

# An assessment of fixed-capacity models of visual working memory

Jeffrey N. Rouder\*<sup>†</sup>, Richard D. Morey\*, Nelson Cowan\*, Christopher E. Zwilling\*, Candice C. Morey<sup>‡</sup>, and Michael S. Pratte\*

\*Department of Psychological Sciences, 210 McAlester Hall, University of Missouri, Columbia, MO 65211; and <sup>‡</sup>Department of Psychology, One Brookings Drive, Washington University, St. Louis, MO 63130

Edited by Richard M. Shiffrin, Indiana University, Bloomington, IN, and approved February 20, 2008 (received for review November 29, 2007)

**Visual working memory is often modeled as having a fixed number of slots. We test this model by assessing the receiver operating characteristics (ROC) of participants in a visual-working-memory change-detection task. ROC plots yielded straight lines with a slope of 1.0, a tell-tale characteristic of all-or-none mnemonic representations. Formal model assessment yielded evidence highly consistent with a discrete fixed-capacity model of working memory for this task.**

working memory | capacity | mathematical models of memory | short-term memory

The study of the nature and capacity of visual working memory (WM) is both timely (1) and controversial (2, 3). A popular conceptualization is that visual WM consists of a fixed number of discrete slots in which items or chunks are temporarily held (2, 4, 5). Nonetheless, there are dissenting viewpoints in which the discreteness is taken as, at most, a convenient oversimplification (6, 7). In this article, we provide a rigorous test of the fixed-capacity model for a visual WM task. Herein, we apply this test to items that differ in color, although the test is suitable to examine the generality of capacity limits across various materials.

We used a common version (8–15) of the task popularized by Luck and Vogel (4, 16) (see Fig. 1A). At study, participants are presented with an array of colored squares. At test, a single square is presented; this square is either the same color as the corresponding square in the study array (a “same trial”) or a novel color (a “change trial”). Participants simply decide whether the test square is the same as or different from the corresponding studied square. In this task, where the color of each square is unique and the colors are well separated, capacity is the number of squares (objects) that may be held in visual WM. This object-based view of capacity is supported by previous research (4), in which performance does not vary with the number of manipulated features per object.

Previous demonstrations of fixed capacity have relied on plotting capacity estimates as a function of the number of to-be-remembered items. Fixed capacity is claimed because capacity estimates tend to asymptote at three to four items for array sizes of four to six items. This approach, however, is not the most rigorous for this model. There are three weaknesses in previous demonstrations: (i) the asymptote of the capacity estimated may be mimicked by models without recourse to fixed capacity; (ii) previous demonstrations are made with aggregate data, and an asymptote in the group aggregate does not necessarily imply asymptotes in all or any individuals; and (iii) the stability of these asymptotes has not been formally assessed. These weaknesses motivate a more constrained test, to be presented subsequently.

**The Fixed-Capacity Almost-Ideal Observer Model.** We define the fixed-capacity ideal observer as one who maximizes the probability of a correct response given the constraint that visual WM is discrete and limited in the number of items that may be held. Here, we derive the ideal observer model and show that it is

closely related to Cowan’s formula (2, 17) for visual WM capacity. Cowan’s formula has been applied in a growing number of studies (9, 18, 19), often in combination with electrophysiological measures (20, 21) or functional neuroimaging (12–14). The measures seem to converge on a human capacity of approximately four simple objects in WM. However, all of this work is tenuous inasmuch as the theoretical assumptions underlying the model have not been examined rigorously. Later in this section, we relax an assumption of the model to approximate ordinary nonideal human decision processes (almost-ideal observer model) and, in *Results*, we include effects of inattention to the display.

The ideal observer conditions her or his response on whether the item is in memory. If so, the ideal observer responds accordingly, and performance is perfect. The probability that a cued square is in memory is a function of capacity, denoted  $k$ , and the number of squares in the study array (array size), denoted  $M$ . If capacity is as great as the number of squares, all may be held in memory. If capacity is smaller, however, only  $k$  may be held, and the probability of any one square being in memory is  $k/M$ . Combining these facts yields:

$$Pr(\text{cued element in memory}) = \min\left(1, \frac{k}{M}\right).$$

If the test item is out of memory, the ideal observer responds “change” only if a change is more likely *a priori*. Let  $g$  denote the probability of a change response when the item is out of memory, and let  $\pi$  be the probability of a change trial. Then,

$$g = \begin{cases} 0 & \pi < .5, \\ b & \pi = .5, \\ 1 & \pi > .5. \end{cases}$$

The parameter  $b$  is a bias that holds if change and same trials are equally likely (i.e.,  $\pi = 0.5$ ) and does not affect the overall probability of a correct answer. The above equation is valid when the probability that the tested item changed ( $\pi$ ) is known. It is not valid for other paradigms in which any one of several items presented at test may have changed (e.g., ref. 22).

Model predictions are easily derived for hit and false alarm rates, the probability of a “change” response for change and same trials, respectively. Let these rates be denoted by  $h$  and  $f$ , respectively. A hit occurs if an item is remembered, or failing

Author contributions: J.N.R., R.D.M., N.C., C.E.Z., C.C.M., and M.S.P. designed research; R.D.M. and C.E.Z. performed research; J.N.R. and R.D.M. analyzed data; and J.N.R. and N.C. wrote the paper.

The authors declare no conflict of interest.

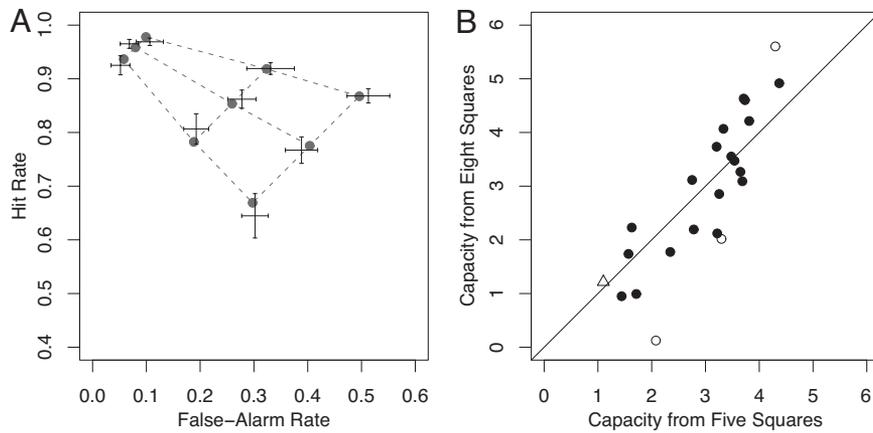
This article is a PNAS Direct Submission.

<sup>†</sup>To whom correspondence should be addressed. E-mail: rouderj@missouri.edu.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0711295105/DCSupplemental](http://www.pnas.org/cgi/content/full/0711295105/DCSupplemental).

© 2008 by The National Academy of Sciences of the USA





**Fig. 2.** Results. (A) ROC results (2D error bars) and predictions (lines and small points) from the five-parameter fixed-capacity model. Error bars denote standard errors of means in each measure. (B) Scatter plot of capacity in the eight-square arrays as a function of that in the five-square arrays. Filled points indicate participants for whom fixed capacity may not be rejected in favor of variable capacity. Open points are exceptions; see text.

below a criterion, respectively. When parameterized in terms of likelihood ratios, the model has a natural selective-influence prediction: (i) set size should affect only the memory strength parameter,  $d'$ , and not the criterion; and (ii) change probability should affect only the criterion and not the memory strength parameter. A seven-parameter selective-influence model is constructed with three  $d'$  parameters (one for each set size), three criteria (one for each change probability), and  $\sigma^2$ . A simpler six-parameter version is constructed by assuming equal variance in mnemonic strength (i.e.,  $\sigma^2 = 1$ ). Derivations for this model are provided as [supporting information \(SI\)](#).

## Results

Averaged hit and false alarm rates are shown in Fig. 2A as two-dimensional error bars. These points fall at the vertices of straight lines. Moreover, the points lie very close to an isosensitivity line of slope 1.0, as predicted by the fixed-capacity theory.

The fixed-capacity model may be formally fit by assuming that the frequencies of hits and false alarms, denoted  $H_{ij}$  and  $F_{ij}$ , respectively, are distributed as conditionally independent binomials:

$$H_{ij} \stackrel{\text{ind}}{\sim} \text{Binomial}(h_{ij}, N_{ij}^{(c)}),$$

$$F_{ij} \stackrel{\text{ind}}{\sim} \text{Binomial}(f_{ij}, N_{ij}^{(s)}),$$

where  $N_{ij}^{(c)}$  and  $N_{ij}^{(s)}$  are the number of change and same trials, respectively, for the  $ij$ th condition. For the experiment with three levels of change probability ( $J = 3$ ), there are four parameters:  $k$ ,  $g_1$ ,  $g_2$ , and  $g_3$ . Estimation is performed by maximizing likelihood (29), and goodness of fit is assessed by comparing the likelihood of the model to that of a vacuous binomial model in which there are no constraints on  $h_{ij}$  and  $f_{ij}$ . For the experiment, the total number of parameters in the vacuous model is  $I \times J \times 2 = 18$ . The substantive model can be tested against the vacuous model with a log-likelihood test statistic (ref. 30) ( $G^2$ ). Given the large number of trials collected per individual, the model may be fit to individual rather than aggregated data.

The four-parameter fixed-capacity model fits poorly for 16 of 23 participants, as indicated by  $G^2$  fit statistics that correspond to  $P$  values  $< 0.05$  [ $G^2(14) > 23.68$ ]. The reason for these poor fits is both easy to diagnose and of secondary importance. One facet of the data is that accuracy with two squares at study is near but not at ceiling. Accuracy averaged 0.945 with none of the 23 participants achieving perfect performance. Any error in the two-square condition implies that capacity for all set sizes is  $< 2.0$ . The data from the higher set sizes, however, are compat-

ible with higher-capacity estimates, which leads to poor fits of the fixed-capacity model.

**Adding Attention.** It seems plausible that below-ceiling performance with two squares may reflect an occasional lapse of attention. The fixed-capacity model, unfortunately, is not at all robust to this misspecification. For example, even for a person with a very high true capacity, the presence of a single error in the two-item condition will result in no likelihood for  $k \geq 2$ . This lack of robustness can easily be modified by assuming that performance on each trial is a mixture of attentive and inattentive states. When the participant is attentive, performance is governed by the fixed-capacity model. When the participant is inattentive, the hit and false alarm rate probabilities are governed by the guessing parameters. The following five-parameter model reflects this role of attention:

$$h_{ij} = a(d_i + (1 - d_i)g_j) + (1 - a)g_j, \quad [6]$$

$$f_{ij} = a(1 - d_i)g_j + (1 - a)g_j, \quad [7]$$

where  $a$  denotes the probability that attention is engaged on a trial. This model also predicts straight-line ROCs with a slope of 1.0. Decreases in  $a$  lower the intercept of the ROC line. Thus, the five-parameter model is still highly falsifiable in the current experiment.

The addition of an attention parameter represents a type of trial-to-trial variability in the underlying process. In this case, the variability is coarse, varying between full attention and no attention. In a more fine-grained view, attention may vary across several levels. The experimental paradigm and subsequent data, however, do not provide sufficient precision to adjudicate between the all-or-none approach to trial-by-trial variability and a more fine-grained view. The current approach is simple and tractable and provides a parsimonious account of the data.

The five-parameter model was fit to each individual by maximizing likelihood. Overall capacity was at 3.35 items, and participants paid attention on 88% of trials. Guessing rates were  $\hat{g} = (0.47, 0.64, 0.79)$  for the base rates of  $\pi = (0.3, 0.5, 0.7)$ , respectively. These guessing rates indicate a biased and probabilistic response strategy when items were out of memory and are approximately concordant with previous probability matching results (23, 24).

The model fit sufficiently well for 20 of the 23 participants, as indicated by  $G^2$  fit statistics that correspond to  $P$  values greater than the 0.05 criterion [ $G^2(13) < 22.36$ ]. Individualized maximum-likelihood estimates of model parameters are used to

**Table 1. Model-selection statistics**

Model-selection statistic	Discrete capacity			Signal detection	
	Full attention, four parameters	Fixed capacity, five parameters	Variable capacity, six parameters	Equal variance, six parameters	Unequal variance, seven parameters
Akaike Information Criterion	10,304.0	9,774.3*	9,775.8	9,791.5	9,796.8
Bayesian Information Criterion	10,674.3	10,237.3*	10,331.4	10,347.1	10,444.9
Normalized Maximum Likelihood	10,505.3	10,025.9*	10,077.7	10,093.4	10,149.0

\*Indicates lowest value (best fit) across models.

derive individualized ROC predictions for each participant. The average of these predictions is also shown in Fig. 2A as the smaller points connected by dotted lines. As can be seen, the averaged predictions are within the standard errors of averaged data. That is, the five-parameter fixed-capacity model does an excellent job of accounting for the selective-influence manipulations.

**Fixed vs. Variable Capacity.** To test the fixed-capacity assumption more critically, we contrasted it to a six-parameter variable-capacity model with separate capacities for each array size. This model had six parameters: three capacities and three guessing parameters.<sup>§</sup> The five-parameter model can be rejected in favor of this six-parameter alternative for 4 of the 23 participants [ $G^2(1) > 3.84$ ]. The nature of these few constant-capacity violations may be seen in the scatter plots of capacity estimates in Fig. 2B for the five- and eight-square arrays. Had capacity been exactly the same across array sizes, then the points would lie on the diagonal. The open circles indicate participants for whom capacity differs between the five- and eight-square arrays. The triangle indicates a participant who also violates fixed capacity; in this case, the participant had a capacity of 1.8 for two-square arrays but a capacity of 1.15 for five- and eight-square arrays. Overall, however, there is no apparent trend away from the diagonal; that is, the distribution of capacities across individuals does not appear to vary across the array size conditions. This fact serves as supporting evidence for the fixed-capacity model.

Even though the fixed-capacity assumption holds across a majority of participants, it does fail for a few. For a few participants, capacity decreases markedly with increasing array size. We suspect these participants may have been intimidated by the larger arrays and failed to encode much of them. Only one participant showed substantially increasing capacity across all three set sizes; perhaps this participant tried harder to encode the study array when more items were presented (31). In sum, fixed capacity is the norm, although subtle individual response characteristics, which seem orthogonal to the process of interest, are observable, too.

**A Comparison to Signal Detection.** We also benchmarked the five-parameter fixed-capacity model against the signal-detection models. The six-parameter (equal variance) and seven-parameter (unequal variance) signal detection models fit well compared with the vacuous model for 19 and 20 of 23 participants, respectively. Whereas the signal-detection and discrete-capacity models are not nested, model comparisons are made with the following three model-selection statistics: Akaike information criterion (32), Bayesian information criterion (33), and an asymptotic approximation to normalized maximum likelihood (34, 35). We computed omnibus model selection statistics by computing the total likelihood of all parameters. For

these three model-selection statistics, lower values indicate better fit. As shown in Table 1, the methods provide for converging results: the fixed capacity model fits best followed by the variable-capacity discrete-model and signal-detection models. As a final check of these model-selection results, we constructed bootstrapped sampling distributions (36) of the difference in deviance between the five-parameter discrete-capacity model and the six-parameter signal-detection model. Two such distributions were constructed, each assuming that one of the models being compared was true. These two distributions were well separated ( $z = 4.9$ ), and the observed difference in deviance favored the fixed-capacity model.

**Discussion**

We have provided strong experimental support for a fixed-capacity model of visual WM in a task in which participants are asked to remember squares of various colors. Observed ROC functions have slopes near 1.0, and the five-parameter fixed-capacity model fits well when compared with a vacuous binomial model, a variable-capacity discrete model, and a variable-capacity signal-detection alternative. Perhaps the greatest advantage of the fixed-capacity model is its simplicity; it explains the extant data with far more parsimony than variable-capacity competitors. The paradigm and model are therefore well suited for exploring more advanced aspects of human visual WM, such as the role of chunking (5).

Although the fixed-capacity model fits well overall, there are violations in some participants. We suspect these may reflect idiosyncratic response characteristics, for example, being intimidated by large array sizes. Researchers need be aware of these possibilities, especially when attempting individual-level capacity estimation.

Software for fitting the fixed-capacity model across several array sizes is available at [web.missouri.edu/~umcaspsychpl](http://web.missouri.edu/~umcaspsychpl).

**Methods**

**Participants.** Twenty-three students from an introductory psychology class at the University of Missouri, Columbia, served as participants.

**Design.** Change probability ( $\pi = 0.3, 0.5, 0.7$ ) and array size ( $M = 2, 5, 8$ ) were manipulated in a within-participant factorial design. Change probability was held constant for blocks of 60 trials, whereas array size varied from trial to trial. The dependent variable of interest was the number of hits and false alarms in each condition.

**Stimuli.** Study arrays were squares whose colors were sampled without replacement from 10 colors (black, white, red, blue, green, yellow, orange, cyan, purple, dark-blue-green). Squares were randomly positioned on the screen as described previously (17). Patterned masks consisted of identical multicolored squares as in Fig. 1A. Stimuli were presented on 17" cathode ray tube monitors (640 by 480 pixels, 120-Hz refresh).

**Procedure.** The structure of a trial is shown in Fig. 1A. Participants depressed one of two keys on a keyboard to indicate whether the test square was the same as or different from the corresponding square in the study array. The experiment was composed of nine blocks of 60 trials each, for a total of 540 trials. To make the change-probability manipulation salient from the outset of

<sup>§</sup>The seven-parameter model with a separate-capacity parameter for each array-size condition and an attention parameter is not identifiable.

a block, participants were first shown a pie chart of the change probability. The change probability was also presented, as a percentage, before every trial. The session took  $\approx 45$  min to complete.

Our experimental procedure had three features that may be necessary to isolate WM capacity. The first feature is that only a single square was presented at test. In a separate pilot experiment, we presented all squares at test and cued a specific target square by encircling it. In this case, there was evidence that capacity was not constant but rose with increasing array size. We attribute this phenomenon to the use of nontested squares as contextual cues, possibly allowing chunking or grouping of squares (15). Participants may quickly learn that squares can be stored and successfully retrieved as a group. The presentation of a single square at test may lessen the benefit of grouping.

The second feature is that each studied square has a unique color within the array. It is our experience that if grouping were to happen, it would be across squares of the same color. Therefore, the constraint of unique colors limits grouping. The final feature was the insertion of a patterned mask between study and test. The patterned mask is useful, because it allows the relative participant to know the relative location of the tested square to the other squares and prevents any residual of perceptual or iconic memory (37) from contributing to the capacity score (18).

**ACKNOWLEDGMENTS.** This research is supported by National Science Foundation Grant SES-0351523, National Institute of Mental Health Grant R01-MH071418, and National Institutes of Health Grant RO1-HD21338.

- Osaka N, Logie RH, D'Esposito M (2007) *The Cognitive Neuroscience of Working Memory* (Oxford Univ Press, Oxford, UK).
- Cowan N (2001) The magic number 4 in short-term memory: A reconsideration of Mental Storage Capacity. *Behav Brain Sci* 24:87–114.
- Miyake A, Shah P (1999) *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (Cambridge Univ Press, Cambridge, UK).
- Luck SJ, Vogel EK (1997) The capacity of visual working memory for features and conjunctions. *Nature* 390:279–281.
- Miller GA (1956) The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychol Rev* 63:81–97.
- Oberauer K, Kliegl R (2006) A formal model of capacity limits in working memory. *J Mem Lang* 55:601–626.
- Wilken P, Ma WJ (2004) A detection theory account of change detection. *J Vis* 4:1120–1135.
- Wheeler ME, Treisman AM (2002) Binding in short-term visual memory. *J Exp Psychol* 131:48–64.
- Fougnie D, Marois R (2006) Distinct capacity limits for attention and working memory: Evidence from attentive tracking and visual working memory paradigm. *Psychol Sci* 17:526–534.
- Olsson H, Poom L (2005) Visual memory needs categories. *Proc Natl Acad Sci USA* 102:8776–8780.
- Cowan N, Morey CC (2007) How can dual-tasks working memory retention limits be investigated? *Psychol Sci* 18:686–688.
- Todd JJ, Marois R (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428:751–754.
- Song J, Jiang Y (2006) Visual working memory for simple and complex features: An fMRI study. *NeuroImage* 30:963–972.
- Xu Y, Chun MM (2006) Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440:91–95.
- Jiang Y, Chun MM, Olson IR (2004) Perceptual grouping in change detection. *Percept Psychophys* 66:446–453.
- Phillips WA (1974) On the distinction between sensory storage and short-term visual memory. *Percept Psychophys* 16:283–290.
- Cowan N, et al. (2005) On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognit Psychol* 51:42–100.
- Saults JS, Cowan N (2007) A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *J Exp Psychol* 136:663–684.
- Awh E, Vartan B, Vogel EK (2007) Visual working memory represents a fixed number of items regardless of complexity. *Psychol Sci* 18:622–628.
- Vogel EK, Machizawa MG (2004) Neural activity predicts individual differences in visual working memory capacity. *Nature* 428:749–751.
- Vogel EK, McCollough AW, Machizawa MG (2005) Neural measures reveal individual differences in controlling access to working memory. *Nature* 438:500–503.
- Pashler H (1988) Familiarity and visual change detection. *Percept Psychophys* 44:369–378.
- Bush RR, Mosteller F (1951) A mathematical model for simple learning. *Psychol Rev* 58:313–323.
- Estes WK (1962) Learning theory. *Annu Rev Psychol* 13:107–144.
- Sternberg S (1969) in *Attention and Performance*, ed Kosner WG (North-Holland, Amsterdam), Vol II, pp 276–315.
- Egan JP (1975) *Signal Detection Theory and ROC Analysis* (Academic, New York).
- Kintsch W (1967) Memory and decision aspects of recognition learning. *Psychol Rev* 74:496–504.
- Green DM, Swets JA (1966) *Signal Detection Theory and Psychophysics* (Wiley, New York).
- Nelder JA, Mead R (1965) A simplex method for function minimization. *Comput J* 7:308–313.
- Bishop YMM, Fineberg SE, Holland PW (1975) *Discrete Multivariate Analysis: Theory and Practice* (MIT Press, Cambridge, MA).
- Wolfe JM, Horowitz TS, Kenner NM (2005) Rare items often missed in visual searches. *Nature* 435:439–440.
- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Aut Control* 19:716–723.
- Schwartz G (1978) Estimating the dimension of a model. *Ann Stat* 6:461–464.
- Myung JI, Navarro DJ, Pitt MA (2006) Model selection by normalized maximum likelihood. *J Math Psychol* 50:167–179.
- Rissanen J (1983) A universal prior for integers and estimation by minimal description length. *Ann Stat* 11:416–431.
- Wagenmakers E-J, Ratcliff R, Gomez P, Iverson GJ (2004) Assessing model mimicry using the parametric bootstrap. *J Math Psychol* 48:28–50.
- Sperling G (1960) The information available in brief visual presentations. *Psychol Monogr* 74 (Whole No. 498).