discs and their locations. In other words, a rapid stimulus presentation is likely to drive a more lossy compression process, whereas longer durations should drive a more lossless compression process.

Overall, the current localization results indicating a lossy compression process do not support a fixed slot theory of short-term memory for spatial memory because the encoding of spatial information is not clearly allocated for each viewed item or groups of items, but rather encoded on a more global scale. This decrease in spatial accuracy based on memory load indicates a shared limited resource that is applied to all items in memory (e.g., see Gorgoraptis et al., 2011). That localization accuracy seems to be near optimal in as little as 200 ms suggests that an overall snapshot of locations is extracted quickly and is used as the primary guide for localization. This supports the idea of hierarchical encoding of scene features (e.g., Brady & Alvarez, 2011), where the location of objects is a global property that is encoded first, separate from individual features.

5. Conclusion

The results from the current study suggest that the greater subitizing range observed in our previous studies (Haladjian & Pylyshyn, 2011; Haladjian et al., 2010) is not due to grouping strategies that could have been used to enhance information processing capacity. Additionally, it appears that even when given plenty of time to study a stimulus array, spatial information is encoded globally in a “snapshot” and not likely encoded separately for each individual item (or each group) into a limited number of slots in short-term memory. The observed lossy-type encoding compression errors support the idea that object locations are not encoded in representations independently of each other (see Brady & Alvarez, 2011). Two-cluster displays may help reduce spatial compression errors, as shown in Experiment 2, but little other benefit from clearly groupable stimuli was observed. Since no evidence for grouping was found, further research is required to determine what characteristics of object arrays may help encode more items during localization, which appears to rely on a mechanism with flexible resources.

Acknowledgments

This research was supported by a grant from the Agence Nationale de la Recherche (Grant # ANR-09-JCJC-0131-01) awarded to Fabien Mathy. Part of this research was conducted during the summer of 2011 at the Université de Franche-Comté thanks to a postdoctoral research grant awarded to Harry H. Haladjian by the Université de Franche-Comté.

Appendix

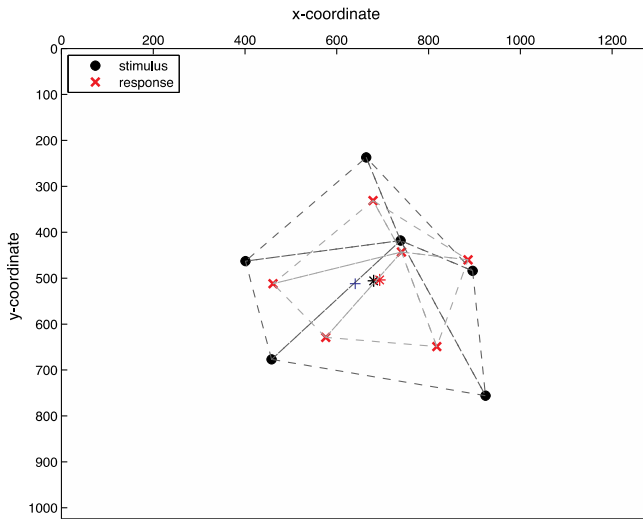


Fig. A. Example of stimulus locations and responses from an actual trial, with compression effects. The dark circles correspond to the stimulus discs and the lighter crosses correspond to the responses (the x and y axes correspond to screen dimensions). The connections between the circles or crosses are the segments identified from the Delaunay triangulation procedure. As this image shows, the distances between the responses are closer to each other than the distances between the stimulus discs, which indicates a spatial compression. The asterisks in the center of the screen correspond to the centroids (color-coded). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

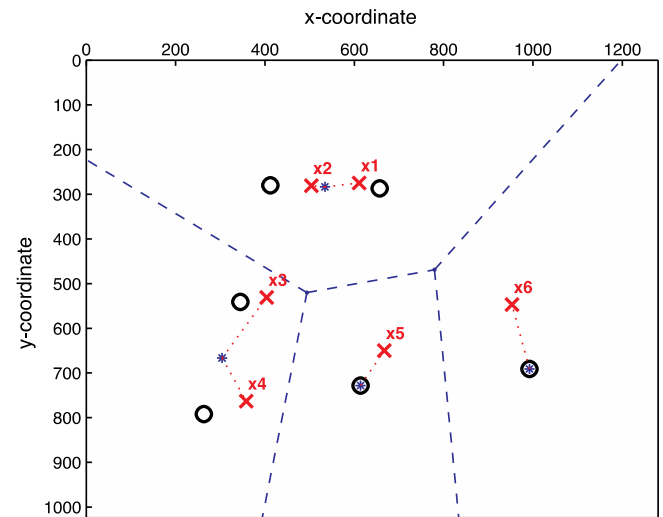


Fig. C. Example of regions created by the k -means clustering for an actual trial in Experiment 1 (where $k = 4$). The black circles represent the locations of the stimulus discs on a display. The lighter crosses correspond to the participant's responses. To determine whether or not a response was made within a region, each response was paired with the nearest centroid of the cluster in a region (designated by an asterisk); the dotted lines indicate to which region centroid a response was linked. Note that regions containing only one disc will share the exact same disc and centroid location. In this trial, a response was made in all four regions.

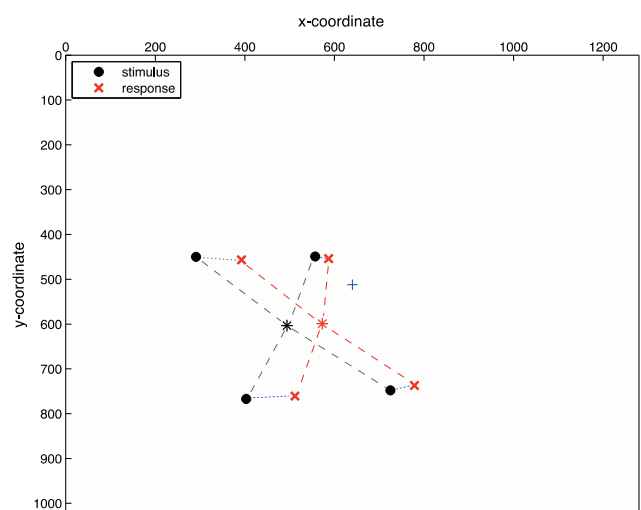
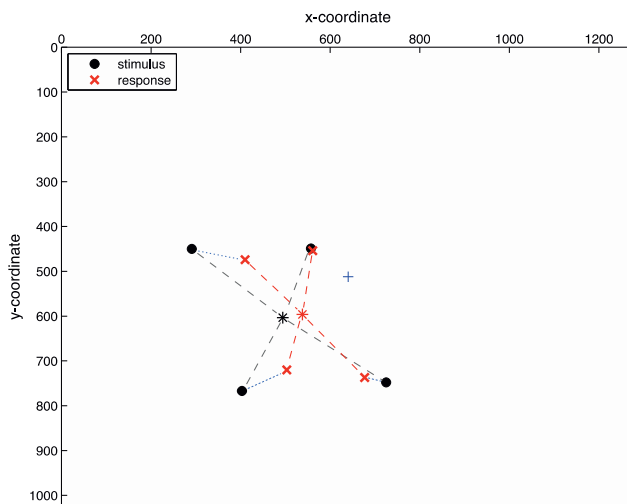


Fig. B. Example of compression errors versus localization errors. The dark dots are the locations of the stimulus discs and the dark asterisk is the centroid of the stimulus discs. The lighter crosses correspond to the participant's responses and the lighter asterisk is the centroid of these responses. The dashed line to the centroid represents the distance used for the compression measure. The dotted lines between the stimulus and response pairs represent the localization error. The left figure shows errors with compression (from an actual trial), and the right figure shows an example of hypothetical responses with the same magnitude of localization errors but without any compression. The errors in the right figure shift all responses in one direction (to the right) instead of shifting toward the centroid in several directions as seen in the left figure.

References

- Aben, B., Stapert, S., & Blokland, A. (2012). About the distinction between working memory and short-term memory. *Frontiers in Psychology*, 3, 301. <http://dx.doi.org/10.3389/fpsyg.2012.00301>.
- Aksentijević, A., Elliott, M. A., & Barber, P. J. (2001). Dynamics of perceptual grouping: Similarities in the organization of visual and auditory groups. *Visual Cognition*, 8(3–5), 349–358. <http://dx.doi.org/10.1080/13506280143000043>.
- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, 15(2), 106–111. <http://dx.doi.org/10.1111/j.0963-7214.2004.01502006.x>.
- Anderson, D. E., Vogel, E. K., & Awh, E. (2013). Selection and storage of perceptual groups is constrained by a discrete resource in working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 39(3), 824–835. <http://dx.doi.org/10.1037/a0030094>.
- Bahcall, D. O., & Kowler, E. (1999). Attentional interference at small spatial separations. *Vision Research*, 39(1), 71–86. [http://dx.doi.org/10.1016/S0042-6989\(98\)00090-X](http://dx.doi.org/10.1016/S0042-6989(98)00090-X).
- Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, 9(10), 7, 1–11. <http://dx.doi.org/10.1167/9.10.7>.
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, 321(5890), 851–854. <http://dx.doi.org/10.1126/science.1158023>.
- Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia*, 49(6), 1622–1631. <http://dx.doi.org/10.1016/j.neuropsychologia.2010.12.023>.
- Berch, D. B., Krikorian, R., & Huha, E. M. (1998). The Corsi block-tapping task: Methodological and theoretical considerations. *Brain and Cognition*, 38(3), 317–338. <http://dx.doi.org/10.1006/brcg.1998.1039>.
- Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science*, 22(3), 384–392. <http://dx.doi.org/10.1177/0956797610397956>.
- Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, 138(4), 487–502. <http://dx.doi.org/10.1037/a0016797>.
- Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, 11(5), 4. <http://dx.doi.org/10.1167/11.5.4>.
- Brady, T. F., & Tenenbaum, J. B. (2013). A probabilistic model of visual working memory: Incorporating higher order regularities into working memory capacity estimates. *Psychological Review*, 120(1), 85–109. <http://dx.doi.org/10.1037/a0030779>.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. <http://dx.doi.org/10.1163/156856897X00357>.
- Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision*, 10(6), 1–10. <http://dx.doi.org/10.1167/10.6.20>.
- Colom, R., Rebollo, I., Abad, F. J., & Shih, P. C. (2006). Complex span tasks, simple span tasks, and cognitive abilities: A reanalysis of key studies. *Memory & Cognition*, 34(1), 158–171. <http://dx.doi.org/10.3758/BF03193395>.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–114. <http://dx.doi.org/10.1017/S0140525X01373922> (discussion 114–185).
- Cutini, S., & Bonato, M. (2012). Subitizing and visual short-term memory in human and non-human species: A common shared system? *Frontiers in Psychology*, 3, 469. <http://dx.doi.org/10.3389/fpsyg.2012.00469>.
- De Lillo, C. (2004). Imposing structure on a Corsi-type task: Evidence for hierarchical organisation based on spatial proximity in serial-spatial memory. *Brain and Cognition*, 55(3), 415–426. <http://dx.doi.org/10.1016/j.bandc.2004.02.071>.
- Dehaene, S., & Cohen, L. (1994). Dissociable mechanisms of subitizing and counting: Neuropsychological evidence from simultanagnosic patients. *Journal of Experimental Psychology: Human Perception and Performance*, 20(5), 958–975. <http://dx.doi.org/10.1037/0096-1523.20.5.958>.
- Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: A tool for unwinding visuo-spatial memory. *Neuropsychologia*, 37(10), 1189–1199. [http://dx.doi.org/10.1016/S0028-3932\(98\)00159-6](http://dx.doi.org/10.1016/S0028-3932(98)00159-6).
- Donkin, C., Nosofsky, R. M., Gold, J. M., & Shiffrin, R. M. (2013). Discrete-slots models of visual working-memory response times. *Psychological Review*, 120(4), 873–902. <http://dx.doi.org/10.1037/a0034247>.
- Dry, M. J., Preiss, K., & Wagemans, J. (2012). Clustering, randomness, and regularity: Spatial distributions and human performance on the traveling salesperson problem and minimum spanning tree problem. *The Journal of Problem Solving*, 4(1) (Article 2).
- Feldman, J. (1999). The role of objects in perceptual grouping. *Acta Psychologica*, 102(2–3), 137–163. [http://dx.doi.org/10.1016/S0001-6918\(98\)00054-7](http://dx.doi.org/10.1016/S0001-6918(98)00054-7).
- Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision*, 11(12). <http://dx.doi.org/10.1167/11.12.3>.
- Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: Competitive content maps for attention and memory. *Trends in Cognitive Sciences*, 17(3), 134–141. <http://dx.doi.org/10.1016/j.tics.2013.01.010>.
- Franconeri, S. L., Alvarez, G. A., & Enns, J. T. (2007). How many locations can be selected at once? *Journal of Experimental Psychology: Human Perception and Performance*, 33(5), 1003–1012. <http://dx.doi.org/10.1037/0096-1523.33.5.1003>.
- Gmeindl, L., Walsh, M., & Courtney, S. M. (2011). Binding serial order to representations in working memory: A spatial/verbal dissociation. *Memory & Cognition*, 39(1), 37–46. <http://dx.doi.org/10.3758/s13421-010-0012-9>.
- Goodale, M. A., & Milner, A. D. (2004). *Sight unseen: An exploration of conscious and unconscious vision*. Oxford: Oxford University Press.
- Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2), 285–339.
- Gorgoraptis, N., Catalao, R. F. G., Bays, P. M., & Husain, M. (2011). Dynamic updating of working memory resources for visual objects. *The Journal of Neuroscience*, 31(23), 8502–8511. <http://dx.doi.org/10.1523/JNEUROSCI.0208-11.2011>.
- Haladjian, H. H., & Pylyshyn, Z. W. (2011). Enumerating by pointing to locations: A new method for measuring the numerosity of visual object representations. *Attention, Perception, & Psychophysics*, 73(2), 303–308. <http://dx.doi.org/10.3758/s13414-010-0030-5>.
- Haladjian, H. H., Singh, M., Pylyshyn, Z. W., & Gallistel, C. R. (2010). The encoding of spatial information during small-set enumeration. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd annual conference of the cognitive science society* (pp. 2839–2844). Austin, TX: Cognitive Science Society.
- Hunter, W. S., & Sigler, M. (1940). The span of visual discrimination as a function of time and intensity of stimulation. *Journal of Experimental Psychology*, 26(2), 160–179. <http://dx.doi.org/10.1037/H0057548>.
- Huttenlocher, J., Hedges, L. V., Corrigan, B., & Crawford, L. E. (2004). Spatial categories and the estimation of location. *Cognition*, 93(2), 75–97. <http://dx.doi.org/10.1016/j.cognition.2003.10.006>.
- Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, 98(3), 352–376. <http://dx.doi.org/10.1037/0033-295X.98.3.352>.
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, 43(3), 171–216. <http://dx.doi.org/10.1006/cogp.2001.0755>.
- Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 683–702. <http://dx.doi.org/10.1037/0278-7393.26.3.683>.
- Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkman, J. (1949). The discrimination of visual number. *American Journal of Psychology*, 62(4), 498–525. <http://dx.doi.org/10.2307/1418556>.
- Kessels, R. P., van Zandvoort, M. J., Postma, A., Kappelle, J., & de Haan, E. H. (2000). The Corsi block-tapping task: Standardization and normative data. *Applied Neuropsychology*, 7(4), 252–258. http://dx.doi.org/10.1207/S15324826AN0704_8.
- Korjoukov, I., Jeurissen, D., Kloosterman, N. A., Verhoeven, J. E., Scholte, H. S., & Roelfsema, P. R. (2012). The time course of perceptual grouping in natural scenes. *Neural Science*, 23(12), 1482–1489. <http://dx.doi.org/10.1177/0956797612443832>.
- Lee, S. A., Shusterman, A., & Spelke, E. S. (2006). Reorientation and landmark-guided search by young children: Evidence for two systems. *Psychological Science*, 17(7), 577–582. <http://dx.doi.org/10.1111/j.1467-9280.2006.01747.x>.
- Lewandowsky, S., Oberauer, K., Yang, L.-X., & Ecker, U. K. (2010). A working memory test battery for MATLAB. *Behavior Research Methods*, 42(2), 571–585. <http://dx.doi.org/10.3758/BRM.42.2.571>.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–281. <http://dx.doi.org/10.1038/36846>.
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, 17(3), 347–356. <http://dx.doi.org/10.1038/nn.3655>.
- Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology*, 111(1), 1–22. <http://dx.doi.org/10.1037/0096-3445.111.1.1>.
- Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition*, 122(3), 346–362. <http://dx.doi.org/10.1016/j.cognition.2011.11.003>.
- Melcher, D., & Piazza, M. (2011). The role of attentional priority and saliency in determining capacity limits in enumeration and visual working memory. *PLoS ONE*, 6(12), e29296. <http://dx.doi.org/10.1371/journal.pone.0029296>.
- Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97. <http://dx.doi.org/10.1037/h0043158>.
- Page, M. P., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, 105(4), 761–781. <http://dx.doi.org/10.1037/0033-295X.105.4.761>.
- Palomares, M., & Egeth, H. (2010). How element visibility affects visual enumeration. *Vision Research*, 50(19), 2000–2007. <http://dx.doi.org/10.1016/j.visres.2010.07.011>.
- Palomares, M., Smith, P. R., Pitts, C. H., & Carter, B. M. (2011). The effect of viewing eccentricity on enumeration. *PLoS ONE*, 6(6), e20779. <http://dx.doi.org/10.1371/journal.pone.0020779>.
- Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition*, 121(1), 147–153. <http://dx.doi.org/10.1016/j.cognition.2011.05.007>.
- Pothos, E. M., & Chater, N. (2002). A simplicity principle in unsupervised human categorization. *Cognitive Science*, 26(3), 303–343. [http://dx.doi.org/10.1016/S0364-0213\(02\)00064-2](http://dx.doi.org/10.1016/S0364-0213(02)00064-2).

- Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, 32(1), 65–97. [http://dx.doi.org/10.1016/0010-0277\(89\)90014-0](http://dx.doi.org/10.1016/0010-0277(89)90014-0).
- Revsin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, 19(6), 607–614. <http://dx.doi.org/10.1111/j.1467-9280.2008.02130.x>.
- Richardson, J. T. E. (2007). Measures of short-term memory: A historical review. *Cortex*, 43(5), 635–650. [http://dx.doi.org/10.1016/S0010-9452\(08\)70493-3](http://dx.doi.org/10.1016/S0010-9452(08)70493-3).
- 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(2), 324–330. <http://dx.doi.org/10.3758/s13423-011-0055-3>.
- Sargent, J., Dopkins, S., Philbeck, J., & Chichka, D. (2010). Chunking in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(3), 576–589. <http://dx.doi.org/10.1037/a0017528>.
- Sheth, B. R., & Shimojo, S. (2001). Compression of space in visual memory. *Vision Research*, 41(3), 329–341. [http://dx.doi.org/10.1016/S0042-6989\(00\)00230-3](http://dx.doi.org/10.1016/S0042-6989(00)00230-3).
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, 138(4), 628–654. <http://dx.doi.org/10.1037/a0027473>.
- Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance*, 19(2), 331–351. <http://dx.doi.org/10.1037/0096-1523.19.2.331>.
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, 101(1), 80–102. <http://dx.doi.org/10.1037/0033-295X.101.1.80>.
- Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, 133(6), 1038–1066. <http://dx.doi.org/10.1037/0033-2909.133.6.1038>.
- Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, 131(1), 48–64. <http://dx.doi.org/10.1037/0096-3445.131.1.48>.
- Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, 4, 1120–1135. <http://dx.doi.org/10.1167/4.12.11>.
- Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, 10(1), 80–87. <http://dx.doi.org/10.3758/BF03196470>.
- Xu, Y. (2002). Limitations of object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), 458–468. <http://dx.doi.org/10.1037/0096-1523.28.2.458>.
- Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, 24(3), 295–340. [http://dx.doi.org/10.1016/0010-0285\(92\)90010-Y](http://dx.doi.org/10.1016/0010-0285(92)90010-Y).
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453(7192), 233–235. <http://dx.doi.org/10.1038/nature06860>.