Perception & Psychophysics
2007, 69 (6), 904-912

Creation of visual long-term memory

Danko Nikolić and Wolf Singer
Max Planck Institute for Brain Research, Frankfurt am Main, Germany
and Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe University, Frankfurt, Germany

Currently, it is not known by which mechanisms novel visual representations are stored in long-term memory (LTM). Here we report evidence that visual working memory (WM) plays an important role in the formation of visual LTM. By varying exposure times and perceptual difficulty of the stimuli, we find that the rate-limiting factor constraining storage in LTM is the amount of information that can be simultaneously kept in WM, whereas the time needed to store this information into LTM is constant irrespective of the size of the WM content. These results support the hypothesis that visual WM serves as a gate for the storage of information into LTM.

The concept of working memory (WM) was introduced as a replacement for short-term memory in order to emphasize the role of information processing—that is, the "working" component of WM—and to deemphasize the formation of long-term memory (LTM) (Baddeley, 1992). Visual WM can store visual information for a short period of time, and the mechanisms underlying encoding and the maintenance of information in visual WM have been studied extensively (Alvarez & Cavanagh, 2004; Luck & Vogel, 1997; Phillips, 1983; Wheeler & Treisman, 2002; Woodman & Luck, 2004). Visual information can be stored also in visual LTM (Phillips & Christie, 1977), but the mechanisms responsible for this storage have not been identified. For verbal material, formation of LTM involves both the resources of attention and those of WM (Greene, 1987; Kane, Hambrick, & Conway, 2005), and it is important to know whether the same is true for storage of visual spatial information. In the present study, we investigated the hypothesis that the time needed to store a given amount of information in LTM depends on how much of this information can be kept simultaneously in WM.

We used a change-detection paradigm with which we could manipulate the amount of information that could be kept in visual WM. Participants were presented with a sample array consisting of two types of elements that were distributed randomly in a two-dimensional matrix (e.g., $\vee$ and $\times$). After an interstimulus interval of 1 sec, this sample was followed by a test array that was identical to the sample, except that one element was missing. Participants needed to identify the type of the missing element (Figure 1). To solve this task, participants had to remember only the locations of one type of element (target) because they could infer that the missing element must have belonged to the type that was not memorized (distractors) if all of the remembered elements were present.

In one condition, the two types of elements differed from each other by color or shape, so that they produced perceptual *pop-out* (i.e., preattentive detection), allowing for an effortless segregation between targets and distractors. In another condition, the targets did not pop out (i.e., $\vee$ vs. $\wedge$). In this case, the detection of targets and their locations was difficult, requiring focused attention and slow serial search (Treisman & Gelade, 1980) (Figure 2A). According to Alvarez and Cavanagh (2004), these two perceptual conditions are associated with different capacities of visual WM. In Experiment 1, we exploited this effect to control and to measure the amount of information that can be retained in WM for the two conditions. In Experiment 2, we required participants to memorize matrices, the size of which exceeded the capacity of WM. By comparing the results from Experiments 1 and 2, we were able to determine whether the time needed to memorize the large stimuli in Experiment 2 could be accounted for by the amount of information retainable in WM that was determined in Experiment 1. Finally, in Experiment 3, we investigated whether the participants actually stored information in LTM when they memorized large stimuli in Experiment 2.

EXPERIMENT 1

Method

Participants and Apparatus. In this experiment we assessed the capacity of WM as a function of the perceptual difficulty of the stimulus. Five female and 2 male participants served in this experiment. Two participants were paid, and the others volunteered. Four different types of stimuli were used; these differed in the degree of pop-out of the array elements. In Condition 1, the elements did not produce perceptual pop-out ($\vee$ and $\wedge$). In the other conditions, the cues that permitted preattentive detection of the elements consisted of color (green $\vee$ and red $\wedge$, Condition 2) or shape differences (e.g., $\vee$ and $\times$, Condition 3) or a combination of the two cues (green $\vee$ and red $\times$, Condition 4) (Figure 2A).

Procedure. The sample stimulus was presented for a period of 1 sec, and the size of the array changed adaptively at each trial and according to the participant's performance: After each correct re-

D. Nikolić, danko@mpih-frankfurt.mpg.de

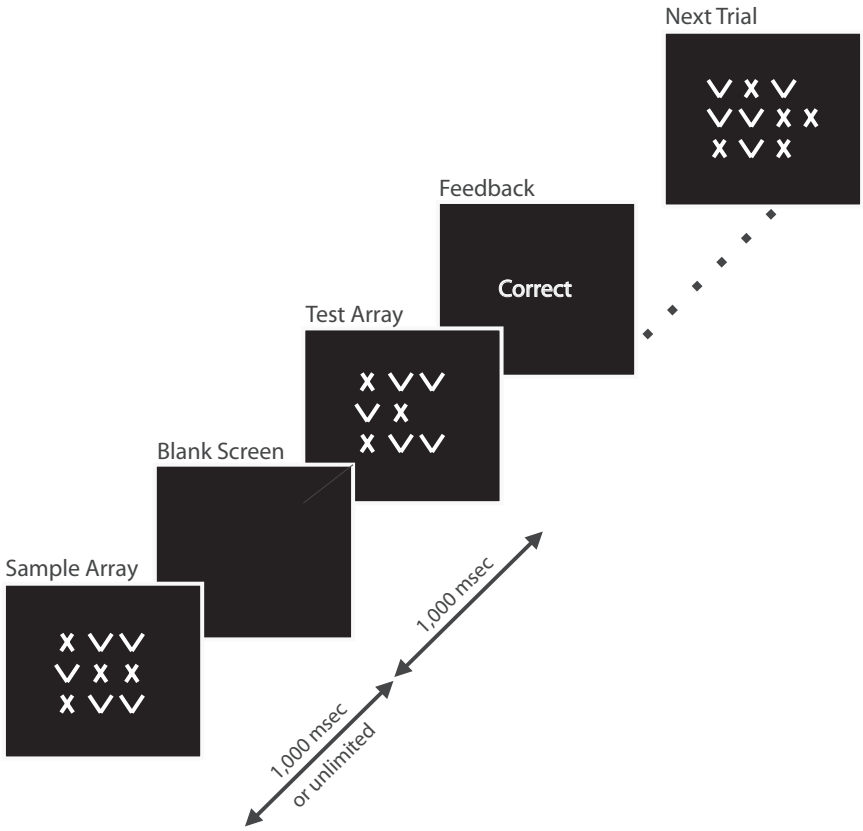

Figure 1. The paradigm used in the present study. Sample arrays consisted of two types of elements and were generated randomly in each trial. Participants needed to identify the missing elements in the test array. To investigate the capacity of visual working memory for these stimuli, the arrays were presented for 1,000 msec in Experiment 1, and to allow near perfect accuracy, participants studied the arrays as long as needed in Experiments 2 and 3. The array size changed adaptively in Experiment 1, which is illustrated by an increase in the array size in the next trial following a correct response.

sponse, the array size increased by one item in the next trial of the corresponding stimulation condition, and after each incorrect response, it decreased by one item.

Participants performed one session with 150 trials per condition, starting with a small array of seven elements. In the pilot study, preceding Experiment 1, participants felt that the task was much easier if they knew which stimulus condition they had to expect next. Thus, in order to provide optimal conditions for encoding, different experimental conditions were presented in a fixed order.

Results and Discussion

The analyzed variables were the growth functions of the stimulus array (Figure 2B). In order to estimate the capacity of WM, we first investigated whether the number of remembered elements was constant across different array sizes. For this purpose, we adapted a probabilistic model relating participants' accuracy, the number of presented elements, and the probability of giving the correct response by chance (Pashler, 1988; see Appendix A). The accuracy of performance, as reflected by the proportion of correct trials in a session, was well below perfect (Condition 1: 57%; Condition 2: 63%; Condition 3: 64%; and Condition 4: 64%) but always above chance level (binomial probabilities, all p values $< .05$, $N = 150$). An ANOVA indicated that the accuracy differed significantly for at least one of the four conditions [$F(3,24) = 3.44$, $p = .033$]. Subsequent paired t tests revealed that only Condition 1 (no pop-out) differed significantly from the others (all t values > 3.15; all p values $< .01$, $df = 6$), whereas there were no differences in accuracy between the various pop-out conditions (Conditions 2–4) (all t values < 1.38; all p values $> .24$, $df = 6$) (see Figure 2B). Therefore, the amount of memorized information depended on the presence of pop-out but not on the number of pop-out inducing cues.

The probabilistic model gave a good fit to the data when it was based on the assumption that, for a particular stimulus type, participants retained always a constant number of elements across different array sizes. The value of the only free parameter in the model, that fitted the data best, was used as the estimate of the number of locations of target elements stored in WM (see Method). This analysis was made by pooling the results from Stimulation Conditions 2–4 and separately for Condition 1 (i.e., pop-out vs. non-pop-out). The results were somewhat counterintuitive, revealing that participants could achieve relatively large sizes of arrays (Figure 2B) even if they remembered the

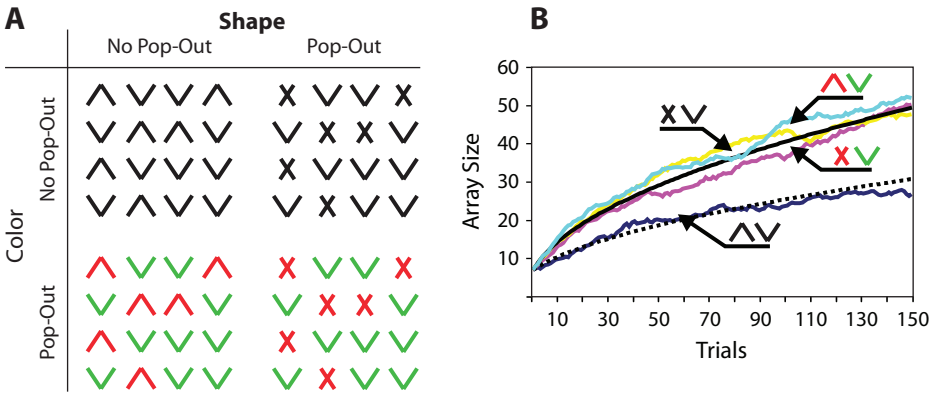

Figure 2. Stimuli and results from Experiment 1, in which we investigated the capacity of visual working memory. (A) The four perceptual conditions in Experiment 1 differed in the degree of perceptual pop-out between the elements. (B) The growth of array size during 150 trials averaged for 7 participants (colored lines). At the beginning of the experiment, the size of the array was always seven elements. Black lines: Theoretical growth of array sizes if participants remembered a fixed number of 4.1 (full line) and 1.4 (dotted line) locations of target elements and guessed otherwise.

locations of only a small number of elements. In Condition 1 (without pop-out), the estimated number of stored locations was 1.4 and in the pop-out conditions it was 4.1 (full and dotted black lines in Figure 2B; standard errors of the fit were 2.2 elements^2 for Condition 1 and 5.7 elements^2 for Conditions 2–4). Thus, when the target elements could be detected directly, the capacity of visual WM corresponded to the typically reported capacity of about four items (Cowan, 2001; Irwin & Andrews, 1996; Luck & Vogel, 1997; Pylyshyn & Storm, 1988; Wheeler & Treisman, 2002), and when the detection of elements required serial search, the number of memorized target locations was drastically reduced. These two estimates were used later in Experiment 2 to investigate whether the capacity of WM predicts the ability of participants to store visual information into LTM.

The narrow WM capacity in the pop-out conditions is in agreement with the experiments on change blindness that also demonstrate that memories for details of complex visual scenes are poor despite the subjective experience of a good perceptual grasp of the stimulus (O'Regan, Rensink, & Clark, 1999). The massive drop of this capacity to only 1.4 items in the non-pop-out condition agrees well with the reports that WM capacity decreases with the increase in the complexity of the stimuli (Alvarez & Cavanagh, 2004) or with the difficulty of categorizing the stimuli (Olsson & Poom, 2005). The narrow WM capacity in the non-pop-out condition is also consistent with the notion that WM cannot be loaded sequentially by adding more locations of targets during serial search. Shifting of the spotlight of attention erases the spatial information previously stored in WM (Awh & Jonides, 2001; Oh & Kim, 2004; Woodman & Luck, 2004).

One important reason for the poor memory performance in the present experiment was the presence of distractors. If stimuli do not include distractors and if participants simply have to memorize the locations of *all* elements, the storage capacity of WM is considerably larger than in the present experiment (Brockmole, Wang, & Irwin, 2002; Hollingworth, Hyun, & Zhang, 2005; Jiang, Olson, & Chun, 2000; Phillips, 1983; Prabhakaran, Narayanan, Zhao, & Gabrieli, 2000; Wheeler & Treisman, 2002). However, it is not immediately clear why distractors also decreased memory performance in the pop-out condition. One possibility is that the distractors prevent strategies such as figural grouping or chunking of the targets, which enhance the capacity of visual WM (Hollingworth et al., 2005; Woodman, Vecera, & Luck, 2003).

EXPERIMENT 2

Method

Participants and Apparatus. In this experiment, participants had the possibility to overcome the limited capacity of WM by storing the information into LTM. They were granted unlimited encoding time and were instructed to achieve near perfect accuracy. Of the 3 female and 3 male participants, 3 were paid for participation, and 2 of them participated in both Experiments 1 and 2. The stimuli were the same as in Experiment 1, but participants signaled the end of the encoding phase with a keypress and the time they needed for encoding was used as a dependent variable. This design is analogous to visual search tasks in which performance is almost always perfect and perceptual strategies are investigated solely on the basis of response times. We fixed the sizes of the arrays at 10, 15, 20, and 25 elements. This required remembering on average 5, 7.5, 10, and 12.5 locations, which exceeded the capacity of WM determined in Experiment 1. Again, we used two sets of elements that required either serial search (∨ and ∧, Condition 1) or produced pop-out due to different shapes (∨ and ×, Condition 2). Each of 6 participants performed two sessions on different days (for individual participants, separated by 1, 4, 6, 7, 11, and 29 days).

Procedure. Participants were instructed to use any strategy that was visual and that would enable them to achieve the required performance, and they were instructed to signal the end of the encoding phase with a keypress, which terminated the sample presentation time. If performance was below 95% for a particular condition, a message was delivered prior to the next trial of the same condition that indicated the type of the next condition and the current performance level (e.g., "Next condition: 4, Performance: 86%"). This message appeared repeatedly in all subsequent trials of the respec-

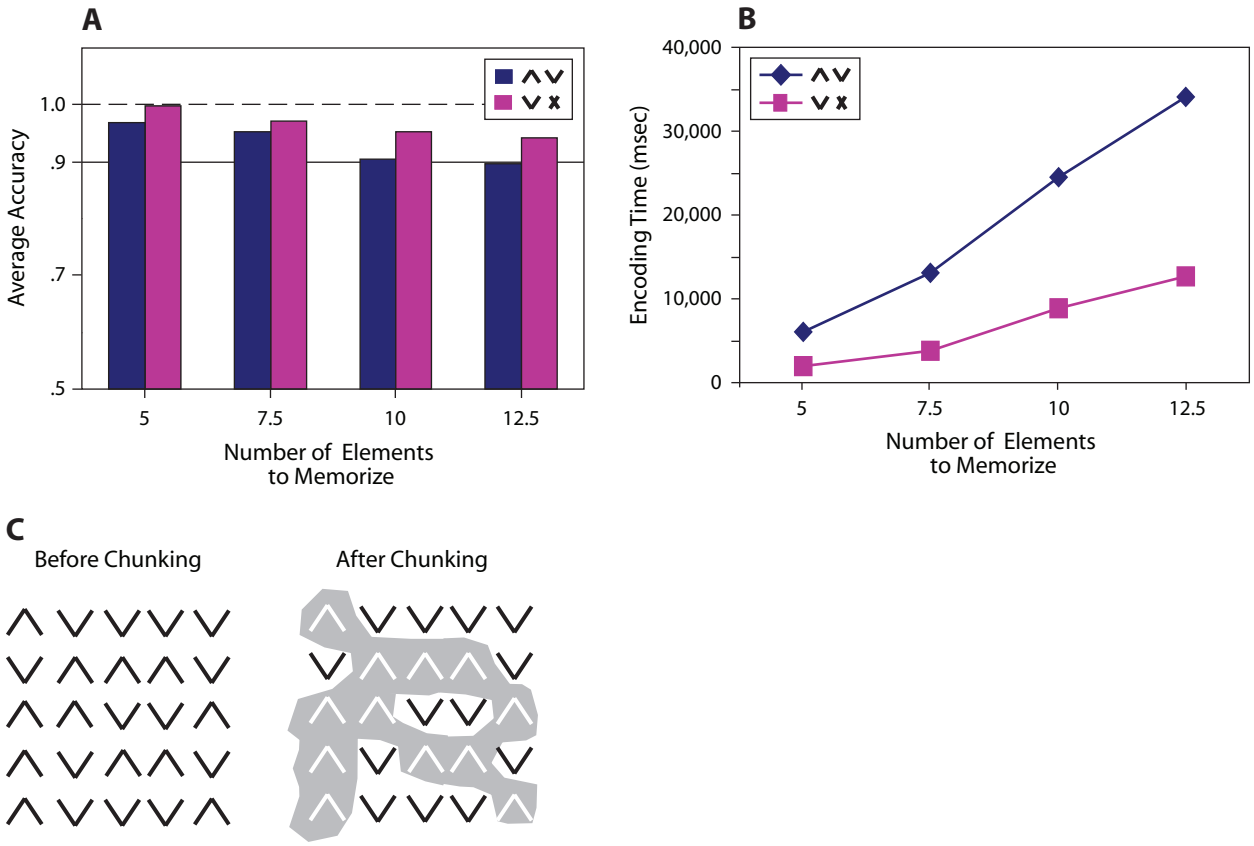

Figure 3. Results from Experiment 2, in which participants memorized large arrays. (A) Accuracy, expressed as proportion correct, for stimuli of different sizes under two different perceptual conditions (chance level: .5). (B) Average encoding times for the same conditions as in (A). (C) Illustration of the participants' strategy to build a figure connecting elements of one class only. If this figure was unchanged, participants could infer that the missing element must have been part of the ungrouped "background" elements.

tive condition until the accuracy reached the 95% level. Participants were also instructed to perform as accurately as possible (i.e., accuracy was emphasized over processing time) and were told that the feedback on low performance before a trial indicated that they had to improve accuracy in this condition by "working harder" and spending more time if necessary. The first 3 participants who took part in Experiment 2 were given 50 trials per condition, and we learned from their responses that 25 trials would have been sufficient, which then became the number of trials given to the remaining 3 participants. The order of conditions was fully randomized. The experiment was followed by an in-depth debriefing procedure in which we obtained information about the participants' experiences and the strategies that they used to cope with the task.

One possibility to solve this task was to simply compare the counts of the elements of one type during the study and the test. We did not mention the counting possibility in the instructions, and we interviewed the participants about this strategy at the end of the experiment. According to participants' responses in the interview, none of them discovered this possibility during the experiment.

Results and Discussion

Participants successfully achieved the main objective of the task; their performance increased to 90% correct or higher in all conditions and for all array sizes (Figure 3A). Thus, their ability to retain the locations of elements was considerably higher than in Experiment 1. Debriefing revealed that participants needed to invest considerable effort when the size of the array increased and, in particular, when the identification of target elements required serial search. According to these subjective reports, all participants used visual strategies and all tried several other strategies before finding the one that worked well. Apparently, this successful strategy was the same for all participants, for both perceptual conditions, and for all array sizes. Participants focused on only one class of elements (the targets) and grouped the selected elements into a figure (or object) by sequentially adding more and more targets until they had bound all elements into a single figure. According to the participants' reports, they could perceive at the end of the encoding phase the shape of the entire figure even if the stimulus did not include perceptual pop-out. This complete percept of the stimulus was their criterion to press the button and to start with the test phase (see Figure 3C and Appendix B for an illustration of the strategy). In the present study, we did not investigate the validity of these subjective reports by independent experimental tests but investigated their chosen encoding times in considerable detail.

Encoding times. In agreement with the reported strategy, the times needed to memorize stimuli increased linearly with the array size in both the pop-out and non-pop-out condition (linear fit, $R^2 = .995$ and $R^2 = .975$ in

Conditions 1 and 2, respectively) (Figure 3B). These times were long and amounted to an average of about 30 sec in the most difficult condition. To investigate the encoding times further, we computed the slopes of the fitted linear functions. With pop-out (Condition 2), participants needed on average 1,345 msec to integrate one additional element into the figure (assuming that participants operated only on one type of element), and this time increased to 3,781 msec per element in arrays lacking perceptual pop-out [$t(5) = 4.37$, $p < .01$]. Thus, in the condition with pop-out, the stimuli were memorized about 2.8 times faster than without pop-out (see Appendix C for details on quantitative methods). Therefore, neither the speed with which stimuli were memorized nor the difference in the speed between the two perceptual conditions (i.e., 2,436 msec/element) could be explained by typical visual search times, which are in the range of 100 msec/element (Treisman & Gelade, 1980) and thus differ by at least an order of magnitude.

Stable interindividual differences. The linear functions describing the encoding times exhibited substantial interindividual differences (coefficient of variation; CV = .99 for Condition 1 and .79 for Condition 2), but for individual participants they remained remarkably stable across sessions (Pearson's $r = .90$, $N = 6$, $p < .01$, and $r = .97$, $N = 6$, $p < .01$, for Conditions 1 and 2, respectively). Moreover, the speeds of memorization with and without pop-out correlated strongly across participants ($r = .88$, $N = 6$, $p < .05$), indicating that despite the large interindividual differences, the lack of pop-out affected the memorization rate of each participant by a similar proportion. This implies that the individual differences in the speed of memory formation are independent of the perceptual conditions.

Relation to WM capacity. The changes in memorization speed associated with the pop-out and non-pop-out condition closely matched the corresponding changes in the capacity of WM that were found in Experiment 1. The 2.8-fold increase in the memorization speed that occurred with pop-out in Experiment 2 matched closely the 2.9-fold increase in the WM capacity that occurred with pop-out in Experiment 1 (4.1 vs. 1.4 elements). This suggests that the rate-limiting factor for formation of memories is the amount of information that can be kept simultaneously in WM, whereas the time needed to form these memories is constant and independent of the size of the WM content. Thus, if the speed of memorization is normalized against the number of target locations that can be stored in WM, the time needed to memorize the respective content of WM is practically identical for the two perceptual conditions; with pop-out, 5,293 msec (3,781 msec/element × 1.4 elements), and without pop-out, 5,514 msec (1,345 msec/element × 4.1 elements), which amounts to a difference of less than 5% (see Appendix C for more detailed quantitative analysis of encoding times). This suggests that there is a close relation between the amount of information available to WM and the time needed to memorize an array whose size exceeds the capacity of WM.

One interesting interpretation of these results is that all of the contents that are simultaneously available in WM can be memorized in a single step (i.e., in parallel) and that serial repetition of such steps is required if the number of elements to be encoded exceeds the capacity of WM. For example, for 12 elements, about 3 memorization steps would be required if 4 locations can be stored in WM, whereas 8 transfer steps would be required if only about 1.5 locations can be stored in WM.

Semantic strategies and depth of processing. Formation of LTM can be facilitated remarkably by the extraction of semantic relations at deeper levels of processing, in which case participants form novel chunks in LTM by relying on expert knowledge and on skills that develop with extensive practice (Chase & Ericsson, 1981; Chase & Simon, 1973; Craik & Lockhart, 1972; Gobet & Simon, 1998). Such quick chunking did not take place in the present experiment, because our participants formed memories slowly, were able to form such memories already at the very beginning of the experiment (i.e., after only a few test trials), and did not report using any semantic strategies. In addition, practice had no effect on the ability of our participants to create memories, as indicated by a nonsignificant difference in the encoding times between the first and the second sessions of the experiment [one-way ANOVA; pop-out, $F(1,10) = 0.30$, $p = .60$; no pop-out, $F(1,10) = 0.20$, $p = .67$]. Therefore, the memorization strategy used in the present task is likely to be based exclusively on visual resources.

Individual differences in integration speed. The interindividual differences in the speed with which participants memorized the present stimuli could not be explained by differences in strategies because the participants' reports obtained during debriefing were very similar. In addition, both the slowest and the fastest participant in the study agreed that the strategy illustrated in Appendix B described well their own experiences. The interindividual differences in memorization speed could not be explained by a trade-off between the speed and the accuracy either, because the participants who were faster in forming memories tended also to be more accurate. The only possible predictor for understanding the reason for the interindividual differences was participant's age, because it correlated negatively (but not significantly) with the integration speed ($r = -.6$, $N = 6$).

EXPERIMENT 3

Method

Participants and Apparatus. In this experiment, we investigated whether the memories for visual stimuli, created in Experiment 2, involved LTM. Five female and 3 male volunteers performed 10 trials identical to those in Experiment 2 with pop-out by color and shape (green ∨ and red ×), followed by 10 trials of an unexpected recognition test probing LTM. This test included five old and five new arrays, and participants needed to indicate those that had been studied previously. Arrays had either 25 elements (on average 12.5 elements to be remembered) or 8 or 9 elements (4 or 4.5 elements to be remembered; 4 participants in each condition) (Figure 4A). The arrays were always created randomly, and each location had the same probability (50%) of being occupied by one of the elements. This made the sequentially presented arrays mutually independent. The generated set of small arrays needed to be prescanned for exact replicas, which sometimes occurred due to the relatively small total number of combinations.

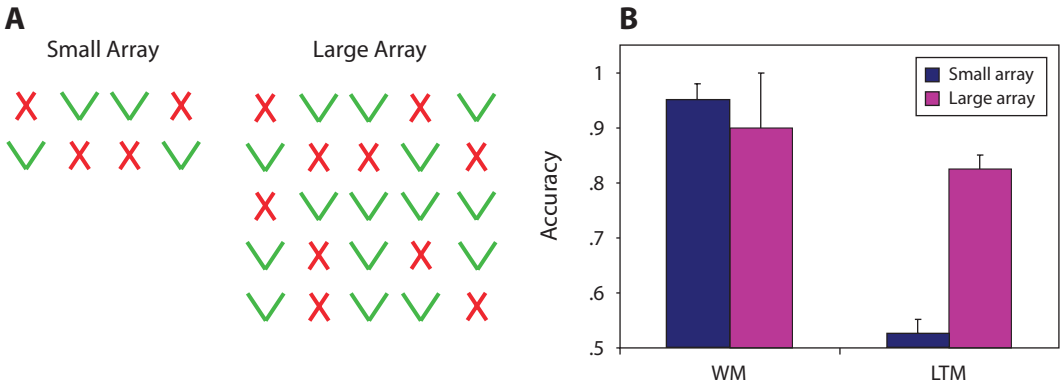

Figure 4. Stimuli and results from Experiment 3, in which we investigated whether the stimuli were stored in long-term memory (LTM). (A) Participants memorized either a small array in which the number of elements did not exceed the capacity of visual working memory (WM) or a large array that exceeded the capacity of WM. (B) Proportion of correct responses for recall of the missing element with 1,000-msec interval between study and test phase (WM) and for recognition of old arrays among new ones with a longer delay that included interference from processing other stimuli (LTM) (chance level: .5). Vertical bars: Standard errors of measurement.

Procedure. Each subject was presented with the same set of arrays, which were presented in a pseudorandom sequence such that at least 6 and on average 9.4 other arrays were shown between presentation of the study and the test array, producing considerable pro- and retroactive interference between the memories of the shapes. The average delay between stimulus presentation and recognition test was about 10 min.

Results and Discussion

The memory performance at the immediate test (1-sec delay) was 90% correct or higher in both conditions. Seven participants reported memorizing large arrays sequentially, and 1 reported memorizing those arrays in parallel. In the later LTM test, all participants who memorized large arrays performed well above chance (binomial probabilities: all $ps < .05$; $N = 10$) and average performance was 83% correct. However, in the LTM tests for the smaller arrays that did not exceed the capacity of WM, participants performed at chance in the later memory test (all $ps > .05$, $N = 10$) with average performance of 53% correct (Figure 4B).

The similarity between the spatial configurations was necessarily much higher in the condition with small rather than large arrays. Thus, the poor recognition of small arrays does not necessarily indicate lack of LTM but could instead reflect the high degree of interference between memories (Murnane & Shiffrin, 1991). Nevertheless, the results indicate that large arrays are stored in LTM. Hence, these results suggest that in Experiment 2 participants needed to rely on LTM in order to achieve memory performance that exceeded the capacity of WM. Thus, the time required for a memorization step can be interpreted as the time required to transfer to LTM the contents kept in WM.

GENERAL DISCUSSION

In the present study, we showed first that with the presence of distractors, the number of items that can be accessed simultaneously and stored in parallel in visual WM is highly limited and, consistent with previous reports (Alvarez & Cavanagh, 2004; Olsson & Poom, 2005), stimulus dependent. Maximum capacity of about 4 target locations was achieved in the pop-out condition, but it dropped to about 1.4 locations when participants had to engage in serial search for the identification of the target elements. In further experiments, we showed that the capacity of visual WM is a good predictor of how fast the same type of information will be stored into visual LTM. The results indicated that the rate limiting factor for the formation of chunks in LTM is not the transfer time from WM to LTM but the capacity of WM. The information seemed to be transferred from WM to LTM with constant speed, which is participant specific and which, in our setting, required on average about 5 sec per one full load of WM. Therefore, the present results indicate that visual WM plays an important role for the formation of visual LTM, a conclusion that is consistent with the reports on the relation between WM and LTM for verbal materials (Greene, 1987; Kane et al., 2005). Our results also suggest the possibility that, when visual LTM needs to be formed and the size of the stimulus exceeds WM capacity, transfer of information from WM to LTM occurs iteratively. Thus, the overall storage time might depend on the number of iterations—that is, the number of items to be stored divided by the number of items that can be kept simultaneously in WM.

The present results are in agreement with the previous findings that the capacity of WM can be expanded by chunking the stimuli on the basis of information stored in LTM (Chase & Ericsson, 1981; Chase & Simon, 1973; Craik & Lockhart, 1972; Gobet & Simon, 1998; Kemps, 2001; Miller, 1956). However, we have no direct evidence that chunking-based integration occurred in the present study. Although participants' subjective reports were consistent with the chunking hypothesis, additional experiments would be needed to prove this claim and to exclude the possibility that the target locations are stored independently.

The long time that our participants needed to form LTM should be expected, because fast encoding that occurs with natural stimuli (Hollingworth, 2004) relies on exper-

tise (Ericsson & Kintsch, 1995), extraction of semantic information, and deeper levels of processing (Chase & Ericsson, 1981; Chase & Simon, 1973; Craik & Lockhart, 1972; Gobet & Simon, 1998). Although naive participants are usually experts for natural images, they could not be experts for the stimuli used in this study. For that reason, the task forced our participants to rely exclusively on low-level visual processing, which is apparently slow in creating novel LTMs.

The long period of about 5 sec, which is needed to transfer information from visual WM to visual LTM, suggests that the "working" component of WM also played an important role in the present task. It would be interesting to know what other mechanisms are engaged during the formation of visual LTM. Substantial evidence suggests close interaction between visual WM and selective visual attention (Alvarez & Cavanagh, 2004; Awh & Jonides, 2001; Luck & Vogel, 1997; Oh & Kim, 2004; Wheeler & Treisman, 2002; Woodman & Luck, 2004). Since selective visual attention plays an important perceptual function for binding elementary visual features (Treisman, 1982; Treisman & Gelade, 1980), one interesting question is whether this function also plays a role in integrating visual information in LTM.

AUTHOR NOTE

This work was partially supported by Alexander von Humboldt-Stiftung and by the Hertie Foundation. The authors thank Bill Phillips, Dragutin Ivanec, Denis Bratko, Ruxandra Sireteanu, Nelson Cowan, David Linden, and Arthur N. Lebedev for help, advice, and fruitful discussions. Correspondence concerning this article should be addressed to D. Nikolić, Max Planck Institute for Brain Research, Deutschordenstr. 46, D-60528 Frankfurt am Main, Germany (e-mail: danko@mpih-frankfurt.mpg.de).

REFERENCES

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**, 106-111.

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, **5**, 119-126.

Baddeley, A. (1992). Working memory. *Science*, **255**, 556-559.

Brockmole, J. R., Wang, R. F., & Irwin, D. E. (2002). Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 315-334.

Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 141-189). Hillsdale, NJ: Erlbaum.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral & Brain Sciences*, **24**, 87-114.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, **11**, 671-684.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, **102**, 211-245.

Gobet, F., & Simon, H. A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory*, **6**, 225-255.

Greene, R. L. (1987). Effects of maintenance rehearsal on human memory. *Psychological Bulletin*, **102**, 403-413.

Hollingworth, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 519-537.

Hollingworth, A., Hyun, J. S., & Zhang, W. W. (2005). The role of visual short-term memory in empty cell localization. *Perception & Psychophysics*, **67**, 1332-1343.

Irwin, D. E., & Andrews, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 125-155). Cambridge, MA: MIT Press.

Jiang, Y. H., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 683-702.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, **131**, 66-71.

Kemps, E. (2001). Complexity effects in visuo-spatial working memory: Implications for the role of long-term memory. *Memory*, **9**, 13-27.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, **63**, 81-97.

Murnane, K., & Shiffrin, R. M. (1991). Interference and the representation of events in memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 855-874.

Oh, S. H., & Kim, M. S. (2004). The role of spatial working memory in visual search efficiency. *Psychonomic Bulletin & Review*, **11**, 275-281.

Olsson, H., & Poom, L. (2005). Visual memory needs categories. *Proceedings of the National Academy of Sciences*, **102**, 8776-8780.

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of "mudsplashes." *Nature*, **398**, 34.

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, **44**, 369-378.

Phillips, W. A. (1983). Short-term visual memory. *Philosophical Transactions of the Royal Society of London: Series B*, **302**, 295-309.

Phillips, W. A., & Christie, D. F. M. (1977). Components of visual memory. *Quarterly Journal of Experimental Psychology*, **29**, 117-133.

Prabhakaran, V., Narayanan, K., Zhao, Z., & Gabrieli, J. D. (2000). Integration of diverse information in working memory within the frontal lobe. *Nature Neuroscience*, **3**, 85-90.

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

Treisman, A. [M.] (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 194-214.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97-136.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, **131**, 48-64.

Woodman, G. F., & Luck, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review*, **11**, 269-274.

Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, **10**, 80-87.

APPENDIX A
The Model for Estimation of WM Capacity

For the modeling of performance functions, we assumed that the probability to give a correct response, P_c, depends on the number of correctly stored elements, N, relative to the total number of elements, S (e.g., $S = 7$ at the beginning of Experiment 1), and on the probability of giving the correct response by guessing, P_g. This results in the following equation for the probability of giving the correct response, $P_c = N/S + P_g (1 - N/S)$. Since there were only two classes of elements, the probability of giving a correct response by guessing was $P_g = .5$. S is half the array size, because participants needed to memorize only half of the elements for correct performance. This model is related to that of Pashler (1988) but assumes that participants use a guessing strategy that is optimal for the present task; that is, if the missing element is not part of the memorized figure, they are not guessing on the basis of the whole array but choose the element of the nonmemorized class. The expected change in array size in a single trial $E\{\Delta S\}$ could be expressed by $E\{\text{Increase}\} + E\{\text{Decrease}\}$, which leads to $E\{\Delta S\} = 2P_c - 1$. It could be shown that with $P_g = .5$ the expected change in array size $E\{\Delta S\} = N/S$. If N is a constant—that is, the same number of elements is remembered regardless of the size of the array—the larger the array in Experiment 1, the smaller $E\{\Delta S\}$ (see the resulting curves in Figure 2B). This model has one free parameter, N, and the value of this parameter that fits the data best can be used to estimate the number of elements stored in WM. Note that if participants are simply guessing (without having any memory), the array size will, in expectation, stay constant.

APPENDIX B
Illustration of the Participants' Chunking Strategy

The following text illustrates the participants' strategies for memorizing larger arrays in Experiments 2 and 3. The text was prepared by 1 participant whose results are not reported (author D.N.) and who made a recollection of his experiences after the completion of the experiment. Thus, the text is an illustration of the experiences rather than an accurate excerpt from an ongoing experimental protocol. In reality, the process is much faster than suggested by these descriptions. Nevertheless, the present text was approved independently by 2 other participants as representing accurately their own experiences. These 2 participants were naive in respect to the hypotheses investigated and the interpretations made in the present study.

> There are so many elements! I could choose to remember the red ones because they somehow seem to make a nicer form. Wait a moment, when I pay attention to the green ones they also seem to make a nice form. Well, I'll take the red elements this time. Ok, let's see. This part here somehow attracted my attention because it seems to be a nice round shape so I could start from here. Now I am looking at it, perceiving clearly the individual elements together with the shape that the elements form and I feel confident that I would remember the position of each of them. But there are a few additional red elements down there that I clearly would not remember because although I see them I do not clearly perceive the shape. Let me pay more attention to them. Now I see clearly the nice "Y" shape. I would remember this, but now I lost hold of my round shape. Let me get back to the round shape. Ok, there it is. I got hold of it quicker this time. Now, how do I get these two parts into one big figure? As I try to switch attention towards the "Y" shape I lose hold of the round shape. But I feel I will get it if I try harder. Let me try again. Hey, now, it is not as bad as the first time. Now, I can partially cover the "Y" shape and still keep the round shape. Let me do this once again. Get the round shape, . . . ok, . . . now slowly widen the attention, . . . yes, that's it, I'm not losing the round shape, . . . now move some more, . . . I feel the big shape is coming, . . . yes, that's it, I got it! I see clearly the entire shape with all details and I feel confident that I'll remember the position of every single red element. I press the button. Wait for one second. . . . Now it is very easy to give the correct response.

(*Continued on next page*)

APPENDIX C
Quantification of the Encoding Time

The present findings suggest that the total time needed for creation of a chunk, T, can be expressed quite accurately by a simple equation (1): $T = c_t + t(n/w)$, where n is the number of elements in a chunk, t is the individual or population time needed to complete one chunking step, w is the number of elements simultaneously available in WM, and c_t is the offset in time consisting of two components: the time needed to start the chunking process, and the time needed for generating the response once the chunking process is completed. If we assume that c_t varies among participants to a much lesser degree than t, it is possible to estimate the value of c_t. This offset can be estimated from individual linear functions relating encoding time and array size by computing their intersection point. The average values of c_t that we obtained were 1,314 msec for lack of pop-out and 1,313 msec for pop-out conditions, indicating that participants needed a constant amount of time of about 1 sec for processes that were unrelated to stimulus size. The same intersection points also provide an estimate of the array size at which participants begin to create chunks, c_w. The values for c_w that we obtained were 4.52 and 4.02 elements for the pop-out and non-pop-out conditions, respectively. This suggests that the time needed for creation of LTM begins to increase linearly with the array size only if more than 4 locations need to be remembered. Surprisingly, c_w does not correspond to the number of elements that can be directly encoded into WM under the respective perceptual condition. Thus, at the beginning of the chunking process WM might be operating at its full capacity (~4 locations) even if there is no pop-out. This suggests that the perceptual limitations for encoding into visual WM do not apply to small stimuli with only a few distractor elements. The implications of this finding exceed the scope of the present study, but it is notable that the present relationships might be expressed more precisely by reformulating Equation 1 so that it takes into account different starting sizes of chunks: $T = c_t + t[1 + (n - c_w)/w]$.

(Manuscript received January 10, 2006;
revision accepted for publication December 21, 2006.)