The Effect of Memory Schemas on Object Recognition in Virtual Environments

Katerina Mania

Department of Informatics

University of Sussex, UK

k.mania@sussex.ac.uk

Andrew Robinson

CAE Systems, UK

Karen Brandt

Department of Experimental Psychology

University of Lancaster, UK

Abstract

Prior theoretical work on memory schemas, an influential concept of memory from the field of cognitive psychology, is presented for application to fidelity of computer graphics simulations. The basic assumption is that an individual's prior experience will influence how he or she perceives, comprehends and remembers new information in a scene. Schemas are knowledge structures; a scene could incorporate objects that fit into a specific context or schema (for instance, an academic's office) referred to as consistent objects but also objects that are not related to the schema in place referred to as inconsistent objects. In this paper, we describe ongoing development of a rendering framework related to scene perception based on schemas. An experiment was carried out to explore the effect of rendering quality and object type on object memory recognition in a room. The computer graphics simulation, rendered using radiosity, was displayed on a Head Mounted Display (HMD) utilizing stereo imagery and head tracking. 36 participants across three conditions of varied rendering quality of the same space were exposed to the computer graphics environment and completed a memory recognition task. Results revealed that schema consistent elements of the scene were more likely to be recognized than inconsistent information. Overall higher confidence ratings were assigned for consistent compared to inconsistent items. Further explorations of the effect of schemas on spatial awareness in synthetic worlds could lead to identifying areas of a computer graphics scene that require higher quality of rendering as well as areas that lower fidelity could be adequate with the ultimate goal being to simulate a perceptual process rather than physics or visual focus.

1. Introduction

It is not computationally feasible to be immersed into an interactive artificial environment which exactly mimics the panoply and complexity of sensory experiences associated with a real world scene. For a start, it is technologically challenging to control all of the sensory modalities to render the exactly equivalent sensory array as that produced by real world interaction. The mapping from the real world environment to a computer graphics environment is mediated by *environmenta*l or *visual* fidelity (Waller, Hunt & Knapp, 1998). The term *visual fidelity* refers to the degree to which visual features in the VE conform to visual features in the real environment. Within this area, one can distinguish between *physical* realism, in which the synthetic scene is an accurate point-by-point representation of the spectral radiance values of the real scene; *photorealism,* in which the synthetic scene produces the same visual response as the real scene even if the physical energy depicted from the image is different compared to the real scene; and finally *functional realism,* in which the same information is transmitted in real and synthetic scenes while users perform visual tasks targeting transfer of training in the real world (Ferwerda 2001). Functional realism embraces many rendering tactics. For example, flight simulators usually provide synthetic scenes that are not physically accurate or photorealistic, however, they are functionally realistic because they transmit similar information to the user as flying a real plane (Waller, Hunt & Knapp 1998). There is always a trade-off between visual fidelity and computational complexity. It is essential to have robust and yet efficient techniques in order to assess the fidelity of VE implementations comprising of computer graphics imagery, display technologies, and 3D interaction metaphors (interaction fidelity) across a range of application fields.

One of the central issues for VE applications is centered on how humans mentally represent an interactive computer graphics world and how their recognition and memory of synthetic worlds of varied fidelity levels correspond to their real-world counterparts. The utility of VEs for any applications for which they are being proposed is predicated upon the accuracy of the spatial representation formed in the VE. Memory tasks are often incorporated in benchmarking processes when assessing fidelity of a VE simulation because spatial awareness is crucial for human performance efficiency of any task that entails awareness of space (Bailey and

Witmer 1994, Bliss, Tidwell & Guest 1997). A commonly employed strategy for assessing the simulation fidelity of a VE is comparing memory performance among VE systems of varied rendering quality or interaction interface often including comparative evaluations in real world settings (Dinh, Waler & Hodges 1999, Mania, Troscianko, Hawkes & Chalmers 2003). Would cognitive systems respond in a similar fashion to a specific scene presented in a VE as they would to that same contextual scene in the real world? And how could we match the capabilities of the VE system related to visual and interaction fidelity to the requirements of the human perceptual and motor systems?

The proposed research suggests a new approach. Being in a certain place, for instance, an academic's office, results in mentally representing spatial elements dependent on their degree of association to that place. These representations could be described as 'assumptions' or spatial 'hypotheses' that humans adopt after a very short exposure to the space. If we were able to exploit existing research on how these 'assumptions' or 'schemas' are formed in real world and then be able to simulate those assumptions in a VE scenario, we would have a powerful new measure of functional realism. Similar information would be transmitted between a synthetic scene and a real world scene, both depicting a specific schema.

2. The Role of Schemas in Memory

The proposed approach is based on classic findings from memory research. Schemas are knowledge structures or cognitive frameworks based on past experience. Schema theories propose that perception, language comprehension, and memory are processes that involve the interaction of new episodic information with old, schema consistent information (Brewer 2000, Kuipers 1975). Schema consistent scene elements are expected to be found in a given context, for instance into an academic's office. Information slots which have not been filled with perceptual information are filled by default assignments based on stereotypic expectations from past experience (Kuipers 1975). The results from the Kuipers 1975 study demonstrated the effect of such stereotypic expectations; if a quick visual scan of a room indicates that there is a clock on the wall, hands may be assigned to it at a memory test after a brief exposure to the room, even though this particular clock did not have hands. Inferences of absent elements of a space are said to occur when memory performance after exposure to a space contains spatial information that was not actually there but was expected to be there based on past associations. This integration may be so complete that episodic information cannot be distinguished from schema consistent information. Alternatively, the two types of information may remain distinct.

It has generally been shown that memory performance is frequently influenced by schema-based expectations and that an activated schema can aid retrieval of information in a memory task (Minsky 1975). There are fundamentally different ways in which schemas might influence memory performance. Schemas could determine which objects are looked at and encoded into memory (e.g., fixation time). They could also guide the retrieval process and also determine what information is to be communicated at output (Brewer & Nakamura 1984). Certain research on schemas in a real world setting has demonstrated that memory performance is better for consistent items; this is known as the consistency effect (Brewer & Treyens 1981). Consistent items are items that are likely to be found in a given environment. Brewer and Treyens 1981 showed that schema expectancy, i.e. how likely an object is to be found in a scene, was positively correlated with recall and recognition scores after a short exposure to a scene, thus, supporting the so-called consistency effect. Participants were taken into what they thought was an academic's office and later were tested for memory of the objects in the room with drawing recall, written recall or verbal recognition. In addition to schema expectancy being positively correlated with recall and recognition, expected objects were inferred in recall. The recall of these items demonstrated the power of the office-schema information in influencing place memory. There is also evidence that inconsistent items provoke better memory performance, known as the inconsistency effect (Lampinen, Copeland, Neuschatz 2001). Different theoretical models support specific hypotheses regarding how schemas influences memory. Pichet's & Anderon's 1966 schema model predicts better memory performance for schema consistent items claiming that inconsistent items are ignored. On the contrary, the contention that schema inconsistent information for an episodic event will be easily accessible and therefore would provoke better memory performance, the so-called dynamic memory model is also supported (Friedman 1979, Schank 1999, Holingworth & Henderson 1998). A review of a large number of independent schema studies conducted by Rojahn and Pettigrew 1992 found small differences between the number of studies supporting better memory performance for consistent objects and studies concluding that inconsistent objects provoke better memory performance. The researchers suggested that this specific outcome of their meta-analysis could be the result of a number of methodological variations such as, methods of presentation of the stimuli, total number of objects in the rooms and time of exposure to the environment. Research investigating how schema theories apply to real versus virtual memories have supported the consistency effect related to recognition scores as well as the inconsistency effect related to confidence ratings thereby demonstrating that synthetic scenes could induce similar responses as real world scenes (Flannery & Walles 2003).

The schema/non schema dichotomy taps into two separate processes each of which plays a role in our encoding of a scene by incorporating all information included in the scene. The appeal of schema theory to our goal of exploring functional fidelity is that it establishes the degree to which the appropriate schema has been activated in the real scene and its VE equivalent. Schema activation occurs when schema consistent or inconsistent information either positively or negatively correlate with memory performance, hence influencing place memory. The higher the similarity between the two schema activations, the higher the functional fidelity of the VE irrespectively of the level of visual or interaction fidelity per se of the VE simulation. This allows us to vary aspects of the VE system investigating how these variations impact on memory performance. How these factors can be investigated experimentally is a challenging research proposition and one which the present paper will be exploring.

The current study explores memory recognition after exposure to a computer graphics environment presented in three varied levels of rendering quality of the same space, displayed on a head tracked, stereo capable Head Mounted Display (HMD). The objective of the present work is to identify the relationship between memory recognition of elements of a scene comprising schema consistent and inconsistent objects and rendering fidelity. The methodology and the scene utilized were similar to that of Brewer and Treyens 1981. The current study was designed to explore the effect of diminished rendering quality, i.e. shadow detail variation, on memory recognition of consistent and inconsistent objects expressed by confidence ratings in an office setting. It was predicted that the results would be similar to those of Brewer and Treyens 1981 in showing the consistency effect. It was also anticipated that rendering quality would not have a strong effect on memory performance especially for objects expected to be found in a specific context.

3. Materials and Methods

3.1 Apparatus

The VEs were presented in stereo at VGA resolution on a Kaiser Electro-optics Pro-View 30 Head Mounted Display with a Field-of-View comprising 30 degrees diagonal. An Intersense Intertrax2, three degree of freedom tracker was utilized for rotation. The viewpoint was set in the middle of the virtual room and navigation was restricted to a 360 degrees circle around that viewpoint and 180 degrees vertically (rotation). Participants were sitting on a swivel chair during exposure. The application ran on a standard PC.

Despite the difference in polygon count the frame rate was retained constant across conditions at 12 frames per second.

3.2 Visual content

According to the group they were assigned to, participants completed the same memory recognition task in one of the following conditions:

1) Using an interactive, pre-computed radiosity simulation of an office on a stereo head-tracked Head Mounted Display (HMD); referred to as the **high-quality condition** (80% radiosity iterations)

2) Using an interactive, pre-computed radiosity simulation of the same office on a stereo head-tracked HMD; referred to as the **mid-quality condition** (40% radiosity iterations)

3) Using a low quality, interactive flat shaded computer graphics simulation of the same office on a stereo head-tracked HMD; referred to as the **low-quality condition**

Each environment varied considerably with regard to the nature of shadows (Figure 1). Radiosity algorithms display view-independent diffuse inter-reflections in a scene assuming the conservation of light energy in a closed environment. All energy emitted or reflected by every surface is accounted for by its reflection from or absorption by other surfaces. Radiosity methods allow any surface to emit light; thus, all light sources are modelled inherently as having area. The surfaces of a scene are broken up into a finite number of *n* discrete patches, each of which is assumed to be of finite size, emitting and reflecting light uniformly over its entire area. The result of a radiosity solution is an interactive three-dimensional representation of light energy in an environment allowing for soft shadows and colour bleeding that contribute towards a near-photorealistic image but without any specular reflections. Setting the subdivision accuracy to varying levels during the calculation of the illumination model produced a radiosity mesh of varying complexity, which in turn produced an illumination solution of varying shading accuracy. This process lead to the production of three distinct environments based around the same basic geometry and materials (Figure 1). The environment of the low-quality condition did not incorporate any shadows. The environment of the mid-quality condition was a result of 40% radiosity iterations. The environment of the high-quality condition was a result of 80% of available radiosity iterations. In all cases, a single ceiling mounted light source was used. The basic model construct was identical and the contents and room layout remained unchanged in each condition. All objects were visible and recognizable in all conditions. The level of luminance of the scene was constant across conditions.

The radiosity rendering process described above resulted in three distinct models of varying polygon count. Each of the three environments was presented in stereoscopic 3D by employing a dual channel video subsystem. The geometric models of the scene were imported into WorldUP - a proprietary VR authoring software package. WorldUP allows simulation of specific behaviour to be added in order to control the interaction with the synthetic scene. Due to the increased polygon count and stereo rendering, the high-quality radiosity environment placed a greater computation demand, therefore, it could not be rendered and displayed in real-time as rapidly as either the mid-quality or low-quality versions. In order to maintain parity with regard to the display and update speed of each environment given the differing levels of computational load, the maximum frame-rate of the high-quality environment was ascertained via the use of a simple frame-rate counter, at 12 frames per second (fps). A simple subsystem calculated the actual frame rate the selected environment was running at, compared this to the desired 12 fps once every frame and paused the simulation for the amount of time corresponding to the differential in frame-rate.

The experimental computer graphics space consisted of a graduate student's office (Figure 1). The objects included in the memory recognition questionnaire fell under five types. They were included in random order in the memory recognition questionnaire:

- Seven present room frame objects (walls, floor, ceiling, light, doorknob, door, light switch, etc.)
- Twenty-seven consistent objects present (computer, monitor, desk, paper bin, etc.)
- Eighteen consistent objects, absent (telephone, pens, computer mouse, etc.)
- Fifteen inconsistent objects present (skull, Viking helmet, etc.).
- Nine inconsistent objects absent (soldering iron, wrench, etc.)

The collection of these objects was largely based on a previous real world study by Brewer & Treyens 1981. Consistent objects are related to the office schema, e.g. it is likely that they are found in a graduate's office. Inconsistent objects are not likely to be found in a graduate's office, therefore, they are not associated to the office schema. This categorisation was the result of a pre-exposure study by Brewer & Treyens 1981. There were a total of seventy six objects listed.

3.3 Participants

Three groups of 12 participants were recruited from the University of Sussex, UK postgraduate population. A between-subject design was utilised balancing groups for age and gender. Participants in all conditions were naive as to the purpose of the experiment. All participants had normal or corrected to normal vision and no reported neuromotor impairment.

3.2 Procedures

The Inter Pupilary Distance (IPD) of each participant was measured prior to exposure and the stereo application's parallax was adjusted accordingly for each individual. The exposure time was 45 seconds across conditions. Participants were led to believe that this was just a practice phase of the main experiment, thus, they were not aware of the experimental task prior to exposure. Participants were given identical instructions across conditions. At the start of the simulation a pop-up window was generated utilised to acquire each participant's ID. Once the ID had been entered, the window was removed and a timer started. When this timer indicated that the 45 seconds of exposure had expired, the simulation was shut down automatically, ensuring that each test participant was restricted to exactly 45 seconds of exposure to the environment. Participants were then asked to indicate on a scale of 1 to 5 whether each object mentioned in the memory recognition list given was present in the environment to which they had been exposed, with a rating of 1 being positively not present and a rating of 5 being positively present. These ratings are referred to as confidence ratings. The room where the experiment was taking place was kept dark during exposure. The amount of time between exposure and memory testing was the same across conditions.

As suggested by Brewer & Treyens 1981, the amount of time spent looking at an individual object may affect memory encoding. Whilst neither the Brewer & Treyens experiment nor this research included systems necessary to track eye movement, a record of each test participant's head movement was monitored through software. Whilst this information is not at a high enough resolution to be useful in determining the time spent looking at each object, the amount and location of participants' idle time was monitored so as to ascertain that it was similar across conditions. This was considered significant in order to meaningfully compare memory recognition scores and confidence ratings across conditions. A measurement was taken once every 4 frames, providing 3 measurements every second across all conditions.

4. Results and Discussion

The participants completed the memory object recognition task across the three conditions. The confidence ratings and derived memory recognition scores were analyzed using analysis of variance (ANOVA).

Participants indicated on a confidence scale of 1 to 5 how strongly they believed each object in the environment was present, with 1 being positively not present and 5 being positively present.

Subsequently, the dataset was recoded to two values:

- 0, if the confidence score for each object was 1,2 or 3
- 1 if the confidence score for each object was 4 or 5

The proportion of correctly identified objects out of the total number of objects in each object category was calculated, signifying recognition scores. The statistical analysis reported below is also supported when confidence ratings are recoded differently, e.g., assigning confidence ratings 1 and 2 as zero and 3,4,5 as one. In the future it would be useful to provide a forced choice memory recognition test to avoid such recodes. The analysis is reported separately for present and absent objects, however, the statistical significances are likewise supported if a single unified analysis was followed.

The recognition scores relating to consistent and inconsistent objects that were present in the scene were subjected to a 3 (viewing condition: high-quality vs. mid-quality vs. low-quality) x 2 (object type: schema-consistent vs. schema inconsistent) mixed model ANOVA. Viewing condition was the between-subjects factor, object type was the within-subjects factor and the percentages of recognized objects was the dependent variable. Table 1 shows the mean recognition scores and standard deviations (in parentheses) as a function of viewing condition and consistent/inconsistent objects. A p-value of .05 was set for all effects. There was a significant main effect of object type, $F(1,33)=77.15$, $p<0.001$ revealing that present consistent objects were more likely to be recognized compared to present inconsistent objects (respective Ms. 50.10 vs. 22.03). There was also a main effect of viewing condition, $F(2,33)=3.27$, $p=0.05$. Post-hoc Tukey tests carried out on this main effect revealed that present objects were more likely to be recognized after participants were exposed to the mid-quality condition compared to the low quality scene, $p<0.05$ (respective Ms. 42.93 vs. 30.58). No other significant effects were observed. The interaction between object type and viewing condition was not significant, $F(2,33) < 1$, ns.

The recognition scores relating to absent consistent and inconsistent objects in the scene were also subjected to a 3 (viewing condition: high-quality vs. mid-quality vs. low-quality) x 2 (object type: schema-consistent vs. schema inconsistent) mixed model ANOVA. Viewing condition was the between-subjects factor, object type was the within-subjects factor and the percentages of recognized objects was the dependent variable. Table 1 shows the mean recognition scores and standard deviations (in parentheses) as a function of viewing condition and consistent/inconsistent objects. There was a significant main effect of object type, $F(1,33)=115.52$, $p<0.001$ revealing that absent consistent objects were more likely to be recognized compared to absent inconsistent objects (Respective Ms. 28.85 vs. 1.54). The main effect of viewing condition was not significant, $F(2,33) < 3$, ns. The interaction between object type and viewing condition also failed to reach significance $F(2,33) < 2$, ns.

Confidence ratings relating to present consistent and inconsistent objects in the scene were subjected to a 3 (viewing condition: high-quality vs. mid-quality vs. low-quality) x 2 (object type: schema-consistent vs. schema inconsistent) mixed model ANOVA. Viewing condition was the between-subjects factor, object type was the within-subjects factor and the confidence ratings was the dependent variable. Table 2 shows the mean confidence ratings and standard deviations (in parentheses) as a function of viewing condition and consistent/inconsistent objects. There was a significant main effect of object type, $F(1,33)=73.79$, $p<0.001$ revealing that confidence ratings were significantly higher for present consistent objects compared to present inconsistent objects (Respective Ms. 3.41 vs. 2.47). The main effect of viewing condition was not significant, $F(2,33) < 1$, ns. The interaction between object type and viewing condition also failed to reach significance $F(2,33) < 1$, ns. Confidence ratings relating to absent consistent and inconsistent objects in the scene were also subjected to a 2 (viewing condition: high-quality vs. mid-quality vs. low-quality) x 2 (object type: schema-consistent vs. schema inconsistent) mixed model ANOVA. This analysis revealed a significant main effect of object type, $F(1,33)=124.91$, $p<0.001$ showing that confidence ratings were higher for absent consistent objects compared to absent inconsistent objects (Respective Ms. 2.81 vs. 1.87). The main effect of viewing condition was not significant, $F(2,33) < 1$, ns. The interaction between object type and viewing condition was also not significant $F(2,33) < 2$, ns.

Interestingly, certain objects with high schema expectancy but which were not present in the room received higher confidence ratings than certain objects with similar schema expectancy that were present. Books, for instance received the highest confidence rating amongst the absent consistent objects with a combined average of 3.9. In addition, objects such as the telephone and pens that were absent received high recognition scores. The recognition of consistent objects which were absent such as books can only be accounted by schema-based knowledge about offices in general becoming integrated with episodic information. The lower recognition frequency of non-schema objects that were absent must be due to the lack of a schema frame supporting accurate recognition.

Monitoring navigational strategies resulted in idle time measurements. The head tracker was recording directional co-ordinates that could vary around 180 degrees, a similar range to the Field-of-View (FoV) of the human visual system. Navigation was monitored for horizontal and vertical actions. When participants were idle, their attention was assumed to be directed to the visible space based on their FoV. The average idle time was 20 seconds. There was not an effect of viewing condition upon idle time and positioning of idle time during exposure, horizontally, $F(2,33) < 1$, ns, nor vertically, $F(2,33) < 1$, ns. Therefore, the amount of time participants spent idle was similar across conditions. Monitoring of idle time is essential when real-time navigation is involved. This dataset could be further complemented with gaze information.

5. Conclusions

Memory research has established that schema-based information may be used in the process of retrieving information from memory (Brewer & Treyens, 1981, Brewer & Nakamura, 1984). However past research has often yielded contradictory results which may have been caused by differences in methodology used or even differences in scene context across experiments (Tojahn & Pettigrew, 1992). These results have sometimes shown better memory performance for objects consistent with the relevant schema but some studies have also shown enhanced recognition for inconsistent objects.

In the present research, we aimed to determine whether the effect of schemas on recognition memory would be analogous to a real-world study by Brewer and Treyens 1981 and a similar methodology was employed. Moreover, the research presented here explored the effect of rendering fidelity on recognition memory based on schemas aiming to identify whether consistent or inconsistent items provoke better memory performance in synthetic worlds and also determine the effect of rendering fidelity on memory recognition in simulated scenes. An important question to address was when and under which circumstances schema activation occurs. A further question to be explored would be what minimal cues relating to visual and interaction fidelity are needed for schema activation to take place.

As predicted, the results demonstrated that participants had better memory recognition performance for consistent objects compared to inconsistent objects. This object consistency effect was found both for objects that were present as well as absent in the scene and was independent of viewing condition, i.e. rendering quality. Overall confidence ratings were also higher for consistent compared to inconsistent objects, irrespective of viewing condition. These findings support the results found in the Brewer and Treyens 1981 study that demonstrated the consistency effect. The effect of rendering fidelity on memory recognition was rather weak and should thus be explored further in future work. Better recognition performance, for the totality of objects in the scene was found for the mid quality condition compared to the low quality condition. It is worth noting that both the mid quality and the high quality scene included shadows whereas these shadows were absent in the low quality scene. Therefore, presence of shadows seems to have induced greater recognition. There was no effect of viewing condition for memory recognition scores between the mid quality and high quality scene which indicates that if shadows were present, then it is immaterial as to whether they are accurately rendered or ‘boxy’-like. These results however, are not universally supported in the analysis, i.e. there was no effect of viewing condition between the low and the high-quality condition and should be explored in the future by introducing a continuous level of detail degradation. Also, shadow variations here were quite subtle and gross manipulations of visual fidelity explored in current research will lead to stronger effects. Moreover, the concept of visual fidelity only indicated the degree to which shadow information is present, hence, it could be expanded to include texture detail, diffuse or specular illumination rendering and interaction variations. Research avenues should identify methodological tactics and scene contexts which might provoke contradictory results Visual display devices, motion bases and related technology are developed to provide perceptual fidelity enhancement for VE systems and their influence on spatial awareness and memory for places should be also explored.

With further development and experimental work, this strategy could be ultimately exploited towards a selective rendering algorithm which would not rely upon either the visual focus of a specific task as in previous research (Cater, Chalmers & Ward, 2003) or rendering in high quality the 2 degrees foveal region of vision and with less detail the periphery of vision based on gaze information (McConcie, G.W., Loschky, L.C. 1997) but upon simulating a perceptual process that occurs in any spatial context, thereby offering greater generality. This goal could be accomplished by assigning a certain degree of visual detail according to the association of an object with the schema in context. It could also be achieved by omitting those objects or parts of objects that would be expected to be present given the nature of the scene (e.g., omitting the hands of a clock). The aim of the present research was to take a first step towards this goal by investigating the effect of rendering quality on object memory recognition by adopting the schema memory model framework.

It is widely recognized that perceptual fidelity is not necessarily the same as physical simulation. Identifying ways to 'induce' reality by possibly distorting physics based on fundamental perceptual processes triggered in real and synthetic worlds rather than simulating the physics of reality is a novel research route worth pursuing.

Acknowledgement

We wish to thank Prof. Bill Brewer, University of Illinois, Champaign, for his comments. This project will be continued based on a 2-year EPSRC funding grant (GR/S58386/01). We wish to also thank the anonymous reviewers for their insightful comments that contributed to the final version of this paper.

References

Bailey, J.H., Witmer, B.G. (1994). Learning and Transfer of Spatial Knowledge in a Virtual Environment. *Proc. of the Human Factors & Ergonomics Society 38th Annual Meeting*, 1158-1162, Santa Monica, CA: Human Factors & Ergonomics Society.

Bliss, J.P., Tidwell, P.D., Guest, M.A. (1997). The Effectiveness of Virtual Reality for Administering Spatial Navigation Training to Firefighters. *Presence: Teleoperators and Virtual Environments*, 6(1), 73-86. MIT Press.

Brewer, W. F. (2000). Bartlett's concept of the schema and its impact on theories of knowledge representation in contemporary cognitive psychology. In Saito (Ed.), Bartlett, culture and cognition, 69-89, Psychology Press.

Brewer, W. F., & Nakamura, G. V. (1984). The nature and functions of schemas. In R. S. Wyer & T. K. Srull (Eds.), Handbook of social cognition (Vol. 1, pp. 119-160). Hillsdale, NJ: Erlbaum.

Brewer, W.F. & Treyens, J.C. (1981). Role of Schemata in Memory for Places. *Cognitive Psychology*, 13, 207-2302.

Cater, K., Chalmers, A., Ward, G. (2003). Detail to Attention: Exploiting Visual Tasks for Selective Rendering. Proc. of the Eurographics Workshop on Rendering, 2003.

Coolican, H. (1999). *Research Methods and Statistics in Psychology*, 3rd edition. Hodder & Stoughton.

Dihn, H.Q., Walker, N., Hodges, L.F. (1999). Evaluating the Importance of Multi-Sensory Input on Memory and the Sense of Presence in Virtual Environments. *Proc. of IEEE VR 1999*, 222-228.

Ferwerda, J. (2001). Hi-Fi rendering. *Proc. of the ACM Siggraph campfire in Perceptually adaptive graphics*, http://isg.cs.tcd.ie/campfire/jimferwerda2.html.

Flannery, K.A., Walles, R. (2003). How does schema theory apply to real versus virtual memories? *Cyberspychology and Behavior,* 6(2), 151-159.

Friedman, A. (1979). Framing pictures: The role of knowledge in automized encoding and memory for gist. *Journal of Experimental Psychology: General,* 108(3), 316-355.

Hollingworth, A. & Henderson, J.M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General*, 127(4), 398-415.

Kuipers, B.J. (1975). A frame for frames: Representing knowledge for recognition. In D.G. Bobrow & A. Collins (Eds.), *Representation and Understanding: Studies in Cognitive Science*. New York: Academic Press, 1975.

Lampinen, J., Copeland, S., Neuschatz, J. (2001). Recollections of things schematic: rooms schemas revisited. *Cognition*, 27, 1211-1222.

Lathrop, W.B., Kaiser, M.K. (2002). Perceived Orientation in Physical and Virtual Environments: Changes in Perceived Orientation as a Function of Idiothetic Information Available. *Presence: Teleoperators and Virtual Enviroments*,11(1), 19-32. MIT Press.

Mania, K., Robinson, A. (2003). Simulating Spatial Assumptions. Technical Sketch. *ACM SIGGRAPH 2003*, San Diego, USA.

Mania, K., Troscianko, T., Hawkes, R., Chalmers, A. (2003). Fidelity Metrics for Virtual Environment Simulations based on Human Judgments of Spatial Memory Awareness States. *Presence, Teleoperators and Virtual Environments*, 12(3), 296-310. MIT Press.

McConkie, G.W., Loschky, L.C. (1997). Human performance with a gaze linked multi-resolutional display. *Advanced Displays and Interactive Displays First Annual Symposium*, 25-34. ARL Federated laboratory advanced displays and interactive displays consortium.

Minsky, M. (1975). A framework for representing knowledge. In P.H. Winston (Ed.), The Psychology of Computer Vision. New York: McGraw-Hill.

Pichert, J.W., Anderson, R.C. (1977). Taking a different perspectives on a story. *Journal of Educational Psychology* 69, 309-315.

Rojahn, K. Pettigrew, T. (1992). Memory for schema relevant information: a meta-analytic resolution. *British Journal of Social Psychology* 31, 81-109.

Schank, R.C. (1999) *Dynamic memory revisited*. Cambridge, UK. Cambridge University Press.

Waller, D., Hunt, E., Knapp, D. (1998). The Transfer of Spatial Knowledge in Virtual Environment Training. *Presence: Teleoperators and Virtual Environments*, 7(2), MIT Press.

N=36	Frame Objects (O=7)	Present Consistent (O=27)	Present Inconsistent (O=15)	Absent Consistent (O=18)	Absent Inconsistent (O=9)
Low	75.71 (15.71)	45.06 (18.87)	16.11 (10.80)	23.14 (16.03)	0.92 (3.20)
Mid	80.00 (17.14)	55.86 (22.19)	30.00 (11.89)	35.64 (16.99)	2.77 (5.02)
High	78.57 (14.28)	49.38 (14.41)	20.00 (10.24)	27.77 (13.40)	0.92 (3.20)
Total	78.09 (15.71)	50.10 (18.77)	22.03 (12.22)	28.85 (15.98)	1.54 (3.89)

Table 1: Recognition scores and standard deviations as a function of viewing condition (N total number of participants, O total number of objects).

N=36	Frame Objects (O=7)	Present Consistent (O=27)	Present Inconsistent (O=15)	Absent Consistent (O=18)	Absent Inconsistent (O=9)
Low	3.76 (0.52)	3.29 (0.57)	2.37 (0.75)	2.69 (0.44)	1.88 (0.74)
Mid	4.08 (0.67)	3.54 (0.75)	2.65 (0.47)	2.93 (0.71)	1.81 (0.58)
High	3.72 (0.57)	3.39 (0.58)	2.39 (0.57)	2.81 (0.48)	1.92 (0.52)
Total	3.85 (0.59)	3.41 (0.63)	2.47 (0.60)	2.81 (0.55)	1.87 (0.60)

Table 2: Confidence ratings and standard deviations as a function of viewing condition (N total number of participants, O total number of objects).

Figure 1: Flat-shaded (top), radiosity mid-quality environment (middle) and high-quality environment (bottom)

Figure 2: Apparatus used in all experimental conditions