

# Visual working memory for simple and complex visual stimuli

HING YEE ENG, DIYU CHEN, and YUHONG JIANG  
*Harvard University, Cambridge, Massachusetts*

Does the magical number *four* characterize our visual working memory (VWM) capacity for all kinds of objects, or is the capacity of VWM inversely related to the perceptual complexity of those objects? To find out how perceptual complexity affects VWM, we used a change detection task to measure VWM capacity for six types of stimuli of different complexity: colors, letters, polygons, squiggles, cubes, and faces. We found that the estimated capacity decreased for more complex stimuli, suggesting that perceptual complexity was an important factor in determining VWM capacity. However, the considerable correlation between perceptual complexity and VWM capacity declined significantly if subjects were allowed to view the sample memory display longer. We conclude that when encoding limitations are minimized, perceptual complexity affects, but does not determine, VWM capacity.

Visual working memory (VWM) allows us to hold visual information in mind for a few seconds (Luck & Vogel, 1997). It helps maintain perceptual continuity across interruptions and saccades of eyes, yet it is severely limited in capacity. Only approximately four visual objects can be retained concurrently (Pashler, 1988; Vogel, Woodman, & Luck, 2001). The basic unit of this capacity is often considered to be a single object. For example, when asked to remember both the colors and the orientations of several objects, people performed as well as when asked to remember only the colors or only the orientations of those objects, even though the former task required the individuals to remember more features (Luck & Vogel, 1997). This finding suggests that VWM capacity is object based.

Although objects are approximate units of VWM (but see Wheeler & Treisman, 2002), the definition of a visual object is not clear-cut (Scholl, 2001). According to one definition, an object is a connected and bounded region of matter that maintains its connectedness and boundaries when it moves (Spelke, Gutheil, & Van de Walle, 1995). These objects, known as *Spelke objects*, are instantiated in the cognition of adults and young children. Because only spatiotemporal properties are critical to this definition, an object can vary in surface complexity. Simple stimuli, such as a yellow Frisbee or a purple plum, are visual objects; so are complex stimuli, such as a person or a shaded cube. If the unit of VWM is a Spelke object, one may expect VWM for Frisbees to be as efficient as that for shaded cubes. But this is not the case.

In a recent study, Alvarez and Cavanagh (2004) investigated the capacity of VWM for simple and complex visual objects, including colors, letters, Chinese characters, random polygons, and shaded cubes. To measure VWM, a sample display of several stimuli of the same type was presented for 500 msec, followed by a retention interval of 900 msec. Then a test display was presented, and subjects had to judge whether it was identical to the sample. Alvarez and Cavanagh found that people could remember twice as many colors as shaded cubes, even though both types of stimuli were bounded visual objects. These results suggest that VWM is sensitive not only to the number of visual objects, but also to the surface complexity of these objects.

A further, striking finding from Alvarez and Cavanagh's (2004) study was that the variability in VWM capacity was highly correlated with a separate measure of perceptual complexity: visual search slope. In the second measure, subjects viewed a display of several items of the same type and searched for a prespecified target. The number of items on the display varied, allowing the slope of search response time (RT) as a function of set size to be estimated. The slope for color search was shallower than that for polygon search, which was in turn shallower than the slope for cube search. The correlation coefficient between the visual search slope and VWM capacity was .996. These results suggest that perceptual complexity, or *informational load*, determines VWM capacity, although an object-based limit sets an upper bound of about four objects on the capacity.

## The Present Study

In this study, we reconsider the role of informational load in determining the capacity of VWM. This is necessary because VWM is often operationally defined by the change detection task, a task that reflects limitations in perceptual encoding, as well as in memory storage (Pash-

---

This study was supported by NSF Grant 0345525 and the Harvard College Research Program. We thank George Alvarez, Patrick Cavanagh, Tram Neill, and Hal Pashler for comments. Correspondence should be directed to H. Y. Eng, c/o Y. Jiang, 33 Kirkland Street, WJH 820, Cambridge, MA 02138 (e-mail: heng@fas.harvard.edu or yuhong@wjh.harvard.edu).

ler, 1988). Although success in change detection tasks indicates that people have effectively encoded and retained information in VWM, failure in these tasks may originate from several factors, including perceptual, memory, and comparison failures (Angelone, Levin, & Simons, 2003). If observers fail to perceive what is on a memory display, they will be unable to detect changes, simply because the necessary information was not encoded into VWM.

A pilot study in our lab suggests that when allowed to view a sample display for as long as they want, observers often spend several seconds viewing random polygons or unfamiliar faces. Limiting the duration of the sample display to 500 msec significantly reduces accuracy. This is noteworthy because in most VWM studies, a duration of 500 msec or shorter has been used. Although 500 msec may be long enough for an observer to perceive a large number of simple stimuli, it is insufficient for a large number of complex stimuli. Thus, visual search slope may be highly correlated with change detection, not because informational load dictates VWM's storage capacity, but because both tasks are perceptually limited. To separate memory capacity limitations from perceptual limitations, one must increase the presentation duration of the memory display.

In this study, we tested subjects in two tasks: visual search and change detection. As in Alvarez and Cavanagh's (2004) study, visual search slope provided a measure of informational load for different types of stimuli. To estimate VWM capacity, we manipulated stimulus type and memory set size in the change detection task. In addition, we varied the duration of the sample display and the retention interval. The retention interval was 300, 900, or 2,000 msec. This manipulation allowed us to determine whether the effect of stimulus complexity on VWM held true across all retention intervals.

The sample display was presented for 500, 1,000, or 3,000 msec. Across all conditions, change detection could fail either because people failed to perceive all the sample items or because they perceived all the items but failed to retain them in VWM. The longer the display duration, the more likely performance would be limited by memory storage, and not by perception. Thus, if VWM was relatively insensitive to informational load, the correlation between the estimated VWM capacity and visual search slope should diminish at longer display durations. Alternatively, if the capacity of VWM was truly determined by informational load, the correlation should remain high, even when perceptual limitations were reduced to a minimum.

## METHOD

**Subjects.** Twenty Harvard students (18–30 years), 6 in Experiment 1A, 6 in Experiment 1B, and 8 in Experiment 1C, participated for payment.

### Experiment 1A: Fixed Memory Display Duration

**Procedure.** The entire experiment consisted of 10 sessions, due to the large number of testing trials. Each session was divided into two phases: visual search and change detection.

**Stimulus types.** In both visual search and change detection tasks, six types of stimuli were used with six exemplars each: colored

squares (red, green, blue, yellow, cyan, and magenta;  $1.4^\circ \times 1.4^\circ$ ), capital letters (A, E, H, K, N, and R;  $1.4^\circ \times 1.2^\circ$ ), random polygons ( $1.8^\circ \times 1.8^\circ$ ), squiggles ( $1.5^\circ \times 1.5^\circ$ ), shaded cubes ( $1.2^\circ \times 1.2^\circ$ ), and unfamiliar faces ( $1.8^\circ \times 1.8^\circ$ ). On a given trial, only one type of stimulus was presented. Figure 1 shows these stimuli.

**Visual search.** On each trial, the subjects searched for a pre-specified target among several distractors of the same type. The target was precued 1,500 msec before the presentation of the search display, which always contained the target. The subjects pressed the space bar upon target detection, providing an RT measure. The display of search items was then replaced by capital letters in the same locations, and the subjects typed in the letter matching the target's location, providing an accuracy measure (Figure 2A).

We varied two factors: stimulus type and display set size. There were six types of stimuli (Figure 1). The display set size was 4, 8, or 12, containing randomly selected items, with the constraint that only one item matched the prespecified target. Items were presented at randomly selected locations from a  $4 \times 3$  invisible matrix ( $16^\circ \times 12^\circ$ ). Trials from different set sizes and stimulus types were randomly intermixed. There were 144 trials per session.

**Change detection.** On each trial, the subjects first viewed a sample memory display that contained several items presented at randomly selected locations from an invisible  $4 \times 3$  matrix ( $16^\circ \times 12^\circ$ ). The memory display was presented for a variable duration and erased. After a variable retention interval, a test screen with two items was shown until a response was made. One item matched the memory item at that location, and the other exemplar was different from before. The test items were accompanied by two digits, 1 and 2, randomly assigned. The subjects pressed the digit next to the changed item (Figure 2B).

We manipulated four factors in orthogonal: stimulus type (six types), memory set size (2, 4, 6, 8, or 10), memory display duration (500, 1,000, or 3,000 msec), and retention interval (300, 900, or 2,000 msec). Overall, these variations yielded 270 disparate conditions, with each condition tested for 20 trials to accumulate a total of 5,400 trials divided into 10 sessions. Trials from different conditions were randomly intermixed.

### Experiment 1B: Flexible Memory Display Duration

This experiment was similar to Experiment 1A, except that we allowed the subjects to control the duration of the sample display. The subjects were tested in only a single session with 288 visual search trials and 360 change detection trials. Visual search was the same as



**Figure 1.** The six types of stimuli used in this study. Each category contained six exemplars. The color squares were shaded with different textures for illustrative purposes; they were in pure color in the actual experiment.

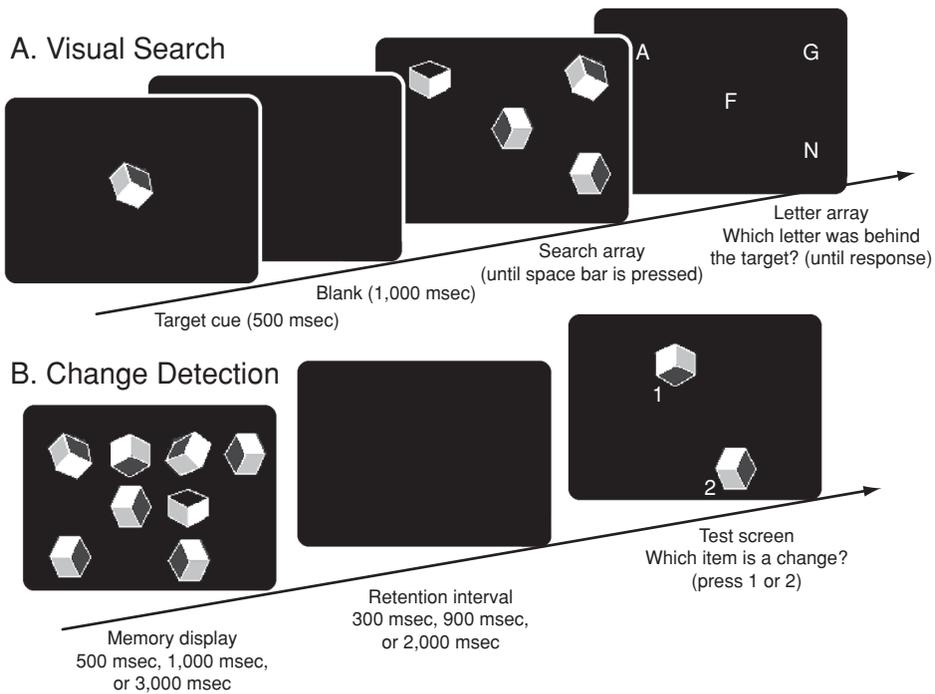


Figure 2. Sample trials tested in the visual search and change detection tasks.

in Experiment 1A. The change detection task included six stimulus types and three memory set sizes (4, 7, and 10). The duration of the memory display was self-paced. The retention interval was always 900 msec. Other aspects of Experiment 1B were identical to those in Experiment 1A.

### Experiment 1C: Concurrent Verbal Working Memory Load

To minimize the contribution of verbal working memory, on each trial of the change detection task, a concurrent verbal load was employed (Vogel et al., 2001). The subjects first viewed six random digits (e.g., “719632”) for 1,000 msec. After the digits had been erased, the sample visual display was presented (1,000 or 3,000 msec), followed by a blank interval (900 msec) and the probe display (until response). Then another string of six digits (e.g., “719832”) was presented. The subjects rehearsed the initial digit string aloud throughout the trial and compared the initial and the final digit strings (same or different).

The subjects performed 720 visual search trials (6 types  $\times$  3 set sizes  $\times$  40 trials) and 720 change detection trials (2 presentation durations  $\times$  6 types  $\times$  3 set sizes  $\times$  20 trials). The experiment was evenly divided into four sessions conducted on consecutive days. Other aspects of the experiment were the same as those in Experiment 1B.

## RESULTS

### Experiment 1A

**Visual search.** Mean accuracy in visual search was at ceiling (99%). We analyzed mean RT for correct trials (Figure 3).

An ANOVA on stimulus type and display set size revealed significant main effects of stimulus type [ $F(5,25) = 67.68, p < .001$ ] and set size [ $F(2,10) = 48.18, p < .001$ ] and a significant interaction [ $F(10,50) = 18.29, p <$

.001]. Overall RT increased from colors to letters, polygons, squiggles, cubes, and faces. The search slope relating RT to display set size also followed that order (in units of milliseconds/item):  $10 \pm 5$  for colors,  $57 \pm 21$  for letters,  $66 \pm 8$  for polygons,  $95 \pm 16$  for squiggles,  $108 \pm 9$  for cubes, and  $177 \pm 24$  for faces. Thus, the informational load, measured from visual search slope, increased from colors to letters, polygons, squiggles, cubes, and faces.

**Change detection.** The change detection task generated a large amount of data, of which only the most relevant are reported here. There was a significant main effect of retention interval [ $F(2,10) = 5.27, p < .05$ ], indicating that memory decayed within 2,000 msec. However, retention interval did not interact with other factors (all  $ps >$

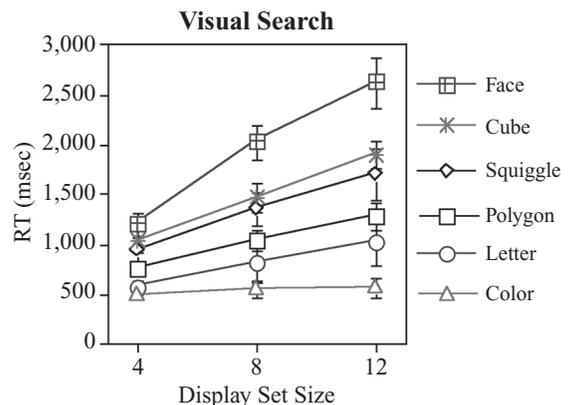
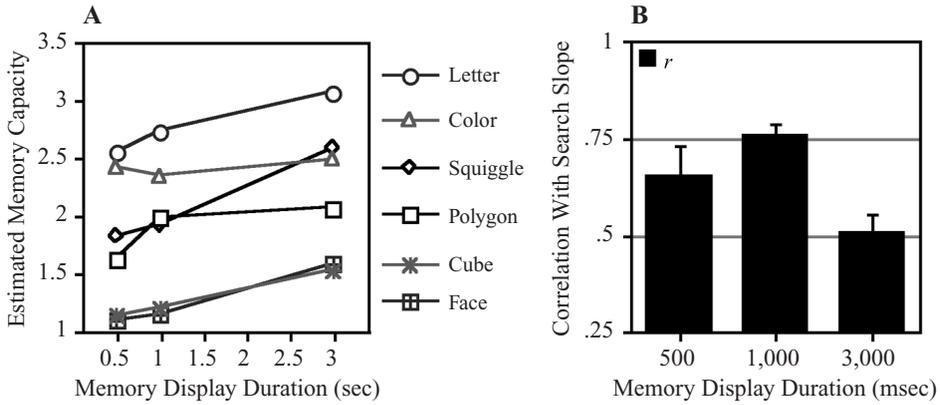


Figure 3. Group mean response time (RT) in the visual search task for different stimulus types and display set sizes.



**Figure 4. (A) Estimated visual working memory capacity in Experiment 1A. (B) Correlation coefficient between visual search slope and estimated memory capacity.**

.15). To simplify data presentation, we pooled results across all retention intervals and estimated the capacity of VWM for each stimulus type and memory duration. A detailed calculation of capacity is shown in the Appendix. Figure 4A shows the estimated VWM capacity.<sup>1</sup>

An ANOVA on stimulus type and presentation duration revealed significant main effects of type [ $F(5,25) = 12.66, p < .001$ ] and memory display duration [ $F(2,10) = 7.04, p < .012$ ] and a significant interaction [ $F(10,50) = 2.10, p < .05$ ]. The estimated capacity was higher for simple stimuli, such as colors, than for complex stimuli, such as cubes and faces. The capacity increased as the memory display duration lengthened, but the increase was more obvious for faces, cubes, and letters than for colors and polygons.<sup>2</sup>

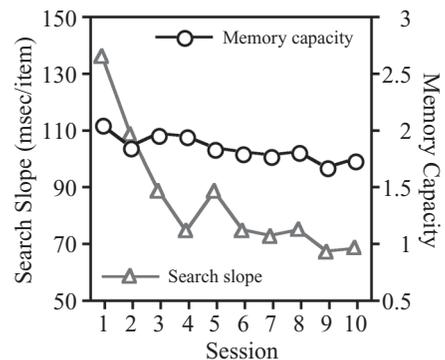
**Correlation analysis.** For each subject, we measured Pearson's correlation coefficient between visual search slope and the estimated VWM capacity across different stimulus types. The group mean is shown in Figure 4B. The correlation was affected significantly by memory display duration [ $F(2,10) = 4.63, p < .038$ ]. In particular, increasing the viewing duration from 1,000 to 3,000 msec reduced the correlation significantly [ $t(5) = 3.35, p < .02$ ]. At 1,000 msec, visual search slope accounted for approximately 64% of the variance in the estimated VWM capacity.<sup>3</sup> This percentage was reduced to 32% when the duration was increased to 3,000 msec. These results suggest that at shorter presentation durations, change detection was limited by perception, as well as by memory. Nonetheless, even with a presentation duration of 3 sec, the correlation remained significant [ $r(5) = -.51, p < .01$ ].

**Practice effects and individual differences.** How does practice affect one's VWM capacity? This is an important question because many aspects of visual cognition are sensitive to learning (Green & Bavelier, 2003). For example, the speed of visual search improved with several practice sessions (Schneider & Shiffrin, 1977), and visual search was more efficient for familiar than for unfamiliar items (Wang, Cavanagh, & Green, 1994). Dividing the experiment into 10 sessions allowed us to examine practice

effects. For simplicity, in this analysis, we averaged across display durations, retention lags, and stimulus types. Figure 5 shows the results.

With practice, visual search slope became shallower [ $F(9,45) = 2.45, p < .023$ ]. Change detection capacity, however, did not change significantly [ $F(9,45) < 1, n.s.$ ]. Five of the 6 observers had no experience in the VWM task prior to this experiment, yet none showed an increase in VWM capacity with practice. This suggests that the capacity of VWM was relatively insensitive to general procedural learning.<sup>4</sup>

Although the observers did not improve with training, there were marked individual differences in overall capacity. Our best subject, a naive observer, had a capacity of 3.2 colors and 2.0 faces, three times better than our worst subject. In addition, an individual who was good at remembering one type was also good at remembering other types. The average pairwise correlation between different stimulus types was .82. What accounts for these individual differences? Did those individuals who had higher memory capacity also have a shallower search slope? The answer, surprisingly, was *no*. Holding stimulus type constant, there was a positive correlation between an in-



**Figure 5. Visual search slope and estimated working memory capacity across 10 testing sessions.**

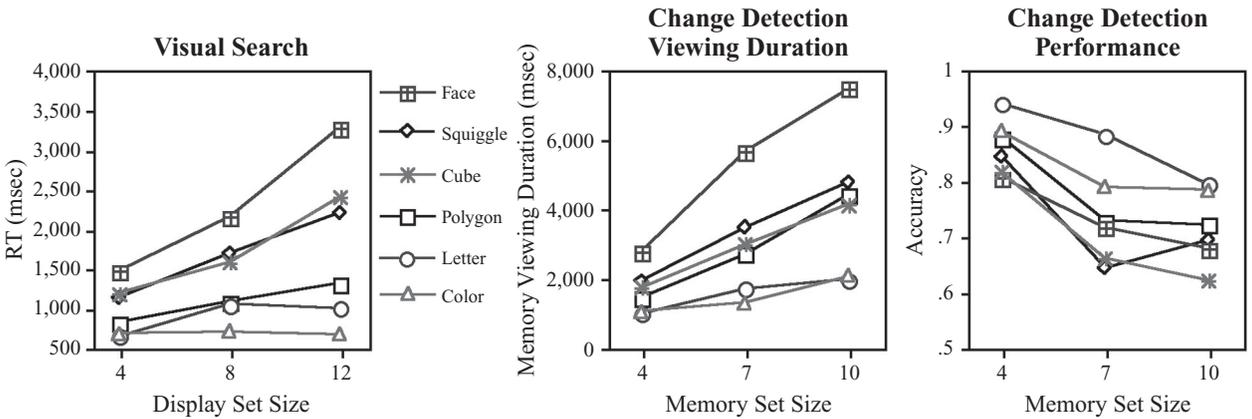


Figure 6. Results from Experiment 1B. The subjects were allowed to view the sample memory display for as long as they wanted.

dividual’s search slope and memory capacity ( $r = .39, p < .05$ ). People whose search slope was shallower tended to have poorer memory capacity. Thus, the inverse relationship between shallow search slope and high memory capacity held true across stimulus types, but not across individuals.

**Experiment 1B**

In Experiment 1A, we allowed the subjects up to 3,000 msec to view a memory display. The estimated capacity had reached plateau for colors and polygons, but not for more complex stimuli, such as cubes and faces. To further eliminate perceptual limitations on change detection of complex stimuli, in Experiment 1B, the subjects viewed the memory display for as long as they wanted.

Not surprisingly, viewing time increased with memory display set size and with stimulus complexity. This pattern mimicked that of visual search speed (Figure 6). As in Experiment 1A, we calculated visual search slope as a measure of informational load. We also calculated the slope of viewing duration as a function of memory set size

in change detection and estimated VWM capacity on the basis of change detection accuracy. There was a significant positive correlation between the visual search slope and the change detection viewing slope [ $r = .57; t(5) = 5.13, p < .004$ ]. In addition, there was a significant negative correlation between the visual search slope and the estimated VWM capacity [ $r = -.52; t(5) = -3.96, p < .011$ ]. The steeper the visual search slope, the smaller the estimated VWM capacity.

**Experiment 1C**

We analyzed change detection restricted to trials with correct verbal responses (93.4% of the trials satisfied this criterion). Mean visual search results and change detection capacity are shown in Figure 7.

Within each subject, the correlation between visual search slope and VWM capacity was calculated across different stimulus types. The mean correlation was  $-.77$  at 1,000-msec and  $-.53$  at 3,000-msec exposure. This difference was significant [ $t(7) = 3.45, p < .01$ ]. Thus, even when verbal WM was minimized, the correlation between

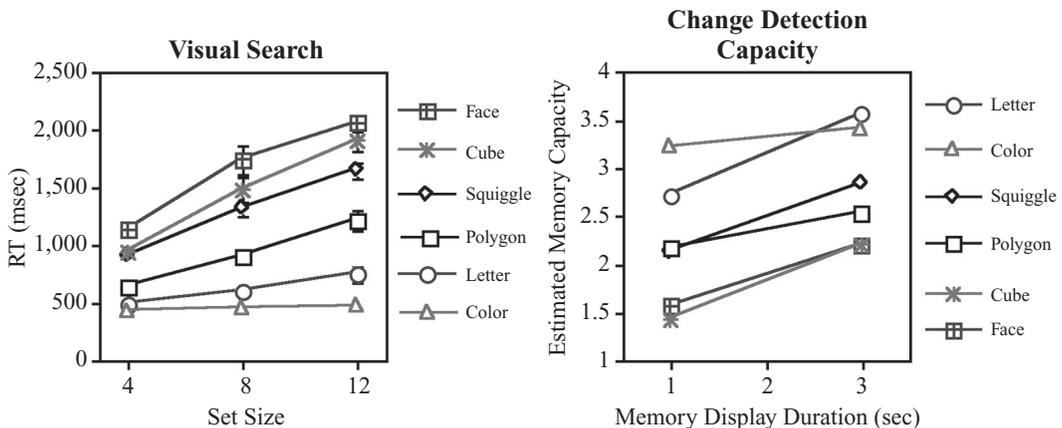


Figure 7. Results from Experiment 1C. A concurrent verbal load was used during the visual change detection task.

VWM capacity and informational load still declined at longer memory exposure durations.

## DISCUSSION

In this study, we investigated how the capacity of VWM would be affected by the perceptual complexity of visual stimuli. Such complexity, known as *informational load*, could be independently quantified by measuring the slope of visual search RT as a function of display set size. The steeper the search slope, the more difficult is visual search, and the heavier is informational load. In Alvarez and Cavanagh's (2004) study, informational load was an excellent predictor of change detection performance, with a correlation coefficient of .996 between visual search slope and estimated VWM capacity.

In this study, we asked whether the observed performance limitation in change detection reflected a limitation in the perceptual encoding or the memory storage process. We hypothesized that the observed limited capacity was, in part, due to an inability to fully encode all the visual stimuli within the typical stimulus presentation duration, especially when the stimuli were perceptually complex. To separate perceptual limitations from storage limitations, we varied the presentation duration of the memory display from 500 msec to as long as the observers wished.

Consistent with Alvarez and Cavanagh's (2004) findings, visual search slopes were shallowest for colors and steepest for cubes and faces. In addition, change detection was generally better for colors than for cubes and faces. There was a significant negative correlation between visual search slope and estimated VWM capacity, suggesting that the heavier the informational load, the lower the VWM capacity. The negative correlation held true whether the memory display was presented briefly (500–1,000 msec), for a moderately long time (3,000 msec), or for as long as the observers wished. Assuming that in the last case, the observers had compensated for the perceptual limitations by viewing more complex stimuli for a longer time, the remaining limitations would primarily reflect VWM storage capacity. The significant correlation between informational load and the estimated VWM capacity confirmed their close relationship.

Our results departed from those of Alvarez and Cavanagh (2004) in two significant ways. First, whereas informational load accounted for about 99% of the variance in VWM capacity in Alvarez and Cavanagh's study, it accounted only for a smaller amount of the variance in ours (30% with a memory display duration of 3 sec or longer). This suggests that although informational load is a significant predictor of VWM capacity, it is not the only predictor. Second, by systematically varying the presentation duration of the memory display, we observed that informational load was a better predictor of the estimated VWM capacity at shorter durations. Because informational load should have a constant influence on the hypothetical VWM capacity, the increased correlation suggests

that at shorter presentation durations, performance was additionally limited by perceptual encoding.

Our findings have significant theoretical implications for studies of VWM. They confirm that VWM is affected not only by the number of objects encoded, but also by the complexity of these objects. We can hold fewer complex objects than simple objects in VWM. However, perceptual complexity only partially determined VWM capacity. What else might affect VWM capacity? We speculate that other perceptual factors, such as the number of objects and their spatiotemporal properties, may also affect VWM. Recent studies suggest that VWM may be characterized as having fixed slots (Zhang & Luck, 2003), flexible slots (Alvarez & Cavanagh, 2004), or both (Song & Jiang, in press). If VWM has components of both fixed slots and flexible slots, informational load will only partially account for VWM capacity.

Although we have emphasized the necessity of minimizing perceptual limitations when estimating VWM capacity, we recognize that prolonging viewing durations could change how items are encoded. Elaborative encoding of a display, such as grouping similar items and storing them as a larger chunk, is more likely to occur at longer durations. Although very brief durations might be insufficient for adequate encoding, extended viewing could result in elaborative encoding. So what is the best viewing duration for a memory display? Although there is unlikely to be a golden standard, we believe that visual search RT can be used as an approximate index. Because search RT corresponds roughly to perceiving half of the items, a viewing duration shorter than search RT is too brief. Conversely, if a memory display lasts longer than twice the search RT, it is perhaps too long. By this rough index, the 3,000-msec duration in Experiments 1A and 1C was too long for colors and too short for faces and cubes, whereas the self-paced viewing duration in Experiment 1B was just about right. A concurrent verbal load task can also help reduce extraneous factors introduced into the task at long exposure durations.

Finally, by testing the subjects in 10 sessions, we were able to estimate the role of procedural learning in VWM. We found that the capacity of VWM was surprisingly stable across sessions, even for naive observers who had nearly no experience with these tasks before. This result stands in contrast with significant practice effects in other aspects of visual cognition, such as visual search and dual-task processing. Consistent with the present observation, Olesen, Westerber, and Klingberg (2004) tested subjects every day for 35 days and found that VWM improved only slightly in one experiment and not at all in another. Olson and Jiang (2004) presented the same memory display 30 times, yet change detection for those displays was not better than that for novel displays. Whether more extensive training can change VWM capacity is an interesting question for future studies.

To summarize, by allowing observers to view the sample memory display for varying durations, we found that change detection was jointly limited by perception and

VWM. At shorter durations, the perceptual complexity of visual objects was highly correlated with change detection accuracy. The correlation decreased but remained significant at longer presentation durations, suggesting that VWM was sensitive to the informational load of visual stimuli. Future studies should address how different perceptual properties of visual stimuli, such as their spatiotemporal and featural properties, contribute to VWM capacity.

## 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.
- ANGELONE, B. L., LEVIN, D. T., & SIMONS, D. J. (2003). The relationship between change detection and recognition of centrally attended objects in motion pictures. *Perception*, **32**, 947-962.
- GREEN, C. S., & BAVELIER, D. (2003). Action video game modifies visual selective attention. *Nature*, **423**, 534-537.
- LUCK, S. J., & VOGEL, E. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **309**, 279-281.
- OLESEN, P. J., WESTERBERG, H., & KLINGBERG, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, **7**, 75-79.
- OLSON, I. R., & JIANG, Y. (2004). Visual short-term memory is not improved by training. *Memory & Cognition*, **32**, 1326-1332.
- PASHLER, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, **44**, 369-378.
- SCHNEIDER, W., & SHIFFRIN, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review*, **84**, 1-66.
- SCHOLL, B. J. (2001). Objects and attention: The state of the art. *Cognition*, **80**, 1-46.
- SONG, J.-H., & JIANG, Y. (in press). Visual working memory for simple and complex features: An fMRI study. *NeuroImage*.

- SPELKE, E. S., GUTHEIL, G., & VAN DE WALLE, G. (1995). The development of object perception. In S. M. Kosslyn & D. N. Osherson (Eds.), *Visual cognition: An invitation to cognitive science* (Vol. 2, pp. 297-330). Cambridge, MA: MIT Press.
- VOGEL, E. K., WOODMAN, G. F., & LUCK, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 92-114.
- WANG, Q., CAVANAGH, P., & GREEN, M. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics*, **56**, 495-500.
- WHEELER, M. E., & TREISMAN, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, **131**, 48-64.
- ZHANG, W., & LUCK, S. J. (2003). Slot-like versus continuous representations in visual working memory [Abstract]. *Journal of Vision*, **3**, 681a.

## NOTES

1. This estimated capacity would correspond to a threshold of 100%. This is why the estimated capacity in this study was lower than that in Alvarez and Cavanagh (2004), who estimated capacity at a 75% threshold. Note, however, using a 100% or a 75% threshold only scaled the capacity value linearly for all conditions.
2. The increase for letters might be attributable to subjects' use of verbal working memory, so we computed the correlation between visual search slope and memory capacity both with letters and without letters. Excluding letters did not change the pattern of results.
3. Individual subjects in Alvarez and Cavanagh's (2004) study produced correlation coefficients ranging from  $-.17$  to  $-.96$ . The average was  $-.80$ , corresponding to 64% of variance in memory capacity accounted for by informational load. This estimation was similar to what we observed here.
4. We also conducted the correlation analysis between search slope and VWM capacity for each session. However, there were too few trials in each session for us to get a stable measure of the search slope. A comparison between the first three sessions and the last three sessions revealed no significant difference in correlation.

## APPENDIX

### Calculation of Memory Capacity

Suppose that for every memory display of  $N$  items, a subject would remember  $C$  items out of the total  $N$  items. For the two test items that were later presented, if one or both test items were among the  $C$  items remembered, the subject's accuracy in identifying the changed item would be 100%, because he or she could match the item(s) with his memory. However, if both test items were not among the  $C$  items remembered, the subject's accuracy would be at chance level, or 50%, because he or she would be making a random guess. Since the likelihood that one test item was not among the  $C$  items remembered was  $(N - C)/N$ , the likelihood that both test items were not remembered would be  $[(N - C)/N] \times [(N - C)/N]$ , or  $[(N - C)/N]^2$ . Therefore, the subject's accuracy could be calculated by the following equation, provided that  $N > C$ :

$$\text{Accuracy} = [(N - C)/N]^2 \times 50\% + \{1 - [(N - C)/N]^2\} \times 100\%. \quad (\text{A1})$$

Because working memory capacity for even the simplest stimulus type (colors) is at or below 4 (Alvarez & Cavanagh, 2004; Luck & Vogel, 1997), we applied the above equation to set sizes from 4 to 10 and then obtained the mean. A similar logic for computing VWM capacity was used in Pashler (1988).