TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 9 (2017) 83–101

ISSN:1756-8757 print / 1756-8765 online
DOI: 10.1111/tops.12248

This article is part of the topic "Best of Papers from the 2016 International Conference on Cognitive Modeling," David Reitter and Frank E. Ritter (Topic Editors). For a full listing of topic papers, see http://onlinelibrary.wiley.com/doi/10.1111/tops.2017.9.issue-1/issuetoc.

Visual Working Memory Resources Are Best Characterized as Dynamic, Quantifiable Mnemonic Traces

Bella Z. Veksler,[a] Rachel Boyd,[a] Christopher W. Myers,[b] Glenn Gunzelmann,[b] Hansjörg Neth,[c] Wayne D. Gray[d]

[a]*Oak Ridge Institute for Science & Education at AFRL*
[b]*U.S. Air Force Research Laboratory*
[c]*Department of Psychology, University of Konstanz*
[d]*Cognitive Science Department, Rensselaer Polytechnic Institute*

Received 25 October 2016; received in revised form 9 November 2016; accepted 15 November 2016

Abstract

Visual working memory (VWM) is a construct hypothesized to store a small amount of accurate perceptual information that can be brought to bear on a task. Much research concerns the construct's capacity and the precision of the information stored. Two prominent theories of VWM representation have emerged: slot-based and continuous-resource mechanisms. Prior modeling work suggests that a continuous resource that varies over trials with variable capacity and a potential to make localization errors best accounts for the empirical data. Questions remain regarding the variability in VWM capacity and precision. Using a novel eye-tracking paradigm, we demonstrate that VWM facilitates search and exhibits effects of fixation frequency and recency, particularly for prior targets. Whereas slot-based memory models cannot account for the human data, a novel continuous-resource model does capture the behavioral and eye tracking data, and identifies the relevant resource as item activation.

Keywords: Visual working memory; Visual search; ACT-R; Eye tracking; Resource allocation

Correspondence should be sent to Bella Z. Veksler, Oak Ridge Institute for Science & Education at AFRL, 711 Human Performance Wing, 2620 Q Street (Bldg 852), Wright-Patterson AFB, OH 45431-7905. E-mail: bellav717@gmail.com

1. Introduction

Working memory (WM) is a limited capacity memory system used for temporarily storing and manipulating information (Baddeley, 2003; Baddeley & Hitch, 1974). Research has demonstrated relationships between WM and a wide scope of intellectual abilities, including fluid intelligence (Fukuda et al., 2010; Unsworth et al., 2014), logical reasoning, and problem solving (Engle et al., 1999).

Baddeley proposed *visual working memory* (VWM) as a distinct process within the working memory system, which he referred to as the visuospatial sketchpad. Visual working memory is a construct hypothesized to be a limited capacity system that maintains representations of visual information temporarily for use and manipulation in the performance of ongoing tasks (Luck & Vogel, 2013). This construct has garnered much attention and has been the focus of many studies and computational models. Even so, answers to fundamental questions, such as its capacity and the precision of its representation, remain elusive (Van den Berg & Ma, 2014).

The dominant approach to studying VWM uses a passive, tachistoscopic version of the change detection paradigm (Alvarez & Cavanagh, 2004). This has been and continues to be the most prominent paradigm in contemporary empirical research on VWM (Donkin, Kary, Tahir, & Taylor, 2016). In this task, a participant typically is instructed to attend to, and remember, information within a stimulus display. The information displayed tends to be a set of unique objects that differ across features, such as shape, color, and/or location. After some time (75–2,000 ms; Bays, Wu, & Husain, 2011; Ester, Drew, Klee, Vogel, & Awh, 2012, respectively), the stimulus disappears for some delayed amount of time (300–3,000 ms; Ester et al., 2012; Rademaker, Tredway, & Tong, 2012, respectively), and the object of the possible change is cued, or a new stimulus appears. If a change occurred, the participant must indicate the change in some manner, either by responding yes/no (c.f., Alvarez & Cavanagh, 2004), by indicating the object (c.f., Anderson, Vogel, & Awh, 2013), location (c.f., Barton, Ester, & Awh, 2009), or some combination thereof. Researchers vary the number of items in a stimulus (i.e., set size) to evaluate VWM capacity and use change identification to evaluate VWM precision.

The initial behavioral findings of change-detection studies were that participants' performance declined as the stimulus set size increased beyond four items (Luck & Vogel, 1997). This behavioral finding has both been successfully replicated (Cowan, Fristoe, Elliott, Brunner, & Saults, 2006; Scolari, Vogel, & Awh, 2008) and failed to replicate (Alvarez & Cavanagh, 2004; Bays et al., 2011). Features of the experiment vary across studies (i.e., presentation time, stimulus attributes, participant report, set size, and data analysis approaches), potentially leading to these disparate results. In spite of (or because of) these discrepancies, two broad classes of competing theories of visual working memory capacity have come to dominate the literature: *slot* and *continuous resource* theories.

1.1. Prominent theories of visual working memory

Slot theories of VWM generally posit a fixed capacity of few discrete items with high to perfect precision (Luck & Vogel, 2013). The trait of discrete items is shared with the long-standing hypothesized structure of the general working memory system (Cowan, 2001). By contrast, *continuous resource* theories of VWM posit a finite resource that can be spread across different areas/items of a scene. This resource is dedicated to VWM and can be flexibly distributed across items in a display (Wilken & Ma, 2004). In the following sections, we briefly introduce, compare, and contrast slot and resource theories (for more detailed reviews, see Brady, Konkle, & Alvarez, 2011; Luck & Vogel, 2013).

Slot theories were the first theories posited to account for VWM capacity. They presume that VWM capacity has a fixed item limit of three to four items. A slot is a memory container, which is filled with a complete object representation which has integrated, bound features that can be accurately recalled independent of visual complexity, be it a single vertical line or a complex Chinese character (Luria & Vogel, 2011). This encoding begins as quickly as visual perception allows, with event-related potential evidence pointing toward accurate yes/no responses for complex objects in natural scenes when presentation is as short as 20 ms (Thorpe, Fize, & Marlot, 1996). Information for an object is not hypothesized to be distributed across slots. When all of the slots are filled (VWM capacity has been reached), no information about other stimuli is stored: no complex items, no conjoined features, nor any single feature at any level or gradient. Further, stimuli are remembered in an all-or-nothing fashion; thus, slot theorists postulate that participants guess when presented with stimuli containing items that surpass VWM capacity.

As experimental paradigms became more sophisticated (Wilken & Ma, 2004), Zhang and Luck (2008) revised their slot theory to the “slots + averaging” model and aimed to measure VWM capacity in terms of objects and the precision of information remembered per object. In this model, if there are more slots than objects to be remembered, objects may be represented again in a free slot, leaving no slot empty. Thus, the model always reaches its capacity, with some of the stored information being redundant. This model predicts that fewer than four objects can be encoded with higher fidelity than four objects. It also indicates that no information about objects upwards of four is stored, meaning that responses should be random after set size has exceeded capacity. They found the probability that an item was stored in memory declined slowly from set sizes one, two and three, and then declined sharply at set size six. Item precision increased as set size increased from one, two, and three items, then remained constant at a set size of six items (Zhang & Luck, 2008). The former result implies that capacity is maximized at three objects, a characteristic of slot theories. The latter result, that item precision did not increase past set size three, was taken as evidence that no further information about items was being stored. However, this revision of slot theory still disregards object complexity and object interaction.

A competing account of visual working memory in the form of *resource theory* proposes to address these shortcomings of slot theory. The “resource” in resource theory refers, somewhat vaguely, to a pool of mental processing power dedicated to visual working memory that can be flexibly distributed across multiple items within a display. Just how mental resources are distributed across items is debated among resource theorists, but the overarching hypothesis that a small number of objects can be encoded with high precision and a large number of objects can be encoded if lower precision is accepted among them. The fewer the objects, the less distributed the memory representations are, leading to the likelihood of more accurate recall. At least some information about all objects displayed is stored in visual working memory, but the information represented may be inaccurate or incomplete. For example, some conjoined features from one object may be stored, along with a complete second object, and a single feature from a third object. As an object can be partially remembered or not fully encoded, resource theorists do not interpret incorrect responses as evidence for guessing, but as a consequence of errors in encoding or retrieval (Bays et al., 2011; Fougnie & Alvarez, 2011; Van den Berg, Shin, Chou, George, & Ma, 2012). This interpretation is supported by evidence that items can interact in memory. For instance, feature information can be "swapped," producing illusory conjunctions (Treisman & Schmidt, 1982) and suggesting that items were either insufficiently maintained or erroneously retrieved (Bays et al., 2011).

Van den Berg, Awh, and Ma (2014) used a factorial comparison of previously reported answers to address core questions of VWM: What is the nature of mnemonic precision, how many items can be remembered, and what effect do spatial-binding errors have on VWM?

The five prominent results to come out of VWM research addressing these three questions were used to generate the scope of corresponding models: (a) the presence of an upper VWM limit (Cowan, 2001; Miller, 1956); (b) memory precision decreases with the amount of information in a scene (Wilken & Ma, 2004); (c) precision comes in stackable quanta (Zhang & Luck, 2009); (d) memory precision varies across trials (Van den Berg et al., 2012); and that (e) incorrect spatial binding leads to localization errors (Wheeler & Treisman, 2002); These five results were organized into three factors that represented disputed answers about VWM precision, capacity, and the potential for spatial-binding errors. The approach resulted in 32 different models, only six of which have been previously reported in the literature. All models were tested on 10 previously published empirical results from a delayed-response, tachistoscopic, change-detection paradigm collected across six different laboratories (resulting in 131,452 trials from 164 participants). The results from these models indicated that a previously unreported model best accounted for the data. The novel model is characterized by a notion of continuous precision that varies across trials along with a capacity that varies across trials combined with the presence of the potential for spatial-binding errors (i.e., variable precision, variable capacity, with spatial-binding errors). Thus, Van den Berg et al.'s (2014) results raise the question of what is leading to variance in VWM precision and capacity.

Donkin et al. (2016) have argued that a VWM system using a continuous resource may appear to support a slot interpretation when the number of items to remember varies from trial to trial. At times, highly precise representations of a small number of objects appear to favor a slot-based model, but when set size is unpredictable participants are biased to focus on a small subset of items, leading to performance suggestive of a slot model. When set size was predictable (the same across multiple trials), resource models best characterized the data. Donkin et al.'s (2016) results seem to support Van den Berg et al.'s (2014) modeling results and explain why precision and capacity may vary across trials.

1.2. Motivation and approach

Visual working memory has been studied extensively within the context of a change detection paradigm, and we present several reasons why basic questions regarding capacity and precision still remain unanswered and contested among slot and resource theorists. There are weaknesses to using passive change detection to understand VWM (Rouder, Morey, Morey, & Cowan, 2011). This paradigm, which consists of delayed responses (2–3 s), does not tap into the functional importance of VWM—to facilitate the accurate completion of an active visual task through the temporary storage of readily available information. Outside the experimental laboratory, visual search does not occur as an isolated and self-contained task, but rather in the context of a task where targets contained in some visual array are distinguished from distractors, and vision is an active process (Findlay & Gilchrist, 2003). Because passive change detection studies typically use tachistoscopic presentation of stimuli, it ensures that participants do not deviate from fixating the center of the screen (Bays & Husain, 2008; Bays, Catalao, & Husain, 2009; Bays et al., 2011). By leveraging eye tracking technology and allowing for overt shifts of attention, we can more readily distinguish the nuances of VWM capacity and precision within this active vision approach, as variability in capacity and precision may be a direct result of where visual attention was allocated during the task.

In addition, this variability may be an adaptation to experimental paradigms that change set sizes from trial to trial, as alluded to by Donkin et al. (2016). There are other potential explanations behind the trial-to-trial variance reported by Van den Berg et al. (2014) other than the randomization of set size across trials put forth by Donkin et al. (2016).

In the current paper, we provide an explanation for the variance in VWM precision and capacity and identify a candidate resource. To do so, we introduce a new eye-tracking paradigm that moves away from the change detection tasks commonly used to investigate VWM. Our new paradigm of *repeated serial search* (Neth, Gray, & Myers, 2006) requires an individual to actively search for different and occasionally repeating targets within a stable visual display. This creates a task that is more realistic and ecologically valid than a passive change detection paradigm because in the real world, a static visual scene rarely changes rapidly and without warning, and humans can move their eyes to explore their visual world. Importantly, it allows us to ask questions of VWM that inform how it drives search behavior and the potential differences in depth of encoding between targets and distractors because we have access to the full history of human fixations. Further, we developed six models in the spirit of the van den Berg (Van den Berg et al., 2014) factorial: three memory types (no memory, slot-based, continuous resource) crossed with two different search strategies (random and closest-to-nearest) to determine which VWM theory best accounts for the human data. In our models, a search strategy is only relied upon if the memory trace is weak or non-existent for the current target of the search.

Previewing the results, our empirical and modeling work leads to five important conclusions: (a) the variability in VWM capacity and (b) precision results from recency and frequency effects from selectively encoding visual information; (c) memory facilitates search behavior; (d) targets have a stronger mnemonic trace than distractors; and (e) the relevant "resource" involved is memory activation. In the following sections, we introduce our paradigm and present empirical results, followed by a model-based analysis of the empirical data.

2. Experiment

To determine the degree to which VWM facilitates visual search, we designed an experiment using a novel *repeated serial search* paradigm. In this paradigm, participants were required to search the same spatial configuration of 10 static items a total of 20 times. This paradigm taps into the VWM construct by motivating participants to retain a maximum amount of information in VWM to facilitate future searches.

Paradigm. On every trial, 10 circular objects with a diameter of 60 pixels each were distributed randomly over a centered white rectangular display on a 17" flat panel screen (measuring 1,270 × 970 pixels). The objects were positioned at least 60 pixels away from any edge, and the distance between the centers of any two objects was constrained to be at least 200 pixels. Each circle contained a hidden label (upper case letter, number, or monosyllabic four-letter word) that specified the target sought by the participant. On any given trial, only one type of label was in the circles (letters, numbers, or words). The order of label types was randomized within each participant's task presentation.

Every trial comprised a total of 20 searches through the display. At the start of each search, the experimental software announced the current target label to the participant (e.g., "cell" in Fig. 1). Participants had to hover with the mouse cursor over a circle to uncover its hidden label. Upon moving the cursor off the circle, the corresponding label was hidden again. Participants were instructed to click on the circle corresponding to the target label. If the clicked circle indeed contained the current target label, a new target was announced. By contrast, clicking on a different circle was recorded as an error and again announced the current target label to provide a reminder to the searcher. Consequently, searchers typically uncovered non-targets (distractors) in the process of searching for targets, and these distractors may turn into targets in subsequent searches.

Participants. A total of 13 Rensselaer Polytechnic Institute undergraduates (3 females) volunteered for course credit. Their mean age was 18.92 years (SD = 1.04).

Procedure. Participants signed informed consent forms, viewed a slideshow of the instructions, and were calibrated to an LC Technologies eye tracker prior to beginning the study. Every participant completed 60 trials in total. Each trial consisted of a series of 20 searches. Every search commenced when a computer-generated voice announced a next target to be found.

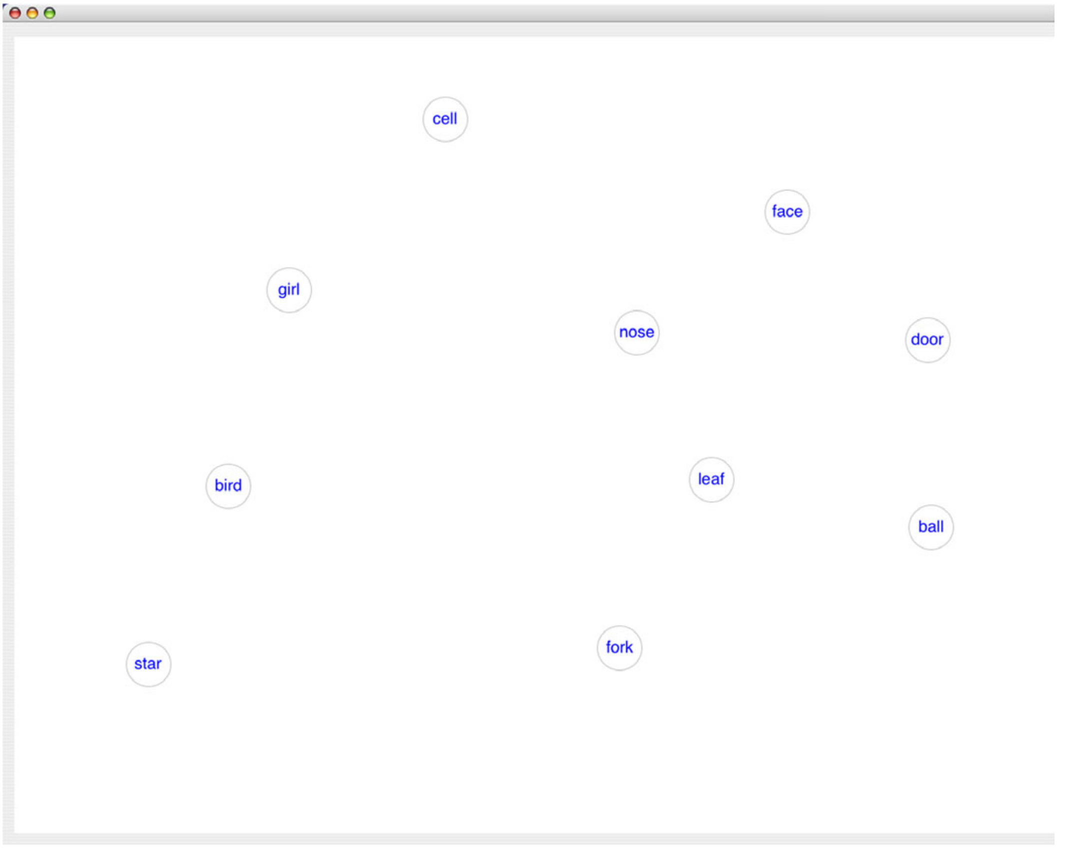

Fig. 1. Example stimulus used in the experiment. Although all labels are visible here, they were hidden from participants' view until a cursor hovered within the circle.

2.1. Results

Visual point of regard and mouse location and click data were collected as participants did the task and then compiled into a sequential history of fixations for each participant. Given this sequence of fixations, at the beginning of a search, we determined how many times each label (item) was fixated (frequency) and how long ago (recency). This allowed us to investigate recency and frequency effects in finding a target. The functional role of labels (i.e., whether labels were previously seen and encoded as targets or as distractors) was also investigated.

2.1.1. Recency effects

For this analysis, we restricted the data to the first two times a label was a target of a search. Fig. 2, left, shows the trend of recency on the number of fixations to find the target as a function of whether the label had previously been a target before (label type). A 2 (label type) × 10 (recency) ANOVA was performed to evaluate the effect of item encoding and recency of an item's last fixation. There was an interaction between whether an item was a target before and how recently it was last fixated, $F(9, 108) = 3.76$, $p < .001$, $\eta^2 = 0.24$. There was also a significant main effect of recency, $F(9, 108) = 11.94$, $p < .001$, $\eta^2 = 0.50$ (see the top line of Fig. 2, left). This effect was greater for labels that had not been previous targets, $F(1, 12) = 73.42$, $p < .001$, $\eta^2 = 0.86$. In general, labels that were prior targets were less impacted by the fixation recency (bottom line of Fig. 2, left).

One plausible explanation for the inverted U-shape of the items that were only distractors prior to the current search is that participants searching through the display are only encoding whether or not the current item is the target, rather than the identity of the item. Encoding recent items as non-targets may result in an inhibition of return effect for more recently fixated items (2–5 fixations ago) leading to longer search times than when the distractor was seen a longer time ago.

2.1.2. Frequency effects

The right panel of Fig. 2 shows the effect of the number of times an item had previously been fixated on how rapidly it is found as a target and as a function of whether it had previously been a target (label type). A 2 (label type) × 7 (frequency) AVOVA was performed to evaluate the effect of label encoding frequency. There was insufficient data in frequency bins 0 and 1 (i.e., in cases where a second search for a target was preceded by zero or one fixations on the item prior to the search), leaving bins 2–8 for analysis. Nonetheless, these bins reflect the general trend in the data. There was a significant interaction between fixation frequency and label type on the number of fixations to find the target, $F(6, 72) = 5.92$, $p < .001$, $\eta^2 = 0.33$, where searches required fewer fixations when a label had been a target before despite being seen <5 times, $F(1, 12) = 38.37$, $p < .001$, $\eta^2 = 0.76$ (bottom line of Fig. 2, right). Further, there was a main effect of

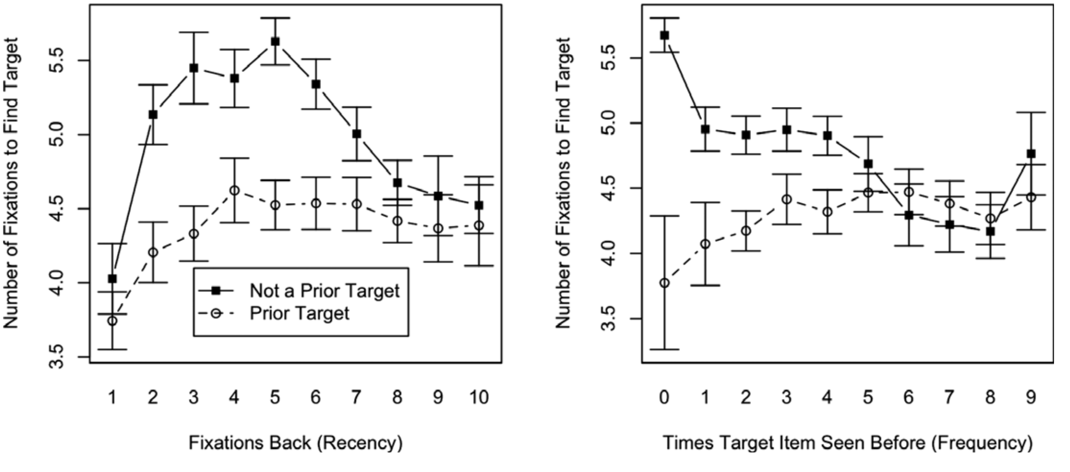

Fig. 2. Mean number of fixations needed to find a target as a function of recency (left) and frequency (right) of seeing the target before. Error bars indicate standard errors.

frequency on number of fixations to find the target, $F(6, 72) = 3.25$, $p < .01$, $\eta^2 = 0.21$. In particular, items that had not been prior targets show a benefit of having seen the item more frequently, whereas previous targets seem to be encoded sufficiently enough that it takes fewer and roughly the same number of fixations to find the target again, irrespective of the number of previous fixations.

2.1.3. Recency and frequency effects

In order to provide a more robust description of the human data, we examined the proportion of all searches in which a target was last seen *R* (Recency) fixations ago or was seen *F* (Frequency) times prior to the search and was found within *N* fixations. Fig. 3 illustrates the respective distributions generated by analyzing the human data in this way. In particular, in the recency graph (on the left), the peak of the distribution shifts to the right (i.e., more fixations to find the target) as *R* increases. Overall, the human data exhibits a curve with an initial peak and gradual decline across all recency values, with the proportion of searches in which a target is found in a higher number of fixations falling off gradually.

In the frequency graph (in the right panel of Fig. 3), the proportion of all searches in which the target is found stays roughly around 10% across all values of *N* when the target has never been fixated before ($F = 0$). By contrast, items which have been fixated more frequently (e.g., $F \geq 1$) show a more pronounced peak at $N = 3$ when compared to items which were fixated less frequently.

Subsequent model runs were compared to these distributions. We wanted to be able to capture both the magnitude of the proportions (in both recency and frequency) as well as the general shape of the distributions as proportions gradually tapered off for higher values of *N*. Note that these distributions are agnostic to whether target labels were previously targets or not.

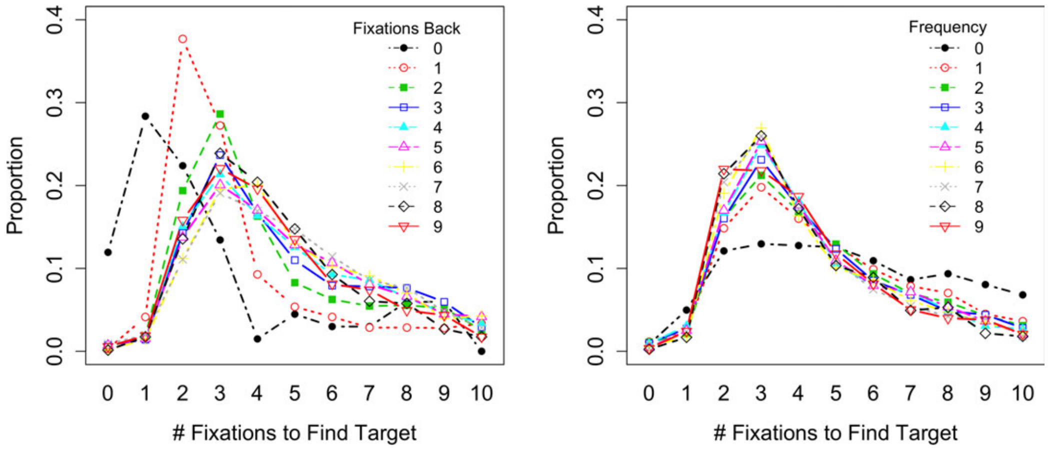

Fig. 3. Proportions of recency (left) and frequency (right) effects in the human data.

2.2. Experiment discussion

The results from the study indicated that the number of fixations to find a target is systematically affected by (a) whether that label had or had not been a prior search target, (b) the recency of a label's previous fixation, and (c) the frequency of a label's previous fixations. Each of these effects contributes to the variability in VWM capacity and precision. An item label more recently encoded will lead to the appearance of a larger VWM capacity and higher VWM precision. Similarly, a label more frequently encoded will lead to the appearance of a larger capacity with greater precision. In passive change detection, the probe is chosen at random and may sometimes select a target that has neither been recently or frequently encoded. This could naturally lead to the perception of capacity and precision variability of VWM. By a model-based analysis that accounts for selective attention processes during search, we can more concretely pinpoint the mechanisms leading to this variability.

3. Model-based analysis

Given the debate in the literature between slot-based and continuous resource models of VWM, we chose to run a factorial combination of models and search strategies. The three classes of models considered were as follows: no memory, slot-based memory, and continuous resource memory. The search strategies were either *nearest first* or *random*. For each memory-strategy combination, scan paths were generated for each of the 20 searches within a trial. An assumption shared by all models was that once an object was visited, it was removed from the set of possible next visits until the next target was announced.

3.1. No memory model

This model served as a theoretical baseline for the other models and searched the display for every search within a trial without any memory for previous targets or distractors. In the *random search* version, the model searched the display in a random fashion. In the *nearest first* version, the model allocated attention to the closest object to the one currently being fixated. No parameters were varied in this model.

3.2. Slot-based memory models

This class of models had a slot-based memory, and the number of slots available ranged from 0 to 10. Slots were instantiated as a queue (FIFO) based on the human fixation history prior to the current search (see Fig. 4). Uncovering a label reinstantiates slot 0 and pushes the label contained in slot i into slot $i + 1$. This corresponds closely to the R denotation in the above analysis of human Recency data. At the start of every search, the model queried its slot-based memory to determine whether the target was already present in one of its slots. If it was, the model immediately directed its attention to the target's

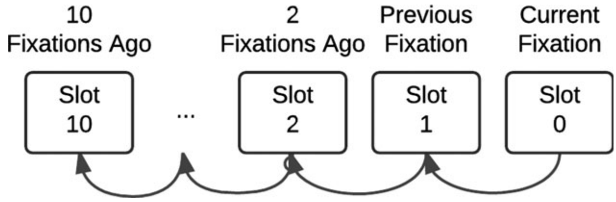

Fig. 4. Slots are instantiated corresponding to the timeline of fixations in the human data.

location. If the target was not stored in a slot, the model searched the display in either a random or nearest-first manner. Only the number-of-slots parameter was varied in this model type.

3.3. Continuous resource memory models

This class of models also relied on the eye fixation history of the trial prior to the current search, but rather than merely considering the order of fixations the continuous resource models incorporate both fixation recency (i.e., the time stamps of when the item was fixated) and frequency (i.e., how many previous fixations were made to the item). We used the declarative memory activation component of the ACT-R cognitive architecture to model memory as a continuous resource (Anderson, 2007). While this activation equation is typically applied to declarative knowledge within the scope of ACT-R modeling, we felt this would be a relevant, useful, and appropriate approximation for the "activation trace" of an item stored in VWM.

The ACT-R memory equation was applied to each of the items on the screen at the beginning of each search. The activation of a given item i in memory is calculated as follows:

$$A_i = \ln \sum_{j=1}^{n} t_{ij}^{-d} + \beta_i + \varepsilon_i \tag{1}$$

where j represents a fixation on the item and t_{ij} a time stamp of how recently the item was seen on fixation j, $-d$ a decay value, β a base-level constant offset, and ε logistically distributed transient noise with a mean of 0 and a standard deviation of σ.

The activation of the target item was recalculated at the beginning of each search and the model checked whether the activation of the target was above the threshold, T, and if so, moved attention directly to the known location of the item. If $A_{target} < T$, the model selected and encoded another item based on either a *random search* or a *nearest-first* strategy. If the target item was still not found, activation was recalculated for the target at each additional movement of attention. Four parameters were varied in the context of ACT-R's memory equation (d, β, T, and σ) to find the best fit to the human recency and frequency data using MindModeling.org (Harris, 2008). In particular, we varied the parameters as follows: d: [0, 1], β: [0, 10], T: [0, 20], and σ: [0, 5]. This created a total of 27,951 parameter combinations for each search strategy.

3.4. Model evaluation

Each of the above models was evaluated on all trials (and searches) obtained from human data.[1] As the ACT-R memory equation uses recency and frequency information as sources of activation for a given chunk in memory, we examined the human data as a function of both the recency and the frequency of previous fixations to current targets. In this case, recency R refers to how many fixations ago an item was last fixated relative to the current fixation. For each parameter set, summary statistics were calculated to determine the percentage of all trials on which the target was seen R fixations ago or was previously fixated F times and found in N fixations. This resulted in 10 distributions for recency and another 10 for frequency, each with 11 data points (one for each N of fixations to find the target, see Fig. 3 for human data). Then the root mean squared error (RMSE) and R^2 scores were calculated for each target recency curve and for each target frequency curve.

A composite goodness-of-fit measure was created to combine the R^2 and RMSE measures to capture both the shape and the magnitude of the differences between human data and model predictions. Because best fits according to R^2 are values closer to 1, and best fits according to RMSE are values closer to 0, we rescaled the R^2 measure ($1 - R^2$) and computed an average of all curves for each parameter setting.

The best-fitting slot-based model was one which contained two slots (i.e., "remembered" the last two items previously fixated; see Table 1). The best-fitting continuous resource model resulted from the following parameter settings: $d = 1$, $\beta = 1$, $T = 10$, and $\sigma = 4.0$ (see Figs. 5 and 6, green dashed line).

The no-memory model established a baseline with which the other memory models could be compared. As can be seen in Table 1 and Figs. 5 and 6, the continuous resource memory model captured human performance much more closely. In particular, whereas a slot-based memory with two slots was the best fitting in this particular class of models, it failed to capture the shape of both the recency and frequency distributions. The

Table 1
Best fits for all model types

Memory	Strategy	Composite Score*
Continuous resource	Nearest first	0.09 (.005)
Continuous resource	Random	0.11 (.006)
Slot (2)	Random	0.35 (.008)
Slot (2)	Nearest first	0.37 (0)
None	Nearest first	0.38 (0)
None	Random	0.39 (.009)

Note. Values in parentheses indicate the *SD* of 1,000 model runs.
*Lower composite scores indicate better model fits.

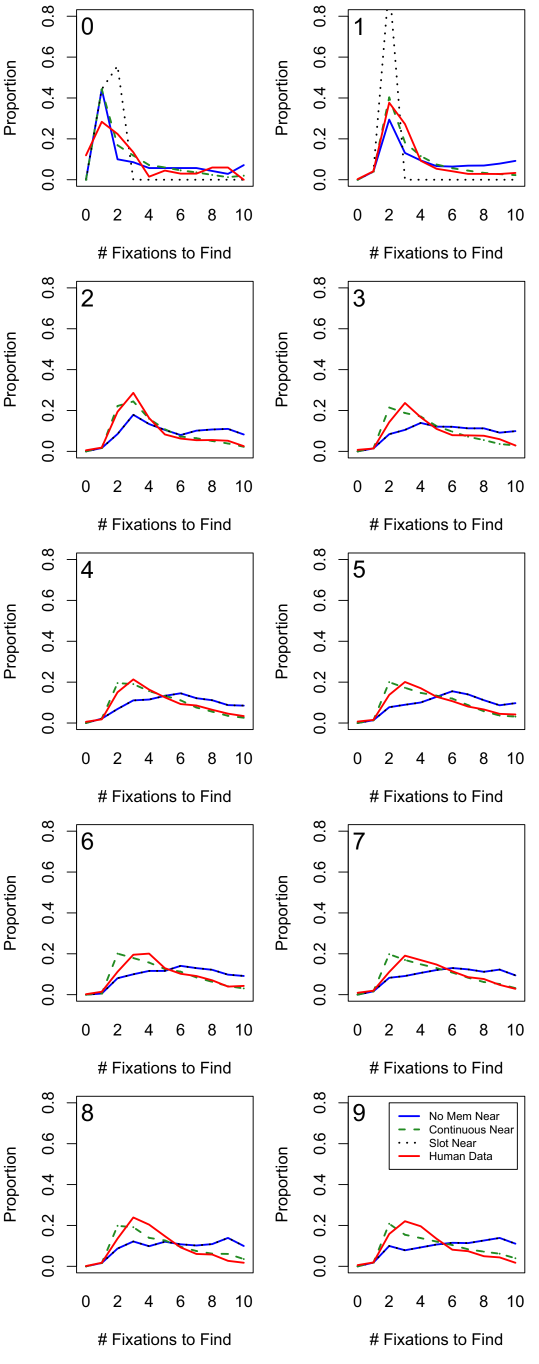

Fig. 5. Model fits of recency for best-fitting slot, continuous resource, and no-memory models. Numbers in upper left corner indicate how long ago item was last fixated (*R*). Each data point in continuous memory model is 1,000 runs of the model.

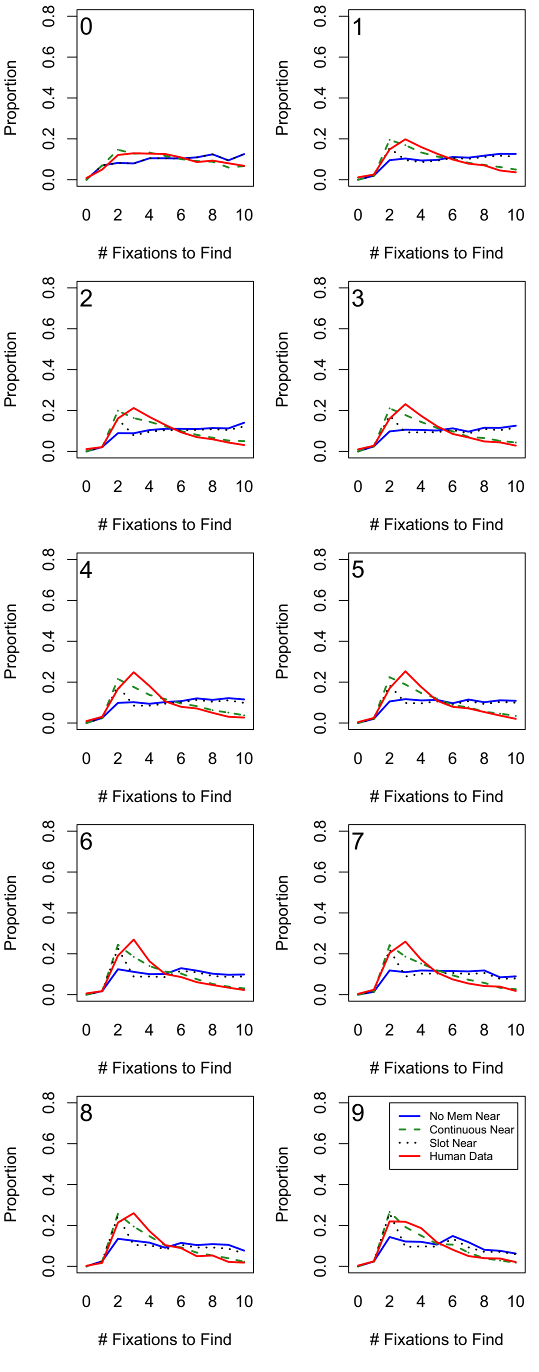

Fig. 6. Model fits of frequency for best-fitting slot, continuous resource, and no-memory models. Numbers in upper left corner indicate how many times an item had previously been fixated (F). Each data point in continuous memory model is 1,000 runs of the model.

continuous resource model, on the other hand, exhibited the same bell-shaped curve with gradual drop-off as the human data for both recency and frequency. Furthermore, a *nearest-first* search strategy was marginally better at capturing the effects than a *random search* model, suggestive of the type of strategy participants may have used as they conducted their search of the display.

We further evaluated the flexibility of all the model types to determine how convincing the fits actually are and whether they could have been achieved merely by searching such a large parameter space. Model flexibility analysis (MFA) was used to calculate the proportion of all empirical outcomes that each model could have potentially fit (Veksler, Myers, & Gluck, 2015). Although the *slot* model only has one parameter (number of slots), it is actually more flexible than the *continuous resource* model with four parameters: MFA revealed flexibility for the *slot* model to be $\phi = .14$; the *continuous resource* model, on the other hand, had a flexibility value of $\phi = .014$. Thus, the *continuous resource* model makes more precise predictions and is less flexible, as the results of the model runs cover less of the potential behavioral space. Furthermore, we ran each model 1,000 times to obtain an estimate of the variability between model runs and found that the performance measures and composite scores are fairly stable, with standard deviation across the 1,000 model runs being on the order of .005–.009 across the various model types.

3.4.1. *Avoidance*

Although the analyses and model results so far indicate that participants find targets faster when they have fixated those labels more recently or multiple times prior to the current search, we also wanted to address the issue of whether VWM is used to facilitate

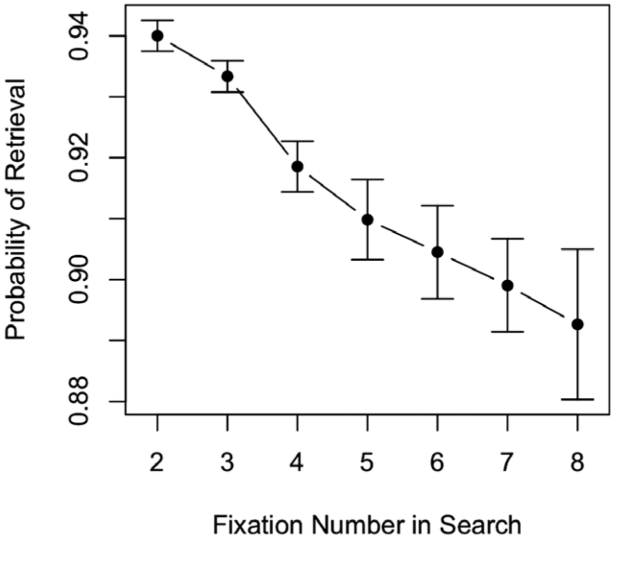

Fig. 7. Probability of retrieval of an item fixated on a given fixation in a search. The first fixation in a search was the target of the previous search.

the search by avoiding labels on the display that had previously been fixated and are current distractors.

If memory was used in this way, we would expect that the first few fixations in a search would be to labels not previously fixated or at least fixated a long time ago and no longer in the memory trace. Using the ACT-R memory equation, we calculated the probability of retrieving a label from memory given its fixation history and found that labels (which are not the current target) fixated first in a search actually had a higher probability of being retrieved, $R^2 = .30$ (see Fig. 7). This suggests that even though those labels could presumably have been recalled and confirmed to not be the target, there was no avoidance strategy to not fixate them throughout the search.

4. Discussion and conclusions

In the current work, we explored why variability in VWM capacity may at times exhibit variable precision and capacity. The new paradigm of *repeated serial search* allowed us to more readily observe the specific shifts of visual attention that occur during natural search. Human data suggest that the variability in VWM precision and capacity is closely tied to selective attention as search progresses.

Selective attention directly affects the ease with which subsequent targets can be found, with both recency and frequency of fixations playing a role in subsequent searches. Items that were previously fixated more recently resulted in faster search times and boosted the likelihood of recalling the location of the target. Likewise, items which had previously been fixated more often were easier to find. Importantly, there was a stronger mnemonic trace for items which were previous targets as these items were found faster than those which were only fixated as distractors during previous searches.

We contrasted two types of models of VWM: a *slot*-based and a *continuous resource*-based model. In the case of the slot-based model, the recency of a label's encoding is taken into consideration to facilitate subsequent searches. However, this was not sufficient to accounts for the human data as it failed to capture the shapes of the distributions in both recency and frequency domains. A continuous resource model, on the other hand, directly incorporated both effects of selective attention. The continuous resource was instantiated as the label's activation, computed by taking into account both the frequency and recency of previous label fixations.

One limitation of the current approach is that neither of the model types explicitly accounts for the stronger mnemonic trace for prior targets. The continuous resource model could potentially account for this difference by including a label's fixation duration in its computation of activation, as targets typically exhibit longer fixations and more opportunity for rehearsal. While such models are beyond the scope of the current work, they are an interesting avenue for future research.

In conclusion, the repeated serial search paradigm elucidates the variability seen in VWM capacity and precision by taking into account a formal notion of selective attention. Future work could apply the same continuous resource model to other data sets to explore the robustness of the model in accounting for various VWM results, as well as incorporating potentially hybrid models which combine memory slots and continuous resources.

Acknowledgments

This work was supported by the Air Force Research Laboratory's Warfighter Readiness Research Division and the Air Force Office of Scientific Research (grant #12RH05COR). This research was supported in part by an appointment to the Postgraduate Research Participation Program at the U.S. Air Force Research Laboratory, 711th Human Performance Wing, Human Effectiveness Directorate, Warfighter Readiness Research Division, Cognitive Models and Agents Branch administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and USAFRL.

Note

1. A Python script of all the models is available as a supplement at http://palm.mind modeling.org/ bveksler/BZV-TopiCS-VSmodel.zip

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*(2), 106–111.

Anderson, J. R. (2007). *How can the human mind exist in the physical universe?* New York: Oxford University Press.

Anderson, D. E., Vogel, E. K., & Awh, E. (2013). Selection and storage of perceptual groups is constrained by a discrete resource in working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 824–835.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews. Neuroscience*, *4*(10), 829–839.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89.

Barton, B., Ester, E. F., & Awh, E. (2009). Discrete resource allocation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(5), 1359–1367.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*(5890), 851–854.

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10), 1–11.

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia*, *49*(6), 1622–1631.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, *11*(5), 1–34.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114.

Cowan, N., Fristoe, N. M., Elliott, E. M., Brunner, R. P., & Saults, J. S. (2006). Scope of attention, control of attention, and intelligence in children and adults. *Memory & Cognition*, *34*(8), 1754–1768.

Donkin, C., Kary, A., Tahir, F., & Taylor, R. (2016). Resources masquerading as slots: Flexible allocation of visual working memory. *Cognitive Psychology*, *85*, 30–42.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309–331.

Ester, E. F., Drew, T., Klee, D., Vogel, E. K., & Awh, E. (2012). Neural measures reveal a fixed item limit in subitizing. *The Journal of Neuroscience*, *32*(21), 7169–7177.

Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Oxford, UK: Oxford University Press.

Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision*, *11*(12), 1–12.

Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology*, *20*(2), 177–182.

Harris, J. (2008). MindModeling@Home: a large-scale computational cognitive modeling infrastructure. In *Proceedings of the sixth annual conference on systems engineering research 2008* (pp. 246–252). Los Angeles, CA.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281.

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400.

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia*, *49*(6), 1632–1639.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97.

Neth, H., Gray, W. D., & Myers, C. W. (2006). Memory models of visual search—searching in-the-head vs. in-the-world. *Journal of Vision*, *5*(8), 8–9.

Rademaker, R. L., Tredway, C. H., & Tong, F. (2012). Introspective judgments predict the precision and likelihood of successful maintenance of visual working memory. *Journal of Vision*, *12*(13), 21.

Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review*, *18*(2), 324–330.

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*(1), 215–222.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522.

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*(1), 107–141.

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26.

Van den Berg, R., & Ma, W. J. (2014). "Plateau"-related summary statistics are uninformative for comparing working memory models. *Attention, Perception, & Psychophysics*, *76*(7), 2117–2135.

Van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(22), 8780–8785.

Van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological Review*, *121*(1), 124–149.

Veksler, V. D., Myers, C. W., & Gluck, K. A. (2015). Model flexibility analysis. *Psychological Review*, *122* (4), 755–769.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*(1), 48–64.

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 1120–1135.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235.

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science*, *20*(4), 423–428.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Data S1. Human data from the visual search study and the corresponding code for all the models in Python 3.