High Regularities in Eye-Movement Patterns Reveal the Dynamics of the Visual Working Memory Allocation Mechanism

Xiaohui Kong,[a] Christian D. Schunn,[b] Garrick L. Wallstrom[c]

[a]*Intelligent Systems Program, University of Pittsburgh*
[b]*Department of Psychology, University of Pittsburgh*
[c]*Department of Biomedical Informatics, University of Pittsburgh*

Abstract

With only two to five slots of visual working memory (VWM), humans are able to quickly solve complex visual problems to near optimal solutions. To explain the paradox between tightly constrained VWM and impressively complex human visual problem-solving ability, we propose several principles for dynamic VWM allocation. In particular, we propose that complex visual information is represented in a temporal manner using only a few slots of VWM that include global and local visual chunks. We built a model of human traveling salesman problem solving based on these principles of VWM allocation and tested the model with eye-movement data. Exactly as the model predicted, human eye movements during traveling salesman problem solving have precise quantitative regularities with regard to both the general statistical pattern of attentional fixations and how they vary across individuals with different VWM capacities. Even though VWM capacity is very limited, eye movements dynamically allocate VWM resources to both local and global information, enabling attention to fine details without loss of the big picture.

Keywords: Visual working memory; Complex problem solving; Computational cognitive model; Global visual information processing

1. Introduction

Contrary to the rich visual world around us, only a minor fraction of available visual information is actually represented in our visual working memory (VWM) for further manipulation (Alvarez & Cavanagh, 2004; Baddeley, 1986; Irwin, 1992; Luck &

Hollingworth, 2008; Todd & Marois, 2004). VWM capacity has been shown to be limited in terms of a fixed number of slots, varying among individuals from only 2 to 5 (Cowan, 2001; Luck & Vogel, 1997; Rouder et al., 2008). On the other hand, humans have displayed a remarkable ability for complex visual problem solving that involves processing a large amount of visual information (MacGregor & Ormerod, 1996; Vickers, Bovet, Lee, & Hughes, 2003; Vickers, Butavicius, Lee, & Medvedev, 2001). For example, humans can find shortest distance routes through 100 cities to near optimal solutions with little effort (Kong & Schunn, 2007). But the question of how such a limited VWM resource is allocated and managed so that it can support complex human visual reasoning abilities is not well understood.

It is clear that global information processing plays an important role in human visual perception (Awh, Barton, & Vogel, 2007; O'Craven, Downing, & Kanwisher, 1999; Sugase, Yamane, Ueno, & Kawano, 1999; Xu & Chun, 2006, 2007) and complex visual problem solving (Graham, Joshi, & Pizlo, 2000; Kong & Schunn, 2007; McNamara et al., 1989; Ormerod & Chronicle, 1999; Pizlo & Li, 2005; Pizlo et al., 2006). As several recent studies have shown, VWM can hold the spatial locations of a certain number of objects regardless of their complexity (Awh et al., 2007; O'Craven et al., 1999; Xu & Chun, 2006). Sugase et al. (1999) suggest that the visual system is able to encode global information directly from visual input. But to study the VWM allocation mechanism directly would require a technology that kept track of the contents of VWM at each moment during a complex visual problem-solving procedure. Unfortunately this technology is yet to be invented. But we can measure the locations of attentional fixations that encode visual information into VWM. As a recent study has shown, VWM resources are allocated by attentional fixations and the contents in VWM are updated through eye movements (Bays & Husain, 2008).

To study the VWM allocation mechanism during complex visual problem solving, we took a ''reverse engineering'' approach. The traveling salesman problem (TSP) requires finding the shortest possible path to visit all the points on a plane and return to the starting point. Because of the simplicity in its definition and the complexity in its problem-solving procedure, the TSP provides us with a very good platform for studying the dynamics of human VWM during complex problem solving. We implemented some principles of VWM allocation into a cognitive model of complex visual problem solving to try to reverse engineer human cognition. In solving the same set of complex visual problems, and by comparing the eye-movement patterns generated by the model with human eye-movement patterns generated by participants with different estimated VWM capacity, we can support or reject the underlying VWM allocation principles.

2. Model

The following principles of VWM allocation were implemented in our model:

Principle 1: The human visual system can represent visual information at different levels of detail (Graham et al., 2000; McNamara et al., 1989; Pizlo & Li, 2005; Pizlo et al.,

2006). Each part of this information, regardless of its level of abstraction, is kept in one slot in VWM (Xu & Chun, 2007).

Principle 2: Information at the most global level is encoded first. This information then guides which local part of visual information should be focused upon and be further expanded into local fine information (de Fockert, Rees, Frith, & Lavie, 2001; Roelfsema, Khayat, & Spekreijse, 2003; Woodman & Luck, 2007), which is also consistent with what the neural measures of visual cortex have suggested: Global information is encoded first and guides the perception of refined information (Sugase et al., 1999). While attending to a local part, one chunk of global information is usually expanded into two or more chunks of local information.

Principle 3: Visual information most relevant to the current part of the task is represented at the most local level. Less immediately relevant visual information can be stored globally in VWM without being expanded. During visual complex problem solving, in order to keep the big picture in mind and focus on key details at the same time, both kinds of information are represented in VWM simultaneously.

Principle 4: Due to a capacity limit in VWM, some global information previously perceived will be lost. But a certain amount of global information is maintained in VWM all the time. Whenever the amount of global information in VWM becomes too low to support problem solving, attentional fixations will be needed to re-attend to the global information and reload it into VWM.

We argue that a VWM allocation mechanism based on those principles, even with only a few slots of VWM to deploy, can represent complex visual information in a temporal manner enabling it to focus on the details without losing the big picture during complex visual problem solving. Meanwhile, these principles of VWM allocation also predict quantitatively different eye movement patterns across individuals with different VWM capacities, as we found in the current experiments.

Those VWM allocation principles were built into a model to predict human performance and eye movement during TSP solving, a classic paradigm for studying complex visual problem solving. Specifically, we use the Euclidean traveling salesman problem, which requires finding the shortest path to visit a set of points on a Euclidean plane, returning to the start. Finding the exact solution of a TSP is an NP-hard problem, so there is no efficient algorithm to exactly solve large instances. Different from existing models of human TSP solving (Graham et al., 2000; MacGregor, Chronicle, & Ormerod, 2004; MacGregor, Ormerod, & Chronicle, 2000), our model takes a VWM capacity limit into account with VWM size as a parameter and a VWM allocation mechanism based on the above principles. A sketch of the model is as follows:

In the model, global information is first perceived by clustering all points into N clusters, where N is the VWM size. Each cluster of points regardless of its size is represented as a chunk of information and occupies one slot in VWM where only the centroid location of those points is kept (Principle 1). In the case of TSP solving, because the cluster containing the current point is most relevant to the next movement decision, our model refines it into subclusters (Principles 2 and 3). The chunk of global information in the VWM slot

containing the current point is now broken down into *M* chunks of refined information, where *M* is a number >2 and smaller than *N*. Because VWM is limited in size and holds only *N* chunks of information, the chunks of global information least relevant to the current decision are lost and replaced by the just perceived information. This process keeps the total number of chunks in VWM under the capacity limit.

Each time VWM updates its contents, the contents are manipulated and sorted based on the centroid locations of the clusters that those chunks are representing, so that these centroids form a path of shortest length when both origin and destination centroid are specified. The first chunk of information is refined recursively until each of the first two chunks contains only one point, where the first chunk contains the current point and the second chunk contains the next point to be visited. After connecting to the next point, the model checks whether there is enough global information in VWM to guide the next decision. When there is enough information, the model starts to refine the global information in the first slot of VWM. Otherwise, the model visually re-attends to the global information at the most global level, by moving the eyes to more remote aspects of problem (Principle 4). When processing global information, we assume eye fixations will be found around the center of gravity (centroid) of cluster of points due to the global effect of saccadic eye movements (Coren & Hoenig, 1972; Findlay, 1982; Findlay & Walker, 1999; Ottes, Van Gisbergen, & Eggermont, 1984). These saccades are required to adequately encode the locations of small objects outside foveal attention.

The above general process was implemented in the model as the following steps:

Step 1. Initialization

The current working set includes all points. The current point is set to be the starting point.

Step 2. Global perception

Points in the current working set are grouped into *M* clusters using the *K*-means clustering algorithm (MacQueen, 1967), where *M* is set to *N* (VWM size) in the first iteration, and the smallest integer greater or equal to $\sqrt{N}$ afterward. The *K*-means clustering algorithm clusters *N* data points into *M* disjoint subsets S_j containing N_j data points so as to minimize the sum of squares criterion:

$$J = \sum_{j=1}^{K} \sum_{n \in S_j} |x_n - \mu_j|^2,$$

where x_n is a vector representing the *n*th point and μ_j is the geometric centroid of the points in S_j. All the centroids are added into the collection of reference points in VWM, which was passed from the previous iteration.

The model then uses a spline curve to connect the current point and all the reference points to sketch a path at a rough scale. The spline curve is hypothesized to be a general smooth route through the centroids, which captures a general tendency in a globally sketched path.

Step 3. Identify current cluster and refine local information

All the points in the current working cluster are projected to its nearest points on the spline curve. If the number of points projected onto the part of the spline curve between the current point and the first reference point is more than 2, let the current working set be this set of points and go back to step 2 for the next iteration. If it is not more than 2, continue to step 4.

Step 4. Move and rehearse global information

If the number of points projected on the spline curve between the current point and the next reference point is <2, connect the current point to those points according to the sequence they projected onto the spline curve. Make the current working set the points projected onto the part of spline between the first and the second reference points. Discard the first reference point from VWM.

If the number of reference points in VWM is ≤2, regroup unvisited points at the most global level and encode those centroids into VWM.

Repeat steps 2–4 until the number of unvisited points is less than the size of VWM. Then find the shortest path for the rest of the points.

3. Statistical eye-movement pattern

To compare the eye-movement patterns the model predicts to those produced by humans, we define the distance of each fixation as the minimum of the following two distances: fixation point to the last visited point and fixation point to the next to-be-visited point (Fig. 1). Our intuition of this definition is as follows: There are two types of fixations during this task.
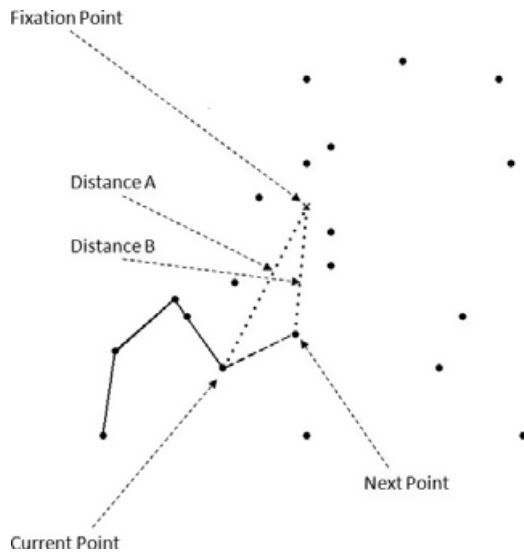

Fig. 1. Fixation distance is the minimum of Distance A and Distance B.

The first type of fixation is used to encode items into VWM for a complex reasoning process before deciding where to go next. The second type of fixation is used to program motor actions to make a mouse click onto the next to-be-visited point after one decides which point it is. Fixations of the second type are usually found near the point that is to be clicked. Before making those fixations, one has already shifted the current goal, as the next point has already been decided. Our definition of fixation distance measures how far away one's attention deviates from the current goal regardless of how far away the next to-be-visited point is from the last visited point.

Generally, our model predicts that fixation frequency decreases regularly as a function of fixation distance using the following logic. Each chunk of local information is an object consisting of only one or two points, while each chunk of global information contains a cluster of potentially many points. Starting from the most global level, the cluster containing the current point will be expanded into finer information. If it were entirely expanded into local information, all global information would be lost. So each level of global information is expanded initially into smaller pieces of the next level of global information. The cluster of points that contains the current point is expanded recursively.

As Fig. 2 illustrates, the closer to the current point, the finer the information perceived and the greater the number of fixations required. In fact, the model makes very precise predictions about the pattern on the relative frequency of eye fixations to increasing distances from the currently related points. The model also predicts that this pattern involving the relative frequency of eye fixations will change quantitatively across individuals with different VWM size. As Fig. 2A–C shows, when global information breaks down into smaller pieces of finer information, global information less relevant to the current goal is discarded from VWM in order to store the finer information. When those pieces of finer information are consumed as the current point moves, one has to re-attend to farther away aspects of the problem to re-acquire the top level global information into VWM in order to enable global planning (Fig. 2D). The first piece of global information is again broken down into finer pieces following the VWM allocation principles. So as the size of VWM increases, one will be going less frequently through this procedure of re-acquiring global information. Thus, fixations with longer distances will be found less often.

4. Methods and materials

4.1. Experiment 1

4.1.1. Participants

Eleven undergraduate students with ages ranging from 18 to 21 volunteered to participate in the study and were compensated up to $20 based on task performance.

4.1.2. Materials

For all participants, we used the same 20 TSPs as the ones used in our previous study (Kong & Schunn, 2007). Ten of them were randomly pregenerated according to a uniform
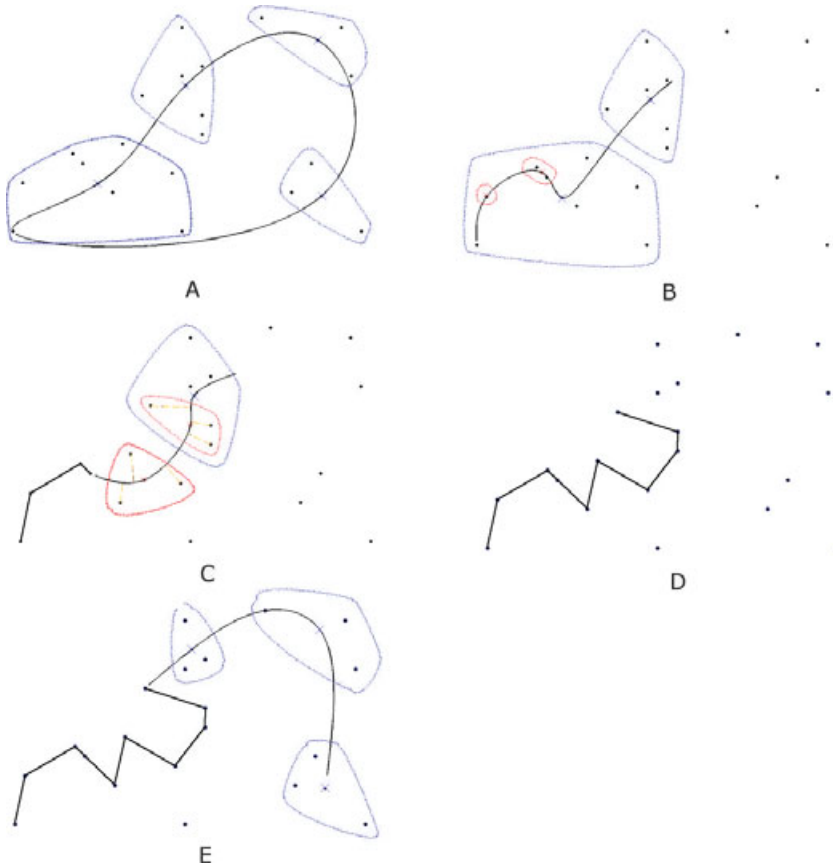
Fig. 2. This figure illustrates a sample visual working memory (VWM) allocation process during a traveling salesman problem (TSP) solving process. Color contours are the objects held in VWM at each moment. Color crosses are the centroids of those clusters. Those centroids represent the locations of the objects indexed in the VWM. (A) Blue contours are the top-level global clusters. The blue crosses are centroids of those clusters. Each of those global clusters was treated as a single object and their location information was encoded in VWM. Our model makes attentional fixations on those blue crosses to encode the position of each object. A rough sketch of the TSP path (black contour) is calculated based on global information. (B) Points between the current point and centroid of the next object were broken down into finer information (red contours). Top-level global information (blue contours) further away from the current point on the rough sketch was discarded from VWM due to VWM capacity limitation. (C) As the next chunk of objects on the sketch contour contains less than two points, the edge between the current point and next point on the sketch contour was connected. (D) At this point of the TSP-solving process, local information was consumed as the edges were connected. There is not enough global information to guide the direction of next move. (E) Global information (blue contours) were re-acquired and encoded into VWM.

distribution with size 10, 15, 20, 25, 30, 40, 50, 60, 70, and 80. The other 10 were downloaded from an online TSP library ''TSPLIB'' (http://comopt.ifi.uni-heidelberg.de/software/ TSPLIB95/) used to test TSP algorithms and most of them are from real-world problems. Those problems have 16, 22, 29, 48, 51, 52, 70, 76, 95, and 100 points each.

4.1.3. Design

The goal of Experiment 1 was to document the exact quantitative pattern of eye movements, especially the relative frequency of global processing eye movements, during visual problem solving.

4.1.4. Methods

Participants were instructed to use a mouse to solve 20 TSPs. Points of 1° visual angle size were displayed on a 30° × 30° region of a 17-inch screen with a white background. A larger point (0.2°) was the starting point. From there, participants used a mouse to left click on the next point to visit, and a line is connected from the last point to the one just clicked. The goal of the task is to find a path as short as possible to visit all the points and return to the starting point (see Fig. 3).

During the experiment, there were 20 problems varying in size from 10 to 100 points. After each problem, performance feedback was displayed as a ratio of their solution length over the optimal solution length.

A Tobii 1750 remote eye-tracker was used to track eye fixations during the problem-solving procedure. All the eye-movement data were recorded by the eye-tracker. All the mouse clicks data were stored by the experiment program in MATLAB (The MathWorks, Inc., Natick, MA).

Participants were paid for up to $20 for this part of the experiment based on their performance—for every problem producing above or equal to average performance, the participant received $1.

4.2. Experiment 2

4.2.1. Participants

Thirty-one non-color-blind volunteers participated in this experiment with ages ranging from 18 to 40. They participated in the study to fulfill a course requirement and were compensated up to $27 depending upon their performance.
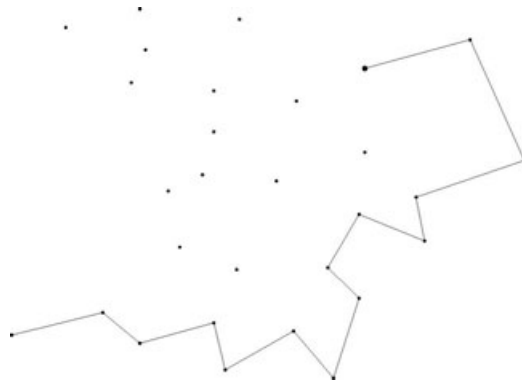

Fig. 3. Example of a traveling salesman problem (TSP) solving trial. Participants left click on the next point to connect to it from the current one.

4.2.2. Materials and methods

4.2.2.1. Design: This experiment was designed to test the relationship between individuals' VWM capacity and their statistical eye-movement patterns during problem solving discussed in the prior section.

Part 1: Estimating individual VWM capacity

We used sample arrays consisting of 1 to 12 colored squares ($3° \times 3°$), each of which were selected at random from a set of seven highly discriminable colors (red, blue, violet, green, yellow, black, and white). All stimulus arrays were presented within a $30° \times 30°$ region of a 17-inch screen with a gray background. The positions of items were randomized in a given array with the restriction that items were separated by at least $4.5°$ (center to center). One item in the test array was different from the corresponding item in the sample array by its color on 50% of trials; the sample and test arrays were otherwise identical. In each trial, the sample array stayed on the screen for 500 ms. Then, after a blank screen of 900 ms, the test array was up on the screen for 2 s. Participants were instructed to press the ''s'' key if the test array was identical to the sample array or the ''d'' key if different within those 2 s (see Fig. 4). During this part of the experiment, there were four sections and each section has 60 trials. Participants were given opportunities to rest between sections. The experiment program in MATLAB recorded all responses. Participants were paid \$3–\$7 based on their performance for this test (if the VWM score was below 2.5, then \$3; for every 0.5 addition score, another \$1 was given).

Part 1 of the experiment was immediately followed by part 2.

Part 2: TSP solving

Part 2 of Experiment 2 (TSP problem solving) was an exact repetition of Experiment 1.
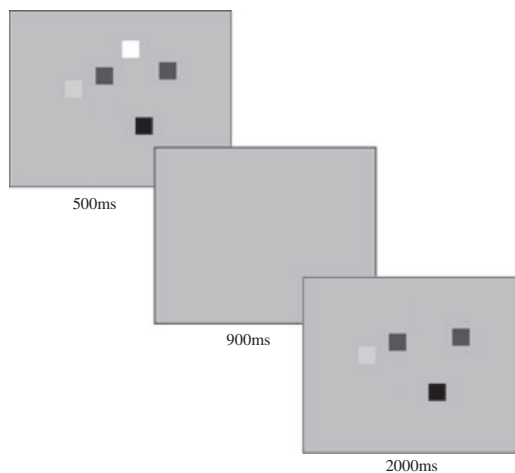

Fig. 4. Example of a visual working memory (VWM) test trial. Memory array stayed on screen for 500 ms. Then after a 900-ms retention interval, test array appear on screen for 2,000 ms, waiting for user response.

4.3. Calculating VWM capacity

We calculated each participant's estimated VWM capacity using Cowan's formula (Cowan, 2001): If a participant can hold K objects in VWM from an array of S items, then on K/S trials the changed item should be one of those being held in VWM. This subject should be able to detect a change on K/S trials in which an item changed. This formula also corrects for guessing. Overall this formula is $K = S \times (H - F)$, where K is the VWM capacity, S is the size of the array, H is the observed hit rate, and F is the false alarm rate.

4.4. Eye-movement data processing

In Experiment 2, we first calculated each of the 31 participants' solution optimality for each of the 20 TSPs. Solution optimality is defined as participant's solution length over the optimal solution length. For each of the 20 TSPs, we then calculated each participant's optimality percentile among all the participants. Then we take the median of this percentile among the 20 TSPs for each participant and excluded those participants whose median optimality is over the 65th percentile among all the participants. The intuition is that we want to exclude participants who performed much worse than the average for a majority of the problems.

4.5. Eye-tracking data analysis

Eye fixation data were exported from the Eye-tracker's CLEARVIEW 2.7 software with a filter setting of 100 ms and 30 pixels, as recommended for mixed contents by the manufacturer (Tobii Technology, Danderyd, Sweden). Fixation distances (in pixels) were then calculated according to our fixation distance definition (Fig. 1). Those fixation distances were then distributed into 43 bins with centers on 10, 25, 40, 55,…,640. According to the manufacturer's manual, the Tobii 1750 has an accuracy of 0.5° to 1.0°. In our experiment setting, 25 pixels roughly equals to 1° of visual angle. The count of the first bin centered at 10 pixels was discarded because it is well below the accuracy of the eye-tracker and is potentially noise. The counts of the rest of the bins and their corresponding fixation distances were fitted with three types of curve: quadratic curve $p_1 x^2 + p_2 x + p_3$, exponential curve $a e^{-bx}$, and power curve $a x^{-b}$. R^2 fitness was calculated for each curve fit. To test whether a curve is fit significantly better by a power curve than an exponential curve, we used a Fisher's Z_r transformation on the Pearson correlation coefficients (Ferguson, 1981). We define the curve type for a given dataset (group or individual) based on the best fitting curve. As it happens, the R^2 fitness for the quadratic curve was significantly worse than both exponential and power curve fitness in all cases. We use the exponential factor of the best fitting exponential curve as the decay rate unless the power curve fitness is significantly better than that of the exponential curve.

5. Results

We first tested the general prediction of the model regarding the distribution of eye fixations in an experiment with 11 participants. They were asked to solve 20 TSPs while their eye movements were recorded. Their fixation distances were extracted and were plotted as a histogram (bins = 30). As predicted by the model, this histogram decreases consistently with increasing fixation distance, and this relationship is closely fit by an exponential curve (R^2 = .97). This finding provides preliminary support for our hypothesis that fine information was frequently examined near the current goal and more global information was examined increasingly less often as distance increases.

We next investigated the relationship between the eye-movement pattern and individual differences in VWM capacity. If our hypothesis about working memory contents and eye movements is correct, then different VWM capacities should generate quantitatively different fixation patterns. Specifically, when VWM capacity is larger, more global information can be kept in VWM while local information occupies the other slots in VWM, thus requiring fewer re-attending saccades to global information. In contrast, when VWM capacity is smaller, global information would need to be attended frequently in order to have sufficient information in VWM to make decisions.

We ran our model with five VWM parameter settings (VWM = 3, 4, 5, 6, 7), 40 times at each setting for all 20 TSPs. Each time an object (a cluster of points) is encoded in VWM, the model makes a fixation to the center of gravity of the object (i.e., the centroid of the cluster). Then we plotted the histograms of fixation distances generated by the model.

Our hypothesized model predicts that the decay rate of the fixation distances histogram would increase as the capacity of VWM increase, because global information would be less-frequently re-attended when there is enough room in VWM to keep it for a longer time. The type of curve that best fits the histogram goes from quadratic to exponential as the VWM size increase from 3 to 7. Our model simulation predicts a trend in the decay rate of fixation distances as a function of VWM size (Table 1).

To test this prediction, we conducted the second experiment in which we first estimated 31 participants' VWM capacity using a change detection task, which has been widely used

Table 1
R^2 fitness of different curves to model simulated fixations histogram (bins = 30) and decay rate of the curve measure by the exponential factor b of the best fit exponential curve ae^{-bx}

VWM	Curve Type			
	Quadratic	Exp.	Power	Exp. Decay Rate b
3	**0.982**	0.954	0.685	.0039
4	0.939	**0.969**	0.831	.0054
5	0.913	**0.970**	0.862	.0058
6	0.903	**0.972**	0.880	.0059
7	0.820	**0.984**	0.906	.0078

Note. Bold indicates best fit among three. VWM, visual working memory; Exp., exponential.

to estimate individual's VWM capacity (Cowan, 2001; Luck & Vogel, 1997; Vogel & Machlzawa, 2004). We used Cowan's formula (Cowan, 2001) to compute the estimated individual VWM capacity as described earlier in the paper. Then we asked participants to solve the 20 TSPs while recording eye movements. To eliminate the noise from individuals who make their decisions based only on local information or used a very different task strategy than the one described in our model, we only analyzed the eye-movement data of the 22 subjects in the group with high overall solution optimality. A histogram of fixation distances was plotted for each subject.

Then we fit three types of curves (quadratic, exponential, and power) to each subject's histogram of fixation distances. All of the individual participant histograms were fit very well by either an exponential curve ae^{-bx} or a power curve ax^b (Rs > .97). Fig. 5 shows examples of curve fits to the fixation distances frequencies of three participants with different VWM size. Exactly as our model predicted, histograms generated from subjects with lower VWM capacity tend to be exponential while individuals with a higher VWM capacity tend to produce a power decay fixation pattern. Whether an individual has an exponential or power eye-movement pattern correlates very well with the individual's VWM size ($R = .61$, $p < .003$, $N = 22$), as predicted by the model (see Fig. 6).

There were sufficient exponential participants to further examine individual differences. For those individuals with exponential eye-movement patterns, we found that the decay rate b of the exponential curves ae^{-bx} also correlates very well with the individual's VWM size ($R = .8$, $p < .001$, $N = 15$; Fig. 7).

We also tested the effect of VWM size on TSP solution optimality. No significant correlation was found between individuals' VWM size and their solution optimality on any of the 20 TSP problems for all the 31 participants (Rs < .29, $p > .12$). So the effect of VWM size on eye-movement pattern is not caused by trimming poorly performing participants. These
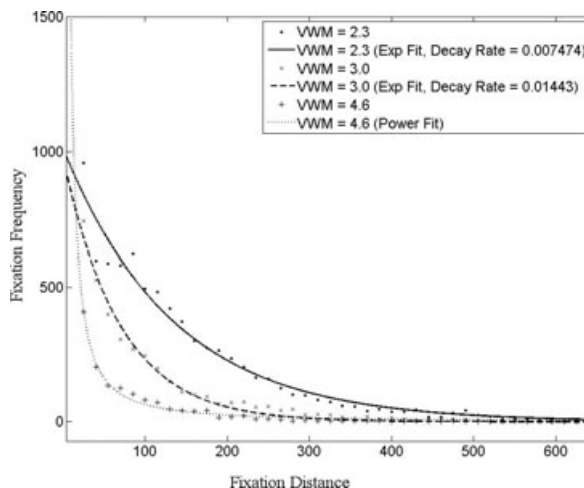

Fig. 5. Examples of fixation distance patterns for three participants with different visual working memory (VWM) capacities.
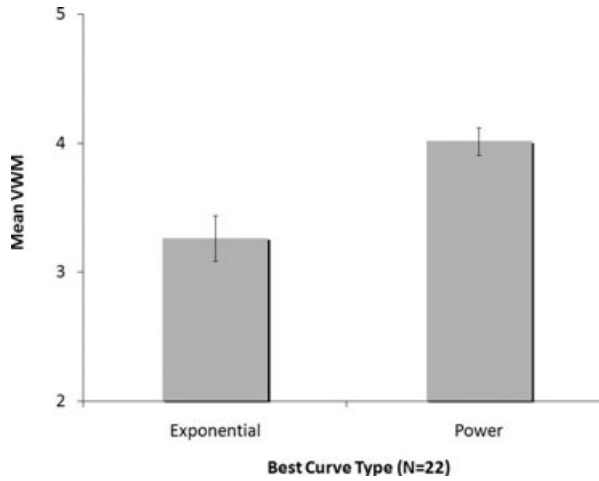
Fig. 6. Histograms of fixation distances fit better to exponential curves than power curves for individuals with smaller visual working memory (VWM) size and fit better to power curves for individuals with larger VWM.
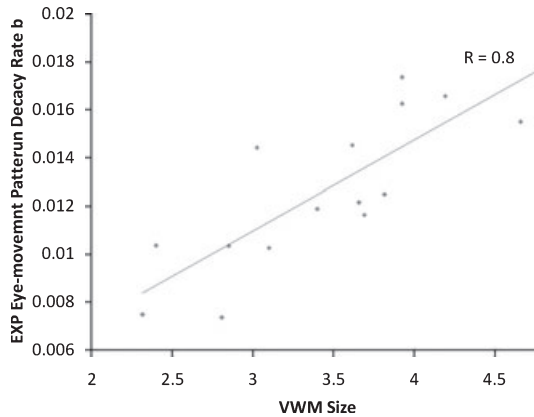

Fig. 7. Correlation between visual working memory (VWM) capacity and decay rate for exponential participants.

results provide strong support for our hypothesis, as the relationship of individual VWM capacity and eye-movement pattern is exactly as predicted by our model.

To examine whether the correlation between eye-movement pattern and VWM size was consistent across different TSPs, we divided our TSPs into two groups using two criteria: large (size > 50) versus small (size ≤ 50); and random (points randomly pregenerated using uniform distribution) versus structured (points from TSPLIB representing real-world cities/problems). Each group has 10 TSPs within our experiment. We calculated the correlation between VWM size and decay rate *b* among the same 15 subjects only using one group of problems. All four groups (large, small, random, and structured) yielded a significant

correlation (R = .56–.66, p < .05, N = 15). This consistency across TSP subtypes demonstrates that the strong relationship between VWM and eye-movement patterns appears not to depend on particular problem type or size.

We also compared the human solution performance with that produced by our model and other existing models of human TSP solving. We calculated Pearson correlation on the mean optimality of the 20 TSPs between 31 participants in the second experiment and those generated by the existing models. Compared to these other models, our model's solution for each problem correlates well with human average solution (our model: R = .73; convex hull, Golden, Bodin, Doyle, & Stewart, 1980: R = .63; sequential convex hull, MacGregor et al., 2000: R = .75; nearest neighbor, Daniel, Richard, Philip, & Ii, 1977: R = .59; pyramid, Graham et al., 2000: R = .6; K-means TSP, Kong & Schunn, 2007: R = .64). Thus, our model adequately captures problem solving as well as eye movements.

6. Discussion

For over a decade, considerable effort had been devoted to investigating the relationship between VWM and eye movements (Bays & Husain, 2008; Epelboim & Suppes, 2001; Hodgson, Dittrich, Henderson, & Kennard, 1999; Postle, Idzikowski, Sala, Logie, & Baddeley, 2006; Soto, Humphreys, & Rotshtein, 2007). Our eye-tracking experiment results for the first time demonstrate a strong correlation between individual VWM capacity and eye-movement patterns in a nontrivial problem-solving scenario. Our model successfully predicated this result as an interaction between global perception, VWM, and visual attention in human complex problem solving.

Although VWM capacity is extremely limited, the human visual system appears to represent global information of different granularity in VWM where only the most relevant information is represented at the finest details. In this way, several chunks of information not only are able to capture the local details but also the big picture. During human complex visual problem solving, the contents of VWM constantly change. The human visual system attends to local details when refining the global information into local ones and puts them into VWM. On the other hand, it also re-attends to global information when global information is needed but not represented in VWM. Individual VWM capacity plays a central role in deciding the precise pattern of visual attention.

Overall, to seek an explanation of the long-standing mystery regarding how such a limited VWM could support such complex human visual reasoning abilities, our results provide strong evidence that the human visual system dynamically allocates VWM resources to represent different granularities of global information and constantly updates its contents via attentional fixations. Although at any given moment the contents in human VWM are no more than several chunks of information, through this mechanism very complex visual information can be represented in a temporal manner to support highly effective complex human visual reasoning.

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.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), 622–628.

Baddeley, A. (1986). *Working memory*. Oxford Psychology Series, No. 11, Working memory, xi, 289 pp. New York: Clarendon Press/Oxford University Press.

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

Coren, S., & Hoenig, P. (1972). Effect of non-target stimuli upon length of voluntary saccades. *Perceptual and Motor Skills*, *34*(2), 499–508.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavior Brain Science*, *24*(1), 87–114; discussion 114–185.

Epelboim, J., & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research*, *41*(12), 1561–1574.

Ferguson, G. A. (1981). *Statistical analysis in psychology and education* (5th ed.). New York: McGraw-Hill.

Findlay, J. M. (1982). Global visual processing for saccadic eye movements. *Vision Research*, *22*(8), 1033–1045.

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, *22*(4), 661–674. discussion 674-721.

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science*, *291*(5509), 1803–1806.

Golden, B., Bodin, L., Doyle, T., & Stewart, W. (1980). Approximate traveling salesman algorithms. *Operations Research*, *28*(3), 694–711.

Graham, S. M., Joshi, A., & Pizlo, Z. (2000). The traveling salesman problem: A hierarchical model. *Memory & Cognition*, *28*(7), 1191–1204.

Hodgson, T. L., Dittrich, W. H., Henderson, L., & Kennard, C. (1999). Eye movements and spatial working memory in Parkinsons disease. *Neuropsychologia*, *37*, 927–938.

Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 307–317.

Kong, X., & Schunn, C. D. (2007). Global vs. local information processing in visual/spatial problem solving: The case of traveling salesman problem. *Cognitive Systems Research*, *8*(3), 192–207.

Luck, S. J., & Hollingworth, A. R. (2008). *Visual memory*. New York: Oxford University Press.

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

MacGregor, J. N., Chronicle, E. P., & Ormerod, T. C. (2004). Convex hull or crossing avoidance? Solution heuristics in the traveling salesperson problem. *Memory & Cognition*, *32*(2), 260–270.

MacGregor, J. N., & Ormerod, T. C. (1996). Human performance on the traveling salesman problem. *Perception & Psychophysics*, *58*(4), 527–539.

MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2000). A model of human performance on the traveling salesperson problem. *Memory & Cognition*, *28*(7), 1183–1190.

MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on mathematical statistics and probability* (pp. 281–297). Berkeley: University of California Press.

McNamara, T. P., Altarriba, J., Bendele, M., Johnson, S. C., & Clayton, K. N. (1989). Constraints on priming in spatial memory: Naturally learned versus experimentally learned environments. *Memory & Cognition*, *17*(4), 444–453.

O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature*, *401*(6753), 584–587.

Ormerod, T. C., & Chronicle, E. P. (1999). Global perceptual processing in problem solving: The case of the traveling salesperson. *Perception & Psychophysics*, *61*(6), 1227–1238.

Ottes, F. P., Van Gisbergen, J. A., & Eggermont, J. J. (1984). Metrics of saccade responses to visual double stimuli: Two different modes. *Vision Research*, *24*(10), 1169–1179.

Pizlo, Z., & Li, Z. (2005). Solving combinatorial problems: The 15-puzzle. *Memory & Cognition*, *33*(6), 1069–1084.

Pizlo, Z., Stefanov, E., Saalweachter, J., Li, Z., Haxhimusa, Y., & Kropatsch, W. (2006). Traveling salesman problem: A foveating pyramid model. *Journal of Problem Solving*, *1*(1), 83–101.

Postle, B. R., Idzikowski, C., Sala, S. D., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. *Quarterly Journal of Experimental Psychology*, *59*(1), 100–120.

Roelfsema, P. R., Khayat, P. S., & Spekreijse, H. (2003). Subtask sequencing in the primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(9), 5467–5472.

Rosenkrantz, D. J., Stearns, R. E., & Lewis, P. M. II (1977). An analysis of several heuristics for the traveling salesman problem. *SIAM Journal on Computing*, *6*(3), 563–581.

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(16), 5975–5979.

Soto, D., Humphreys, G. W., & Rotshtein, P. (2007). Dissociating the neural mechanisms of memory-based guidance of visual selection. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(43), 17186–17191.

Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature*, *400*(6747), 869–873.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, *428*(6984), 751–754.

Vickers, D., Bovet, P., Lee, M. D., & Hughes, P. (2003). The perception of minimal structures: Performance on open and closed versions of visually presented Euclidean travelling salesperson problems. *Perception*, *32*(7), 871–886.

Vickers, D., Butavicius, M., Lee, M., & Medvedev, A. (2001). Human performance on visually presented traveling salesman problems. *Psychological Research/Psychologische Forschung*, *65*(1), 34–45.

Vogel, E. K., & Machlzawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751.

Woodman, G. F., & Luck, S. J. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search?. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(2), 363–377.

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*(7080), 91–95.

Xu, Y., & Chun, M. M. (2007). Visual grouping in human parietal cortex. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *104*(47), 18766–18771.