Quantifying fidelity for virtual environment simulations employing memory schema assumptions

NICHOLAOS MOURKOUSSIS

University of Sussex
FIONA M. RIVERA

University of Sussex
TOM TROSCIANKO

University of Bristol
TIM DIXON

University of Bristol
RYCHARDE HAWKES

Hewlett Packard Laboratories
and
KATERINA MANIA

University of Sussex

In a virtual environment (VE), efficient techniques are often needed to economize on rendering computation without compromising the information transmitted. The reported experiments devise a functional fidelity metric by exploiting research on memory schemata. According to the proposed measure, similar information would be transmitted across synthetic and real world scenes depicting a specific schema. This would ultimately indicate which areas in a VE could be rendered in lower quality without affecting information uptake. We examine whether computationally more expensive scenes of greater visual fidelity affect memory performance after exposure to immersive VEs, or whether they are merely more aesthetically pleasing than their diminished visual quality counterparts. Results indicate that memory schemata function in VEs similar to real world environments. "High-level" visual cognition related to decisions based on visual information is unaffected by ubiquitous graphics manipulations such as polygon count and depth of shadow rendering; "normal" cognition operates as long as the scenes look acceptably realistic. However, when the overall realism of the scene is greatly reduced, such as in wireframe, then visual cognition becomes abnormal. Effects that distinguish schema-consistent from schema-inconsistent objects change because the whole scene now looks incongruent. We have shown that this effect is not due to a failure of basic recognition.

---

1. INTRODUCTION

Authors' addresses: Nicholaos Mourkoussis, Fiona M. Rivera and Katerina Mania are with the Department of Informatics, University of Sussex, UK, BN1 9QH; e-mail: N.Mourkoussis@sussex.ac.uk, F.M.Rivera@sussex.ac.uk, K.Mania@sussex.ac.uk; Tom Troscianko and Tim Dixon are with the Department of Experimental Psychology, University of Bristol, UK, BS8 1TU; e-mail: tom.troscianko@bristol.ac.uk, Timothy.Dixon@bristol.ac.uk.; Rycharde Hawkes is with Hewlett-Packard Labs, Bristol, UK, BS34 8QZ; e-mail: rycharde.hawkes@hp.com.

The entertainment world appears to consider highly realistic visual quality one of the keys to success, with cinematic quality graphics claimed for the next generation of gaming consoles. On the other hand, when interactive immersive Virtual Environments (VEs) are implemented for training rather than entertainment purposes, visual quality might not be as significant. If the training is to be effective, the skills acquired must transfer into the real world at appropriate levels of performance. A VE with maximum visual and interaction fidelity would result in a transfer of information equivalent to real world training since the environments would be indistinguishable [Mania et al. 2004]. Visual fidelity refers to the degree to which visual features in the VE conform to visual features in the equivalent real environment [Waller et al. 1998; Rodger and Browse 2000]. Functional realism refers to the communication of similar information in the real and virtual world rather than the aesthetics or physics, in the sense that users are able make the same judgments and perform the same tasks as in the real world [Ferwerda 2001; Nemire et al. 1994].

It is tempting to replicate the real world as accurately as possible in order to provide equivalent experiences [Liu et al. 2005]. Whilst arguably ideal, it is not yet computationally feasible for this to occur. Trade-offs between visual/interaction fidelity and computational complexity should be applied to a simulation system without detracting from its training effectiveness [Mania et al 2003; Wanger 1987]. There is, therefore, a call for efficient techniques assessing the fidelity of a VE and determine its relationship with performance in order to economize on rendering computation without compromising the level of information transmitted (functional realism) [Ferwerda 2003]. Whilst existing techniques aim to address this issue, it can be argued that no one technique provides a comprehensive approach. Previous research suggested rendering in high quality the fovea region based on gaze information [McConcie and Loschky 1997; Watson et al. 1997]. Gaze-dependent rendering can economize on computational complexity by rendering in high resolution only those parts of an image that are at the focus of attention. However, such an approach might encounter difficulties of maintaining updates of the multi-resolution display without disturbing the visual processing after a fast (~4ms) eye saccade.

Following a different approach, it has been proposed to assign selective high quality rendering in the visual angle of the fovea (2o) centred on the users' task focus [Cater et al. 2003]. In this manner, inattentional blindness may be exploited to accelerate rendering by reducing quality in regions unrelated to the task focus. This approach, however, cannot be applied when there is no overt task to be conducted and even when there is, we cannot predict exactly where each task-relevant saccade will land. Haber et al. [2001] suggested rendering the informative areas of a scene in varying quality based on saliency models, where the most salient areas are rendered at higher quality. Such bottom-up visual attention models do not always predict attention regions in a reliable manner [Marmitt and Duchowski 2002]. Indeed, Land [1999] argues that salience models do not account for any fixations when carrying out a real-world task such as making a cup of tea. Correlation between actual human and artificially presented scanpaths was found to be much lower than predicted. Moreover, we have no generally accepted model of comparing scanpaths.

Given the limitations of the existing selective rendering techniques as discussed, an approach is needed which can be applied in more general settings, irrespective of the task at hand, the hardware equipment and bottom-up processes such as saliency models. The reported experiments attempt to explore how functional fidelity is communicated by exploiting existing research on memory schemata which could ultimately guide selective rendering based on spatial cognition processes. Schemata are cognitive structures based on past experience representing how humans mentally achnowledge the information

communicated by a given stimulus. In terms of real world scenes, schemata represent the general context of a scene such as 'office', 'theatre' etc. and the memory of the objects in a given context according to their association with the schema in place.

The contribution of this paper is an innovative approach to measuring functional realism, influenced by schema memory theory applied to the simulation arena. According to the proposed measure, similar information should be transmitted between synthetic and real world scenes, depicting the same schema. This would, in due course, enable reduction of computational complexity by indicating which objects or areas in a synthetic scene could be rendered in lower quality without affecting information uptake. The discussed experiments examine whether computationally more expensive scenes containing a higher polygon count, or superior rendering quality and thus greater visual fidelity, affect memory performance after exposure to immersive VEs displayed on a stereo capable, head-tracked Head Mounted Display (HMD), or whether they are merely more aesthetically pleasing than their diminished visual quality counterparts.

The paper is organized as follows: in section 2, we provide a literature review of the effect of schemata on memory performance in real and synthetic environments. Section 3 discusses the common methods and materials in both reported experiments. Section 4 presents the first experiment that aims to identify the effect of polygon reduction on memory performance and presents the results. Section 5 discusses the second experiment that aims to study the effect of two extremes type of rendering (wireframe vs. radiosity) on memory performance and includes relevant results. Finally, in section 6 we present a comprehensive discussion of the implications of these experiments on assessing the fidelity of a simulation.

2. THE ROLE OF SCHEMATA IN MEMORY

In order to investigate the tradeoffs between visual fidelity and computational complexity, the studies in this paper put forward an approach based on classic findings from memory research in which schemata explain memory processes. Schemata are knowledge structures of cognitive frameworks drawn from experience that facilitate the interpretation of events, and influence memory retrieval. Memory schemata were introduced to psychology by the work of Bartlett, who stated that:

> *"Remembering is not the re-excitation of innumerable fixed, lifeless and fragmented traces. It is an imaginative reconstruction, or construction, built out of the relation of our attitude towards a whole active mass of organized past reactions or experience, and to a little outstanding detail which commonly appears in image or in language form." [Bartlett 1932]*

Bartlett believed that memory is a gist-oriented function that aims to interpret stimuli based on previous built stereotypes. Various experimental studies have investigated the effect of schemata on memory performance and although they agree that schemata affect the perception of stimuli, they disagree about precisely how. Studies in real-world settings, and in virtual environments have both produced contradictory results. Some have found that memory performance is better for objects that are congruent with the schema (consistent objects – consistency effect) rather than objects which do not appear to belong to the schema (objects that are inconsistent with the schema – inconsistency effect). Brewer and Treyens [1981] asked participants to wait in an academic's office for 35 seconds before being subjected to a memory test for recall and recognition. They found that both recall and recognition were significantly better for the objects consistent with and academic's office. They also found that participants falsely recalled and recognized consistent objects that were absent from the room. Memory performance has also been found to be better for objects that are incongruent with the schema, e.g.

inconsistent objects. In a replication of Brewer and Treyens [1981], recall results significantly correlated with the salient objects irrespective of their schema-expectancy [Saab et al. [1984]. They assumed that this deviation occurred because their participants were recruited from an introductory psychology course and probably lacked a clear schema for a lab assistant's office. Pezdek et al. [1989] found that memory performance was better for schema-inconsistent objects after participants had been exposed to an experimental waiting area.

Studies on pictures or drawings of real world scenes demonstrated similar contradictory results. Loftus & Mackworth [1978] instructed participants to look at pictures without being aware that they would be completing a recognition test. Results indicated that participants fixate earlier, more often, and with longer duration on inconsistent objects and memory performance is better for those objects. In a similar experiment participants had better memory performance for consistent objects in inconsistent but possible (novel) locations in the scenes rather than either consistent objects in consistent locations, or consistent objects in inconsistent but impossible locations, or, finally, unorganized inconsistent objects [Hock et al. 1978]. Research concluded that schemata function as semantic pattern detectors [Friedman 1979]. Perceptual knowledge in schemata could be adopted for relatively automatic pattern recognition and comprehension. According to this work, participants were able to identify the consistent objects quickly. In contrast, identification of inconsistent objects required more analysis of local visual details. Participants became aware of the changes that had occurred to inconsistent objects and often did not notice when consistent ones were deleted or replaced with different consistent objects. Goodman [1980] found that whilst participants were better able to recall or recognize consistent objects, they were better able to recognize figurative details of the inconsistent objects.

Flannery and Walles [2003] explored the role of schemata in memories acquired through real and virtual environments. They found that participants had better recognition for consistent objects, but were more confident for the recognition of the inconsistent objects. In a preliminary virtual environment experiment, Mania et al. [2005] investigated the effect of memory schemata on object recognition in VEs. Participants were instructed to explore a VE for 45 seconds, unaware that a memory performance task would follow. Results revealed better recognition and confidence scores for the schema-consistent objects. . These studies did not investigate the effect of varied aspects of visual fidelity such as style of rendering and polygon detail on object recognition in scenes of varied context.

Given these conflicting results, Rojahn & Pettigrew [1992] conducted a review of 60 independent studies and suggested that such effects could be due to methodological variations. The study concluded that there is a small difference between the number of studies supporting stronger memory performance for schema-consistent objects (the consistency effect) and better memory performance for schema-inconsistent objects (the inconsistency effect). Both recognition tests corrected for guessing and recall tests, revealed better memory performance for schema-inconsistent information. However recognition tests uncorrected for guessing, consistently uncovered better memory performance for schema-consistent information. Furthermore, it was concluded that the inconsistency effect corrected for guessing is supported: when the task is less complex, therefore, more processing capacity is left of encoding inconsistent information; when the proportion of irrelevant/inconsistent items is small; when the inconsistent objects are highly inconsistent; when the category is of moderate or high importance to the participants.

Despite contradictory results in literature as detailed above, existing research exploring the effect of schemata in perception, comprehension and memory agree that

schema consistent elements are 'expected' to be found in a given context. It seems that information slots that have not been filled with perceptual information are likely to be filled by default assignments based on stereotypical expectations. Attention in a given stimulus is fixated more times on the inconsistent information; therefore, there is better memory performance for the figurative details of inconsistent information. Similar consistent information may activate the same schema at different times.

Of particular interest to this present study is the moderating variable of the proportion of schema inconsistent items. According to Rojahn and Pettigrew [1992], the smaller the proportion of schema inconsistent items the better the memory performance for this category of objects. However, results of experimental studies for real world scenes and VEs are contradictory. For example, in [Brewer and Treyens 1981; Mania et al. 2005] the experimental rooms used to evoke the schemata were similarly populated with far more consistent than inconsistent objects. Results of these studies indicated that memory performance was better for the schema consistent information. These results seem to disagree with the results of the experimental study of [Pezdek et al. 1989], who they found better memory performance for the inconsistent objects.

The consistent/inconsistent dichotomy seems to tap into a single memory process, 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 either schema consistent or schema inconsistent information positively or negatively correlates with memory performance, hence influencing place memory by exploiting schema-related expectations and prior knowledge. The closer the similarity between schema activations between the real and virtual settings, the higher the functional fidelity of the VE irrespective of the level of visual or interaction fidelity per se. This allows us to vary aspects of the VE system investigating how these variations impact memory performance. How these factors can be investigated empirically is a challenging research proposition and one that the present paper will explore.

Previous work included a preliminary investigation of the effect of object type (consistent vs. inconsistent) and shadows (flat-shaded scene vs. radiosity scene) on object memory recognition in an immersive VE [Mania et al. 2005]. The flat-shaded scene contained no shadows. Radiosity algorithms display view-independent diffuse inter-reflections in a scene assuming the conservation of light energy in a closed environment. It produces colour-bleeding effects from one surface to another, shades inside the shadow area and creates soft-edge shadow with penumbrae along shadow boundaries imitating the physical propagation of light in the real environment. Results revealed that schema consistent elements of the scene were more likely to be recognized than inconsistent information. In addition, higher confidence ratings were assigned to consistent rather than inconsistent items. Total object recognition was better for the scene including rough shadows compared to the flat-shaded scene, however, higher shadow quality did not make any difference. Thus, low quality of rendering was adequate for better memory recognition of consistent objects. This study was limited to the investigation of subtle shadow variations whilst the visual quality of the objects themselves remained identical.

The experiments discussed in this paper aim to further investigate the effect of variation of visual fidelity, plus differing object ratios on object recognition in immersive simulations. The first study explores the effect of diminished rendering quality produced by reducing the polygon count of the objects that comprise a scene. It has been shown that human observers are sensitive to changes in polygon degradation when still images of single objects were viewed [Watson et al. 2001]. The initial study presented here questions whether human observers' visual preference for the higher fidelity images

comprising a higher object polygon count translate into enhanced memory performance influenced by a particular schema in immersive real-time, interactive VEs. Experiment 2 explores the effect of diminished rendering quality produced by two types of rendering (wireframe vs. radiosity) on memory schemata and memory recognition. In addition, it also explores the effect on memory performance of varying the ratio of consistent versus inconsistent objects within a VE. For both experiments, memory recognition accuracy is accompanied by self-report of confidence levels which provides extra information on recognition upon guessing. Such recognition can be removed in the analysis resulting in a higher probability of identifying variations of responses.

3. OVERVIEW OF THE EXPERIMENTS

3.1 General Methods
This section describes the general methods applicable to both experiments to follow. Study-specific methods are discussed in the subsequent sections.

3.1.1 Participants. Participants were recruited from the University of Sussex, UK undergraduate, postgraduate and other associated research staff populations. Participants in both experiments were naïve as to the purpose of the experiment. All participants had normal or corrected to normal vision and no reported neuromotor impairment. Participants were balanced across groups for age, background and gender.

3.1.2 Stimulus and Apparatus. The interactive simulated scenes were displayed at stereoscopic VGA resolution in head-tracked HMDs by employing a dual channel video subsystem. An Intersense Intertrax2 three degree of freedom (DOF) tracker sensed motion of participants' head movements. The viewpoint was set in the middle of the synthetic rooms with horizontal navigation of 360o and vertical rotation of 180o provided. Participants sat on a swivel chair during exposure. Each room contained both consistent and inconsistent objects. In order to maintain parity with regard to the display and update speed of each environment given the different levels of computational load, the maximum frame rate of each environment was ascertained by monitoring the system's built-in frame rate counter. Parameters were then set to restrict and equate the frame rate across experimental conditions. The simulation system ran on a standard PC.

3.1.3 Procedure. Both experiments were conducted in three consecutive stages: training, exposure to stimulus, and memory test.

During the training phase, the interpupilary distance (IPD) of each participant was measured with a common ruler prior to exposure and the stereo application's parallax was adjusted accordingly to reduce visual stress. The training simulation consisted of an empty room in which each participant navigated while the correct fit of the HMD was ensured. When the participant indicated that he/she was familiar with the equipment in use, the experimenter informed the participant that the next training scene would be loaded. Participants were explicitly instructed that they were still in training. However, they were subsequently exposed to the main experimental stimulus. A pop-up window was generated at the start of the main simulation to acquire each participant's ID. Once the ID had been entered, the window was removed and a timer started. When the timer indicated that the study specific exposure time had elapsed, the simulation was automatically shut down, ensuring that each participant was restricted to the same preset time of exposure to the environment. Participants were given identical instructions across conditions. The real-world room where training and exposure to stimulus took place was kept dark during exposure. The amount of elapsed time between exposure and memory testing was within 2 minutes.

After exposure, participants completed a subject-paced memory recognition questionnaire including a list of objects. Participants indicated whether they believed that each object was actually present in the VE by circling 'YES' or whether it was absent by circling 'NO'. Their responses are referred to as recognition data. Participants also indicated on a scale of 1 to 5 their confidence concerning the presence or absence of each object, where 1 indicated no confidence, and 5 absolute certainty. These ratings are referred to as confidence ratings.

As reported in existing literature, the amount of time spent gazing at an object may affect memory encoding [Loftus and Mackworth 1978; Flannery and Walles 2003; Rojahn and Pettigrew 1992]. A record of each participant's head movement was recorded. It was not possible to track actual eye movement. Whilst this information is not at high resolution to be useful in determining the time spent fixating at each object, the amount and location of the participants' idle time was calculated to ascertain that it was similar across conditions. This is significant in order to meaningfully compare memory recognition scores and confidence ratings. A measurement was acquired on every frame, providing 30 measures every second for the first study, and 20 measures every second for the second study in this paper.

4 FIRST EXPERIMENT: THE EFFECT OF POLYGON COUNT

This experiment was designed to explore the effect of memory schemata and diminished rendering quality, i.e. high and low polygon count of objects, on memory recognition of consistent and inconsistent objects in a domestic bathroom setting. It was predicted that the higher polygon count scene would provoke better memory performance for the inconsistent objects. The level of saliency of the inconsistent objects is higher; therefore, they may require a higher level of visual detail to be perceptually acknowledged. Consistent objects are somehow expected to be found in a given scene, therefore, because of such preconceptions, visual detail may be less essential for memory recognition of these objects.

4.1 Method
4.1.1 Design. The experiment had two independent variables in a mixed subjects design. The first between-subjects variable was 'polygon count' comprised of two levels: 'low-poly' (6,000 polygons) and 'high-poly' (100,000 polygons). The second within-subjects independent variable was 'object type' comprised of two levels: 'consistent' and 'inconsistent'. The dependent variables were 'hit rate', i.e. rate of successful recognition of present objects and 'false alarm rate', i.e. rate of incorrect recognition of absent objects.

According to the group they were assigned to, participants were exposed to an interactive pre-computed radiosity simulation of a domestic bathroom in one of the 'polygon count' conditions (Fig. 1).

Fig. 1. Low-polygon (left) and high-polygon (right) environment,

The room viewed contained 30 objects, 15 inconsistent and 15 consistent. Consistent and inconsistent items were selected in subjectively visually matching pairs where possible, i.e. each having roughly the same physical scale, and level of geometric detail. Examples of such pairings included a bathtub and a beer barrel, a toothbrush and a fork, a mirror and a dartboard. There was no duplication of objects besides the taps that were obviously consistent, placed in two pairs at two separate locations, by the sink and bathtub. In an attempt to counterbalance any biases relating to the doubling of the number of the taps in the scene, two sets of darts perceived as inconsistent objects were included in the scene, one pair on the dartboard near the sink, and one on the bathtub.

4.1.2 Participants. 32 participants were recruited as discussed in section 3.1.1. 50% were male and 50% female. Age ranged mainly from 18-25. Participants formed two balanced groups of 16 corresponding to the two polygon count conditions (between-subjects).

4.1.3 Apparatus and Stimuli. The technical apparatus has been described under General Methods (cf. Section 2.1.2). The VEs were presented in stereo at VGA resolution on a Rockwell Collins Pro-View 30 HMD with a diagonal field of view (FOV) of 30 degrees. The frame was constant across conditions at 30 frames per second (fps), and rotational navigation was enabled.

Stimulus for experiment 2 was an academic office following the common choice of scene/context selection in literature [Brewer and Treyens 1981; Saab et al. 1984; Pezdek et al. 1989; Mania et al. 2005]. Stimulus for experiment 1 differed due to preliminary modelling tests which indicated that an office scene would not be ideal one to vary polygon count on the consistent objects. This is because the schema consistent objects such as a desk, computer monitor and books could be represented using basic rectangular shapes and a low polygon count. Modelling such objects with additional polygons made little visual difference on such objects. Thus, there would be insufficient visual differences between the low polygon version and the high polygon version of the consistent objects. Alternative scenes were considered, such as a school classroom, bathroom, dining room, garage, restaurant, child's bedroom, and living room. Most of these contexts were discounted because there was no clear-cut discrimination by participants between consistent and inconsistent objects who stated that they kept bicycles and other potentially 'inconsistent' items in their living or dining rooms. A domestic bathroom was preferable since the pilot study revealed clear-cut impressions of what to expect to be present/absent in a bathroom. Furthermore bathroom objects such as bathtubs or sinks, tend to be curved compared to the rectangular shapes found in an office. Preliminary modelling tests confirmed that the variation of polygon count for such objects resulted in greater visual differences. Thus using a domestic bathroom scene enabled the polygon count of both the consistent and inconsistent objects to both be manipulated in a similar manner. Experiment 2 still used academic's office scene since is

the more common choice from literature mentioned above and wireframe v radiosity rendering could be used to vary visual fidelity without raising the same issues as varying polygon count.

The base models consisted of the minimum geometry required to form a basic shape that could be enhanced in order to create the high-poly scene. This was achieved by applying a NURMS (Non-Uniform Rational Smooth) subdivision surface smoothing mesh with 1 or 2 iterations, depending on the object. The NURMS modifier added additional polygons for each base model in order to achieve a more rounded appearance. The aim was to model objects with little or no visible jagged edges whilst retaining the scene's total polygon count under 100,000 due to software and hardware limitations. This approach differed from Watson et al. [2001] that initially simplified each of the objects' polygon count down to the number contained in the model of the object with the fewest polygons. Since the objects selected for the study presented in this paper ranged in complexity from simple objects such as a bar of soap, to more complex objects such as a bathtub or sink, this approach would not be suitable. Objects that are more complex would appear much coarser than their simple counterparts would without allocating additional polygons. Preliminary modeling tests confirmed this. Reducing polygon counts by a consistent percentage produced substantially poorer visual appearance in the simple objects compared to the complex ones. To maintain consistency of appearance throughout the condition, therefore, the polygons in each object were reduced on an individual basis until it was felt that the object contained the minimum number of polygons required to retain the object's recognizable shape. Thus, it was decided to create scenes in which all objects communicated the same perceived visual quality per condition, regardless of polygon count per object, as determined by detailed pilot studies.

The low-poly scene was created by applying the 3ds Max MultiRes modifier to reduce the polygons on an object-by-object basis, starting from the high-poly scene. Although other polygon reducing tools were considered, the MultiRes modifier was selected largely due to its ease of use and ability to decrease polygons in models by a user-defined percentage. Although it was initially evident that by applying default uniform grey shading influencing factors such as colour could be avoided, subsequent pilot studies revealed that certain grayscale objects could not be recognized when viewed through the HMD. Colour was consequently introduced simply to facilitate identification of the objects. Blinn shaders of basic diffuse colour and specular values were applied to all objects. Textures were not used because texture mapping could hide the underlying geometry of the objects. A tiled-effect texture was applied only on the floor and walls in order to enhance the realism of the scene, and to help establish the room's 'identity'.

The 60 objects included in the memory recognition questionnaire fell under 4 distinct categories; 15 schema-consistent objects that were present in the scene (e.g. sink, bathtub, toothbrush); 15 schema-inconsistent objects present in the scene (e.g. toaster, banana); 15 schema-consistent objects that were absent (e.g. towel, bath mat); 15 inconsistent objects absent (e.g. iron, sword).

A preliminary list of objects was assembled based on an initial pilot study which explored which objects were expected to be found in a bathroom and which were not. Bathtub, sink and toilet were the most popular as expected items and a kettle, toaster and electrical appliances as strongly unexpected. Five participants subsequently ranked the objects on the list. The consistency of each item was rated on a scale from 1 to 6 according to whether they expected to find each object in a typical domestic bathroom, or not with 6 being the most expected, and 1 being the least. Based on these ratings, consistent objects were selected from the high end of the scale, and the inconsistent ones from the low end.

4.1.4 Procedure. As detailed in Section 3, participants completed a memory questionnaire after 60 seconds of exposure to either the low or high poly condition. The time of exposure was defined after control studies indicating that participants might omit viewing parts of the environment if exposed to the scene for less time.

4.2 Results and Discussion

The hit and false alarm rates obtained were used to calculate the signal detection constituents d' (sensitivity) and β (bias), as outlined in by Wickens [2001], with d' being the distance between signal and noise distributions in participant responses (i.e. identifying correctly when a object was present or not), and β being the bias towards replying 'yes' or 'no'. The d' is calculated by transforming hit and false alarm rates into z-scores, which give the position of a score in relation to the number of standard deviations it is above or below the mean. The d' score equals the hit ratio z-score minus the false alarm ratio z-score: d' = z(HIT) − z(FA). The β value is the ratio of the height of the Hit and False Alarm rate distribution curves: β = h(HIT)/h(FA). These values were calculated for each condition in the experiment, and scores were compared using a 2 x 2 mixed Analysis of Variance (ANOVA), using log β, as the β distribution is not statistically normal.

When participants were asked to recognise which items were present in the scene, they were also asked to rate how confident they were of their responses on a scale of 1 to 5. These confidence ratings were subsequently used to correct the recognition responses

Table 1. Exp 1 – Uncorrected confidence results

N = 32	Cons. *d'* (O = 15)	Inc. *d'* (O = 15)	Cons. *β* (O = 15)	Inc. *β* (O = 15)
High Poly	1.36 (.65)	1.47 (.58)	-.03 (.67)	.91 (.61)
Low Poly	.98 (.55)	1.38 (.55)	-.30 (.40)	.71 (.62)
Total	1.17 (.62)	1.43 (.56)	-.16 (.56)	.81 (.62)

d' and β scores and standard deviations (in brackets) as a function of viewing condition
(N Total Number of Participants, O Total Number of Objects).
Cons = Consistent, Inc. = Inconsistent.

for possible guessing. This provided three sets of d' and β results: one of uncorrected confidence; one of moderate confidence (i.e. corrected so that only recognition scores with corresponding confidence ratings of 4 and 5 were used); and one set of high confidence (i.e. only confidence scores of 5 were used).

4.2.1 Uncorrected Confidence Results. The descriptive statistics for the uncorrected d' and β scores are shown in Table 1.

Note that the greater the d' score, the easier it was for participants to distinguish between signal and noise (i.e. recognise correctly present objects), whilst the closer a β score is to 0, the less bias was found in participant responses, with positive scores indicating a bias towards responding 'no', and negative scores a bias towards responding 'yes'.

ANOVA analysis of the d' results revealed no significant main effects of viewing condition (F (1,30) = 2.61, p = .117) or object type (F (1,30) = 3.06, p = .090), and no interaction (F (1,30) = .961, p = .335). However, the object type results were not significant but close to significance, considering a two-tailed test was used. Therefore, with no correction the inconsistent objects are close to having a significantly more correct recognition scores than the consistent ones. As discussed in the literature [Rojahn and

Pettigrew, 1992], when recognition tests are uncorrected for guessing, memory performance usually does not shift towards either the consistent or the inconsistent objects.

The β results revealed no significant main effect of viewing condition (F (1,30) = 2.11, p = .157), a significant main effect of object type (F (1,30) = 59.3, p < .001), but no interaction (F (1,30) = .065, p = .801). This indicates that the inconsistent responses were significantly more biased to responding 'no' than the consistent responses, which themselves had close to no bias at all. As discussed in the literature, when participants attempt to remember if an inconsistent object was present, they tend to answer 'no' (item was not present) if they are not certain about the object's presence. For example in this

Table 2. Exp 1 – Moderate confidence results

N = 32	Cons. *d'* (O = 15)	Inc. *d'* (O = 15)	Cons. *β* (O = 15)	Inc. *β* (O = 15)
High poly	1.41(.64)	1.59(.56)	0.92(.62)	1.11(.47)
Low poly	1.17(.40)	1.49(.60)	0.84(.57)	1.07(.49)
Total	1.29(.54)	1.54(.57)	0.88(.59)	1.09(.48)

d' and β scores and standard deviations (in brackets) as a function of viewing condition
(N Total Number of Participants, O Total Number of Objects).
Cons = Consistent, Inc. = Inconsistent.

particular experiment, a beer barrel was present in the domestic bathroom. If participants on the recognition test were not sure of the object's presence they would most probably answer 'no' assuming that if the object was present they would remember it since it stands out of the context of the scene.

4.2.2 Moderate Confidence Results (4-5 confidence). Table 2 shows the descriptive statistics for the confidence scores of 4-5 corrected results on a rating scale from 1 to 5.

The 4-5 corrected d' scores for the first experiment found a main effect of object type (F (1,30) = 4.33, p = .046), but not of viewing condition (F (1,30) = 1.20, p = .282), and no interaction (F (1,30) = .315, p = .579). The effect of object type indicated that inconsistent objects (d' = 1.55) were significantly more accurately identified than the consistent objects (1.30).

The β results for the 4-5 corrected responses found no effect of object type (F (1,30) = 3.20, p = .084), viewing condition (F (1,30) = .155, p = .697) and no interaction (F (1,30) = .036, p = .852), suggesting that bias was not significantly altered by any of the conditions when corrected to this level.

4.2.3 High Confidence Results (5 confidence level). Table 3 presents the corrected descriptive statistics for the confidence scores of 5 on a rating scale from 1 to 5.

The d' results revealed no main effect of viewing condition (F (1,30) = .001, p = .982), an almost significant main effect of object type (F (1,30) = 4.14, p = .051), and no interaction (F (1,30) = .603, p = .443). Interestingly, the p value of the object type finding has dropped from the 3-4-5 correction, suggesting that correcting to such a degree is causing a loss of power to the overall analysis, as some of the data is not being used.

The β results revealed no main effect of viewing condition (F (1,30) = .204, p = .655), nor object type (F (1,30) = .225, p = .639) and no interaction (F (1,30) = 1.07, p = .309), suggesting that correcting for guessing to this extent has caused some other systematic trends to occur. This indicates that people who were very sure of their responses were uniformly biased towards answering 'no'.

5. SECOND EXPERIMENT: EFFECT OF TWO EXTREMES TYPES OF RENDERING (RADIOSITY VS. WIREFRAME)

Experiment 1 clearly supported the inconsistency effect besides the uncorrected for guessing confidence condition. Inconsistent objects were more correctly recognized as being present or being absent than the consistent objects. In order to more closely examine the effect of diminished quality, Experiment 2 was designed to explore the effect of memory schemata, diminished rendering quality (i.e. wireframe and radiosity) and varied ratios of consistent/inconsistent objects in the scene on memory recognition of consistent and inconsistent objects, in an academic's office setting. It was predicted that retrieval of spatial information in a VE operates according to the schemata cognitive framework in real-world studies where an increase in the ratio of consistent versus inconsistent objects available in a VE will provoke better memory performance towards inconsistent rather than consistent objects. It was also anticipated that rendering quality would not have an effect on memory recognition for the consistent objects, but it would have an effect for the inconsistent objects.

5.1 Method

Participants were exposed to six different instances of a scene in terms of rendering, i.e. wireframe and radiosity, and varied ratio of consistent versus inconsistent objects and were required to recognise objects present in the environment.

5.1.1 Design. The experiment had three independent variables in a mixed subjects design. The first between-subjects variable was 'rendering quality' that comprised of two

Table 3. Exp 1 – High confidence results

N = 32	Cons. *d'* (O = 15)	Inc. *d'* (O = 15)	Cons. *β* (O = 15)	Inc. *β* (O = 15)
High poly	1.31 (.59)	1.45 (.69)	1.23 (.49)	1.19 (.54)
Low poly	1.22 (.38)	1.54 (.40)	1.22 (.41)	1.34 (.34)
Total	1.26 (.49)	1.49 (.56)	1.23 (.45)	1.27 (.45)

d' and β scores and standard deviations (in brackets) as a function of viewing condition
(N Total Number of Participants, O Total Number of Objects).
Cons = Consistent, Inc. = Inconsistent.

options: 'wireframe' and 'radiosity'. The second between-subjects independent variable was 'object ratio' (ratio of consistent versus inconsistent objects in the scene) that comprised of three options: '35 consistent/15 inconsistent', '25 consistent /25 inconsistent' and '15 consistent /35 inconsistent'. The third within-subjects variable was 'object type' comprised of two options: 'consistent' and 'inconsistent'. The dependent

variables were 'hit rate', i.e. rate of successful recognition of present objects and 'false alarm rate', i.e. rate of incorrect recognition of absent objects.

Participants were exposed to an interactive simulation of an academic's office in one of the conditions below. Each scene (Fig. 2) varied considerably with regard to the rendering algorithm (i.e. wireframe vs. radiosity) and the proportion of consistent vs. inconsistent objects available.

• A wireframe rendered simulation including either 35/15, 25/25 or 15/35 consistent/inconsistent present objects, referred to as the '35/15, 25/25 or 15/35 wireframe' condition.

• A radiosity rendered simulation including either 35/15, 25/25 or 15/35 consistent/inconsistent present objects, referred to as the '35/15, 25/25 or 15/35 radiosity' condition.

In each scene, 50 target present locations were defined. Thirty of those were common across conditions. Fifteen were of consistent object type and 15 of inconsistent object type. The remaining of the target present locations was populated with appropriate objects in order to comply with the defined object ratio. Consistent and inconsistent targets were selected in subjectively visually matching pairs where possible, i.e. each having roughly the same physical scale, and level of geometric detail.
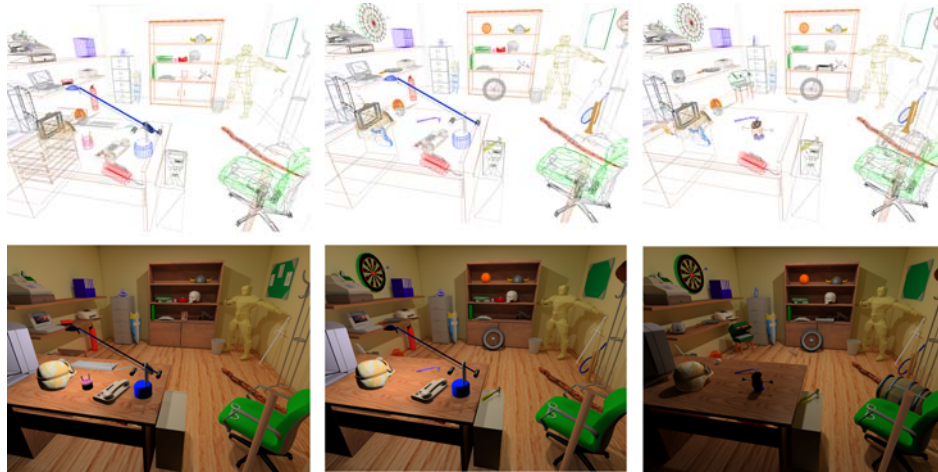

Fig. 2: 35/15 wireframe environment (top-left), 35/15 radiosity environment (bottom-left), 25/25 wire-frame environment (top-middle), 25/25 radiosity environment (bottom-middle), 15/35 wireframe environment (top-right), 15/35 radiosity environment (bottom-right)

5.1.2 Participants. 120 participants were recruited, as discussed in section 2.1.1. 50% were male and 50% female. Age ranged mainly from 18-25. The 120 participants formed 6 balanced groups of 20 corresponding to the six viewing conditions (between subjects).

5.1.3 Stimuli and Apparatus. The technical apparatus has been described under General Methods (cf. Section 3.1.2). The VEs were presented in stereo at VGA resolution on a Kaiser Electro-optics Pro-View XL50 HMD with a diagonal field of view (FOV) of 50 degrees. The frame rate was constant across conditions at 20 frames per second (fps).

The level of luminance of the scene was constant across the radiosity conditions. Wireframe can be defined as a display type that shows the geometric object made up of its edges and drawn as lines resembling a model made of wire. All edges or contour lines are displayed without removing invisible or hidden parts or filling surfaces. In wireframe rendering every section of the mesh of each object was coloured in with the average colour of its material or texture. This comes in useful for quick previews. As in

experiment 1. Colour was introduced simply to facilitate identification of the objects. Hidden line removal was not performed and objects appeared transparent.

In all cases, a single ceiling mounted light source was used. The basic model construct was identical (i.e. ten room frame objects: four walls, floor, ceiling, light, doorknob, door, light switch) and common objects such as the desk and general room layout remained unchanged. All objects were visible and recognizable in all conditions as verified by pilot studies. Two groups of 4 participants were initially recruited. The first group viewed the wireframe scenes whereas the second group viewed the radiosity scenes. Each participant explored the scenes displayed on an HMD naming the objects identified. Participants were asked to verbally indicate whether they were certain or whether they were guessing. Objects that could not be identified were subsequently modified and more groups were invited to complete the same task. This task was repeated, recruiting new groups of participants until all objects were identifiable.

The 100 objects included in the memory recognition questionnaire fell under 4 distinct categories: schema-consistent objects present in the scene (e.g. computer screen, desk, bookcase, books, folders, swivel chair); schema-inconsistent objects present in the scene (e.g. android, skull, snake, spade, shield, cash-register, dartboard); schema-consistent objects absent from the scene but included in the memory recognition questionnaire (e.g. clipboard, computer speakers); schema-inconsistent objects absent but included in the memory recognition questionnaire (e.g. bathtub, dog, faucet/tap, gas mask, donkey).

A preliminary list of objects was assembled based on an initial pilot study's results which explored which objects were expected to be found in a office and which were not. A list of 130 candidate objects was initially constructed. This initial list was modified taking into account similar lists used in real-world studies [Brewer and Treyens 1981; Saab et al. 1984; Pezdek et al. 1989] or in VEs [Flannery and Walles 2003; Mania et al. 2005]. The schema expectancy and saliency of each object was rated on a scale from 1 (low probability) to 6 (high probability) by 10 participants. The saliency rating indicated how noticeable an object was in the context of the room. Results of the schema expectancy rating identified which objects were strongly schema-consistent (rating greater than or equal to 4.5), schema-inconsistent (rating less than or equal to 2.5), or schema-irrelevant, i.e. neutral objects that could be either consistent or inconsistent (rating between 2.5 and 4.5).

Out of the 100 objects remaining, 50 were consistent and 50 were inconsistent. Each scene was divided in four distinct hypothetical 'object regions' comprising of the front, left, right and back wall. 13 objects were placed by the front and back wall regions and 12 objects were placed by the remaining two. Schema consistent and schema inconsistent objects were evenly spatially allocated by each wall region. For example, in the '35/15wireframe' condition, 9 schema consistent objects were placed in each of the front, left and back wall regions and 8 by the right wall region. Additionally, 4 schema inconsistent objects were placed by the front-, back- and right- wall region and 3 by the left-wall region. Similarly sized objects were placed in the same target locations. There were no objects placed in close proximity, i.e. right below or above the participant's viewpoint. The position of common objects was kept the same across all conditions. These objects were also spatially balanced. Pilot studies confirmed that participants could identify and name correctly all objects included in the scene when rendered in wireframe and modified the content of our scenes accordingly. Moreover, we provided a number of alternative names for certain words included on the list, especially when two different words indicate the same object e.g. faucet and tap. These tests were subject paced.

5.1.4 Procedure. As detailed in Section 3, participants completed a memory questionnaire after 55 seconds of exposure to either the low or high poly condition. The

Table 4. Exp 2 – Uncorrected confidence results

N=120	Rendering Quality	Object Ratio (consistent-inconsistent)	Cons. *d'* (O = 15)	Inc. *d'* (O = 15)	Cons. *β* (O = 15)	Inc. *β* (O = 15)
	Wireframe	35/15	1.22 (.58)	1.39 (.49)	.44 (.74)	.78 (.67)
	Radiosity	35/15	1.12 (.43)	1.85 (.54)	.46 (.66)	.69 (.73)
	Wireframe	25/25	1.17 (.35)	1.31 (.58)	.34 (.57)	.66 (.72)
	Radiosity	25/25	1.16 (.61)	1.72 (.63)	.53 (.61)	.74 (.77)
	Wireframe	15/35	1.10 (.57)	1.62 (.39)	.45 (.73)	.88 (.72)
	Radiosity	15/35	1.07 (.73)	1.41 (.63)	.41 (.80)	.66 (.69)
Total			1.14 (.55)	1.55 (.57)	.44 (.67)	.74 (.71)

time of exposure was defined after pilot studies allowed participants to view the scene for 35, 45 or 55 seconds.

5.2 Results and Discussion

As in the previous experiment reported in this paper, The hit and false alarm scores were again used to calculate d' and β. In addition, the confidence scores allowed correcting for possible guessing as in Section 4.2. All resulting d' and β scores were then analysed using a 2 x 3 x 2 mixed ANOVA.

5.2.1 Uncorrected Confidence Results. The uncorrected for guessing results shown in Table 4 revealed a range of interesting findings.

The ANOVA of the d' scores revealed no significant main effect of rendering quality ($F (1,114) = 1.07$, $p = .302$), no main effect of object ratio ($F (2,114) = .418$, $p = .660$), but did find a significant main effect of object type ($F (1,114) = 45.5$, $p < .001$). This indicates that, as in Experiment 1, inconsistent objects had a much greater d' than consistent objects and are thus allowing better recognition. However, this result is worthy of note since it would be expected that memory performance in an uncorrected-for-guessing recognition test would not shift either the consistent or inconsistent objects [Rojahn and Pettigrew 1992]. On the other hand, there were studies, as reported in the literature, which presented better memory performance for the inconsistent objects even when the recognition tests were uncorrected for guessing [Pezdek et al. 1989]. In addition, there was a significant two-way interaction between rendering quality and object type ($F (1,114) = 4.81$, $p = .030$), as well as a significant three-way interaction between rendering quality, object ratio and object type ($F (2,114) = 3.52$, $p = .033$).

The significant interactions were explored further with the use of Tukey's Honestly Significant Difference (HSD) post-hoc testing. The rendering quality-object type interaction revealed that there was a significant difference between consistent and inconsistent object type in the wireframe quality condition (1.17 vs. 1.44, HSD = .265, p = .05), as well as the inconsistent objects being significantly higher in the radiosity condition (1.11 vs. 1.66, HSD = .323, p = .01), as shown in Fig. 3. This would indicate that when radiosity rendering is used, there is a greater difference between the consistent and inconsistent object types than when wireframe rendering is used, with inconsistent

having significantly greater d' than consistent in radiosity than in wireframe. This can thus be seen to qualify the main effect of object type described above. Therefore, when VEs are rendered in radiosity, participants remember the inconsistent objects significantly better than the consistent. The poor quality in wireframe forces participants to spend more time encoding all objects irrespective of their 'object type' value since they might need to devote more effort in identifying each object.
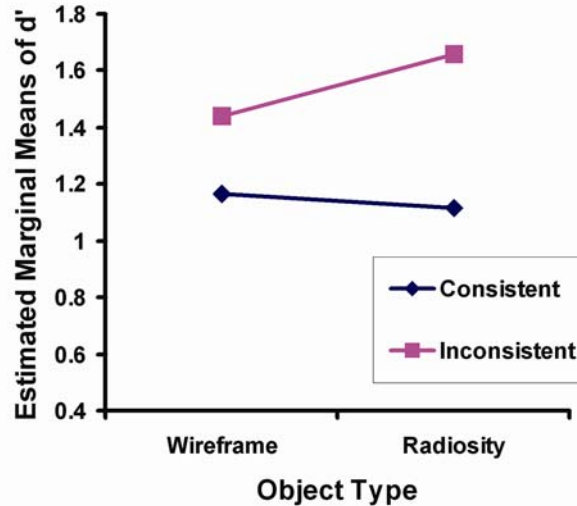

Fig. 3. ANOVA Interaction between Rendering Quality and Object Type

The three-way interaction between all variables was also explored with Tukey's HSD, by first analysing the potential two-way interactions that make up the three-way. The rendering quality-object ratio interaction produced no significant differences, whilst the object type-rendering quality interaction is described above. The object ratio-object type interaction showed that there was a significant difference between consistent and inconsistent objects at the 35/15 ratio level (1.17 vs. 1.62, HSD = .428, p = .01), between consistent and inconsistent at the 15/35 ratio (1.09 vs. 1.52), but no significant difference between object type at the 25/25 ratio (1.17 vs. 1.51).

It can therefore be posited that it is those two two-way interactions in particular that are mediating the three-way interaction, as shown in Fig. 4. This suggests that the three-way interaction arises thus: in the wireframe 35/15 condition, there is no significant difference between consistent and inconsistent object type (1.22 vs. 1.39, HSD = .575, p = .05), whilst when considering the radiosity 35/15 condition, there is a significant difference between consistent and inconsistent (1.12 vs. 1.85). Moreover, when considering the wireframe 15/35 condition, this pairing is approaching significance (1.10 vs. 1.62), whilst the corresponding radiosity 15/35 comparison is not significantly different between object type (1.07 vs. 1.41). The overall indication is that in the wireframe condition, greater ratio of consistent items leads to no significant difference between consistent and inconsistent object types, whilst a greater ratio of inconsistent objects in the scene leads to greater recognition of inconsistent over consistent objects. In contrast, in the radiosity condition when there is a greater level of consistent items in the object ratio there is significantly greater recognition of inconsistent items, and when the ratio is reversed this significant difference is lost. These results are worthy of a further discussion and contrast with existing literature. In Rojahn & Pettigrew [1992], it was stated that memory performance should shift towards the inconsistent objects the smaller the ratio of the inconsistent objects in the scene. The results above partially agree with this statement and quality of graphics seems to have significant impact on the effect of

the ratio of consistent/inconsistent objects on memory performance. Whilst the radiosity condition results correlate with this statement, the same does not occur for the wireframe condition.
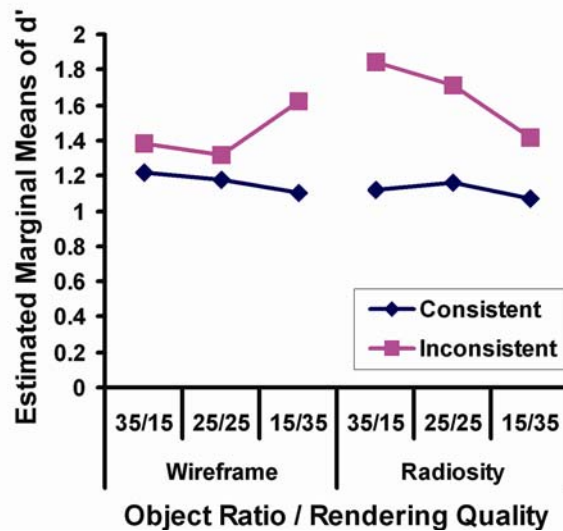

Fig. 4. ANOVA Interaction between Object Ratio, Rendering Quality and Object Type (ratios representing consistent/inconsistent objects in each scene)

The radiosity condition could be perceived as a real life office. When there are only a few inconsistent objects in the room they are noticeable because they stand out as being inconsistent. As the ratio of inconsistent to consistent objects increases, the room is no longer an office of inconsistent and consistent objects: it becomes an office full of objects, which are all more or less inconsistent. In the wireframe condition, as we do not perceive the world in wireframe, each room itself becomes inconsistent. We just look at the objects and remember them regardless of whether they are consistent or not as we are not looking at the room in that way. As the ratio becomes larger including more inconsistent objects we remember them as interesting objects.

The β results were also analysed using a 2 x 3 x 2 mixed ANOVA. This revealed no significant main effect of rendering quality (F $(1,114)$ = .009, p = .923), no main effect of object ratio (F $(2,114)$ = .041, p = .960), but did locate a main effect of object type (F $(1,114)$ = 15.8, p < .001). No interactions were found between the variables. This main effect of object type indicates that, as in Experiment 1, participants were significantly more biased towards saying 'no' for the inconsistent objects than the consistent. Therefore, participants' better recognition for the inconsistent objects, revealed by the main effect of d, can be explained due to their increased bias towards rejecting an inconsistent object as being present. For example, participants that might not have been certain about the presence of an object, they would more easily consider an inconsistent object as not being present rather than a consistent one.

5.2.2 Moderate Confidence Results (4-5 confidence). As shown in Table 5, when the results were corrected for guessing the subsequent signal detection values changed to quite a degree.

The Experiment 2 results for d' 4-5 corrected revealed a main effect of object type (F $(1,114)$ = 16.3, p < .001), but no effect of object ratio (F $(2,114)$ = 1.44, p = .241) nor quality (F $(1,114)$ = 1.91, p = .170), revealing the inconsistent objects (d' = 1.51) to be more accurately identified than accurate (1.25). No interactions were found between conditions.

Table 5. Exp 2 – Moderate confidence results

N=120	Rendering Quality	Object Ratio (consistent/ inconsistent)	Cons. d' (O = 15)	Inc. d' (O = 15)	Cons. β (O = 15)	Inc. β (O = 15)
	Wireframe	35/15	1.25(.46)	1.45(.47)	1.09(.55)	1.24(.50)
	Radiosity	35/15	1.23(.56)	1.79(.58)	1.03(.55)	1.17(.50)
	Wireframe	25/25	1.21(.35)	1.33(.60)	1.05(.52)	1.09(.52)
	Radiosity	25/25	1.54(.40)	1.57(.65)	1.35(.44)	1.13(.73)
	Wireframe	15/35	1.11(.75)	1.57(.37)	0.93(.73)	1.29(.57)
	Radiosity	15/35	1.13(.52)	1.32(.63)	1.09(.54)	1.1(.59)
Total			1.25(.53)	1.50(.57)	1.09(.57)	1.17(.57)

d' and β scores and standard deviations (in brackets) as a function of viewing condition (N Total Number of Participants, O Total Number of Objects).
Cons = Consistent, Inc. = Inconsistent.

The β results revealed no main effect of object type (F (1,114) = 1.63, p = .204), nor object ratio (F (2,114) = .136, p = .873), nor quality (F (1,114) = .135, p = .714), as well as no interactions. This suggests that bias was not significantly altered by any of the conditions when corrected to this level.

5.2.3 High Confidence Results (5 confidence level). The corrected results, only including confidence scores of '5' are shown in Table 6.

Table 6. Exp 2 – High confidence results

N=120	Rendering Quality	Object Ratio	Cons. d' (O = 15)	Inc. d' (O = 15)	Cons. β (O = 15)	Inc. β (O = 15)
	Wireframe	35/15	.92 (.45)	1.28 (.44)	1.05 (.49)	1.26 (.43)
	Radiosity	35/15	1.09 (.45)	1.58 (.65)	1.14 (.50)	1.29 (.49)
	Wireframe	25/25	.96 (.53)	1.03 (.69)	1.09 (.45)	1.06 (.69)
	Radiosity	25/25	1.31 (.33)	1.53 (.59)	1.46 (.24)	1.26 (.56)
	Wireframe	15/35	.89 (.47)	1.29 (.38)	1.03 (.50)	1.28 (.47)
	Radiosity	15/35	.96 (.51)	1.22 (.60)	1.07 (.52)	1.18 (.47)
Total			1.02 (.47)	1.32 (.59)	1.14 (.47)	1.22 (.52)

d' and β scores and standard deviations (in brackets) as a function of viewing condition (N Total Number of Participants, O Total Number of Objects).
Cons = Consistent, Inc. = Inconsistent.

ANOVA testing of the d' results revealed a main effect of rendering quality (F (1,114) = 7.54, p = .007), no main effect of object ratio (F (2,114) = 1.03, p = .360), and a main effect of object type (F (1,114) = 35.1, p < .001). This indicates that high levels of correction lead to an additional significant result, with radiosity rendering achieving a significantly greater difference between signal and noise distributions in recognition scores than the wireframe rendering quality. Additionally, as in Sections 5.2.1 and 5.2.2, inconsistent type objects have a significantly greater d' than consistent objects, showing that this holds no matter the level of correction.

There was also an interaction that was not significant but close to significance at the .05 level, between object ratio and object type (F (2,114) =2.81, p = .064). As in Section 5.2.2, this was explored descriptively in Fig. 5. This figure suggests that the interaction occurred in the inconsistent results being much greater than the consistent in the 35/15 and the 15/35 conditions, but not in the 25/25 condition, indicating that the mid-ratio condition did not lead to as great differences in object types as the other ratio conditions.
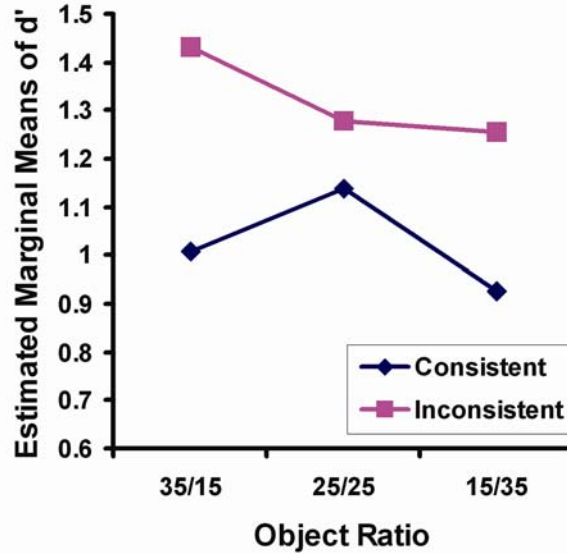

Fig. 5. ANOVA Interaction between Object Ratio and Object Type (ratios representing consistent/inconsistent objects in each scene)

The β results also revealed a different pattern of results to prior analyses, with ANOVA revealing no main effect of rendering quality (F (1,114) = 1.86, p = .176), nor object ratio (F (2,114) = .330, p = .719), nor object type (F (1,114) = 2.94, p = .089). This seems to suggest that correcting so severely for guessing is again leading to somewhat skewed results, as in Experiment 1. There was however an interaction between object ratio and object type (F (2,114) = 4.14, p = .018).

Tukey HSD testing revealed no significant differences between the conditions of the interaction, although this is explained by the nature of the interaction, as shown in Fig. 6. The interaction simply means that inconsistent object type bias is greater in the 35/15 and 15/35 ratios, whilst the consistent bias is greater in the 25/25 ratio. The Tukey test is unable to dissect this interaction, as it is too conservative to pick up the small but regular variations shown in the results and in Fig. 6.
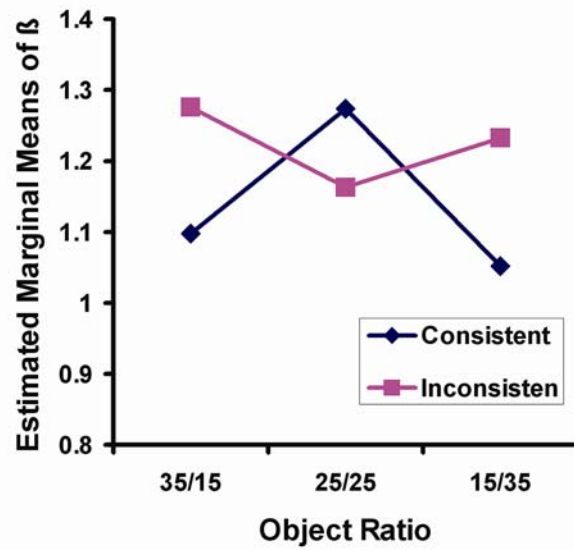
Fig. 6. ANOVA Interaction between Object Ratio and Object Type (ratios representing consistent/inconsistent objects in each scene)

6. GENERAL DISCUSSION

Memory research has established that schema-based information is drawn in the process of retrieving information from memory. The studies presented in this paper investigated the effect of visual quality on information uptake by participants when exposed to immersive VEs. The scope of this work has been to identify regions of a scene that could be rendered in lower quality without affecting spatial information uptake. Recognition of objects of varying quality could be affected according to whether they 'fitted' the context of the scene. Participants' memory performance was assessed after exposure to scenes of varying rendering quality through recognition of present objects in a scene as well as inferred objects. Two memory schemata moderating variables, those of 'object type: consistent vs. inconsistent' and 'ratio of consistent/inconsistent' objects in a scene were varied in order to validate this hypothesis.

6.1   The effect of 'object type'

Findings of both experiments reported clearly supported the inconsistency effect. Inconsistent objects were more correctly recognized as either being present or being absent than the consistent objects. There was a slight variation of results, in terms of levels of confidence between the two studies. It was expected that memory performance would not be significantly different for either consistent or inconsistent objects when the recognition scores were uncorrected for guessing [Rojahn and Pettigrew 1992]. That was evident, however, only in the first experiment, whereas in the second experiment memory performance was better for the inconsistent objects at all three confidence levels. Confidence scores revealed significantly higher confidence ratings for inconsistent objects compared to the consistent ones. Coupled with verbal comments by participants who stated that they had paid more attention to the inconsistent objects, this implies that some element of guesswork had taken place regarding the consistent objects during the first experiment. Moreover, participants stated that they had assumed that the consistent objects were present, whereas they had specifically observed the inconsistent items.

Similar effects have been reported in memory literature, notably schema activating occurring while assigning hands to a clock after exposure to a scene although no hands were present for instance [Kuipers, 1975]. However, taking into account relevant β results, there was no bias towards saying 'yes' for the consistent objects.

6.2 The effect of the ratio of consistent versus inconsistent objects

According to existing research in real-world scenes, variations in the ratio of consistent/inconsistent objects in a scene should produce an effect in memory performance. In particular, it is reported that memory would shift towards the inconsistent effect, the smaller the ratio of those objects in the scene [Rojahn and Pettigrew 1992]. Experiment 2 investigated the effect of this moderating variable in immersive simulations. There was no main effect of this ratio irrespective of confidence level. Despite that, an interaction was revealed between object ratio, viewing condition and object type. The overall indication is that in the wireframe condition, a greater ratio of inconsistent objects in the scene leads to greater recognition of inconsistent over consistent objects. In contrast, in the radiosity condition, there is significantly greater recognition of the inconsistent items when there is a lower proportion of inconsistent items. The latter is in line with the hypothesis mentioned above. However, the former is opposite to this hypothesis. This controversy could probably be explained taking into account the quality variation coupled with comments by participants who stated that they had found it challenging to identify the objects present in the wireframe condition. Recognition while exposed to the wireframe scenes was probably based on the episodic assimilated information rather than in an activated office schema.

6.3 The effect of quality of graphics

A main effect of rendering quality on memory recognition was revealed when the recognition results were reported with high confidence (in Experiment 2). In particular it was shown that memory performance is better for the inconsistent objects in the radiosity rendering condition compared to the wireframe condition. Consistent objects could, therefore, be rendered at lower quality, whereas inconsistent objects require higher levels of graphics quality. Selective rendering algorithms require decisions to be taken in terms of which areas of a scene could be rendered in high quality and which ones in lower quality. In past research, such decisions were based on eye gaze (requiring eye tracking equipment) or rendering in high quality specified task areas. Based on these results, this work theoretically formalizes the following eye gaze and task independent idea which is a leap forward in this area: in any space, we could utilize the lowest quality possible in order to render objects which are consistent to each context without any difference on memory recognition in relation to the fully fledged rendering solution. Therefore, in a bathroom, the basin or bathtub could be rendered in lower quality since they are 'expected' as part of every bathroom. Therefore, neither object ratio nor rendering quality make a difference for consistent objects – these objects are "recognized" at the same level regardless. In contrast, for inconsistent objects, both factors make a difference. When the inconsistent objects are rendered in high quality, recognition is best when there are only a few inconsistent objects, and falls as the number of inconsistent objects increases. The opposite is true when the inconsistent objects are rendered as wire frames: recognition is relatively poor when there is a small number of inconsistent objects but increases as there are more inconsistent objects. The practical use of this approach remains to be tested when such a system materializes; however, its theoretical underpinnings have immediate applicability towards a complete selective rendering system. Moreover, it advances significantly past approaches, endeavouring simulation of a cognitive process rather than simulation of physics.

Both reported studies demonstrated that memory schemata aid memory retrieval in VEs in a way similar to existing real-world literature. In both experiments memory performance shifted towards the inconsistency effect. Varying quality of computer graphics provoked an effect on memory recognition when the perceptual distance between the scenes was large (Experiment 2) and the recognition data were of high confidence. In this case, it seems that it would be possible to render in lower quality the consistent information in a VE and still achieve the same information uptake. The findings of this study, besides the particular case discussed above, suggest that simulations can survive degraded visual fidelity in the form of polygon count and rendering method, without clear loss of memory performance. When interactive immersive VEs are used for training rather than entertainment purposes visual quality is not as significant.

High level visual cognition is that which takes place late in the HVS, that is parietal and temporal cortex and into the frontal lobes when decisions based on visual information need to be made. Low level visual cognition is that which takes place in the occipital lobe, early on in the visual processing 'stream', e.g. the visual signal is received in the retinae, and initially passed through the Lateral Geniculate Nucleus to the occipitial lobe at the back. Thus, these higher decision processes are unaffected by changes in the experiments, whilst normal (visual) cognition occurs as long as the scene is realistic. Taken together, the results suggest that "high-level" visual cognition is generally unaffected by ubiquitous graphics manipulations such as polygon count and depth of shadow rendering; "normal" cognition operates as long as the scenes look acceptably realistic. On the other hand, when the overall realism of the scene is greatly reduced, such as in the wireframe condition, then visual cognition becomes abnormal. Specifically, effects that distinguish "schema-consistent" from "schema-inconsistent" objects change because the whole scene now looks incongruent, and we have shown that this effect is not due to a failure of basic recognition. Thus, a recipe for anyone wishing to use such displays in studies of visual cognition is to construct environments which look acceptably realistic in terms of polygon count but need not be of very high quality. This relates to the kinds of high-level visual cognitive effects we have studied here, such as object congruence. Lower-level effects, such as recognition, can be dissociated from these high-level effects (but make their presence felt when the scene is further degraded, e.g. when colour is removed from the wireframe scenes). Thus, "high-level" processes need somewhat greater realism than "low-level" ones.

ACKNOWLEDGMENTS

REFERENCES

BARTLETT, F.C. Remembering. Cambridge University Press, Cambridge, England, 1932, p-213
BREWER, W.F. TREYENS, J.C. 1981 Role of schemata in memory for places, Cognitive Psychology, 13, 207-230
CATER, K. CHALMERS, A. WARD, G. 2003 Detail to attention: exploiting visual tasks for selective rendering, in Eurographics Workshop on Rendering, 25 - 27 June 2003, 270–280
FERWERDA, J. 2001 Hi-fi rendering, Campfire in Perceptually Adaptive Graphics, 26 – 29 May 2001, http://isg.cs.tcd.ie/campfire/jimferwerda2.html
FERWERDA, J. 2003 Three varieties of realism in computer graphics, Proc. SPIE Human Vision and Electronic Imaging '03, 20 -24 January 2003, 290-297
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 automatized encoding and memory for gist, Journal of Experimental Psychology: General, 108, 3, 316-355

GOODMAN G. S. 1980 Picture memory: how the action schema affects retention, Cognitive Psychology, 12, 473-495

HABER, J. MYSZKOWSKI, K. YAMAUCHI, H. SEIDEL, H.P. 2001 Perceptually guided corrective splatting, Computer Graphics Forum, 20, 3, 142-152

HOCK, H.S. ROMANSKI, L. GALIE, A. WILLIAMS, C.S. 1978 Real-world schemata and scene recognition in adults and children, Memory & Cognition, 6, 4, 423-431

KUIPERS, B.J. 1975 A frame for frames: representing knowledge for recognition, in representation and understanding: studies in cognitive science, BOBROW, D.G. and COLLINS, A. (Eds.), New York: Academic Press, 1975

LAND, M. F. 1999 Motion and vision: why animals move their eyes, Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioural Physiology, 185, 4, 341-352

LIU, G. AUSTEN, E.L. BOOTH, K.S. FISHER, B.D. REMPEL, M.I. ENNS, J.T. 2005 Multiple object tracking is based on scene, not retinal, coordinates, Journal of Experimental Psychology: Human Perception and Performance, 31, 2, Apr. 2005, 235-247

LOFTUS, G. R. MACKWORTH, N. H. 1978 Cognitive determinants of fixation location during picture viewing, Journal of Educational Psychology: Human Perception and Performance, 4, 4, 565-572

MANIA, K. ADELSTEIN, B. ELLIS, S.R. HILL, M. 2004 Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity, in Proceedings of the 1st Symposium on Applied perception in graphics and visualization 7 - 8August 2004, 39-47

MANIA, K. ROBINSON, A. BRANDT, K. 2005 The effect of memory schemata on object recognition in virtual environments, Presence, Teleoperators and Virtual Environments, 14, 5, 606-615

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

MARMITT, G. DUCHOWSKI, A.T. 2002 Modeling visual attention in VR: measuring the accuracy of predicted scanpaths, in Eurographics 2002, Short Presentations, 2- 6 September 2002, 217-226

MCCONCIE, G. W. LOSCHY, L.C. 1997 Human performance with a gaze linked multi-resolutional display, in. First Advanced Displays and Interactive Displays Annual Symposium, 25-34.

NEMIRE, K. JACOBY, R.H. ELLIS, S.R. 1994 Simulation fidelity of a virtual environment display, Human Factors, 36, 1, 79-93.

PEZDEK, K. WHETSTONE, T. REYNOLDS, K. ASKARI, N. DOUGHERTY, T. 1989, Memory for real-world scenes: the role of consistency with schema expectation, Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 4, 587-595

RODGER, J.C. BROWSE R A. 2000 Choosing rendering parameters for effective communication of 3D shape, IEEE Computer Graphics and Applications, 20, 2, 20-28

ROJAHN, K. PETTIGREW, T. F. 1992 Memory for schema-relevant information: a meta-analytic resolution, British Journal of Social Psychology, 31, 2, 81–109

SAAB, Z. TROTTIER, J. WALL, H.M. 1984 Memory for places: role of schemata-related expectancy and saliency in recall and recognition, Psychological reports, 54, 1984, 607-610

WAGNER, L. 1987 The effects of shadow quality on the perception of spatial relationships in computer generated imagery, in Proc. of Symposium on Interactive 3D Graphics, 1987, 39-42.

WALLER, D. HUNT, E. KNAPP, D. 1998 The transfer of spatial knowledge in virtual environment training, Presence: Teleoperators and Virtual Environments, 7, 2, 129 – 143

WATSON, B. FRIEDMAN, A. MCGAFFEY, A. 2001 Measuring and predicting visual fidelity," proc. twenty height annual conference on computer graphics and interactive techniques, 2001, 213 – 220.

WATSON, B.A. WALKER, N. HODGES, L.F. REDDY, M. 1997 An evaluation of level of detail degradation in head-mounted display peripheries, Presence: Teleoperators and Virtual Environments, 6, 6, 630–637

WICKENS, T.D. Elementary Signal Detection Theory. Oxford University Press, New York, 2001