Lyon, D. R., Gunzelmann, G., & Gluck, K. A. (2006). An ACT-R Model of Spatial Visualization Capacity. In D. Fum, F. Del Missier & A. Stocco (Eds.), *Proceedings of the 7th International Conference on Cognitive Modeling* (pp. 381-382). Trieste, Italy: Edizioni Goliardiche.

Key Components of Spatial Visualization Capacity

Don R. Lyon (don.lyon@mesa.afmc.af.mil)
L3 Communications at Air Force Research Laboratory
6030 S. Kent St., Mesa, AZ 85212-6061 USA

Glenn Gunzelmann (glenn.gunzelmann@mesa.afmc.af.mil)
Kevin A. Gluck (kevin.gluck@mesa.afmc.af.mil)
Air Force Research Laboratory
6030 S. Kent St., Mesa, AZ 85212-6061 USA

Introduction

Spatial visualization is ubiquitous in human cognition. People visualize the spatial aspects of things in a variety of situations, ranging from mentally repositioning furniture to solving complex scientific and engineering problems. However, as anyone who has listened to complex directions over the phone can attest, our capacity to visualize spatial information is limited. Here we describe an ACT-R model (the *spatial field model*) of the underlying processes that limit visualization capacity, and apply this model to a measurement task called *path visualization* (Lyon, 2004). Like some other visualization tasks, (e.g., Brooks, 1968; Attneave & Curlee, 1983; Kerr, 1993) path visualization requires participants to imagine a series of movements (up, right, down…) in a two-dimensional or three-dimensional array of locations. Unlike these other tasks, path visualization requires participants to decide whether each new path segment intersects with any part of the existing path. This requirement to detect intersections forces a visualization strategy, rather than, for example, a verbal rehearsal strategy. As the path gets longer, load increases, visualization capacity is exceeded, and accuracy declines.

The spatial field model predicts this decline very accurately by positing three key components of visualization capacity. These components are implemented using ACT-R's declarative memory processes. The model test reported here uses no free parameters, because (1) the values for all but two parameters are standard in ACT-R memory models, and (2) the values for the two remaining parameters were set previously using data from a different set of participants.

Method

Thirteen paid participants (6 women and 7 men) were each given five 30-trial path visualization sessions. During each trial, the participant saw a sequence of 15 text phrases presented on a CRT. Each phrase described the direction and distance (e.g. 'Left 1') of a segment of a path. There were six possible directions (Right, Left, Forward, Back, Up, Down); all distances were one. Each phrase was presented for 2000 msec, followed by a blank screen for 133 msec, then the presentation of the next phrase. As each new phrase was presented, the participant decided whether or not the endpoint of the new path segment intersected with any previously presented part of the path, and pressed a key with the right index finger to indicate 'yes', or the left index finger to indicate 'no'. In the rare event that no key was pressed during the presentation of a text phrase, 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.

All paths started at the center of an imaginary 5 x 5 x 5 cube (Figure 1). Segment directions were relative to a fixed frame of reference, so, for example, 'Left 1' was always toward the left side of the cube as depicted in the figure, regardless of the direction of the previous segment. No two successive segments could be on the same axis, so the path always turned to a new axis with each new text phrase. Paths were randomly generated with the restrictions that (1) a near-balance of intersection and no-intersection segments existed in the entire corpus of paths, and (2) 50% of the paths were restricted to a 2D plane; 50% were 3D. Although a picture similar to Figure 1 was shown in the instructions, no image was displayed during performance of the task. The path had to be visualized from the sequence of text phrases.

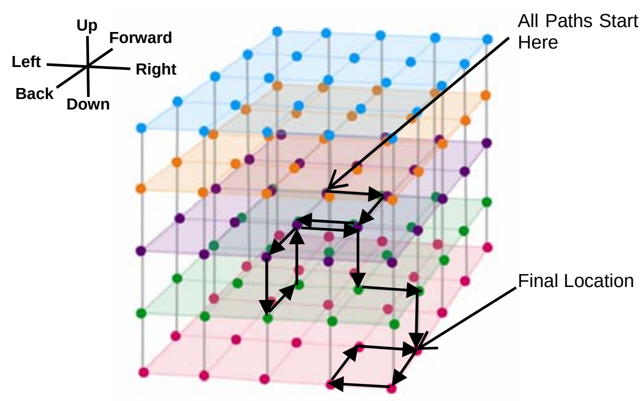

Figure 1: Depiction of a 3D path in an imaginary space.

Spatial Field Model

Previous research with path visualization (Lyon, Gunzelmann, & Gluck, 2004) suggested the existence of a very strong spatial interference effect. When sections of a path were clustered together, accuracy in detecting whether or not an intersection occurred was substantially worse than when the current segment was relatively far away from most of the rest of the path. We were able to emulate this spatial interference effect using the similarity mechanism in ACT-

R's declarative memory representation. Similarity decreased exponentially as a function of the Euclidian distance between locations *in a 3D space*. Thus, prior path segments at locations very close to the current one had a greater chance to be retrieved through partial matching (error of commission), producing more false intersection responses.

The model works as follows. Each time a new segment description is presented, the model reads the description and sets a goal to determine the location to which this new segment points. After this goal is accomplished, a chunk representing this location is stored in declarative memory. As each new segment is processed, the model attempts to retrieve a chunk from memory that represents the same location. If a chunk is successfully retrieved, the model decides that an intersection with a prior part of the path has occurred and responds 'yes', otherwise it responds 'no'. We present path segments rapidly, leaving little time for rehearsing the previous path, for recoding the path in a different representation, or any other mnemonic device. Therefore the model does not include such processes.

The spatial field model embodies the following three-component theory of visualization capacity limits:

1. *Visualized elements in nearby locations interfere with each other* (Lyon, Gunzelmann, & Gluck, 2004). Spatial visualization mimics real space in the sense that interference between visualized elements follows roughly the same spatial pattern as the elements themselves would form in real space.
2. *Visualized elements become less available with time.* In ACT-R, this results from activation decay in declarative memory.
3. *'Re-visiting' a visualized location makes it easier to access.* ACT-R embodies this aspect of the theory with its base-level learning mechanism.

The spatial field model uses commonly accepted values for most of ACT-R's parameters, so we needed values for only two parameters: the new spatial interference parameter, and retrieval threshold. These values were obtained using data from a prior study with different participants.

Results and Conclusion

Figure 2 shows the decline in visualization accuracy as additional segments are added to the path, reflecting the limited capacity of human spatial visualization. The predictions of the spatial field model matched the participant data well ($r = 0.95$, $RMSD = 0.023$), with no parameter adjustments or any other tailoring of the model to this dataset. Without the spatial interference mechanism, the model fits neither the segment load effect nor several other aspects of the data (e.g. accuracy for intersection versus no-intersection cases), even when retrieval threshold is allowed to vary to optimize fit.

The tight fit of the spatial field model to human data suggests that the new spatial interference process it contains, combined with ACT-R's existing memory activation processes, can account for the effects of exceeding human visualization capacity in this difficult task.

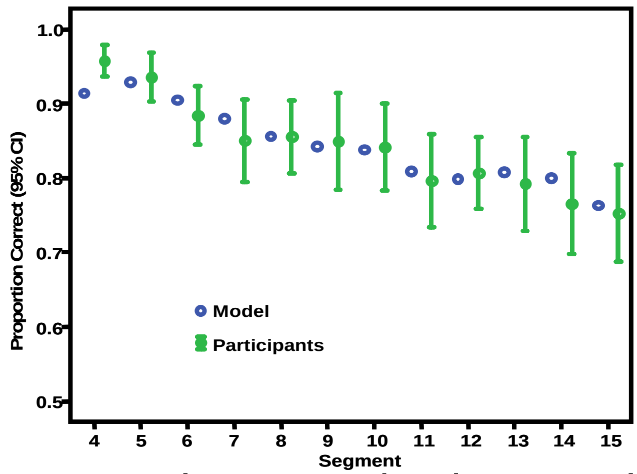

Figure 2. Visualization accuracy by path segment, model predictions and human data.

The spatial field model suggests a particular view of spatial visualization capacity. Capacity may not be best conceived as a 'number of items' limit. Instead, capacity limitations are viewed in this model as the result of a combination of decay and interference mechanisms that influence the probability of successful recall. In this respect, our model of spatial visualization is similar to ACT-R models of verbal declarative memory tasks.

Acknowledgments

We thank Ben Sperry and Jerry Ball for software development, Christy Caballero for research assistance, and the U. S. Air Force Office of Scientific Research for support (Grant # 02HE01COR).

References

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.

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

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

Lyon, D. R. (2004). *Measuring visuospatial working memory using path visualization.* Technical Memorandum 23-1, Air Force Research Laboratory, Mesa, Arizona.

Lyon, D. R., Gunzelmann, G., & Gluck, K. A. (2004). Emulating a Visuospatial Memory Field Using ACT-R. In M. Lovett, C. Schunn, C. Lebiere, and P. Munro, *Proceedings of the Sixth International Conference on Cognitive Modeling* (pp. 368-369). Mahwah, NJ: Lawrence Erlbaum Associates.