

The Number and Quality of Representations in Working Memory

Weiwei Zhang and Steven J. Luck

Center for Mind & Brain and Department of Psychology, University of California, Davis

Psychological Science
22(11) 1434–1441
© The Author(s) 2011
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797611417006
http://pss.sagepub.com



Abstract

Flexible-resource theories characterize working memory as a flexible resource that can store either a large number of low-quality representations or a small number of high-quality representations. In contrast, limited-item theories propose that the number of items that can be stored in working memory is strictly limited and cannot be increased by decreasing the quality of the representations. We tested these fundamentally different conceptualizations of working memory capacity by determining whether observers could trade quality for quantity in working memory when given incentives to do so. We found no evidence that observers could increase the number of representations by decreasing their quality in working memory, but observers could make such a trade-off at earlier processing stages. Our results show that the capacity limit of working memory is best characterized as a limit on the number of items that can be stored and not as a limit on a finely divisible resource that simultaneously determines the number and quality of the representations.

Keywords

visual memory, short-term memory, visual attention, visual perception

Received 3/30/11; Revision accepted 5/29/11

The limited capacity of working memory is a key concept in both classic and contemporary theories of cognition (Cowan, 2005; Miller, 1956). Individual differences in working memory capacity predict differences in a wide range of higher-order cognitive abilities. This predictive ability appears to be particularly strong for visual working memory (VWM) capacity, which accounts for more than 40% of the variance in global fluid intelligence (Fukuda, Vogel, Mayr, & Awh, 2010) and almost 80% of the variance in overall cognitive performance (Gold et al., 2010). VWM capacity varies across the life span (Gazzaley, Cooney, Rissman, & D'Esposito, 2005) and is reduced in people with psychiatric disorders (Gold et al., 2010). An understanding of the limits of VWM capacity is therefore essential for an understanding of the human mind.

Two broad classes of theories of VWM capacity have been proposed (for a review, see Luck, 2008). *Flexible-resource theories* propose that VWM capacity reflects the flexible allocation of limited cognitive resources. According to these theories, allocating more resources to an item will allow it to be represented with greater quality or precision (Bays & Husain, 2008; Palmer, 1990; Wilken & Ma, 2004). That is, resources can be focused on a small number of items to create high-quality representations or distributed among a large number of items to create low-quality representations. In contrast, *limited-item theories* propose that the number of items in memory (i.e., capacity, K) is strictly limited and cannot be increased by

decreasing the precision of the representations (Anderson, Vogel, & Awh, 2011; Zhang & Luck, 2008). In these theories, VWM capacity is analogous to a set of slots rather than a pool of resources.

These theories can be directly tested only with tasks that involve remembering simple features, for which the concept of precision can be unambiguously operationalized. Researchers have used this approach in several recent studies by testing memory for arrays of single-feature items and measuring how precision and K vary as the set size (the number of items to be remembered) varies. In some of these studies, K increased and precision decreased as the set size increased to 3 or 4 items, at which point both K and precision reached an asymptote (Anderson et al., 2011; Zhang & Luck, 2008). The findings from these studies support limited-item theories of VWM capacity. However, results from other studies, in which no asymptote for either K or precision was found (Bays & Husain, 2008; Wilken & Ma, 2004), support flexible-resource theories.

In all of these studies, the researchers varied the number of items in the displays, assuming that observers would attempt to store as many items as possible in VWM. However, it is

Corresponding Author:

Weiwei Zhang, UC-Davis Center for Mind & Brain, 267 Cousteau Place, Davis, CA 95618

E-mail: wwzhang@ucdavis.edu

possible that observers would strategically devote all of their resources to a limited number of items when the set size is large even if they could, in principle, store a low-quality representation of every item. We therefore took a different approach in the present study to assess the most fundamental difference between flexible-resource and limited-item theories, which is whether people can increase the number of items stored by decreasing the quality of the representations when motivated to do so.

We report the results of four experiments. Participants in all experiments were University of California, Davis, students between the ages of 18 and 30; all had normal color vision and normal or corrected-to-normal visual acuity. In Experiments 1, 2, and 3, we used three different methods to motivate observers to store a large number of low-precision representations. All three of these approaches failed to produce an increase in the number of items stored in VWM. In Experiment 4, we demonstrated that trade-offs between K and precision can be produced in iconic memory, which indicated that the failure to increase the number of items stored in VWM in the earlier experiments was not caused by weak manipulations of motivation. Thus, our manipulations were strong enough to influence iconic memory, but they could not induce a trade-off between quality and quantity in working memory.

Experiment 1

To measure K and precision independently, we used a VWM recall task in which observers were presented with a set of colored items and then reported the color of one item—indicated by a probe at the time of test—by clicking on a color wheel (Fig. 1a; see Zhang & Luck, 2008). If the probed item was present in VWM, then the observer's response would be near the correct color, and errors would be distributed normally around the correct color. If, however, the probed item was absent from memory, then the observer would have to guess randomly, and the result would be a uniform distribution of errors. If the observer remembered the probed item on some trials and guessed randomly on other trials, the overall distribution of errors would consist of a mixture of a normal distribution and a uniform distribution, which is identical to a normal distribution with a vertical offset. It is possible to mathematically decompose such a mixture into its components. The standard deviation of the normally distributed portion of the mixture reflects the precision of the representation when the probed item is present in memory. The amount of vertical offset can be used to determine the probability that the probed item was present in memory, and K is the set size multiplied by this probability. This analytic approach has been used in several recent studies (Anderson et al., 2011; Zhang & Luck, 2008, 2009).

We manipulated the amount of precision needed to perform the experimental task by varying the number of distinct colors in the color wheel. In the high-precision condition (see Fig. 1a), the color wheel contained 180 equally spaced color

values. After participants indicated their response on each trial, they were given feedback about the distance between the color of the probed item and the reported color (i.e., the magnitude of the recall error). A precise representation was needed to minimize the error in this condition. In the low-precision condition (Fig. 1b), the color wheel contained a small set of discrete spokes. A relatively imprecise representation was sufficient to produce a correct response in this condition, and storing a large number of low-quality representations in VWM would therefore maximize overall performance. If observers were able to increase K by decreasing the precision of representations, then they should have been able to store more items in the low-precision condition than in the high-precision condition.

Method

Two separate groups of 13 observers participated in Experiments 1a and 1b. Stimuli were presented on a CRT monitor with a gray background (15.1 cd/m^2) and a continuous fixation point at a viewing distance of 57 cm. Each trial began with the presentation of a sample array for 200 ms. This array consisted of four colored squares ($2^\circ \times 2^\circ$). The colors were selected from a master set of 180 evenly distributed and isoluminant hues on a circle in the perceptually homogeneous Commission Internationale de l'Eclairage Lab color space (for details, see Zhang & Luck, 2008).

After a 1,000-ms delay, a test array was presented. The test array consisted of four outlined squares at the locations of the four squares in the sample array and a color wheel (8.2° diameter; 2.2° thickness). The outline of the probe square was thicker than that of the other squares in the test array; observers were instructed to report the color of the corresponding sample square. In the high-precision condition, the color wheel consisted of all 180 colors in the master set. In the low-precision condition, it consisted of 9 (Experiment 1a) or 6 (Experiment 1b) colored spokes (0.29° wide); one of the spokes was the same color as the probed sample square, and the colors of the other spokes were at 40° (Experiment 1a) or 60° (Experiment 1b) increments from this color. In both conditions, observers were instructed to indicate which color exactly matched the color of the probed sample square by clicking a mouse; observers were given as much time as they needed to respond. Each of the colors in the sample array was one of the colors present in the array of spokes used in the low-precision condition; consequently, the colors in the sample array were separated by multiples of 40° (Experiment 1) or 60° (Experiment 2) in both the low-precision and the high-precision conditions. Feedback consisted of an arrow that pointed to the correct answer and a cross positioned at the reported color; the angular difference between these two markers represented the magnitude of recall error. Each observer completed 24 practice trials and 150 experimental trials in each condition; the order of conditions was counter-balanced across observers.

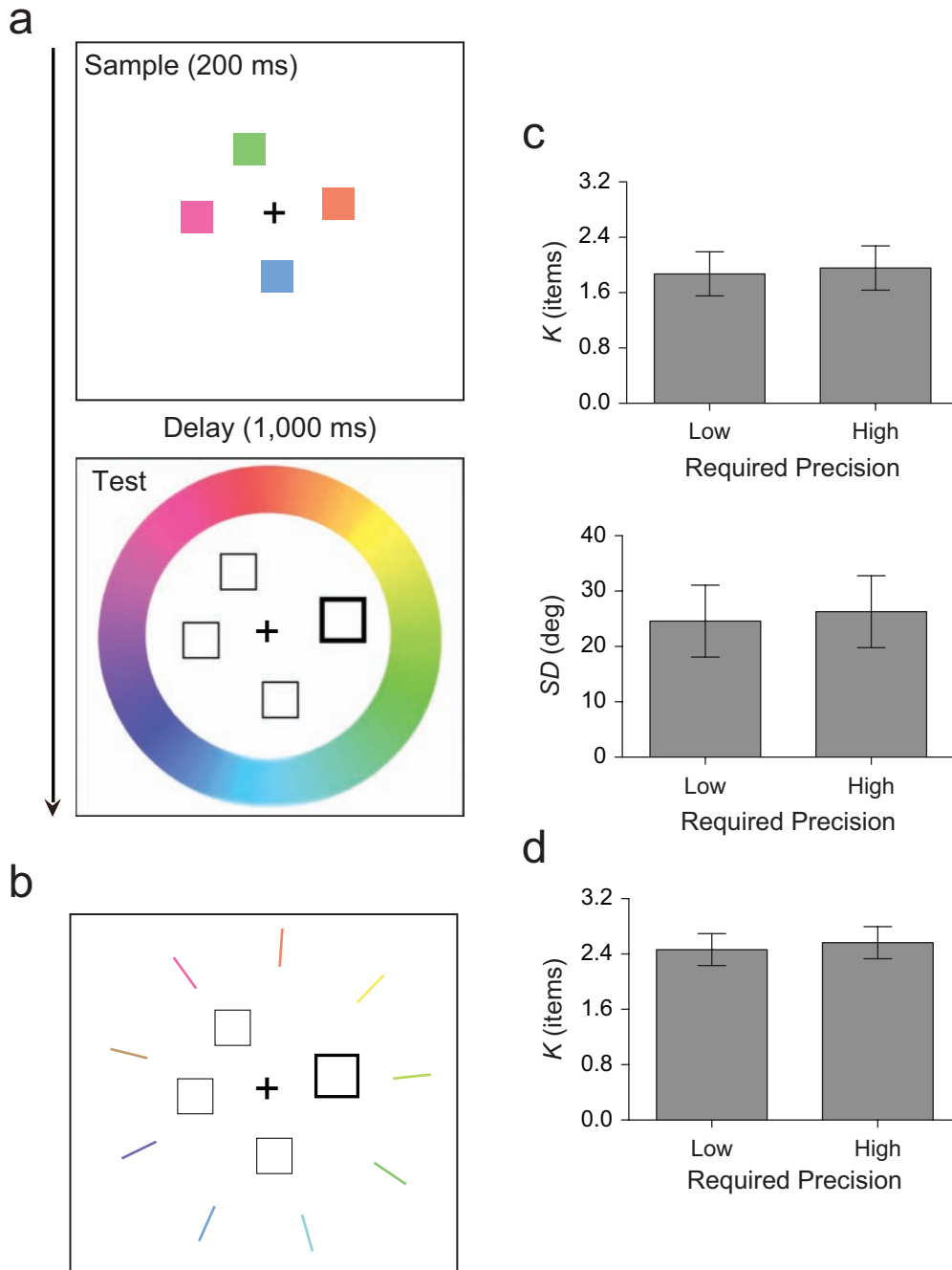


Fig. 1. Trial structure and results from Experiment 1. In the high-precision condition of Experiments 1a and 1b (a), a sample array of four colored squares was presented for 200 ms. After an interstimulus interval of 1,000 ms, the test array was presented; this array consisted of a color wheel containing 180 equally spaced color values and four outlined squares at the locations of the four squares in the sample array. One of the squares in the test array had a thicker outline, and observers reported their memory for the color of the corresponding sample-array square by clicking on the color wheel. In the test array for the low-precision condition of Experiment 1a (b), the color wheel contained nine spokes. The graphs show (c) the mean number of items stored in memory (K) and the standard deviation (inversely related to the precision of the representations) for each condition in Experiment 1a and (d) K for each condition in Experiment 1b. For Experiment 1b, we present only results for K because standard deviations could not be estimated reliably in that experiment given the granularity of the test displays. Error bars in all graphs represent within-subjects 95% confidence intervals (Cousineau, 2007).

Maximum likelihood estimation was used to determine the SD and K parameters (for details, see Zhang & Luck, 2008). Note, however, that none of the conclusions we draw from this

study depend on the use of this specific quantitative model because similar results were obtained when we simply compared the raw distributions of recall errors across conditions

using analysis of variance (see the Supplemental Material available online for details).

Results and discussion

Figure 1c shows the mean K and SD estimates for each condition in Experiment 1a, and Figure 1d shows the mean K estimate for each condition in Experiment 1b (see the Supplemental Material for raw data and goodness-of-fit results). In both experiments, K was nearly identical between the low-precision and high-precision conditions.¹ Paired t tests showed that there were no significant differences between the low-precision condition and high-precision condition in either experiment—Experiment 1a: $t(12) = 0.80, p = .43$; Experiment 1b: $t(12) = 1.20, p = .25$. A Bayes factor analysis (Rouder, Speckman, Sun, Morey, & Iverson, 2009) indicated that the null hypothesis (no difference between the low-precision and high-precision conditions) was 3.57 times more likely to be true than the alternative hypothesis (a difference between conditions) in Experiment 1a and was 2.52 times more likely to be true than the alternative hypothesis in Experiment 1b. In addition, confidence interval analyses indicated that we can be 95% confident that K was increased in the low-precision condition relative to the high-precision condition by no more than 0.24 items in Experiment 1a and no more than 0.08 items in Experiment 1b. Thus, although it is impossible to prove the null hypothesis, these results show that the null hypothesis was substantially more likely to be true than the alternative hypothesis and that, even if there was a real effect, it was very small (less than a quarter of an item's worth of increased capacity in the low-precision condition).

We also found no indication of reduced precision (increased SD) in the low-precision condition relative to the high-precision condition of Experiment 1a (SD could not be meaningfully estimated for Experiment 1b because there were only six possible responses). The difference in SD between conditions was not significant, $t(12) = 0.65, p = .53$, and a Bayes factor analysis indicated that the null hypothesis was 3.96 times more likely to be true than the alternative hypothesis. In addition, a confidence interval analysis indicated that we can be 95% confident that the SD was no more than 4.82° larger in the high-precision condition than in the low-precision condition.

These results indicate that observers cannot increase the number of items in VWM or decrease the precision of the representations even in a task that requires only low precision.

Experiment 2

In Experiment 2, we used a different method to encourage observers to reduce precision and thereby increase the number of items stored in VWM. Rather than being provided with exact feedback about the difference between the reported color and the color of the probed item, observers were simply told whether their response was correct or incorrect. In the low-precision condition, a response was considered correct if it was within 60° of the correct color value; in the high-precision condition, a response was considered correct if it was within 15° of the

correct color value. Thus, if it is possible to trade off precision for the number of items stored in VWM, observers could have maximized their performance in the low-precision condition by maintaining a large number of imprecise representations.

Our method in Experiment 2 eliminated the physical differences between the color wheels in the two conditions of Experiment 1. This manipulation can be conceived as requiring different decision criteria in the different conditions, an approach that has been used for decades to distinguish between continuous and discrete processes in signal detection theory (Atkinson & Juola, 1973, 1974; Mandler, 1980).

Method

The procedure of Experiment 2 was the same as that of the high-precision condition of Experiment 1a with the following exceptions. A new sample of 14 observers was tested. After each response, we provided observers with feedback using a white arc that measured either 30° (high-precision condition) or 120° (low-precision condition). This arc was superimposed on the color wheel so that it was centered at the location of the correct color. In principle, observers could have used the midpoint of the arc to determine the distance between their response and the color of the probed item, but they were instructed solely to make sure that their response was within the range indicated by the arc, and they were not informed that the midpoint of the arc denoted the location of the correct color.

Results and discussion

Figure 2a shows the mean K and SD values for each condition. Observers stored an average of 2.44 items in both the low-precision and the high-precision conditions. These values were not significantly different, $t(13) = 0.39, p = .71$, and the Bayes factor analysis indicated that the null hypothesis was 4.64 times more likely to be true than the alternative hypothesis was. In addition, a confidence interval analysis indicated that we can be 95% confident that K was no more than 0.40 items greater in the low-precision condition than in the high-precision condition. These results provide converging evidence against flexible-resource theories.

Responses were slightly more precise in the high-precision condition than in the low-precision condition, and this small difference between conditions was marginally significant, $t(13) = 2.12, p = .054$. However, the Bayes factor analysis indicated that a difference between conditions was only 1.27 times more likely to be true than the null hypothesis was. Moreover, the 2° difference we observed was tiny relative to the 90° difference in the precision requirements of the low-precision and high-precision conditions.

Experiment 3

Because it is possible that the observers in Experiments 1 and 2 were not sufficiently motivated to decrease the precision of the stored representations, we provided them with monetary

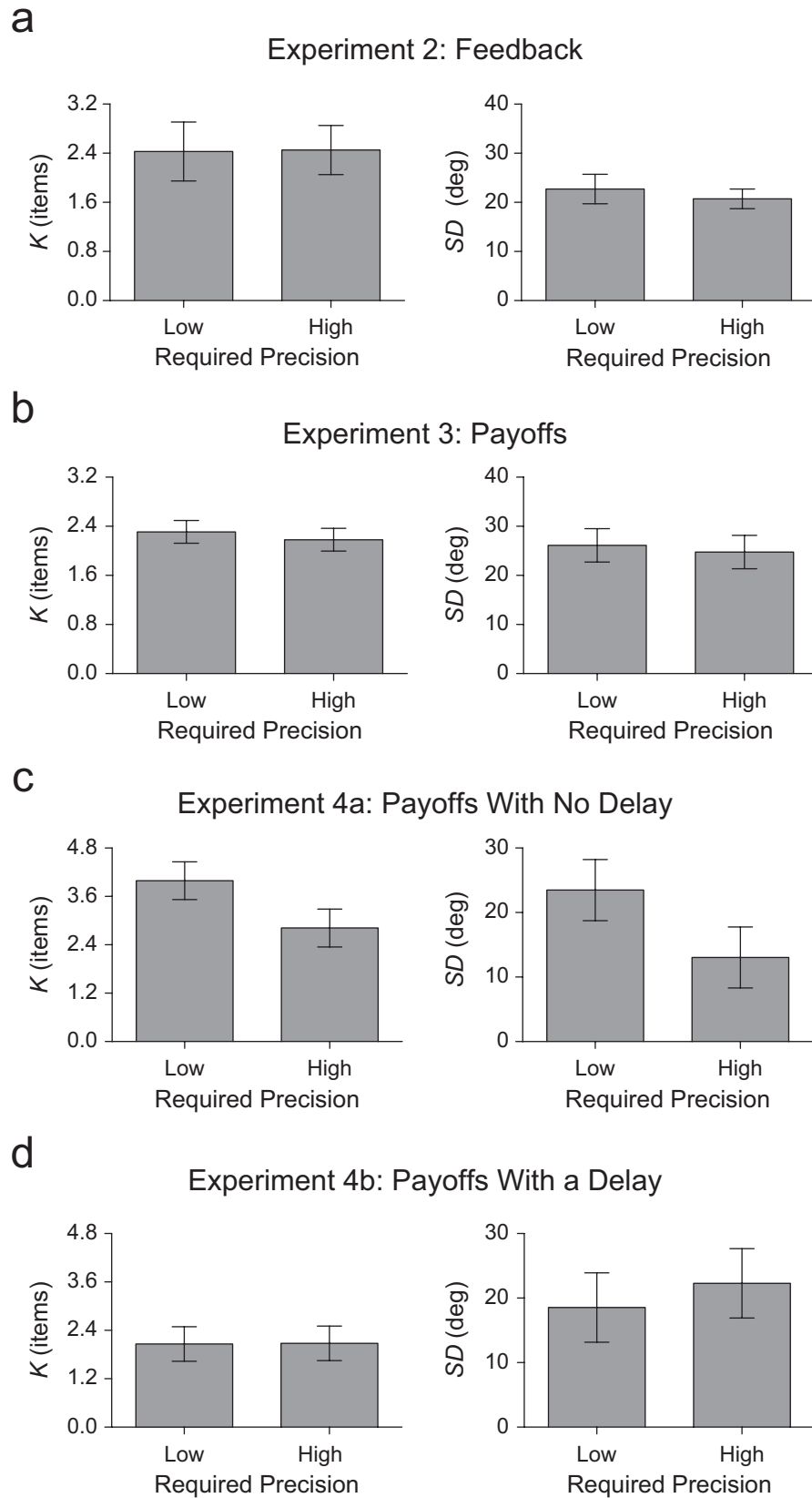


Fig. 2. Results from (a) Experiment 2, (b) Experiment 3, (c) Experiment 4a, and (d) Experiment 4b. The graphs show the mean number of items remembered (K ; left column) and their standard deviation (inversely related to precision; right column) for the low-precision and high-precision conditions in each experiment. Error bars represent within-subjects 95% confidence intervals (Cousineau, 2007).

incentives in Experiment 3. Specifically, observers earned money for responses that were within 20° (high-precision condition) or 60° (low-precision condition) of the correct color. Observers could earn more money in the low-resolution condition by reducing their precision and increasing the number of items stored in VWM.

Method

The procedure of Experiment 3 was the same as that of the high-precision condition in Experiment 1a, with the following exceptions. A new sample of 10 observers was tested. In the high-precision condition, observers earned \$0.06 if their response fell within 20° of the color of the probed item and earned nothing otherwise. In the low-precision condition, observers earned \$0.04 if their response fell within 60° of the color of the probed item and earned nothing if the response fell 60° to 100° from the correct value. To encourage observers to store in memory at least some information about every item in the low-precision condition, we penalized them \$0.02 for wild guesses (responses that were more than 100° from the correct color). In addition, observers received a base payment of \$10.00. The observers were fully informed of the payment contingencies. The amount of money earned on each trial and the total earnings were displayed at the center of the screen.

Results and discussion

Figure 2b shows the mean K and SD values for each condition. The K values in the low-precision and high-precision conditions were not significantly different, $t(9) = 1.39$, $p = .20$. A Bayes factor analysis indicated that the null hypothesis was 1.87 times more likely to be true than the alternative hypothesis was. In addition, a confidence interval analysis indicated that we can be 95% confident that K was no more than 0.32 items greater in the low-precision condition than in the high-precision condition. The payoff manipulation had no significant impact on SD , $t(9) = 0.64$, $p = .54$. The null hypothesis for SD was 3.56 times more likely to be true than the alternative hypothesis was, and we can be 95% confident that the SD was no more than 4.90° larger in the low-precision condition than in the high-precision condition.

Observers earned an average of \$13.69 in both the low- and the high-precision conditions. For an observer to earn the maximum amount of money, his or her optimal SD would have been 25.2° in the high-precision condition and 38.8° in the low-precision condition (assuming the K - SD trade-off proposed by Bays & Husain, 2008; for details, see the Supplemental Material). The observed SD of 24.7° in the high-precision condition was close to the optimal SD . However, the observed SD of 26.2° in the low-precision condition was significantly lower than the optimal value of 38.8°, $t(9) = 3.35$, $p = .004$;

observers would have increased their incentive-based earnings by 38% if they had been able to achieve the optimal SD in this condition. Thus, even in the presence of financial incentives, observers were unable to strategically increase the number of items stored in working memory by reducing the precision of the representations. Experiment 4 was designed to determine whether our incentive structure was sufficiently powerful to influence trade-offs under conditions that required minimal VWM involvement.

Experiment 4

Research in computational neuroscience has suggested that the functional properties of neural systems underlying VWM necessarily limit the number of discrete representations (Rafone & Wolters, 2001; Wang, 2001). However, there is no reason to believe that earlier stages of visual representation are subject to the same limits. In fact, trade-offs between the number of attended items and their resolution have been demonstrated in several studies of perception (Alvarez & Franconeri, 2007; Eriksen & Yeh, 1985; Horowitz & Cohen, 2010; Howard & Holcombe, 2008; Shulman, Wilson, & Sheehy, 1985; Treisman & Gormican, 1988). Therefore, in Experiment 4, we sought to determine whether payoff manipulations could influence performance under conditions that were not limited by the constraints of VWM storage. In Experiment 4a, we eliminated the delay between the offset of the sample array and the onset of the test array, allowing observers to select information from higher-capacity, more fragile memory systems (Landman, Spekreijse, & Lamme, 2003; Sperling, 1960). We manipulated the payoffs as we had in Experiment 3, predicting that in this situation, observers would be able to trade precision for the number of items stored. Experiment 4b was identical to Experiment 4a except that the delay between the offset of the sample array and the onset of the test array was reinstated so that observers would be forced to use VWM.

Method

The procedure of Experiment 4 was the same as that of Experiment 3, with the following exceptions. To avoid ceiling effects, we increased the set size to six items. To avoid masking of iconic memory, we replaced the test array with a single arrow adjacent to the location of the probed item. To further minimize masking, we presented not only the test array but also the sample array with a color wheel, which was rotated randomly. In addition, observers earned points instead of money because pilot testing indicated that monetary incentives were unnecessary. We eliminated the delay between the offset of the sample array and the onset of the probe in Experiment 4a to minimize observers' use of VWM, but reinstated the delay in Experiment 4b. New samples of 9 and 12 observers were tested in Experiments 4a and 4b, respectively.

Results and discussion

The mean K and SD values for each condition are shown in Figures 2c and 2d. The results from Experiment 4b were similar to those from Experiment 3, with no difference in K or SD between the low-precision and high-precision conditions— K : $t(11) = 0.06, p = 0.95$; SD : $t(11) = 0.57, p = 0.58$. A Bayes factor analysis favored the null hypothesis by a factor of 4.65 for K and 4.00 for SD , and we can be 95% confident that any increase in the low-resolution condition was no more than 0.24 items for K and no more than 1.62° for SD . Thus, we found that when performance was limited by VWM, observers could not trade precision for an increase in the number of items stored in memory.

In Experiment 4a, which minimized VWM involvement, K increased significantly from 2.76 items in the high-precision condition to 3.96 items in the low-precision condition, $t(8) = 3.76, p = .006$, and this effect was accompanied by a significant increase in SD from 13.0° in the high-precision condition to 23.4° in the low-precision condition, $t(8) = 3.31, p = .011$. Thus, providing incentives for low precision led to a nearly 2-fold increase in precision and a 43% increase in K . The alternative hypothesis (a difference between the low-precision and high-precision conditions) was 10.7 and 6.1 times more likely to be true than the null hypothesis for SD and K , respectively. Even though the incentives were not monetary, observers could trade precision for the number of items stored when performance was based primarily on processes that precede VWM encoding.

Statistical comparisons of Experiments 4a and 4b were performed with a mixed-model, two-way analysis of variance. Because there were differences in K and SD between the low- and high-precision conditions in Experiment 4a but not in Experiment 4b, there was a significant interaction between experiment and condition for both K , $F(1, 19) = 9.18, p = .007$, and SD , $F(1, 19) = 4.53, p = .04$. Thus, incentives had no impact on K or SD when the task stressed VWM, but they had a large effect when the task stressed perception and iconic memory.

General Discussion

Our results from multiple experiments provide converging evidence that VWM capacity is characterized by a limit on the number of items that can be stored rather than by a finite pool of continuously divisible resources. We repeatedly found that observers could not increase the number of representations in VWM by reducing the precision of the representations, whether observers were motivated to do so by the number of response alternatives, by the granularity of the feedback, or by direct incentives. In contrast, we found that observers could trade the precision of representations for storage of a greater number of items at an earlier stage of representation; this finding is consistent with results of previous psychophysical and electrophysiological studies.

Our results should not be taken to imply that precision in VWM is completely inflexible. Other research has shown that precision and number of items can be traded off in VWM when the number of items is below the item limit: Precision can be increased when attention is focused on a single item by a spatial cue (Zhang & Luck, 2008) or by payoffs (Zhang & Luck, 2011). Our results are consistent with studies showing that precision increases as set size decreases below the item limit (Anderson et al., 2011; Zhang & Luck, 2008). Thus, observers can strategically increase precision by focusing resources onto a smaller number of items (see Zhang & Luck, 2008, for a quantitative model of resource allocation in VWM). However, there is a limit on the number of items that can be stored, and observers cannot exceed this limit by reducing the quality or complexity of the representations (Alvarez & Cavanagh, 2004; Awh, Barton, & Vogel, 2007; Barton, Ester, & Awh, 2009).

The item limit may arise from the need to segregate individual VWM representations to avoid interference that would cause them to collapse (e.g., Raffone & Wolters, 2001). Consequently, the overall storage capacity of VWM may be modifiable by factors that influence the segregation of representations. For example, extensive training may increase K by optimizing segregation processes (Klingberg, 2010; Scolar, Vogel, & Awh, 2008).

K may also be increased by means of chunking strategies that allow more efficient use of each representation (Miller, 1956). Results from previous studies in which observers appeared to trade off precision and capacity (e.g., Bays & Husain, 2008; Wilken & Ma, 2004) may be explained by the observers' use of chunking strategies. For example, an observer who is shown 20 haphazardly scattered dots might organize the dots into four clusters and store the centroid of each cluster in VWM. Only four representations would be present in VWM, but all 20 dots would contribute to these representations. If asked to report the remembered location of a single dot, the observer could use the stored cluster centroids to make an informed guess, and the precision of this response would be related to the number of dots in each cluster. In this manner, observers could use a small set of discrete representations to flexibly represent a large number of items at the cost of reduced precision. Precision and number of items would be traded off from the perspective of performance, but there would still be a strict item limit from the perspective of the underlying representational structure and neurobiology.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This study was made possible by a postdoctoral fellowship from the National Institutes of Health to W. Z. (1F32MH090719) and by a grant from the National Institutes of Health to S. J. L. (R01MH076226).

Supplemental Material

Additional supporting information may be found at <http://pss.sagepub.com/content/by/supplemental-data>

Note

1. The K values were considerably lower in Experiment 1a than in Experiment 1b, but these results likely reflect random differences between the participants in the two experiments. More generally, the K values observed with the color-wheel recall paradigm used in this study tend to be lower than those observed with change-detection paradigms. However, the K values are highly correlated across paradigms (Gold et al., 2010; Zhang & Luck, 2008).

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.
- Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track? Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision, 7*(13), Article 14. Retrieved from <http://www.journalofvision.org/content/7/13/14>
- Anderson, D. E., Vogel, E. K., & Awh, E. (2011). Precision in visual working memory reaches a stable plateau when individual item limits are exceeded. *Journal of Neuroscience, 31*, 1128–1138.
- Atkinson, R. C., & Juola, J. F. (1973). Factors influencing the speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 583–612). New York, NY: Academic Press.
- Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology: Vol. 1. Learning, memory and thinking* (pp. 243–290). San Francisco, CA: Freeman.
- Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18*, 622–628.
- Barton, B., Ester, E. F., & Awh, E. (2009). Discrete resource allocation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1359–1367.
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321*, 851–854.
- Cousineau, D. (2007). Confidence intervals in within-subjects designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1*, 42–45. Retrieved from <http://www.tqmp.org/content/vol01-1/p042/p042.pdf>
- Cowan, N. (2005). *Working memory capacity*. New York, NY: Psychology Press.
- Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 583–597.
- Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review, 17*, 673–679.
- Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience, 8*, 1298–1300.
- Gold, J. M., Hahn, B., Zhang, W. W., Robinson, B. M., Kappenman, E. S., Beck, V. M., & Luck, S. J. (2010). Reduced capacity but spared precision and maintenance of working memory representations in schizophrenia. *Archives of General Psychiatry, 67*, 570–577.
- Horowitz, T. S., & Cohen, M. A. (2010). Direction information in multiple object tracking is limited by a graded resource. *Attention, Perception, & Psychophysics, 72*, 1765–1775.
- Howard, C. J., & Holcombe, A. O. (2008). Tracking the changing features of multiple objects: Progressively poorer perceptual precision and progressively greater perceptual lag. *Vision Research, 48*, 1164–1180.
- Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences, 14*, 317–324.
- Landman, R., Spekreijse, H., & Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research, 43*, 149–164.
- Luck, S. J. (2008). Visual short-term memory. In S. J. Luck & A. Hollingworth (Eds.), *Visual memory* (pp. 43–85). New York, NY: Oxford University Press.
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*, 252–271.
- 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.
- Palmer, J. (1990). Attentional limits on the perception and memory of visual information. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 332–350.
- Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience, 13*, 766–785.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225–237.
- Scolari, M., Vogel, E. K., & Awh, E. (2008). Perceptual expertise enhances the resolution but not the number of representations in working memory. *Psychonomic Bulletin & Review, 15*, 215–222.
- Shulman, G. L., Wilson, J., & Sheehy, J. B. (1985). Spatial determinants of the distribution of attention. *Perception & Psychophysics, 37*, 59–65.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74*, 1–29.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95*, 15–48.
- Wang, X. J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends in Cognitive Sciences, 24*, 455–463.
- Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision, 4*(12), Article 11. Retrieved from <http://www.journalofvision.org/content/4/12/11>
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*, 233–235.
- Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science, 20*, 423–428.
- Zhang, W., & Luck, S. J. (2011). *Visual working memory representation in high resolution*. Manuscript in preparation.

Supplementary Materials for The Number and Quality of Representations in Working Memory

Parameter Fits

Supplementary Figure S1 shows the observed pattern of responses in each experiment, quantified as the distance between the reported color and the actual color, along with curves showing the estimated values from the parameter fitting procedure. Supplementary Table S1 shows the quality of the fits, estimated with adjusted R^2 .

Model-Independent Analyses

It should be noted that the conclusions drawn from this study do not depend on our specific quantitative model of VWM. That is, the raw data from which the parameters were estimated show no differences between the low- and high-precision conditions in any of the experiments except Experiment 4a (in which the K and SD parameters were found to vary as well). This can be seen visually in the figure, and it was supported by ANOVAs conducted for each experiment with factors of condition (low or high precision) and error bin (with 6 bins for Experiment 1b and 9 bins for all other experiments). A significant interaction between condition and error bin was found for Experiment 4a [$F(8,64)=3.22$, $p=0.004$], but not for any other experiment. Thus, the distribution of errors was not influenced by the precision needed for the task in any of the experiments in which performance was limited by VWM, but the distribution was significantly influenced by the precision manipulation when performance was not limited by VWM.

Optimal Performance in Experiment 3

If precision can be traded for capacity, then observers should have attempted to increase the number of representations by decreasing precision in the low-precision condition of Experiment 3. However, this assumes that the payoff contingencies used in this experiment were sufficient to motivate a change in precision. To assess this, we computed the amount of money that would

have been expected as a function of the K/SD tradeoff in each condition, using the tradeoff ratio proposed in a leading resource model (Bays and Husain, 2008). Figure S2 shows the expected earnings as a function of the SD, along with the optimal and observed SD values. The observed SD was near the optimal value in the high-precision condition (because we designed the contingencies to reflect typical SD values), but it was far from the optimal value in the low-precision condition.

Figure S1. Probability distribution of error magnitude (difference between reported color and actual color) in the low- and high-resolution conditions of each experiment.

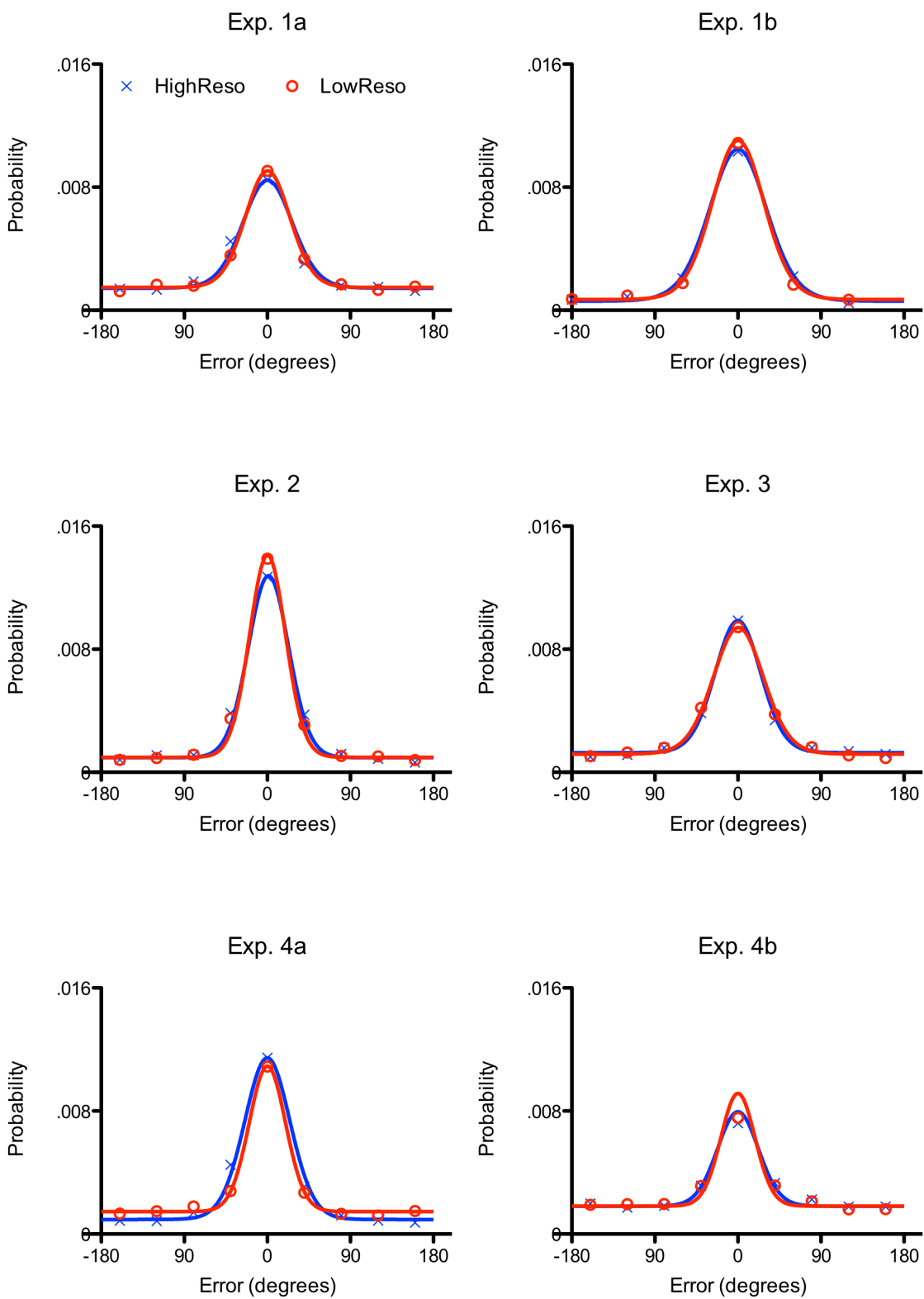


Figure S2. Expected reward as a function of the precision (standard deviation) adopted by the observer in the low- and high-resolution conditions of Experiment 3.

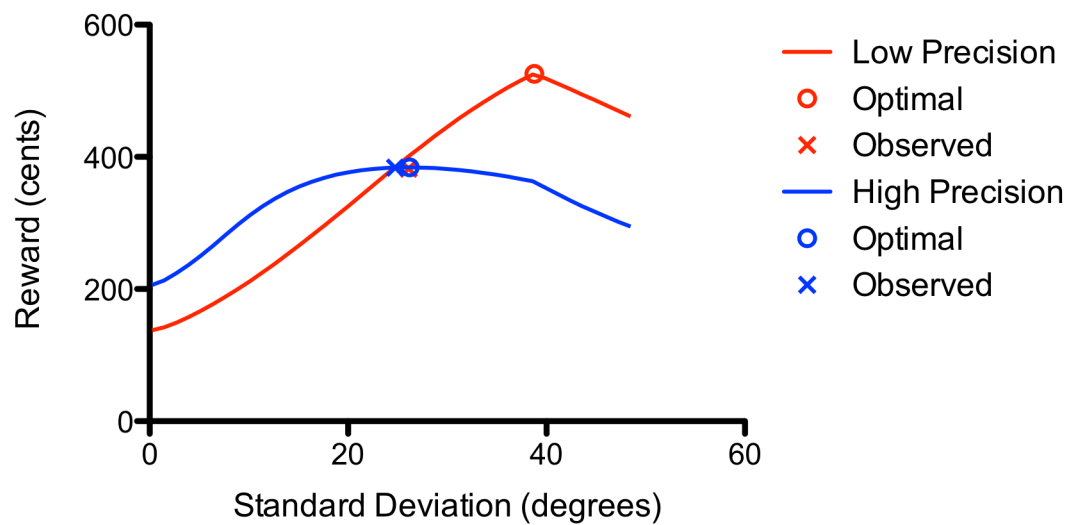


Table S1. Percentage of variance ($100 \times \text{adjusted } R^2$) explained by the quantitative model in each experiment.

	Exp. 1a	Exp. 1b	Exp. 2	Exp. 3	Exp. 4a	Exp. 4b
Low-precision	88%	82%	95%	94%	94%	89%
High-precision	96%	95%	96%	95%	91%	90%