Available online at www.sciencedirect.com
ScienceDirect
Cognitive Psychology 57 (2008) 122–152
Cognitive Psychology
www.elsevier.com/locate/cogpsych

A computational model of spatial visualization capacity ☆

Don R. Lyon [a,*], Glenn Gunzelmann [b], Kevin A. Gluck [b]

[a] *L3 Communications at Air Force Research Laboratory, 6030 South Kent Street, Mesa, Arizona 85212-6061, USA*
[b] *Air Force Research Laboratory, Mesa, Arizona, USA*

Accepted 7 December 2007
Available online 7 March 2008

Abstract

Visualizing spatial material is a cornerstone of human problem solving, but human visualization capacity is sharply limited. To investigate the sources of this limit, we developed a new task to measure visualization accuracy for verbally-described spatial paths (similar to street directions), and implemented a computational process model to perform it. In this model, developed within the Adaptive Control of Thought-Rational (ACT-R) architecture, visualization capacity is limited by three mechanisms. Two of these (*associative interference* and *decay*) are longstanding characteristics of ACT-R's declarative memory. A third (*spatial interference*) is a new mechanism motivated by spatial proximity effects in our data. We tested the model in two experiments, one with parameter-value fitting, and a replication without further fitting. Correspondence between model and data was close in both experiments, suggesting that the model may be useful for understanding why visualizing new, complex spatial material is so difficult.

Keywords: Spatial visualization; Visuospatial working memory; Mental imagery; ACT-R; Computational model

☆ This research was supported by the U.S. Air Force Office of Scientific Research (Grant 02HE01COR), and the Air Force Research Laboratory's Human Effectiveness Directorate (Contracts F1624-97-D-5000 and FA8650-05-D-6502). Portions of this research have been presented at the International Conference on Cognitive Modeling and the Annual Meeting of the Cognitive Science Society. We thank David Irwin and an anonymous reviewer for suggesting the analysis of model variants; Jerry Ball, Michael Krusmark and two anonymous reviewers for other helpful comments; Ben Sperry for software development; and Christy Caballero and Lisa Park for research assistance.
* Corresponding author. Fax: +1 480 988 2230.
E-mail address: don.lyon@mesa.afmc.af.mil (D.R. Lyon).

0010-0285/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.cogpsych.2007.12.003

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JUN 2008	**Journal Article**	**01-01-2006 to 30-06-2008**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
A Computational Model of Spatial Visualization Capacity	**FA8650-05-D-6502**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER **61102F**

6. AUTHOR(S)	5d. PROJECT NUMBER
Don Lyon; Glenn Gunzelmann; Kevin Gluck	**2313**
	5e. TASK NUMBER **AS**
	5f. WORK UNIT NUMBER **2313AS02**

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
L-3 Communications, ,6030 South Kent Street,Mesa,AZ,85212-6061	**; AFRL-RH-AZ-JA-2008-0003**

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
Air Force Research Laboratory/RHA, Warfighter Readiness Research Division, 6030 South Kent Street, Mesa, AZ, 85212-6061	**AFRL; AFRL/RHA**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) **AFRL-RH-AZ-JA-2008-0003**

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Published in Cognitive Psychology, 57(2), 122-152 (2008)

14. ABSTRACT

Visualizing spatial material is a cornerstone of human problem solving, but human visualization capacity is sharply limited. To investigate the sources of this limit, we developed a new task to measure visualization accuracy for verbally-described spatial paths (similar to street directions), and implemented a computational process model to perform it. In this model, developed within the Adaptive Control of Thought-Rational (ACT-R) architecture, visualization capacity is limited by three mechanisms. Two of these ("associative interference" and "decay") are longstanding characteristics of ACT-R's declarative memory. A third ("spatial interference") is a new mechanism motivated by spatial proximity effects in our data. We tested the model in two experiments, one with parameter-value fitting, and a replication without further fitting. Correspondence between model and data was close in both experiments, suggesting that the model may be useful for understanding why visualizing new, complex spatial material is so difficult.

15. SUBJECT TERMS
Visualization; Spatial Ability; Problem Solving; Measures (Individuals); Memory; Models; Human problem solving; Spatial paths; ACT-R; Declarative memory; Associative interference; Decay;

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Public Release**	**30**	

1. Introduction

Spatial visualization is ubiquitous in human cognition. People visualize the spatial aspects of situations ranging from mentally repositioning furniture to solving complex scientific and engineering problems. However human capacity to visualize spatial information is limited. For example, consider visualizing driving directions given over the phone. If the route is too complex, the number of turns and segments will exceed people's ability to create a clear and accurate mental image of the route. In this paper, we describe a new method for studying spatial visualization that is similar to route visualization. Data obtained using this method suggest a particular model of the underlying processes that limit human visualization capacity. We have developed and tested a computational implementation of this model to assess its validity.

Our model concerns the visualization of *verbally-described spatial information*, as opposed to memory for visually-presented pictures. While the capacity of picture memory may be related to spatial visualization capacity, picture memory may also use episodic encodings of depictive iconic information. By using verbal descriptions, we can be sure that spatial visualization performance has not been augmented with traces of pictorial information from the stimulus.

Even after restricting the focus to verbal description, there are a large number of existing studies that require some form of spatial visualization (c.f. Franklin & Tversky, 1990). However, very few of these have directly addressed the capacity of the spatial visualization system. In the remainder of the introduction, we review some previous attempts to develop an accurate view of the limitations of this ability.

Kosslyn (1980, 1994) discussed visualization capacity in the context of a general theory of mental imagery. One facet of his theory is the hypothesis that visualization uses structures normally employed for vision. In particular, mental images are distributed across a retinotopically-mapped visual buffer that is part of the visual processing stream. An aspect of this theory that is relevant to spatial visualization capacity is that this buffer has limited spatial dimensions. Kosslyn (1978) presented evidence to support this, showing that information can be lost from the image due to buffer overflow if the image is larger than the visual buffer's dimensions, which correspond to the visual field of view. Kosslyn (1980) also proposed that the visual buffer has limited resolution, another kind of capacity limit. A third aspect of Kosslyn's theory that bears upon visualization capacity is the notion that the visualization buffer is two-dimensional. One might infer from this that visualizing three-dimensional spatial material may require additional processing capacity. Potential evidence against this latter idea is the Shepard and Metzler (1971) finding that mental rotation in depth is as fast as rotation in the picture plane. However there is some uncertainty about the implications of the Shepard and Metzler finding for the dimensionality of visualization space, as subsequent studies have shown that apparent rotation rate depends on a variety of stimulus, task, and strategic considerations (e.g. Just & Carpenter, 1985; Just, Carpenter, Maguire, Diwadkar, & McMains, 2001; Shepard & Metzler, 1988).

Attneave and Curlee (1983) directly addressed the issue of visualization capacity using a variation of a verbal description task developed by Brooks (1968). In the Brooks task, participants visualized information in particular locations of a rectangular grid. Attneave and Curlee (1983) developed a variation of this task where participants were asked to visualize movements of an imaginary spot within a matrix of cells. They were given a starting loca-

tion in the (imaginary) matrix, then were given a verbal sequence of movements (left, right, up, down. . .; 0.75 s. per item), and had to point to the final visualized location. The size of the matrix was varied from 3×3 to 8×8. Attneave and Curlee found that accuracy dropped dramatically between 3×3 and 4×4 matrices, suggesting that even the 4×4 matrices exceeded visualization capacity.

A capacity of 3×3 locations may seem low, even for an ephemeral representation like spatial imagery. However, as Attneave and Curlee point out, this refers only to the capacity of the visualization workspace to place an undifferentiated marker into distinct location slots. Familiar material, such as well-learned routes and spatial information related to rich contextual visual information, can presumably be retrieved and visualized as a single unit or a few subunits. Nevertheless, even with these considerations there would still be a limited amount of information that could be maintained in visuospatial working memory at any given time.

Attneave and Curlee went on to extend this paradigm to test Kosslyn's idea that limited visualization capacity may, in part, be explained in terms of a fixed-size image buffer that can 'overflow' when one is attempting to visualize a large image. They attempted to look for overflow effects by presenting matrices that subtended both large and small visual angles. They found that performance did not differ as a function of visual angle, thus providing no evidence that visualizations were extending beyond some 'field of view' limitation in their participants. Instead, Attneave and Curlee concluded that human visualization is limited in the number of locations that can be represented, an idea similar to Kosslyn's notion that the resolution of the spatial buffer may be limited.

Kerr (1987, 1993) extended the Attneave and Curlee task further to investigate the difficulty associated with visualizing three-dimensional space, using physical (cardboard or wood) displays of the matrices. When she compared accuracy for 2D and 3D arrays of various sizes ($2 \times 2 \times 2$, $3 \times 3 \times 3$ vs. 2×2, 3×3, or 5×5), she found that, for arrays of three or fewer locations per dimension, accuracy was equally high for 2D and 3D arrays, and accuracy for the $3 \times 3 \times 3$ array was higher than for the 5×5 array. Following Attneave and Curlee (1983), Kerr suggested that visualizing three locations per dimension is within human visualization capacity, regardless of whether the array is 2D or 3D. However Diwadkar, Carpenter, and Just (2000), using a similar task, found that participants were consistently slower to verify final object location after movements through a $3 \times 3 \times 3$ array than for a 5×5 array. Here, the overall size of the space is similar, and performance was worse when those locations were within a 3D space. This discrepancy between the Kerr and the Diwadkar et al. findings may have resulted from differences in the nature of the materials (cardboard 3D models vs. computer-generated arrays) and/or the dependent measure being assessed (accuracy versus response time).

To summarize, research using variants and extensions of Brooks's (1968) original task show that spatial visualization accuracy usually drops when the visualization space is larger than three locations per dimension. However there may be reason for caution in generalizing this result beyond the location-tracking task (e.g. Lyon, Gunzelmann, & Gluck, 2006a). Barshi and Healy (2002) note that this task does not require that the entire path be simultaneously visualized, since participants only report the final location. Therefore, under some conditions, it might be possible to use a coordinate-tracking strategy. Participants might learn to compute and remember numerical or verbal descriptors that repre-

sent the coordinates of the current object location. Such a strategy could effectively eliminate the need to use spatial visualizations at all. This strategy is easier for the 2D case than for the 3D condition, since only two coordinates need to be tracked to maintain an accurate representation of the location. Also, there is good reason to suspect that differentiating three values on each axis may be relatively easy. This is because qualitative references can easily be applied (e.g., left, right, middle), a strategy that has been found for other spatial tasks (e.g., Gunzelmann & Anderson, 2006). Thus, it is conceivable that a non-spatial strategy could have been used by at least some participants in the location-tracking research presented so far. This limits the certainty with which the findings can be attributed to mechanisms of the human visualization system.

Barshi and Healy (2002) attempted to overcome the issues associated with final-location reporting by asking participants to recall the entire verbally-described path. Unfortunately, there are also potential problems with this method. For example, a participant could rehearse the sequence of segments verbally, and report them back without ever actually constructing a mental visualization of the path. Asking people to draw the path may allow this same strategy, and might also introduce spatial interference between the response and the visualization if a participant does try to visualize the path (Brooks, 1968). In view of these methodological issues, we designed a new task that requires a spatial judgment that strongly encourages the construction of a spatial representation. This task is described next.

2. The path visualization task

In the task used by Attneave and Curlee (1983) and Kerr (1987, 1993), a single response is obtained only after an entire path (sequence of locations) is presented. The researcher can measure the accuracy of this final response, but not the accuracy with which the sequence of steps leading to the final location is represented. Noting this, Attneave and Curlee stated that "...we should like to be able to calculate a less arbitrary measure such as the probability of a correct internal response on an individual step." (p. 23). The path visualization task addresses this measurement problem, providing both accuracy and response time for making a spatial judgment as each segment is added to a visualized path.

In path visualization, participants listen to a sequence of verbally-described movements through an array and make the following speeded yes-no decision *after each move:* Does the current location coincide with *any* previously visited location? A correct decision requires that the points on the path be represented in a mental form that can support this re-visitation judgment. Instead of requiring only a single location to be tracked, path visualization requires the participant to attempt to represent the entire path that has been presented, so that a revisit to any part of it can be detected. Many variations of this basic paradigm are possible (e.g. Lyon, Gunzelmann, & Gluck, 2006b). The details regarding the particular variant used in this research are described below, in the context of the empirical research.

Note that verbal rehearsal of the list of movement descriptions presented ('right', 'up', 'back', 'left', etc.) is insufficient to decide whether each movement results in a revisit to a previous location, The participant must build a cognitive representation of the path that is being described. As noted above, this is not necessarily true of other tasks. The location-tracking task, for example, might be amenable to the strategy of updating and remembering the numerical coordinates of the final location. In path visualization, this strategy

would be far more difficult to employ because, instead of tracking one set of coordinates, one would need to calculate and remember the coordinates of every location in the path. For 3D paths, at the 15th segment, this would require retrieving a match to three coordinates from a list of fourteen other three-coordinate sets. It is conceivable that, with extensive practice, participants could develop unitary representations of each of the 125 three-coordinate sets representing locations of the space. However even in this case, a participant would need to calculate each new segment's coordinates while keeping a list of other sets in mind, do the retrieval, and respond quickly (there is a 2-s deadline). No participant in our studies reported using any variation of the coordinate conversion strategy, and we have been unable to employ the strategy ourselves.

Because it is so difficult to use non-visual strategies in path visualization, this task may provide an advantage over some existing tasks for studying spatial visualization. In our view, the natural strategy for accomplishing path visualization is to attempt to visualize the path. This means that by understanding path visualization performance, we can hope to gain insight into the nature of visualization capacity. We attempt to do this by developing a computational cognitive model to perform the path visualization task and comparing its performance to human participant performance. The cognitive model and the theory it embodies are described next.

3. Modeling visualization capacity

To fully understand capacity limits in the context of the path visualization task, it is necessary to accurately model not only the processes of generating, maintaining, and using visualized spatial information, but also other critical components of human performance on this task. These include the perceptual processes of encoding the stimulus information from the screen, the motor mechanisms for responding, as well as the cognitive mechanisms that are used to reason about the stored information to determine whether to respond 'yes' or 'no.' To capture all of these components of performance requires a more comprehensive theory of the human cognitive architecture than merely an account of visuospatial working memory capacity and processes alone. We chose to embed our model within a cognitive architecture in order to enhance its psychological plausibility by leveraging mechanisms that have been shown to provide effective accounts of human performance across a broad range of domains and tasks. In particular, we use ACT-R (Adaptive Control of Thought-Rational), which contains an extensively tested set of mechanisms that have been successful in modeling other cognitive processes (Anderson, 2007; Anderson et al., 2004; Anderson & Lebiere, 1998).

ACT-R is a cognitive architecture, implemented in software, which contains general mechanisms to account for human cognitive performance on a variety of tasks (see Anderson & Lebiere, 1998 for a review). Central cognition in ACT-R is represented by a serial production system, which integrates the outputs from specialized processing modules containing mechanisms responsible for different aspects of human cognitive performance. For instance, there is a vision module, with mechanisms for directing attention and encoding visual information from the computer screen. There is also a declarative module, which functions in the storage and retrieval of knowledge in the form of 'chunks.' Each of these modules operates serially (i.e., process a single request at a time), but the set of modules operates in parallel. For this research, it is notable that there is no 'visuospatial module' or mechanisms to handle the generation, encoding, and manipulation of internally generated

visual images. For that reason, in the model we describe below, we actually utilize the existing declarative module, with its associated mechanisms, for implementing and testing our account.

Although we employ ACT-R's declarative memory in our model, we make no claim regarding the nature of the system that implements these dynamics in the brain for spatial visualization tasks. Neither our data nor our model provide firm evidence for either 'propositional' or 'depictive' representations of spatial visualizations. Indeed, we accept the argument (Anderson, 1978) that either kind of representation, when combined with an appropriate set of processes operating on it, could produce a particular computational result.

However, this does not mean that spatial visualization and declarative memory for verbal material must use identical cognitive processes. Indeed, the results of our experiments suggest that at least one additional parameter must be added to standard ACT-R declarative memory operations in order to account for human path visualization accuracy. An accumulation of this kind of evidence could lead to modifications of the architecture to better account for spatial visualization, and perhaps other spatial aspects of cognition, a possibility we have begun to explore (Gunzelmann & Lyon, 2008).

ACT-R's perceptual and motor modules give it the ability to interact directly with software-based tasks under realistic timing constraints based on psychophysical research. Thus, the model we describe here acts just like a participant in the experiment by encoding the description of the path segment, processing that information to determine where in the space that segment leads and whether it revisits a location on the previous portion of the path, and finally eliciting a virtual keypress to indicate its response. The portion of the model that we focus on in this paper is the process of determining whether or not a new segment revisits a previous path location. Other descriptions of ACT-R provide detailed accounts of the functioning of the perceptual and motor portions of the architecture (e.g. Byrne & Anderson, 1998). In this model, we use default mechanisms and parameter settings for encoding the stimulus from the monitor and for generating responses once the response choice has been made.

The keys to the model's performance are the representation used to identify the locations visited in the space as segments are added to the path during a trial, combined with quantitative activation mechanisms borrowed from ACT-R's declarative memory, which influence the availability of information about previously visited locations. Our account is based on the notion of a 'spatial field,' which is an internally-generated space used to represent the path as additional segments are added. The directions of individual path segments are represented in absolute terms with reference to a $5 \times 5 \times 5$ externally-viewed space, rather than in egocentric terms from the point of view of an observer on the path. We chose this representation because, in the task itself, segment directions are presented in absolute terms. 'Left' always denotes the left side of the space viewed externally, not the left-hand side of a viewer on the path. In addition, participants saw a representation of this space at the beginning of the experiment, which reinforced this reference frame (see Fig. 1). As a consequence, the natural representation for the paths is an allocentric one. There is no need for either a participant or the model to perform the additional step of translating a segment description to an egocentric perspective, and no-one reported doing so.

After each new segment description is presented, the model generates the location of the new end of the path, and tags it using a 3-digit number that corresponds to the coordinates

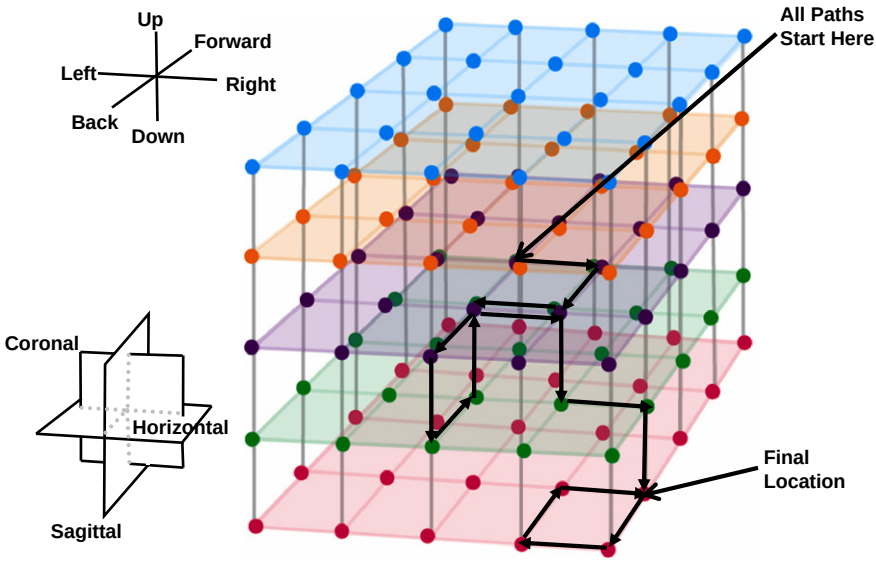

Fig. 1. Example of a path to be visualized. Participants did not see this picture during the task. Instead they visualized paths described by a sequence of verbal directions. All paths started at the center of the space. 3D paths could wander throughout the space, as illustrated. 2D paths were constrained to coronal, sagittal or horizontal planes.

of that location in the space. For example, if the current endpoint is at the center of the $5 \times 5 \times 5$ space (tagged '333') and the path grows by extending to the right, a new location would be visualized, and given the tag '343'. These tags are for computational and descriptive convenience only. That is, they are utilized in computing Euclidean distances between points, but the model does not reason explicitly about the numerical values or relations.

This location generation process represents the first step in making a judgment as to whether or not the segment revisits a location. It is always accurate in the model, reflecting a simplifying assumption that the model always has an accurate representation of the meaning of each new segment description relative to its current location in space. For example, it never erroneously generates a new location to the *right* of the current one when the segment description says '*Left*'.

Note that evidence from Kerr (1987, 1993) may be relevant to the validity of this simplification, since she showed that accuracy drops on a location tracking task for spaces this large. However, there were other methodological differences between Kerr's studies and ours, including a faster presentation rate in her studies. Time pressure could lead to increased errors in correctly updating spatial information based on verbal directions such as 'right' and 'left'. However when Kerr used a presentation rate of 1 s (still twice as fast as the rate in our experiments), overall accuracy was above 90%, regardless of the space being used, so if there were descriptor translation errors in Kerr's data, they were not an overwhelming factor, even at this fast presentation rate.

Another factor that we believe would tend to minimize misinterpretation of path descriptors is our use of allocentric descriptors, as noted above. It would have been more difficult to interpret terms such as 'left' and 'right' if they referred to an egocentric heading that was not the same as the viewer's perspective on the space. By defining the terms rel-

ative to a particular natural external perspective, we hoped to minimize the frequency of such errors. Although it is possible that an occasional translation error occurred even with allocentric descriptors, we believe that the assumption of accurate translation of the path descriptors is appropriate for this initial modeling effort.

Because the model accurately updates its current location within the space, the difficulty in the task stems from determining whether the current location corresponds to any previous position on the path. It is in this process that mechanisms associated with ACT-R's declarative memory module are utilized to capture the critical components of our 'spatial field' account of human performance. These components of the architecture, and their relevance to the model's performance, are described next.

3.1. Critical ACT-R mechanisms

The declarative memory component of ACT-R has been extensively validated to account for a variety of memory phenomena, from list learning (Anderson, Bothell, Lebiere, & Matessa, 1998), to the fan effect (e.g., Anderson, 1974; Anderson & Reder, 1999), to the representation of arithmetic facts (Lebiere, 1999). Notably, ACT-R's declarative memory mechanisms have not been used to address phenomena associated with spatial visualization. However, we hypothesized, based on our early results (Lyon et al., 2006a), that many of the same mechanisms might be appropriate for representing the availability of visualized spatial information. The key processes for our model that are implemented in ACT-R's declarative memory are practice-based increases in activation, delay-based decay of activation, spreading activation, and similarity-based partial matching. The interaction of these processes provides extensive explanatory power in traditional memory paradigms, and we believe the mechanisms capture important dynamics associated with visuospatial working memory as well.

Each of these declarative memory mechanisms contributes to the predictive utility of our account. First, repetition leads to higher levels of activation of declarative 'chunks' (units of knowledge), which makes them more accessible both in terms of ability to retrieve them as well as how quickly the information can be retrieved. Conversely, not using a chunk allows its activation to decay. These mechanisms result in the Power Law of Practice and the Power Law of Forgetting in ACT-R. Second, spreading activation allows the current context to influence the level of activation of the chunks in declarative memory, by boosting the activation of related chunks in memory. And, third, partial matching gives ACT-R the ability to retrieve chunks from declarative memory that are not exact matches to the particular requests that are made, but which are similar on various dimensions. The equations governing these processes are presented in Appendix A.

The interaction of ACT-R's declarative mechanisms provides an account for the detailed dynamics of human memory, including both errors of commission and errors of omission. In path visualization, both of these are important, since participants sometimes fail to recognize revisits (an error of omission), and also erroneously indicate that a revisit has occurred when it has not (an error of commission). The similarity mechanism is particularly important in the context of this model because we use similarity to establish the relationship between locations in the space to create the spatial field. As noted above, each segment of the path ends at a particular location. We use the similarity mechanism in ACT-R's declarative memory to represent the 'closeness' of those locations in 3-dimen-

sional space. A Euclidean distance[1] (D) is calculated between points, and this value is scaled by a spatial interference parameter (SI) to compute match similarity (M), using Eq. (1):

$$M = \mathrm{SI} * \left(\frac{1}{1+D} - 1\right) \tag{1}$$

This equation is used to define the similarity between two spatial location slot values in chunks in declarative memory, where the maximum value of zero denotes identical values, interpretable as representing no mismatch between the values. As illustrated in Eq. (A3) in Appendix A, the similarity value impacts the likelihood of a chunk being retrieved by determining the 'mismatch penalty' that is assessed to chunks with slot values that do not exactly match the retrieval request.

The impact of the similarity mechanism is to make it increasingly unlikely that a particular chunk will be retrieved as its disparity from the request made to the declarative memory module increases. On the other hand, this mechanism allows similar chunks to be retrieved in response to requests, even if they do not match exactly. We now describe how declarative memory mechanisms are applied in the current model to instantiate our theory of visualization capacity.

3.2. Model design

To understand the way in which ACT-R's declarative mechanisms were used to capture processes in visuospatial working memory, it is necessary to describe in some detail the process that the model uses to make its judgment as to whether a revisit occurred or not. As noted above, the current path location within the $5 \times 5 \times 5$ space is tagged with a 3-D coordinate, with the center of the space being '333.' Path visualization trials always start at this center point. After a move command (e.g., "Up 1") is given, a declarative chunk representing the new location (after adding the described segment) is stored in declarative memory, creating an episodic trace of a segment to that location. These chunks accumulate as additional segments are presented, and serve as the information on which a judgment is made as to whether the location in the space has been visited previously or not. In the current model, this judgment is made by requesting a retrieval from declarative memory for a chunk that indicates the same point in the space as the current location. If this retrieval request results in a chunk from declarative memory being successfully retrieved, then the model responds 'yes.' Otherwise it responds 'no.'

This decision-making process is modulated by the declarative memory mechanisms described above. Similarity is influential in this process, because locations that are nearby to the current location have representations that are more similar than locations that are farther away. Thus, if the newest segment of the path does not produce a revisit, but visits a point in the space that is nearby to other points that have been visited, there is a greater likelihood that the model will erroneously respond that a revisit has occurred, as compared

[1] Given that the task involves city-block progressions through the space, one might suspect that city-block distances would work better in evaluating the impact of crowding. Unfortunately, Euclidean and city-block metrics only start to really diverge when distances exceed 1 (i.e. non-adjacent visits), and non-adjacent visits don't have much effect on accuracy. Thus, this particular version of the task does not provide a strong means for evaluating the relative merit of Euclidean versus city-block metrics for measuring distance in visualization space.

to segments that lead the path into a portion of the space that has not been visited previously. As will be demonstrated in the empirical data below, this is a key phenomenon that is observed in human performance.

In addition to the errors of commission just described, the model produces errors of omission. That is, it is possible for the model to retrieve *nothing* on a particular segment, even when the point in the space has been visited previously. This is because ACT-R contains a retrieval threshold, which sets the minimum level of activation that must be attained for a chunk to be retrieved. If no chunk exceeds the retrieval threshold, then no chunk is retrieved. In the model, this is used as evidence that the point in the space has not been visited before, and often this is the right conclusion. However, decay in activation with the passage of time and stochastic noise in activation values mean that sometimes the point has been visited, but the chunk representing that location in memory is not available.

An important consideration is that, although this model uses retrieval and decay mechanisms from ACT-R's verbal declarative memory, it uses them to emulate spatial visualization, rather than a verbal mediation strategy. When the model attempts to retrieve a particular prior visit, one way to interpret this retrieval is that it represents the process of visualizing a location to 'see' if anything is there from the prior path. Other interpretations may be possible, but under any interpretation, an important feature of our model is that its performance is *not* produced by reasoning about either numerical coordinates or verbal descriptions such as 'near right', 'far above', etc.

We believe that this model accurately explains the major influences on human performance in the path visualization task. It embodies the following three-part theory of the nature of capacity limits in spatial visualization:

1. *Visualized elements become less available with time*. Our preliminary path visualization data showed that the probability of detecting a revisit declines with the time between visits to a location. In ACT-R, such effects are accounted for through the activation decay mechanism.
2. *Activation is shared among similar elements*. This mechanism is part of most ACT-R memory models, and accounts for phenomena such as the fan effect.
3. *Visualized elements in nearby locations interfere with each other*. An implication of this is that spatial visualization mimics some aspects of real space. We propose a new, proximity-based mechanism to account for this effect.

The remainder of the paper presents two experiments conducted to test our instantiation of this theory in a computational process model implemented in ACT-R. Experiment 1 tests the model's ability to mimic human visualization capacity limits, as evidenced by a drop in revisit detection accuracy as a function of path length. The data from this experiment were also used to obtain best estimates of the model's parameters. Experiment 2 is a test of the model on a new dataset, where the model was used to make predictions about human performance using the parameter values derived from Experiment 1.

Implementing the spatial field model using an architecture such as ACT-R, in which many other cognitive tasks have been modeled, has the virtue of connecting the processes involved in spatial visualization with those involved in other cognitive processes, thus working toward an integrated and more generalizable model of cognition. However in any ACT-R model there will be several global parameters reflecting the generality of

the model across a wide variety of tasks. Varying all of these parameters could result in overfitting a particular dataset, and could make extending, or even interpreting, the model's performance difficult (e.g., Roberts & Pashler, 2000). Therefore most of the parameter values for the spatial field model were left unchanged from their default values, established over the history of applying ACT-R to a wide range of tasks.

Only two parameters of the model were initially varied to improve the quantitative fit to the data. These are: (1) a new spatial interference parameter, which has no established values from previous research, and (2) the retrieval threshold, which is typically varied in memory experiments to "...map activation levels onto performance" (Anderson et al., 1998, p. 250) to calibrate for specific task and stimulus characteristics. The new spatial interference parameter affects the probability of retrieving from memory a prior visit to a *nearby* location in space while trying to retrieve a visit to the current location. It does not directly affect the probability of retrieving a prior visit to the current location, if there has been one. In contrast, the retrieval threshold parameter affects the *overall* probability of retrieving an item from memory. Changes to this parameter will affect the model's overall relative proportion of revisit versus no-revisit responses.

Our primary interest was to model the accuracy of path visualization, so the parameter estimation was based only on accuracy data from the experiments. We also collected response time data as a check on the possibility of a speed/accuracy tradeoff under certain conditions.

4. Experiment 1

As noted above, the spatial field model embodies a three-part, activation-based explanation of visualization capacity. One prediction derived from these mechanisms is that revisit detection accuracy should decline substantially as more and more segments of a long path must be visualized. To evaluate the extent of this decline in humans, we compared the performance of the model to human visualization accuracy using very long (15-segment) random paths that are extremely difficult to visualize correctly.

4.1. Methods

4.1.1. Participants

Fifteen people with normal vision (eight men and seven women) were paid to participate in the study for a total of 5 hours, 1 hour per day.

4.1.2. Materials

In this experiment, participants completed a version of the path visualization task using text-based descriptions of paths in a $5 \times 5 \times 5$ imaginary space. Paths began at the center of the space, and were described as a sequence of segments. Each segment was presented as a text phrase consisting of a direction and a length (the length was always 1 unit). There were six possible directions (Right, Left, Forward, Back, Up, Down). Directions were given in absolute terms, as shown in Fig. 1. So, "Forward" always corresponded to movement along the horizontal plane, 'into' the space, from the viewpoint assumed in Fig. 1. No two successive segments could be on the same axis, so each new segment resulted in a 90-degree change in direction in the path. Since the directions referred to an external view of the space, the direction of a prior segment did not affect the description of the next

segment. For example, the descriptor 'Right' always referred to a segment drawn toward the right-hand side of the space (viewed externally), *not* to the egocentrically-defined right of a hypothetical traveler on the path. Fig. 1 shows a sample path in this space. Participants were shown an image of the space prior to beginning the experiment, but it was not available during the trials. Participants were asked to visualize the path growing through the space as the trial continued.

Each path was 15 segments long, and consisted of a quasi-random sequence of 90-deg turns. Unfortunately, fully random three-dimensional paths contain a very low proportion of revisits (about 20%). In preliminary experiments, this led some participants to exhibit a bias to respond 'no'. Therefore in this experiment we used a set of 3D paths with a somewhat higher proportion of revisits (30%), and we included a high proportion (50%) of paths that were restricted to one of three 2D planes (horizontal, coronal or sagittal, see Fig. 1). Since random 5×5 2D paths have 50% revisits, we were able to raise the overall proportion of revisits in the study to about 40%. It would have been desirable to have an overall proportion of revisits of 50%. Unfortunately, sets of 3D paths that approach 50% revisits have many paths with either long single-plane sequences or repeating loops, and are therefore quite unrepresentative of 3D paths in general.

4.1.3. Procedure

The experiment consisted of ten half-hour sessions, with two sessions per day. For each participant, before the first session, a set of 1000 quasi-random paths was generated. This set contained 50% 3D paths and 50% 2D paths, the latter divided equally between horizontal, coronal and sagittal planes. A session consisted of 30 trials, so only 300 paths from each path set were ever presented. For each trial, a path was randomly selected without replacement from that participant's 1000-path set. Each path required 15 revisit judgments. Within a trial, each segment description was presented for 2000 ms, followed by a blank screen for 133 ms, and then the presentation of the next segment description. The segment descriptions (for example, "Right 1") were presented as large yellow text on a blue background, using a standard PC monitor. To facilitate visualization, stimuli were presented in a nearly dark room, with faint illumination sufficient to locate the response keys. Response times were measured from the onset of the stimulus screen to a key-press on a standard PC keyboard. Viewing distance was approximately 70 cm.

As each new segment description was presented, the participant decided whether or not the endpoint of the new path segment was a location that had been part of the path defined by the prior segments. Participants were instructed to consider the starting location for each trial as part of the path. If the participant believed that the segment resulted in a revisit of a location in the prior path, the right arrow key was pressed with the right index finger; if not, the left arrow key was pressed with the left index finger. In the rare event (0.7% of responses) that no key was pressed during the presentation of a text phrase (a 2 s response window), the response was scored as incorrect, and the presentation of the next phrase proceeded normally. Participants were instructed to respond as accurately and quickly as possible. Small bonuses were paid for maintaining high overall accuracy and low response time.

4.1.4. Data analysis

Since it is impossible in this task to return to the same position in the array in less than four segments, no revisits can occur for the first three path segments. Therefore partici-

pants were instructed that 'no' would always be the correct answer for Segments 1–3. Accuracy and response time for each of the remaining twelve segments were computed across all paths for each participant.

For initial model testing, the primary result of interest was the extent to which revisit detection accuracy drops as path length increases, increasing the load on the visualization system. Human data and model predictions will first be compared on this aspect of the data, both for the initial experiment (Experiment 1) and for the replication described in Experiment 2. Following this, a more detailed analysis of other aspects of the data will be presented.

4.2. Results

4.2.1. Human data

The data confirm that long (15-segment) paths exceed human visualization capacity. Revisit detection accuracy declined substantially as the path length increased ($F(11,154) = 56.9$, $p < 0.001$). Detection accuracy at Segment 15 was only 78%, whereas accuracy was 93% for Segment 4. This decline was not due to a speed-accuracy tradeoff, since response time increased with increasing path length ($F(11,154) = 7.21$, $p < 0.003$).

4.2.2. Model fit

Because we used standard values for most of the model's parameters, we could fit the model by varying only the spatial interference (SI) and retrieval threshold (T_r) parameters. Predictions were generated by presenting the model with the same path sequences that were given to the participants.[2] The model was run once for each of the 300 paths per participant, for a total of 4500 model runs. Fig. 2 shows the accuracy of the model's responses using the best-fitting set of parameter values ($SI = 2.1$, $T_r = -0.9$), matched against human data ($RMSD = 0.020$, $r = 0.95$). For comparison, the mean 95% confidence interval for the human data points, CI_{data}, was 0.109.

4.3. Discussion

The results of Experiment 1 confirm that, in a revisit detection task that cannot easily be performed without actually visualizing the path, accuracy declines substantially as the load on the visualization system increases. The spatial field model accounts for this decline as a joint effect of spreading activation, activation decay, and spatial interference acting on the base-level activation of a visualized item. As is evident from Fig. 2, the model predictions fall well within the range of accuracy shown by the participants.

The results of the experiment confirm the predictions of the model and provide support for the mechanisms we have proposed as determinants of spatial visualization capacity. We find it encouraging that limitations in spatial visualization can be captured using a combination of mechanisms that have been extensively validated for declarative memory, plus a new spatial interference process. Regardless of how different kinds of information are represented in the brain, it seems reasonable to suspect that the common neural instantiation would give rise to at least some similar processes in both verbal and spatial memory

[2] Each participant saw a different set of randomly generated paths. However, because of potential effects associated with the statistics of the paths (e.g., probability of a revisit), we did not generate new random paths for evaluating the performance of the model, but rather used the same paths that each participant used.

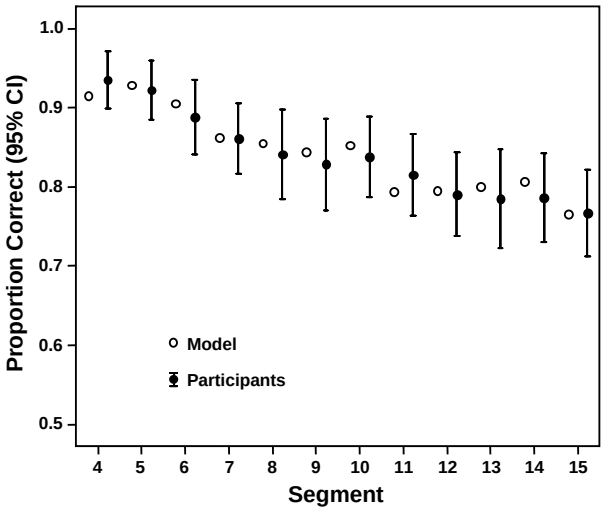

Fig. 2. Human and model accuracy for individual path segments in the path visualization task, Experiment 1.

systems. However, to increase our confidence in the account we have proposed, we conducted an additional evaluation of the model. Experiment 2 provides a replication of the essential features of Experiment 1, using different participants and different paths. Thus, we can ensure that the performance observed so far is not the result of unique features of either the participants or the stimuli tested in Experiment 1. Also, it would be desirable to assure that the model makes accurate predictions for another dataset without varying any of the parameter values. We therefore generated predictions for Experiment 2 using the parameter values derived from Experiment 1.

5. Experiment 2

Experiment 2 was a replication of Experiment 1 using different participants and paths. It was performed in the context of a larger study comparing path visualization for allocentric path descriptors (as used here) and egocentric (observer-on-path) path descriptors. Data from the latter condition are neither reported nor modeled here, since allocentric and egocentric perspectives are cognitively distinct (e.g. Avraamides, 2003; Gugerty & Brooks, 2004) and are therefore likely to require different cognitive processes.

5.1. Method

Details of the method were similar to Experiment 1. Thirteen paid participants with normal vision (six women and seven men) were given ten 30-trial path visualization sessions, two sessions per day. Each day, one session was an exact replication of Experiment 1, using the same apparatus, materials and procedure. A second daily session provided data (not reported here) on the other variation of the path visualization task mentioned above. Thus, the data described below is from 5 half-hour sessions (150 trials) for each participant. The participants were unaware of the purpose of the study. None of them had participated in any similar study before.

5.2. Results

5.2.1. Human data

As expected, revisit detection accuracy declined substantially ($F(11,132) = 35.0$, $p < 0.001$) and response time increased ($F(11,132) = 30.1$, $p < 0.001$) as the path length increased.

5.2.2. Model predictions

Fig. 3 shows human and model accuracy by path segment. Model predictions again matched the human data (RMSD = 0.023; $CI_{data} = 0.095$; $r = 0.95$). The model accounts for the decline in revisit detection accuracy as path length increased, with no parameter adjustments or any other modification. The accuracy of these zero-free-parameter predictions shows that the success of the model in accounting for the data in Experiment 1 was not due to overfitting. Moreover, potential issues of transfer and practice that might have arisen because participants also performed an egocentric version of the task did not noticeably affect the model fit. Therefore, the model is able to capture human performance on this task when the participants and stimuli vary. We acknowledge that this replication does not address the generality of the model across variations in the task. That issue is an area of focus in our current research efforts.

6. Analysis of combined accuracy data

Since the fit of the model to the path length accuracy data was nearly equal for both experiments (RMSD: 0.020 vs. 0.023; $r = 0.95$ in both cases), additional tests of the model's performance were conducted using the combined data from all 28 participants. In particular, we examined model vs. human data on 2D vs. 3D paths and revisit vs. no-revisit

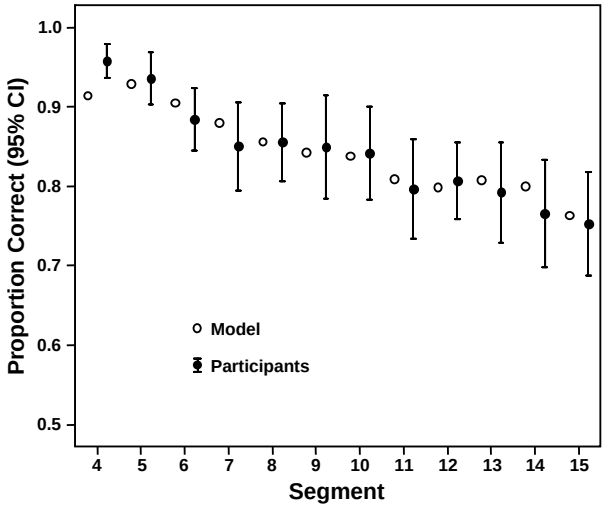

Fig. 3. Human and model accuracy for individual path segments in the path visualization task, Experiment 2.

path segments. We also tested the model's predictions regarding the effects of number of segments intervening between revisits to a location, which addresses the issue of how decay may impact performance. Finally, we looked at the interfering effects of spatial proximity, or crowding, of path segments.

6.1. Predictions for 2D vs. 3D paths

In the introduction, we cited some studies that examined the visualization of both 2D and 3D spatial information. An issue that arises in such studies is whether 3D information is processed differently than 2D information. There are proponents on both sides of this issue. One view is that visual processing in the brain reflects the inherent dimensionality of the world, and visual representations are (in important ways) analogous to 3D space (a 'sandbox in the head;' Attneave, 1972). As noted earlier, the influential early results of studies comparing rotation in depth to rotation in the picture plane (e.g. Shepard & Metzler, 1971) seem to support this view.

The alternate view is that visualizing 3D versus 2D stimuli involves fundamentally different representations and/or processes. Just et al. (2001) argued that mental rotation performance supports this view under some conditions. As noted earlier, Diwadkar et al. (2000), using a location tracking task, found that verifying final object location was slower for $3 \times 3 \times 3$ arrays than for 5×5 arrays. Moreover, fMRI measures showed that activation in parietal cortex was significantly greater for the 3D condition. Diwadkar et al. interpreted these results to suggest that 3D space is more difficult to represent than 2D space.

Unfortunately it is difficult to address this issue definitively by comparing performance on 2D and 3D materials. For example, in path visualization, if differences in accuracy between 2D and 3D paths are observed, one cannot necessarily ascribe them to the dimensionality of the paths *per se*, because there may also be inherent differences in path statistics such as proportion of revisits and mean proximity of path segments. Alternatively, explaining an absence of performance differences in the two cases faces all of the challenges associated with accepting the null hypothesis, in conjunction with alternative explanations based upon stimulus properties like those just mentioned.

While we recognize the inherent difficulty of settling this issue experimentally, the computational cognitive model we have described implicitly embodies the view that there is no qualitative difference in how 2D and 3D information is processed. That is, the spatial field model proposes no special cognitive processes for dealing with either 2D or 3D paths. The most relevant spatial parameter, spatial interference, is based on Euclidian distance, with the same computational mechanism and parameter value for paths of two and three dimensions. Therefore to the extent that the spatial field model makes different predictions for 2D and 3D paths, these differences would be due to path statistics, rather than to inherently different cognitive mechanisms. On the other hand, if people's performance does not match the model's predictions for either 2D or 3D paths, then we must consider the possibility that, unlike the model, people use somewhat different processes for 2D vs. 3D material.

Fig. 4 shows model predictions and human data for the 2D vs. 3D path conditions. The model, which invokes no additional cognitive processes for 3D paths, fits the data reasonably well for both kinds of paths (*2D*: RMSD = 0.019; $CI_{data} = 0.074$; $r = 0.97$; *3D*:

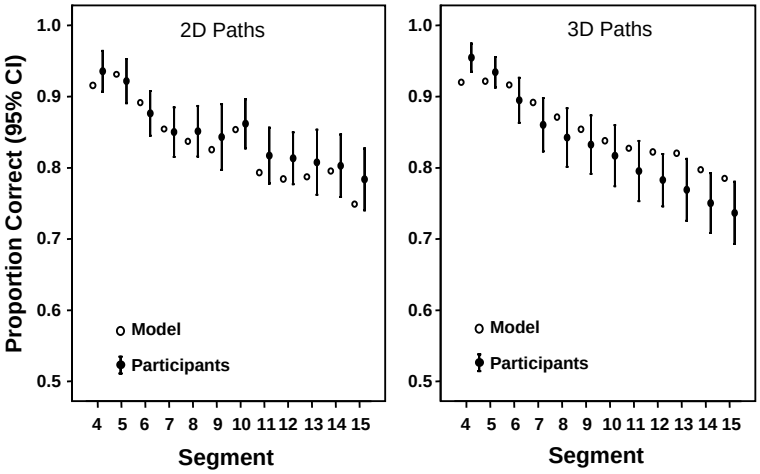

Fig. 4. Human and model accuracy for 2D paths (left panel) and 3D paths (right panel), Experiments 1 and 2 combined.

RMSD = 0.035; CI_{data} = 0.073; r = 0.97). This result could be viewed as support for the notion that processing of complex 2D material encounters essentially the same sources of capacity limits as the processing of 3D material. However another interpretation is that the model predicts well for both 2D and 3D paths in this paradigm because all of the paths occur in a pre-defined 3D space, and therefore people recruit cognitive processes relevant to 3D material even for 2D paths. Thus, our results are not conclusive with regard to the existence of different underlying cognitive processes for 3D and 2D material. However the data do confirm the success of the spatial field model in explaining visualization capacity limits for both 2D and 3D complex paths.

Our results for 2D and 3D paths are potentially relevant to another issue: the effect of the size of the visualization space on visualization accuracy. As noted earlier, studies using the location tracking task (e.g. Attneave & Curlee, 1983; Kerr, 1987, 1993) found that accuracy is generally lower for larger spaces, at least for spaces with more than three locations per dimension. Since our 2D paths wander in a plane containing many fewer locations than our 3D space, one might wonder why this does not produce an advantage for 2D paths in our data. One possible explanation was mentioned earlier—that participants use a 3D visualization space even for 2D paths. Another consideration is that, in the location tracking task, if one becomes lost, the probability of guessing the final location will depend directly on the number of possible choices—the size of the space. However, in path visualization, one can fail to retrieve part of the prior path, but nevertheless continue to visualize the succeeding segments, and detect revisits that occur within these later segments. Unlike the location tracking task, in path visualization one does not have to choose from among all locations in the space in order to make a response. Our model reflects this aspect of path visualization. 'Empty' locations do not compete with previously presented path segments in memory. Consequently the model does not predict an effect of the size of the *space* for the path visualization task; performance is only affected by the size and nature of the *path*.

6.2. Predictions for revisit vs. no-revisit cases

In the spatial field model, the mix of influences on performance is somewhat different for path segments that do not result in a revisit, and those that do. The likelihood that the model will answer correctly in a revisit case depends largely on how many segments have intervened since the location was last visited. The model's accuracy in no-revisit cases depends strongly on the number of previous visits to nearby locations. Therefore it is important to test whether the model correctly captures human data for these two decision cases. Fig. 5 shows model predictions and human accuracy data separately for no-revisit cases (RMSD = 0.045; $CI_{data} = 0.072$; $r = 0.94$, left panel) and revisit cases (RMSD = 0.103; $CI_{data} = 0.149$; $r = 0.64$, right panel). While the model accurately captures the trends in the data as a function of path length, it predicts a larger accuracy difference between revisit and no-revisit cases than was observed. It underpredicts accuracy for no-revisit cases, and overpredicts for revisits.

One possible explanation for this pattern of prediction error has to do with ACT-R's activation decay parameter. This parameter influences the drop in base-level activation of an item in memory over time. Close examination of the mechanics of the spatial field model reveals that activation decay has opposite effects on the model's decision for no-revisit and revisit cases. An increase in the rate of decay will reduce the activation of the chunks in memory that represent path segments. For no-revisit cases, errors are caused by the accidental retrieval of nearby path segments. The less active these interfering segments are, the less likely such errors will be. Therefore, increasing the rate of activation decay actually improves the model's accuracy in no-revisit cases. However, this comes at the cost of accuracy in revisit cases, since reduced levels of activation for previous segments also reduces the odds that a previous segment at the current location will be correctly retrieved. Therefore a general increase in the value of the activation decay parameter improves the model's prediction in both cases, by both reducing revisit accuracy

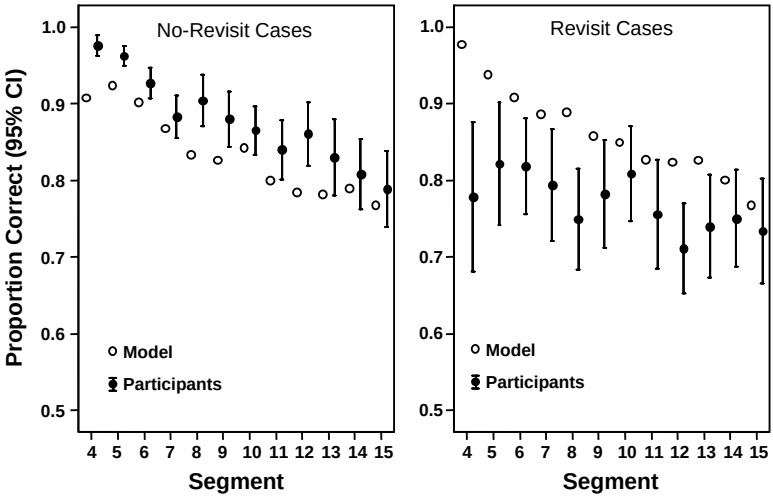

Fig. 5. Human and model accuracy for no-revisit segments (left panel) and revisit segments (right panel), Experiments 1 and 2 combined.

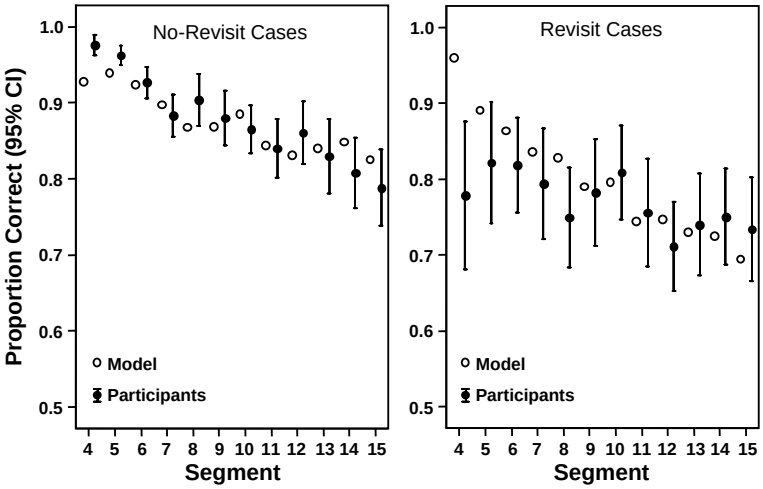

Fig. 6. Human and model accuracy for no-revisit segments (left panel) and revisit segments (right panel) after increasing the model's activation decay rate parameter from 0.5 to 0.55, Experiments 1 and 2 combined.

and increasing no-revisit accuracy. Fig. 6 shows the results of increasing the decay rate parameter in the model by 10%, from 0.5 to 0.55. With this modification, the model's predictions are indeed closer to the data (No-Revisit: RMSD = 0.027; $CI_{data} = 0.072$; $r = 0.90$; Revisit: RMSD = 0.096; $CI_{data} = 0.149$; $r = 0.66$), although the predictions are not within confidence intervals for all path lengths.[3] Since increasing decay affects revisit and no-revisit accuracy in opposite directions, it has little effect on the model's overall proportion of correct responses ($Pcorr_{0.5} = 0.846$; $Pcorr_{0.55} = 0.841$), hence the model's fit to the average accuracy loss with increasing path length remains good (RMSD = 0.010; $CI_{data} = 0.070$; $r = 0.99$).

Unfortunately, changing the value of the decay parameter violates the goal of keeping constant the values of parameters that are hypothesized to capture the fundamental characteristics of human memory as modeled in ACT-R. However, there are two potential considerations here. First, it may prove to be the case that small changes in the value of some parameters will be necessary to account for individual differences in various cognitive processes. If so, then perhaps different samples of individuals may also be best fit by slightly different parameter values. Second, the default parameter values in declarative memory are derived largely from studies using tasks like verbal list learning (Anderson et al., 1998). It is possible that representations of visualizations like those required for this task may be more susceptible to decay. This gets back to an issue we addressed above. While we are utilizing ACT-R's declarative memory as a means of implementing our theory, that should not be viewed as a theoretical claim that verbal and visualized material share the same cortical *representation*. Rather, we regard our work as an illustration that

[3] In particular, human accuracy for the seemingly easy 4-segment paths was well below that of the model. A revisit in a 4-segment path can only be created by a 'box' pattern that returns to the starting location. We suspect that sometimes the participant may forget that the starting location counts as a 'visited' location, and may therefore fail to detect this first revisit.

the same *processes* may operate in both representational systems. We address this issue further in the General Discussion.

6.3. Effect of intervening segments

The ACT-R activation decay mechanism, when applied in the spatial field model, makes a strong prediction about the likelihood of correctly detecting a revisit—the more segments that intervene between visits to a location, the less likely revisit detection should be. This is because each intervening segment takes time, during which activation of the chunk representing the visited location decays, making it less likely to be retrieved. Although we did not design our study explicitly to test this prediction experimentally, it can be examined by plotting revisit detection accuracy as a function of the number of path segments that intervene between visits to a location.

The results (Fig. 7) show a substantial decline in accuracy with number of intervening segments ($F(5, 135) = 22.7$; $p < 0.001$). After 11 intervening segments, revisit detection has declined to near chance. The model with decay set to 0.55 approximates this general accuracy decline (RMSD = 0.08; $CI_{data} = 0.153$; $r = 0.90$).

A potential problem for this analysis is that number-of-intervening-segments might be expected to correlate with path length. If so, other processes besides decay that are affected by path length but not intervening segments (associative interference, for example) could be contributing to this accuracy decline. However the path-length and number-of-intervening-segments variables actually share very little variance over the set of paths ($R^2 = 0.106$). This is because, although occasionally there will be a revisit with, say, a path length of four and three intervening segments, this occurrence is relatively unlikely. The probability of revisit builds up substantially with path length for all intervening-segments values.

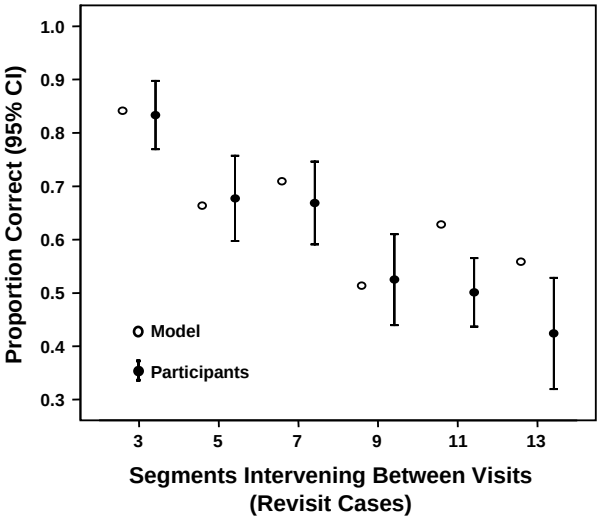

Fig. 7. Human and model accuracy as a function of number of segments intervening between visits to a location, Experiments 1 and 2 combined.

The fact that the number-of-intervening-segments variable shares little variance with path length suggests that it reflects a cognitive process that is distinct from memory load *per se*. This is confirmed by a partial correlation analysis of the set of mean accuracies for each combination of path length and number-of-intervening-segments. The resulting (partial) correlation between intervening-segments and accuracy, controlling for path length, is substantial ($r(32) = 0.92$, $p < 0.001$). This confirms that number-of-intervening-segments is a relatively distinct component of visualization accuracy, as predicted by the model.

6.4. Effects of crowding

As noted earlier, the existence of a spatial interference process in path memory was suggested by the influence of prior nearby path segments on path visualization accuracy. Informal evidence for this effect comes from an *adjacent segments analysis*, calculated as follows: For each new path segment we counted the number of previous segments that ended in an adjacent location, where adjacency was defined as one unit away from the current segment on any axis or diagonal. We call these "near visits." Fig. 8 shows the resulting mean accuracy by number of near visits for the combined data of Experiments 1 and 2, together with the predictions of the model with the decay parameter set to 0.55.

These data are consistent with a spatial interference process. Accuracy drops substantially ($F(12, 324) = 182$; $p < 0.001$) as number of near visits increases, that is, as the area adjacent to the current segment becomes more crowded with previously visited locations. Of course some decline in accuracy might be expected in any case because the number of near visits tends to increase with the number of segments in the path. However the two variables are easily distinguishable, sharing only slightly more than a third of their variance across the set of paths ($R^2 = 0.36$). Moreover, a partial correlation analysis yields

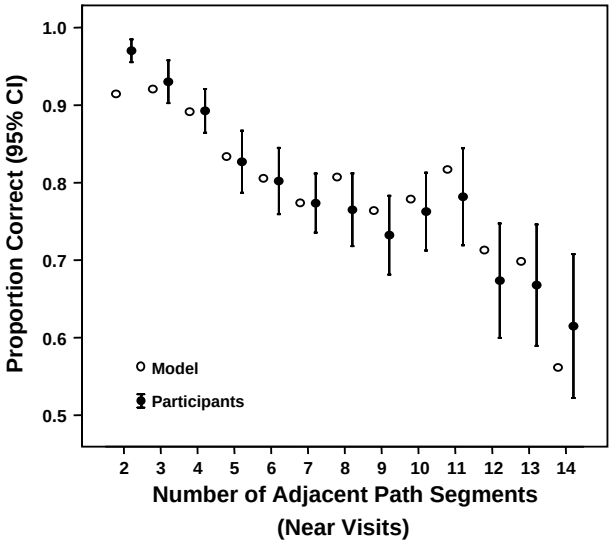

Fig. 8. Human and model accuracy as a function of number of prior path segments to adjacent locations, Experiments 1 and 2 combined.

a highly significant partial ($r(99) = -0.79$; $p < 0.001$) between near-visits and accuracy, with number of path segments controlled.[4]

Fig. 8 also shows the spatial field model's predictions. The model fits well (RMSD = 0.033; $CI_{data} = 0.102$; $r = 0.95$) and even exhibits nonlinearities similar to those in the human data. These nonlinearities are strongly related to a particular characteristic of our set of paths. We computed the mean number of path segments intervening between visits to a location, for each near-visit value, over the revisit cases in the entire set of paths. We found that, generally, as near visits increase, number-intervening tends to increase also, but there are pronounced dips at near-visit values of 8, 10 and 11. These are precisely the points at which both model and human accuracy improves. This improvement is not surprising given the results of the intervening-segments analysis presented above, in which accuracy tends to be better with fewer intervening segments. The aspect(s) of path geometry that produces dips in mean-segments-intervening at these points (at least for our sample of paths) is unclear. Nevertheless, the excellent fit of the model suggests that its spatial interference mechanism is a reasonable representation of a key determinant of accuracy in path visualization.

6.5. Other model variants

We have presented evidence for the usefulness of decay, associative interference and spatial interference in accounting for data on human spatial visualization. However it is reasonable to ask whether these mechanisms are absolutely necessary. Could models lacking one or another of these mechanisms also account for the data?

Answering this question is slightly more complicated than simply disabling a mechanism in the model and keeping all other parameters (particularly the retrieval threshold parameter) unchanged. This is because each of these mechanisms affects the level of activation of information in the model, so removing any of them will change the net activation of visualized items, changing the model's propensity to respond 'yes'. Therefore, a valid test of these model variants requires that we remove a mechanism from the model and then optimize the retrieval threshold parameter so that the overall proportion of 'yes' responses approximates that of the full spatial field model (and the human data). Using this technique, we tested three variants of the spatial field model—associative interference removed; decay removed; and spatial interference removed.

The results for each model variant are shown in Fig. 9. As the figure shows, each of the variants produced a greater proportion of correct responses than the human participants. This is to be expected since each variant removes a process that, in the full model, reduces visualization accuracy. It confirms that all three mechanisms have a role in limiting visualization performance.

Clearly, none of the model variants fit the human data as well as the full model did (c.f. Figs. 2 and 3 above). The model variant without associative interference provided the best fit (RMSD = 0.038; $CI_{data} = 0.070$; $r = 0.97$), followed by the model without decay (RMSD = 0.058; $r = 0.98$). The model without spatial interference produced the worst

[4] It is not possible to also 'partial out' the number-of-intervening-segments variable described in the previous section, because the deleterious effect of near visits occurs largely for no-revisit cases, whereas number-of-intervening-segments is only defined for revisit cases.

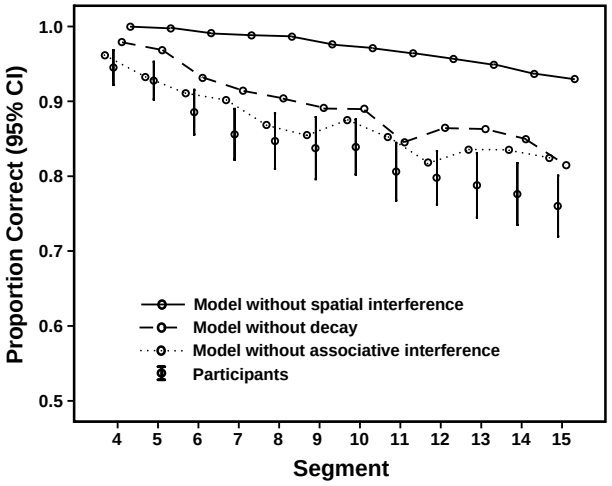

Fig. 9. Human and model accuracy for variants of the model from which a single component process (either spatial interference, decay or associative interference) was removed. No variant matched human accuracy as closely as the full model (Fig. 3) did.

fit (RMSD = 0.136; $r = 0.94$) due to its failure to replicate human performance on path segments that did not revisit a prior location. Without spatial interference, this model variant always responded correctly on no-revisit segments, whereas humans, on average, miss 12% of them.

7. General discussion

The goal of this investigation was to better understand people's ability to accurately visualize complex spatial material. Earlier studies (e.g. Attneave & Curlee, 1983; Kerr, 1987, 1993) tested visualized locations in 2D or 3D space, and found that visualization capacity is sharply limited. The results of our experiments are consistent with this finding, and provide additional details that were used to test a computational model of spatial visualization in the context of a new visualization measurement technique, path visualization. The requirements of our task provide an advantage over similar tasks by reducing the utility of verbally-based strategies that could influence performance independently of spatial visualization processes. The empirical results demonstrate that this task is cognitively challenging for participants, and we believe that it taps important mechanisms associated with spatial visualization. Additionally, we have demonstrated that a computational model based on the idea of a *spatial field* can account for patterns of accuracy in visualization performance associated with different aspects of the task and stimuli.

Given that the spatial field model can account for our visualization accuracy data, what can we say about the nature of human visualization capacity? In particular, what causes visualization capacity limits? The answer, according to the spatial field model, is that limited visualization capacity is due largely to three interacting processes: (1) activation decay; (2) associative interference; and (3) spatial interference. We now consider each of these processes in more detail.

7.1. Activation decay

In the spatial field model, when a new segment of a path is presented, the model determines the location to which it points. This process results in the creation of a unit ('chunk') containing the new location. This chunk is created with a particular base level of activation, but this activation immediately begins to decay exponentially (c.f. Anderson & Lebiere, 1998). We have found that the ACT-R default activation decay rate accounts for many, but not all aspects of path visualization accuracy. A small increase in decay rate was required to best capture the phenomena in the data. The ACT-R memory mechanisms were derived largely from experiments using non-spatial verbal materials, so it is perhaps unsurprising that the default decay rate is not an exact fit for a visualization task. However, the 10% increase in decay rate that we used may not be large enough to propose different underlying decay processes for spatial and verbal material.

7.2. Associative interference

In ACT-R, the activation of a unit in memory depends not only on its base-level activation (subject to decay), but also its momentary associative activation. For our model, associative activation works as follows. As the model reads a new segment description, the chunk in the *goal buffer* is modified to represent the new location. In ACT-R, activation spreads from the components of the chunk in the goal buffer to all related items in memory, where 'related' means an *exact* match to one or more features of the goal chunk. However the total amount of activation to be spread is fixed. Therefore the greater the number of related items in memory, the less activation each one will get. As noted earlier, this mechanism accurately models many declarative memory phenomena, particularly the fan effect (Anderson, 1974). The consequence for the path visualization model is that, for each path, as more segments are presented, the chance that they will share properties with earlier segments will increase, and therefore, on the average, less activation will be spread to any particular prior segment.

Two aspects of associative activation in path visualization should be noted. First, dividing associative activation among more chunks in memory reduces accuracy only on revisit cases. Indeed, anything that tends to lower activation values will tend to produce more 'no' responses, which are only errors when a revisit has occurred. Second, associative activation, by itself, is not explicitly spatial. If two path segments in memory have the same end location as the goal segment, each one will benefit from the activation spreading from the goal, but to a lesser extent than if there were only one matching segment in memory. No activation spreads to chunks representing other locations (either nearby or farther away), since this mechanism requires an exact match. Also, since other information, like the segment description, is represented in the goal, the impact of spreading activation will be considerably diminished for later segments in the path. This contributes somewhat to lower accuracy for revisit cases as path length increases.

7.3. Spatial interference

Spatial interference is a critical process in the model we have proposed. It is the process that most clearly differentiates spatial visualization from declarative memory for verbal materials. An account of the impact of nearby visited locations is not easily accommo-

dated using a verbal memory explanation, and our attempts to fit the data using such a representation in ACT-R were unsuccessful. Therefore we turned to the notion of a spatial field—that somewhere in the brain there are structures or processes that represent at least some aspects of space isomorphically, so locations that are nearby in real space will show evidence of proximity in imaginary space. As noted earlier, the ACT-R architecture does not currently have a built-in process that would correspond to a spatial memory field. However we were able to emulate the operation of a spatial field using ACT-R's partial matching mechanism. When a path segment is presented and the model must decide whether or not the location it points to has been visited before, the model attempts to retrieve a location from memory, under the assumption that if anything is retrieved, then the current location is familiar and therefore has been visited before. However, even if the current location has *not* been visited, partial matching allows a path segment to another location to be retrieved if it is active enough. The spatial aspect of the model is that the activation of all prior-segment chunks that do not exactly match the current location is reduced in proportion to their distance from the current location. This is achieved by calculating a mismatch penalty that is a function of Euclidian distance from the current location (Eq. (1)). So, locations that are nearby are assessed a smaller mismatch penalty and are more likely to be erroneously retrieved than locations that are relatively farther away.

Previous uses of the similarity-based partial matching mechanism in ACT-R have involved verbal declarative memory paradigms such as forward or backward serial recall, recognition, and free recall (see Anderson et al., 1998). In these applications, similarity-based partial matching has been used to represent proximity in terms of presentation order, semantic meaning, or some other dimension. Our model demonstrates that this mechanism may be an appropriate means of characterizing proximity in visualization space. However the fact that our model is implemented using mechanisms developed in the context of declarative memory should not be construed to imply that we believe spatial working memory is merely the application of verbal declarative memory to spatial material. In fact, there is substantial evidence for differences between the processing of verbal and spatial material, including research on dual-task interference (Logie, 1995), hemispheric specialization (Casasanto, 2003), mental imagery (Kosslyn, 1980, 1994), and individual differences (Hegarty, Montello, Richardson, Ishikawa, & Lovelace, 2006). The real claim embodied by our account is that the mechanisms that act on spatial visualizations and those that act on verbal declarative knowledge seem to have some similar characteristics (c.f. Anderson & Paulson, 1978). Thus, whereas we acknowledge that spatial visualizations may be represented within a structure that is specialized for spatially distributed material (such as the image buffer proposed by Kosslyn, 1980, 1994), we suggest that the information represented in that structure may be subject to some of the same activation dynamics that determine the availability of verbal declarative knowledge.

Finally, note that the three processes that drive our model—decay, associative interference, and spatial interference—do not operate in isolation. For example, if a prior visit to a location adjacent to the current one occurred many segments ago, the activation of the chunk representing the prior visit will be relatively low due to decay. Thus, retrieval of this nearby chunk (spatial interference) will, on average, be less likely than if the visit had been recent. The embedding of these processes in a cognitive architecture allows them to interact in order to jointly influence performance.

Although our results suggest a role for three processes—decay, associative interference and spatial interference—in determining visualization capacity, we acknowledge that in

more complex visualization tasks, additional cognitive processes will interact with the underlying representation, helping to determine visualization accuracy. An example of this is the presentation of path segments from an egocentric perspective. In other studies, we are assessing path visualization accuracy with egocentric path descriptions. Our initial results (Lyon, Gunzelmann, & Gluck, 2007) suggest that egocentric descriptions are difficult to translate to an allocentric perspective, perhaps because this translation requires one to keep track of facing direction (and, in three dimensions, virtual body orientation), and to transform this information to an allocentric framework. We are currently attempting to model these processes in order to compare the spatial field model to human performance for egocentrically-described paths.

Modeling path visualization given different kinds of descriptors is but one aspect of the higher-level objective of moving from a solid understanding of the basic processes that limit visualization capacity to a model of how those processes contribute to performance in more complex and realistic spatial tasks. Another subgoal in this endeavor is to understand how various kinds of spatial knowledge affect visualization. In the design of the path visualization task, we wanted to eliminate the possibility of using pre-existing route knowledge so that the limits of the visualization system itself could be studied. Nevertheless such knowledge—landmarks, previously learned patterns, maps, diagrams and more—will need to be part of a complete model of human spatial problem solving.

Clearly one kind of knowledge utilization—the grouping of sequences of path segments into meaningful units—is at least sometimes used in remembering routes, and could potentially be applied to the path visualization task. Our participants have reported only one such grouping—the formation of a square with four successive segments. It is, of course, possible that participants chunk segment sequences in other ways without being aware of doing so; however our model fits the data for these experiments without the need to add a sequence-chunking mechanism. Nevertheless, such a mechanism would clearly be important if, for example, paths were not quasi-random but followed recognizable patterns.

Other research using ACT-R modeling has looked at the general issues of recognizing patterns from past experience (e.g., Gonzalez, Lerch, & Lebiere, 2003; Gunzelmann & Lyon, 2006; Lebiere, Gonzalez, & Martin, 2007; West & Lebiere, 2001). In particular, Lebiere and his colleagues (Lebiere et al., 2007; West & Lebiere, 2001) explore how game players may represent patterns in their opponents by using instance-based learning. In this research, sequences of moves are stored coherently as chunks, providing insight into patterns that players produce. However the number of possible sequences in these tasks was relatively small. In contrast, there are many potential sequences in the path visualization task, particularly when longer sequences are considered. Thus, utilizing such a strategy to improve visualization accuracy is likely to require substantial amounts of practice with the task.

Another way to extend our analysis of visualization capacity would be to apply it to the understanding of individual differences in cognition. In our view, the path visualization task is a potentially useful measure of the ability to visualize complex spatial material. A comparison of computational models for path visualization and other related tasks might help define the common cognitive processes underlying spatial ability.

8. Conclusion

Our goal in this paper was to come nearer to an understanding of why human visualization capacity is limited. We chose to focus on visualizing *verbal* descriptions, so that

memory for pictorial information from the stimulus would not augment the participant's self-generated visualizations. We developed a task, path visualization, in which participants attempt to detect revisits in complex paths, thereby forcing them to generate an internal spatial representation of the path, rather than simply reproduce a list of verbal descriptions.

We examined patterns of errors that participants made while trying to determine whether each segment of a path revisited a prior path location. An ACT-R model of path visualization successfully accounted for these error patterns, specifically: (1) the increase in error frequency as paths get longer; (2) the similarity in error frequency for 2D and 3D paths; (3) the error/load curves for revisit vs. no-revisit segments; (4) the effect of intervening path segments on revisit detection accuracy; and (5) the effect of crowding (proximity of path segments) in imaginary space.

Our model suggests a particular view of spatial visualization capacity. Capacity is not best conceived as a 'number of items' limit. Instead, capacity limitations are viewed as the result of processes that affect the activation of visualized items. If activation is insufficient to allow all items to be successfully maintained, then 'capacity' has been exceeded. For example, suppose you are hearing a segment-by-segment description of an (unfamiliar) route over the phone and trying to visualize a mental map of it. According to our model, several factors will affect your capacity to visualize the entire route. The first segments of the route will decay over time, reducing the chance that they can be retrieved. In addition, all segments will suffer associative interference as more and more segments are presented. (In this respect, our model is similar to other ACT-R models of verbal declarative memory tasks.) Finally, our data suggest that there is yet another limitation that applies in this situation. To the extent that parts of the route crowd closely together, there will be interference between adjoining parts of visualization space itself. According to our model, all of these processes, operating together, determine the point at which your capacity to visualize the route accurately will be exceeded.

We have shown that a proximity-based spatial interference process, when combined with ACT-R's standard memory activation equations, can account for patterns of errors in visualizing complex spatial material in both two and three dimensions. However, at this point there are several viable theories regarding the origin and potential neural substrate of this spatial interference process, including different theories regarding how the information in spatial visualizations is represented. For example, we know that interference between spatial locations is a characteristic of the human visual system at many different levels, including simple-cell receptive fields (Hubel & Wiesel, 1977), the formation of simple visual groups (Lyon, 1992), and lateral interference in character strings (e.g. Bouma & Leigen, 1977). So one might interpret our evidence of spatial interference in visualization as support for the notion that when people visualize, they are (weakly) activating structures normally used for vision (e.g. Kosslyn, 1980, 1994). Our participants report the subjective experience of 'seeing' a 'mental picture' of the visualized paths, but of course this is insufficient to establish that actual visual structures are involved.

On the other hand, perhaps spatial proximity acts entirely within declarative memory, as just another type of semantic similarity, one that has no necessary connection to the visual system. Several ACT-R models have invoked non-spatial kinds of similarity to account for various other declarative memory phenomena (e.g. Anderson et al., 1998). Perhaps our data indicate a similar phenomenon that happens to involve spatial rather than semantic features.

A third possibility is that the spatial proximity effect is generated within a structure that is specialized for some aspect of spatial processing, but is not part of the early vision system *per se*. For example, several areas within the parietal lobes appear to be involved in spatial processing (Jonides & Smith, 1997; Raffi & Siegel, 2005). There are also areas in prefrontal cortex that have been associated with spatial working memory (Funahashi, Bruce, & Goldman-Rakic, 1993).

This uncertainty about the source of the spatial interference effect points to the need for further research. However our data do provide evidence for an important point. With respect to spatial interference, the imaginary space of human visualization seems to function as if it were a real space. 'Imaginary' spatial proximity has real effects on performance, and the spatial field model captures these effects quite accurately. Thus, it provides at least an initial basis for studying—within a cognitive architecture—the space-like character of human visualization.

Appendix A. Summary of memory activation processes in the spatial field model

The spatial field model uses the mechanisms embodied within ACT-R 5.0's declarative memory module. Each new segment of a path is encoded as a unit ('chunk') in declarative memory. It is assigned an initial *base-level activation* value, including the addition of stochastic noise, which then decays exponentially with time. If the path retraces this segment, the activation of the corresponding chunk is strengthened by repetition. The residual activation of chunk i (B_i) after n revisits, each at time t_j, is described by Eq. (A1), in which d is the activation decay parameter (normally set to 0.5):

$$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right) \tag{A1}$$

In addition to base-level activation, path-segment chunks receive temporary activation when they match the values of slots of the chunk that is currently in the goal buffer (the 'goal chunk') through the spreading activation mechanism in ACT-R. In the spatial field model, the goal chunk is usually the most recently presented segment of the path. Therefore this chunk will activate chunks representing previous segments to the extent that they match the start-location slot and end-location slot of the current segment. The total activation emanating from the goal chunk is W. We used the default value of 1.0 for W. The proportion of the total goal-chunk activation that is sent to chunks matching each slot j is W_j.[5] This activation is further split among all of the chunks in memory that match the j slot value (for example, all the segments that contain a reference to the same location in the space as the new endpoint). So the proportion of activation that a given chunk i gets by matching a slot j in the goal chunk is S_{ji}.[6] Finally, random noise ε is added to the temporary increment in activation given by the goal chunk. So the total activation of a chunk in declarative memory, with partial matching off, is given by Eq. (A2):

[5] By default, W_j is equal to W/c, where c is the number of slots in the goal.

[6] S_{ji} is equal to $W_{j/x}$, where x is the number of chunks in declarative memory with a slot value for j that matches the value for that slot in the goal.

$$A_i = B_i + \sum_j W_j S_{ji} + \varepsilon \tag{A2}$$

Eq. (A2) captures two of the mechanisms we have mentioned as being critical to our account, but by itself it is insufficient to explain our finding of 'crowding' in imaginary space. Therefore a spatial field was emulated using ACT-R's partial matching mechanism, which introduces similarity-based interference. After each path segment is presented, the model attempts to retrieve another segment that ends at the same location (a 'revisit'). Partial matching allows chunks to be retrieved that are not an exact match to the retrieval probe. In the spatial field model, the retrieval probe is a chunk that contains only slot values for those aspects of the current path segment that need to be matched (i.e., the location of the end of the segment). The model places this retrieval probe chunk in the retrieval buffer and attempts a retrieval. Chunks in memory that do not match the probe perfectly are assessed an activation penalty that increases as the value in the relevant slot (s) become less similar to the slot value (s) in the retrieval probe. Eq. (A3) shows the activation of chunk i with partial matching enabled:

$$A_i = B_i + \sum_j W_j S_{ji} + \sum_k P_k M_{ki} + \varepsilon_1 + \varepsilon_2 \tag{A3}$$

In this equation, B_i is the base-level activation, $W_j S_{ji}$ is the temporary activation due to the goal chunk, summed over j slots, and $P_k M_{ki}$ is the partial matching penalty (value zero or negative) obtained by summing the similarities (M_k) between slot(s) in the retrieval probe and the slots in memory chunk i. In our model, M_k is assigned based on the Euclidean distance between the current endpoint of the path, and the location encoded in the chunk being evaluated in declarative memory. This value is calculated using Eq. (1) above. These similarities can be modified by differential weights (P_k), however we simply assigned the default value of 1.5 to P for all k. Finally in Eq. (A3) we show separately the two sources of activation noise noted previously—ε_1 is 'permanent noise' added when a chunk is created, and ε_2 is temporary noise that accompanies temporary activation from the goal chunk.

References

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 5*, 451–474.

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review, 85*, 249–277.

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

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*(4), 1036–1060.

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38*, 341–380.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah N.J.: Erlbaum.

Anderson, J. R., & Paulson, R. (1978). Interference in memory for pictorial information. *Cognitive Psychology, 9*, 178–202.

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General, 128*, 186–197.

Attneave, F. (1972). Representation of physical space. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 283–306). Washington D.C.: V.H. Winston & Sons.

Attneave, F., & Curlee, T. E. (1983). Locational representation in imagery: A moving spot task. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 20–30.

Avraamides, M. N. (2003). Spatial updating of environments described in text. *Cognitive Psychology, 47*, 402–431.

Barshi, I., & Healy, A. F. (2002). The effects of mental representation on performance in a navigation task. *Memory and Cognition, 30*, 1189–1203.

Bouma, H., & Leigen, C. P. (1977). Foveal and parafoveal recognition of letters and words by dyslexics and by average readers. *Neuropsychologia, 15*, 69–80.

Brooks, L. R. (1968). Spatial and verbal components in the act of recall. *Canadian Journal of Psychology, 22*, 349–368.

Byrne, M. D., & Anderson, J. R. (1998). Perception and action. In J. R. Anderson & C. Lebiere (Eds.), *The atomic components of thought* (pp. 167–200). Mahwah, N.J: Erlbaum.

Casasanto, D. (2003). Hemispheric specialization in prefrontal cortex: Effects of verbalizability, imageability and meaning. *Journal of Neurolinguistics, 16*, 361–382.

Diwadkar, V. A., Carpenter, P. A., & Just, M. A. (2000). Collaborative activity between parietal and dorso-lateral prefrontal cortex in dynamic spatial working memory revealed by fMRI. *NeuroImage, 12*, 85–99.

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General, 119*, 63–76.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1993). Dorsolateral prefrontal lesions and oculomotor delayed-response performance: Evidence for mnemonic"scotomas". *Journal of Neuroscience, 13*, 1479–1497.

Gonzalez, C., Lerch, F. J., & Lebiere, C. (2003). Instance-based learning in real-time dynamic decision making. *Cognitive Science, 27*(4), 591–635.

Gugerty, L., & Brooks, J. (2004). Reference frame misalignment and cardinal direction judgments: Group differences and strategies. *Journal of Experimental Psychology: Applied, 10*, 75–88.

Gunzelmann, G., & Anderson, J. R. (2006). Location matters: Why target location impacts performance in orientation tasks. *Memory & Cognition, 34*(1), 41–59.

Gunzelmann, G., & Lyon, D. R. (2008). Mechanisms of human spatial competence. In T. Barkowsky, C. Freksa, M. Knauff, B. Krieg-Bruckner, & B. Nebel (Eds.), *Spatial cognition V, lecture notes in artificial intelligence #4387* (pp. 288–307). Berlin, Germany: Springer-Verlag.

Gunzelmann, G., & Lyon, D. R. (2006). Qualitative and quantitative reasoning and instance-based learning in spatial orientation. In R. Sun & N. Miyake (Eds.), *Proceedings of the twenty-eighth annual meeting of the cognitive science society* (pp. 303–308). Mahwah, NJ: Lawrence Erlbaum Associates.

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence, 34*, 151–176.

Hubel, D. H., & Wiesel, T. N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London B, 198*, 1–59.

Jonides, J., & Smith, E. E. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology, 33*, 5–42.

Just, M., & Carpenter, P. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review, 92*, 137–172.

Just, M. A., Carpenter, P. A., Maguire, M., Diwadkar, V., & McMains, S. (2001). Mental rotation of objects retrieved from memory: A functional MRI study of spatial processing. *Journal of Experimental Psychology: General, 130*, 493–504.

Kerr, N. H. (1987). Locational representation in imagery: The third dimension. *Memory and Cognition, 15*, 521–530.

Kerr, N. H. (1993). Rate of imagery processing in two versus three dimensions. *Memory and Cognition, 21*, 467–476.

Kosslyn, S. M. (1978). Measuring the visual angle of the mind's eye. *Cognitive Psychology, 10*, 356–389.

Kosslyn, S. M. (1980). *Image and mind*. Cambridge: Harvard University Press.

Kosslyn, S. M. (1994). *Image and brain*. Cambridge: MIT Press.

Lebiere, C. (1999). The dynamics of cognition: An ACT-R model of cognitive arithmetic. *Kognitionswissenschaft, 8*, 5–19.

Lebiere, C., Gonzalez, C., & Martin, M. (2007). Instance-based decision making model of repeated binary choice. In R. L. Lewis, T. A. Polk, & J. E. Laird (Eds.), *Proceedings of the 8th International Conference on Cognitive Modeling* (pp. 67–72). Ann Arbor, MI: University of Michigan..

Logie, R. H. (1995). *Visuo-spatial working memory*. Hove: Erlbaum.

Lyon, D. R. (1992). Position-linked interference in forming simple visual groups. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 1139–1157.

Lyon, D. R., Gunzelmann, G., & Gluck, K. A. (2007). Visualizing egocentric vs. exocentric path descriptions. In D. S. McNamara & G. Trafton (Eds.), *Proceedings of the twenty-ninth annual meeting of the cognitive science society*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lyon, D. R., Gunzelmann, G., & Gluck, K. A. (2006a). Key components of spatial visualization capacity. In *Proceedings of the seventh international conference on cognitive modeling* (pp. 381–382). Trieste, Italy.

Lyon, D. R., Gunzelmann, G., & Gluck, K. A. (2006b). Virtual travel does not enhance spatial working memory for landmark-free paths. In R. Sun & N. Miyake (Eds.), *Proceedings of the twenty-eighth annual meeting of the cognitive science society (p. 2550)*. Mahwah, NJ: Lawrence Erlbaum Associates.

Raffi, M., & Siegel, R. M. (2005). Functional architecture of spatial attention in the parietal cortex of the behaving monkey. *Journal of Neuroscience, 25*, 5171–5186.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review, 107*(2), 358–367.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701–703.

Shepard, S., & Metzler, D. (1988). Mental rotation: Effects of dimensionality of objects and type of task. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 3–11.

West, R. L., & Lebiere, C. (2001). Simple games as dynamic, coupled systems: Randomness and other emergent properties. *Journal of Cognitive Systems Research, 1*(4), 221–239.