Visual working memory is better characterized as a distributed resource rather than discrete slots

Liqiang Huang

A recent important debate in the field of visual working memory has focused on whether it represents a small set of high-precision representations (the "slot" model) or all items in parallel (the "resource" model). When faced with a large number of items, the slot model claims that high-precision representations of several items are stored and no information is retained about the other items, whereas the resource model claims that some imperfect information about each of the items can be stored. In this study, the observers tried to memorize and then recall six (out of eight possible) colors. The distribution of their scores (i.e., the number of correct responses) was modeled, and the empirical pattern of distribution fitted precisely with the prediction of the resource model but clearly differed from that of the slot model. Dependence analysis also revealed that the reports of items were approximately independent of each other, suggesting that all of the items were represented in parallel, as predicted by the resource model but not by the slot model. Overall, the data favored the resource model, not the slot model.

Keywords: visual working memory, visual attention, resource model


Introduction

Traditional models of visual working memory have usually assumed that while observers can memorize some information about what they have perceived (Phillips, 1974), there is a capacity limit to this memory, and therefore, only a fixed number of objects can be stored (Luck & Vogel, 1997; Pashler, 1988). The recent “slot” model has explicitly advocated the notion that visual working memory is composed of a fixed number of slots for “discrete fixed-precision representations” (Zhang & Luck, 2008). On the other hand, another recent model, the “resource” model, claims that, instead of there being a fixed limit on the number of objects stored in visual working memory, a fixed amount of representational resource can be flexibly allocated to represent either a small number of objects with high precision or a large number of objects with lower precision (Bays & Husain, 2008; see also Frick, 1988; Wilken & Ma, 2004).

The critical difference between these two models is the number of objects that can actually be represented in the visual working memory. The resource model suggests that when observers are faced with a large number of objects (e.g., six), the visual working memory can store (imprecise) representations of all of the objects, whereas the slot model suggests that it can only store (fairly precise) representations of a small number of objects and retains absolutely no information whatsoever about the remaining objects. However, these theories cannot be distinguished simply by estimating, from the accuracy in tasks such as comparing a probe with a memorized set of items (e.g., Cowan, 2001; Pashler, 1988), the capacity limit of visual working memory, as, even if many items have been represented, we would expect an underestimation in the number of memorized objects if the representations have low precision (Alvarez & Cavanagh, 2004; Awh, Barton, & Vogel, 2007).

In Experiment 1, the observers were asked to memorize six (out of eight possible) colors. They were then tested individually on all six colors, and the total number of correct responses for each trial was counted as their score for that trial. If the slot model is correct, we would expect their scores to be fairly constant from trial to trial. For example, if there are three slots, the observers should be able to make fairly accurate responses for three out of the six colors but, obviously, could only make random guesses for the other three; therefore, their scores should frequently be close to three and there will be little uncertainty about the outcome. Clearly, as some of the observers’ guesses may be correct or their judgments may be less than perfect (even for an object stored in a slot), their scores could be, respectively, either above or below this figure. Nevertheless, in terms of the distribution of scores, if the slot model is correct, we would expect to see an obvious peak at three. On the other hand, if the resource model is correct, the resource is evenly distributed to all six colors; therefore, the observers should have some imprecise memory about each of the six colors. If this is the case, all of their six responses, although imperfect, should be better than random guessing and should therefore result in a much wider distribution of scores than that predicted by the slot model. In other words, the observers will have some information about
each of the colors but will not really be sure about any of them. The outcome will therefore largely depend on luck: sometimes most of the reports will be correct; sometimes most of them will be incorrect. Overall, there will be considerable uncertainty about the outcome.

Experiment 1: Better empirical support for the resource model than the slot model

Observers

University undergraduate students, all of whom had normal or corrected-to-normal vision, participated in this project. There were 17 observers in Experiment 1.

Apparatus

The stimuli were presented on a 1,024 × 768 pixels CRT color monitor. The observers viewed the display from a distance of about 60 cm and entered their responses using a mouse. The program was written in Microsoft Visual Basic 6.0 and was run on Microsoft Windows XP using timing routines tested with the Blackbox Toolkit.

Stimuli

In the stimuli display (e.g., the left panel of Figure 1a), six disks (diameter of 1.04 cm) of different colors, randomly selected from eight possible colors (red, green, blue, yellow, purple, cyan, black, and white), were presented on a gray background. The six disks were evenly placed in a circle around the center of the screen (see Figure 1a) at angles of 30°, 90°, 150°, 210°, 270°, and 330°, respectively (right side = 0°), and in positions that were 5.2 cm from the center.

Procedure

The observers were asked to respond as accurately as possible but were under no time pressure (i.e., unspeeded response). At the start of each trial, a small white fixation cross was presented in the center of the screen for 400 ms. This was followed by a short blank interval (400 ms), after which the stimuli display was presented. The stimuli were presented for 800 ms and then disappeared. Immediately following the offset of the stimuli, six arrays of buttons were presented, each placed in a position corresponding to the outer side of one of the six disks (e.g., the right panel of Figure 1a). Each array consisted of eight colored buttons corresponding to the eight possible colors. The observers were required to click the button in an array that corresponded to the color of the disk in that particular position in the stimuli display; for example, if an observer believed that the disk at the top of the screen was red, he or she would click the red button of the correspondingly positioned array. After each response, the array which includes the clicked button was completely disabled so that only one response could be made for each disk. In addition, in all of the remaining arrays, the buttons of the reported color were disabled so that a color could not be reported more than once in each trial. A disk of the reported color was presented in the chosen location to allow the observer to see the choice he or she had made.

Figure 1. (a) Method of Experiment 1. (Left) The observers were asked to memorize six colors. (Right) They were then tested individually on all six colors by being required to click the appropriate button in each of the six arrays. To determine the occurrence frequency of each score, the total number of correct responses was counted as the score for each trial. (b) Results of Experiment 1. The distribution of the scores in Experiment 1, with error bars representing the 95% confidence intervals, was plotted and compared with the predictions of the slot model and the resource model. The distribution of the scores was clearly consistent with the resource model and inconsistent with the slot model.
The observers were free to report the six colors in any order. After the observers had made responses in all six arrays, a digit flashed in the center of the screen to show the score (i.e., the number of correct responses), and the next trial began 400 ms later. A sample trial, illustrating the stimuli and the procedure for making responses, is shown in Movie 1. Each observer completed 10 blocks (20 trials per block); the first block was regarded as practice and excluded from the analysis.

**Details of the modeling**

Monte Carlo simulations were used to generate predictions from both the resource and slot models. The simulation program was written in Visual basic.Net and the source codes are included in the Supplementary material. In both models, there is one parameter for the precision of representation, namely the probability of correctly memorizing a color (i.e., being able to select it correctly from among all eight choices) when it is represented; henceforth, this is referred to as P_report. The number of slots is a parameter in the slot model; henceforth, this is referred to as N_slot. In the slot model, when a color is not represented with a slot, the observer has no information at all about that color and can only make random guesses, whereas in the resource model, all of the colors are always represented (i.e., N_slot = 6). The selections of P_report and N_slot are elaborated in detail below.

In this study, the concrete procedure of the Monte Carlo simulation ran as follows:

1. **In each simulated trial, P report was used to randomly and independently determine whether each of the six colors was “memorized” or “lost.”** For example, if P_report was 0.75 for one particular color, a random real number was generated between 0 and 1, and if the random number was less than 0.75, then that color was marked as memorized; however, if the random number was greater than 0.75, then that color was marked as lost. In the resource model, as all six colors are considered to be represented, all of them would go through this “memorized/lost classification” step. In the slot model, only N_slot colors (i.e., the colors that are represented with slots) would go through this step; the other 6 − N_slot colors would be marked as “lost” straightaway.

2. **After all six colors had been placed into the two categories (i.e., memorized or lost), correct responses were simulated for all of the memorized colors, and then random responses were simulated, one after another, for the lost colors.** In this sequence of simulated responses, a color that had already been chosen (either because it was a memorized color or because it was a guess made on a lost color) was excluded from further responses.

3. **After all six responses had been made, the number of correct responses (memorized colors and lucky guesses) was counted as the score for this simulated trial.**

For each P_report value, 1,000,000 simulated trials were run to generate a “predicted distribution of the scores,” and a Pearson’s $\chi^2$ goodness-of-fit index (i.e., $\chi^2 = \sum_{i=0}^{6} \frac{(\text{Prediction}_i - \text{Observation}_i)^2}{\text{Prediction}_i}$) between this predicted distribution of scores and the observed distribution of scores in Experiment 1 was calculated to describe how good the fit was between the model and the empirical data. A greater value of this $\chi^2$ indicated a worse fit between the observed and predicted distributions. Also, this $\chi^2$ index differs from the typical use of $\chi^2$ on that it measures percentages rather than frequencies of occurrences so that its value will be much smaller.

**Comparing the predictions of the models with the results of Experiment 1**

To execute the models, the parameters had to be determined. For the slot model, N_slot can be calculated from the parameter pm reported in Zhang and Luck (2008). In their study, this parameter pm is used to describe the portion of colors that was retained in the slots. When the observers attempted to memorize six colors, pm is approximately 0.4 in their study; in other words, it was assumed that 60% of the six colors were completely lost, whereas 40% of the six colors were retained in the slots. The number of slots (i.e., N_slot) should be reasonably determined as $0.4 \times 6 = 2.4$. Therefore, N_slot should be set as 2 or 3 in the slot model. P_report was left as a free parameter in both the resource and slot models. In other words, I tried to find the P_report that would give the best fit. This best fitting P_report was found by a gradient descent optimization method with a minimum step of 0.01.

Figure 1b shows the quantitative predictions about the distribution of scores from both the resource model and the slot models and then compares these with the empirical data from Experiment 1. In terms of the predictions of the models, the distribution of scores was clearly consistent with the resource model ($\chi^2 = 0.0089$, obtained when P_report = 0.45) and inconsistent with the slot model for both N_slot = 2 ($\chi^2 = 1.89$, obtained when P_report = 0.92) and N_slot = 3 ($\chi^2 = 0.33$, obtained when P_report = 0.79).

All of the modeling discussed so far fitted the aggregated distribution of the scores of all 17 observers to the models. To assess the models further, I fitted the distribution of the scores of each individual observer to both the resource and the slot models and found that the fit was significantly worse for the slot models of both N_slot = 2 and N_slot = 3 than for the resource model (N_slot = 3, $t(16) = 3.47$, $p < 0.002$; N_slot = 2, $t(16) = ...
4.10, $p < 0.001$). This individual-$\chi^2$-based comparison will hereafter be used to assess the difference between the models.

**Results of Experiment 1:**  
**Dependence analysis**

A dependence analysis could also shed light on the issue of distinguishing between the resource model and the slot model. Concretely, I picked $X$ of the six colors ($X = 1, 2, 3, 4, 5$) and divided the trials into $X + 1$ groups according to how many correct responses ($0 \sim X$) had been made for the $X$ colors, and then the average score of the remaining $6 - X$ colors was calculated for each group. The predictions of the slot model ($N_{\text{slot}} = 3$) and resource model were also made by running dependence analysis on the simulated trials; these predictions were then plotted for comparison. The slot model predicts that the representations of the colors will depend (negatively) on each other (i.e., a downward trend in these graphs), whereas the resource model predicts that representations of the colors will be roughly independent of each other (i.e., a flat trend in these graphs). The resource model was supported and the slot model was rejected.

Figure 2. A dependence analysis. I picked $X$ of the six colors ($X = 1, 2, 3, 4, 5$). The trials were divided into $X + 1$ groups according to how many correct responses ($0 \sim X$) had been made for the $X$ colors, and then the average score of the remaining $6 - X$ colors was calculated for each group. The predictions of the slot model ($N_{\text{slot}} = 3$) and resource model were also made by running dependence analysis on the simulated trials; these predictions were then plotted for comparison. The slot model predicts that the representations of the colors will depend (negatively) on each other (i.e., a downward trend in these graphs), whereas the resource model predicts that representations of the colors will be roughly independent of each other (i.e., a flat trend in these graphs). The resource model was supported and the slot model was rejected.

score of the remaining $6 - X$ colors for each group. If the slot model is correct, then the following logic is reasonable: Only two or three colors are stored in high-precision slots; if the score is greater on the chosen $X$ colors, then it is likely that more of the high-precision slots have been spent on these $X$ colors, and we would expect a lower score for the remaining $6 - X$ colors. On the other hand, if the resource model is correct and some amount of resource is spent on representing each color, it is plausible that the representations of all of the items are independent of each other, and we would expect a constant score for the remaining $6 - X$ colors, regardless of the score on the chosen $X$ colors.
Figure 2 plots the results of dependence analysis for $X$ as 1, 2, 3, 4, and 5. The predictions of the slot model ($N_{\text{slot}} = 3$) and the resource model were also made by running a dependence analysis on the simulated trials; these predictions were then plotted for comparison. As illustrated above, the slot model predicts that the score for the remaining $6 - X$ colors will depend negatively on the score for the chosen $X$ colors. Interestingly, the resource model actually predicts a small positive dependence. This is because making correct responses helps subsequent responses by removing incorrect colors from the choices available for these subsequent responses. For example, if the first response is a random guess, it only has a 1/8 probability of being correct; however, after five correct responses have been made, the sixth response, if it is a random guess, has a 1/3 probability of being correct. The empirical data from Experiment 1 clearly fit well with the predictions of the resource model but not with the predictions of the slot model, thus again supporting the resource model and rejecting the slot model. The individual-$\chi^2$-based comparison as described above also confirmed that the fit was systematically worse for the slot model ($N_{\text{slot}} = 3$) than for the resource model (dependence on 1 color, $t(16) = 2.21, p < 0.05$; dependence on 2 colors, $t(16) = 3.20, p < 0.01$; dependence on 3 colors, $t(16) = 4.15, p < 0.001$; dependence on 4 colors, $t(16) = 4.62, p < 0.001$; dependence on 5 colors, $t(16) = 4.45, p < 0.001$).

**Experiment 2: Multiple responses had little impact on the scores**

Experiment 2 addressed one potential concern of the results of Experiment 1. An important difference between the present Experiment 1 and previous studies is that here the observers were required to make multiple responses, whereas in previous studies only one response was usually required for each trial. Therefore, it might be argued that the wide distribution observed in Experiment 1 was a result of the interference caused by making multiple responses; in other words, the observers may have forgotten some of the colors when they were making their responses and this added variations to the scores. In Experiment 2, the observers were tested on all six items in half of the blocks (i.e., as in Experiment 1), but, in the other half of the blocks, they were only tested on one item. By comparing the performance between the six-item test and the one-item test, we can know whether interference caused by multiple responses would make a significant difference.

**Methods**

There were 10 observers in Experiment 2. The method of Experiment 2 was identical to that of Experiment 1 except that, in half of the blocks, the observers were tested on only one item. The six- and one-item test blocks alternated and the starting block was counter-balanced across observers. The observers completed 10 blocks (25 trials per block); the first two blocks were regarded as practice and were therefore excluded from the analysis.

**Results and discussion**

The average scores were, respectively, 3.18 and 0.47 in the six-item and one-item tests. After considering and removing the contribution of lucky guesses, the probability of successfully memorizing a color (i.e., $P$) was 0.43 and 0.40 in the six-item and one-item tests, respectively. Therefore, the interference and forgetting caused by making multiple responses clearly had little impact.

Experiment 2 also helped address another potential concern on the results of Experiment 1. One may point out that in Experiment 1, the arrays of response buttons immediately followed the offset of the stimuli, and therefore, the iconic (sensory) memory could potentially contribute to the performance. Experiment 2 argues against this possibility. If iconic memory did make a significant contribution to the performance in the present method, then we would expect the one-item test to be substantially better than the six-item test in Experiment 2. This is because in the one-item test the observers can always read out the information of the single color from iconic memory, but in the six-item test, they could only read out the information of the first of the six colors. Therefore, the result of Experiment 2 ensures that iconic memory made no significant contribution in the present method.

**General discussion**

**Individual differences**

One possible objection to the present results is that the large variation in the distribution of scores in the data was mainly the result of individual differences; for example, some observers may have had one slot and others five slots, resulting in the distribution of scores being spread out. However, this possibility can be ruled out, as the mean scores were fairly consistent across the observers: in the empirical data of Experiment 1, the average score of the 17 observers was 3.295, with a standard deviation (as a description of the individual differences) of only 0.433. To estimate the effect of these slight individual differences, a slot model was simulated in which half of the observers had two slots (i.e., $N_{\text{slot}} = 2$) and the other half had three slots (i.e., $N_{\text{slot}} = 3$). The resulting standard deviation on the score was 0.436, which was similar to what was actually observed in Experiment 1. The predicted distribution from this model is plotted in the
blue curve of Figure 1b: the distribution is indeed slightly stretched out, but it clearly still does not fit the empirical data in Experiment 1 ($\chi^2 = 0.52$, obtained when $P_{\text{report}} = 0.89$), and the fit is statistically worse than the resource model ($t(16) = 3.60$, $p < 0.002$).

The dependence analysis also helped rule out the alternative explanation based on individual differences, because, in this analysis, the scores were first calculated for each individual observer before being summarized into the overall average. Therefore, individual differences, no matter how large, would not eliminate the negative dependence predicted by the slot model.

A slot model with a variable number of slots from trial to trial

The slot model could perhaps be defended by suggesting that the number of slots is not fixed and could vary randomly from trial to trial. For example, Luck (2008) speculated that a fixed quantity of the neural mechanism is used to represent the items and that the system always strives to keep a “fixed precision” for the represented items; therefore, when the noise is high, the neural mechanism will spend more representational mechanism on each item and cover less items, and, when the noise is low, it will spend less representational mechanism on each item and cover more items. So the system would work as fixed precision slots, but the number of slots could vary randomly from trial to trial, depending on the level of noise in the neural system.

This model only differs from the resource model on a minor point, namely that it ensures that a sufficient amount of resource is spent on the represented items for fixed precision, whereas the resource model does not have this constraint. Although this model is still called a “slot” model by Luck (2008), it seems that it should be more fairly referred to as one type of resource model (i.e., a fixed precision resource model) because it explicitly embraces the central idea of “resource,” which is that a whole amount of representational resource (or representational mechanism) can be allocated to more or less items depending on the need. On the other hand, it is very unlike the typical concept of slots (i.e., a fixed set of vacancies that can be filled in); therefore, calling this model a “slot model,” as if it is clearly against the concept of resource, is conceptually misleading. Nevertheless, the naming of a model is not very important, and, hereafter, I will refer to this model neutrally as the “fixed precision model” so that it will not be confused with either the slot model or the resource model.

I modeled this fixed precision model to see how it fitted the distribution of scores in Experiment 1. As mentioned above, Zhang and Luck (2008) indicated that the average N_slot should be 2.4, so the N_slot in an individual trial was simulated as a normal distribution centered on 2.4. The standard deviation of this distribution, $\sigma_N$, is a new parameter of this model. Reasonably, the N_slot should be assumed to be a Gaussian distribution on a logarithmic scale. So, the number of slots in each trial is generated by randomly picking a sample value from $N_{\text{slot}} \cdot \text{Normal}(0, \sigma_N)$ and then rounding it to the closest integral number. Both $\sigma_N$ and $P_{\text{report}}$ are free parameters. As above, the best fitting $\sigma_N$ is also found by a gradient descent optimization method, with a minimum step of 0.01. The best $\chi^2$ was found to be 0.098, obtained when $\sigma_N = 0.56$ and $P_{\text{report}} = 0.92$. Even at this point, the fit is still significantly worse than the resource model ($t(16) = 3.86$, $p < 0.001$).

Another important disadvantage of the fixed precision model is that it has two free parameters ($\sigma_N$ and $P_{\text{report}}$), whereas the resource model has only one parameter ($P_{\text{report}}$). Therefore, by the principle of Occam’s razor, the resource model should be preferred due to its parsimony. Indeed, in the recent methods (e.g., Bayes factors) that have been put forward to compare models, such a difference in complexity can often be decisive in favoring one model over another when two models are, in other respects, similarly good.

A slot model with large N_slot

So far, we have only considered slot models with an N_slot of no more than 3. This estimation is reasonable because the N_slot obtained from Zhang and Luck (2008) is merely 2.4 and other previous estimations (e.g., Cowan, 2001) have rarely been higher than 3.

Nevertheless, if the slot model is allowed to have more slots, it will naturally become more similar to a resource model, and therefore, the fit will improve. Eventually, if the slot model is allowed six slots, it will become completely identical to the resource model. In Figure 3, I plotted the $\chi^2$s for models with different number of slots for both the slot model and the fixed precision model. For the slot model, the fit is significantly worse than the resource model when $N_{\text{slot}} = 4$ ($\chi^2 = 0.057$, obtained when $P_{\text{report}} = 0.64$, $t(16) = 1.87$, $p < 0.05$) and becomes indistinguishable from the resource model when $N_{\text{slot}} = 5$ ($\chi^2 = 0.011$, obtained when $P_{\text{report}} = 0.53$). For the fixed precision model, the fit is significantly worse than the resource model when $N_{\text{slot}} = 3$ ($\chi^2 = 0.032$, obtained when $\sigma_N = 0.44$ and $P_{\text{report}} = 0.82$, $t(16) = 2.58$, $p < 0.02$) and becomes indistinguishable from the resource model when $N_{\text{slot}} = 4$ ($\chi^2 = 0.011$, obtained when $\sigma_N = 0.31$ and $P_{\text{report}} = 0.66$), although, as mentioned above, by the principle of Occam’s razor, the fixed resolution model is still less desirable at this point.

All in all, it is clear that neither the slot model nor the fixed resolution model can be as good as the resource model unless the number of slots is rather large. Therefore, at the very least, if the visual working memory indeed has a limit on its number of slots, the number of
Alternative account for the all-or-none finding

I have shown above that the slot model is ruled out by the data. The fixed precision model is also not very plausible. To disprove the notion of “fixed precision slots,” one needs to explain the all-or-none findings reported in Zhang and Luck (2008). Indeed, this finding has been challenged on an empirical basis in further studies. For example, Bays, Catalao, and Husain (2009) pointed out that the correct report of colors depends not only on a memory for colors but also on a memory for the locations of these colors. If observers are fairly good at memorizing colors but make a lot of errors in memorizing their locations, then they could very often report the color of a wrong item. The colors of different items are uncorrelated with each other; therefore, if one only plots the distribution of reports in relation to the target color, the reports caused by the “color of wrong items” would appear to be pure random guessing (i.e., the flat line on the distribution of reports). Bays et al. (2009) went further and showed that when observers were asked to memorize six items, the data analysis method used by Zhang and Luck (2008) indicated that 48% of the items were completely lost in the memory and that the responses were random guesses; however, there was strong evidence to indicate that these random guess responses were often actually caused by the memory of other nontarget items (i.e., items on the wrong locations). After location errors were appropriately considered and compensated for, the portion of random guess responses dropped dramatically to only 14% of all responses. In brief, the majority of the responses that Zhang and Luck (2008) attributed to random guessing could actually be caused by observers misremembering which color was at the tested location.

Conclusion

In this study, observers were asked to report six memorized colors and their scores (i.e., number of correct responses) were plotted. We compared the predictions of different models with this distribution of scores. The results of this comparison favored the resource model and clearly ruled out the slot model as typically conceptualized. A slot model with more and variable slots could fit the data but, overall, is still less desirable than the resource model. Furthermore, there are reasons to believe that the all-or-none finding upon which the slot model is based could be explained by alternative accounts (e.g., memory errors about the locations of colors).

Acknowledgments

This research was supported by RGC Grant 444108 from the Hong Kong government. I thank two anonymous reviewers for their very helpful comments on an early draft of this article.

Commercial relationships: none.

Corresponding author: Liqiang Huang.
Email: lqhuang@psy.cuhk.edu.hk.
Address: Department of Psychology, The Chinese University of Hong Kong, Hong Kong, The People’s Republic of China.

Footnote

1My assumption regarding the even distribution of resource does not oppose the idea that observers can, voluntarily, pay attention to a subset of objects and have a better representation of these objects than others. In the present study, the task required the observers to memorize all of the colors presented, so an even distribution is
plausible. However, if a task required observers to voluntarily memorize only one subset and to ignore the others, then it is quite possible that the representational resources would only be allocated to that subset.

References


