Walking through virtual doors: A study on the effects of virtual location changes on memory

P. Watson\textsuperscript{1} and S. E. Gaudl\textsuperscript{1}

\textsuperscript{1}University of Plymouth, UK

Abstract

The spaces we inhabit can shape and influence the way in which we learn or reinforce information. Virtual reality (VR) is a technology that allows us to alter and create designed environments with great freedom over the visual, audio, and narrative elements. This freedom would benefit from further guidelines that detail approaches and implementations to best achieve desired information delivery goals. In this paper, we present findings of a study that applies location-based memory strategies to VR environments, with the aim to aid word list recall without the subjects being required to apply any memory strategy themselves. Our findings suggest that VR may interfere with the incidental processing of multiple rooms and potential aid to recall as demonstrated in real-world studies [PTT\textsuperscript{16}] [Smi82].

CCS Concepts

- Human-centered computing → Virtual reality; • Applied computing → Interactive learning environments;

1. Introduction

A key design consideration for Virtual Reality (VR) experiences is how to pass on information to the user. In some cases, this will be an important requirement of a VR tool. To illustrate, in education and training scenarios, users may need to retain a certain amount of information before they apply it to a given scenario. For example, recalling the action sequences and safety information before applying first aid. The incidental processing of a physical environment has been linked to key processes of memory and learning [CVMP14], [Bro11]. How we separate information between locations [PTT\textsuperscript{16}], and how we travel between these areas can have a significant impact on what is recalled [SV01]. This suggests that memory recall can be improved without the application of memory strategies by the user. VR is an ideal platform to apply and test these concepts as users are able to traverse infinite locations without needing to leave their real-world rooms. This study explores...
if separating information between immersive virtual rooms can aid recall, as has been observed in real world trials [PTT*16].

2. Background

Event cognition research emphasises the role of location in recalling a day’s activities. To sort through the daily stream of experiences, information is often categorised into events. An event can be simply described as an amount of time, within a location that is perceived to have a beginning and an end [ZT01]. The mental representation and simulation of these events are called event models.

An important construct of an event model is the spatial location and time of day that an event takes place [RZ11]. To recall the events and associated experiences, one may first recall the spatial location of an event model. Event models therefore act as a structure to recall both the chronology and content of information that an individual is exposed to. Although the boundary of where one event model ends and another starts can be established by any significant focus of attention within a location, prominent changes in the environment are thought to define the placement of event boundaries, for example walking through doorways.

Walking through doorways has shown to both aid and interfere with recall [RZ17]. When information is shared between rooms (and therefore event models), both event models are activated on recollection, compete and cause retrieval interference [PR16]. This has been observed in experiments where participants are asked to recall or recognise objects they are carrying or have just carried between rooms compared to within a single room [PR18]. Conversely, if information remains unique between rooms, then walking through doorways can aid recall. This has been observed that presenting wordlists between two rooms compared to one (but same walking distance) [PTT*16]. Much like breaking down one large list into two smaller ones, the event models act as memory hooks to smaller sets of information which can be brought together to make a larger body of knowledge. Similarly, to event cognition work, environmental context dependent memory research has also found evidence of improvements to recall of word lists when encoding occurs between multiple room locations (multiple contexts) [SGB78], [Smi82]. Showing more changes in location (therefore separating a body of information between more locations) can aid memory recall and recognition displayed in a variety of media. For example: video content [GZF17], imagined narratives through reading, and recalling a word list from a desktop screen [PTT*16]. The recall enhancement appears to hold over longer durations, with improved information recall at one week [FBEZ17] and two weeks [RAF17]. Event frameworks will be robustly remembered after a month but information within will have significantly reduced [FBEZ17]. This body of work suggests that recall can be improved by segmenting information between locations whether within the physical world, imagined narratives or on media platforms.

Work that has observed an enhancement to recall by segmenting word lists between rooms has used real world methodologies. Smith [Smi82], had participants learn four word lists in either a single room, between 2 rooms, or between four rooms. Participants would wait in a hallway between location changes. Recall significantly increased as room numbers increased. More recently, Pettijohn [PTT*16], exposed participants to two word lists in either a single room or two room condition. In each case, participants would walk towards the designated room location and have the first word list read to them. Subjects would subsequently walk towards the next location for the second word list. Using multiple rooms (subjects walking through a doorway between word lists) significantly increased recall.

Virtual reality is a promising platform to test & apply theories of event cognition on information recall. Firstly, it affords spatial understanding of a virtual location and a psychological sense of “being there” and thus creates the illusion that the user is in a new reality [Sla14]. Head movement is tracked in 3D space enabling a more natural scanning process of spatial cues, compared to desktop displays. Allocentric (objects in relation to each other) spaces may also benefit from more immersive displays. When cave experts were asked a series of questions based on 3D virtual map of a cave, tasks that involved finding map details and comparative measurements were answered more quickly and with greater accuracy using higher immersive displays (larger field of view, stereoscopic rendering turned on) [SB07]. Therefore, an immersive display can aid the egocentric mapping of an environment, which will also help build the allocentric representation and in turn a cognitive map of the environment [EPJS17]. For example, users had better spatial perception of a virtual architectural models when using an HMD [PAI17] and increased recall of spatially placed information [KPV19]. Secondly, content delivered on VR is not restricted by time and space, and therefore can potentially use the principles of location-based event cognition to design experiences that facilitate recall. For example, segmenting a body of information between X number of rooms.

The incidental processing of a physical environment has been linked to key processes of memory, including the episodic structuring of events over time, and information retrieval. In such examples, the environment acts as a mental structure that is associated to information experienced within. By separating information between multiple locations or local areas, more information may be recalled compared to using a single location. If the separation of information between environments aids recall in the real world, then this concept could be a guiding blueprint to the design of better information recall within VR environments. To our knowledge, no articles can be found that explore the use of virtual reality on improving memory recall by physically walking through virtual doorways into room spaces.

H1: Within an immersive virtual environment, separating wordlists between two rooms divided by a doorway will aid recall of word lists compared to a single room.
H0: Within an immersive virtual environment, separating wordlists between two rooms divided by a doorway will not aid the recall of word lists compared to a single room.
3. Methodology

This study received approval from the Research Ethics & Integrity Committee, University of Plymouth. For this experimental study, 29 (13 female) participants were recruited. The participants came from two distinct backgrounds, computer science and psychology undergraduate students. Participants were rewarded with credits towards class systems or free food. Participants took part in this experiment one at a time.

This study employs and builds on an existing methodology [PTT’16] from event cognition work and was adapted to the VR platform. The same within-subjects design was used whereby participants would listen to two word lists in one of two conditions. Either a single room condition (1R), with both word lists in the same room, or in a two-room condition (2R), with one word lists in each room, separated with a doorway (Figure 2). After listening to both word lists in a condition, participants would take off their VR headset and be presented with a distraction task (a set of maths questions for 1 minute) and then asked to recall and write on a piece of paper as many words from the previous condition as they could. Once complete, participants would repeat this process for the alternate condition. Starting position (and therefore wordlist), and condition were counterbalanced across the trials. The recall of these two conditions were then compared.

For our study the procedure was as follows: After signing a consent form, participants completed a demographic survey, and were given a briefing on the upcoming task to explain that they would be listening to speakers and shown images of what the speakers looked like. These steps took no more than 5 minutes. Participants were then fitted with the HTC Vive HMD (Head Mounted Display) [Viv21], to provide the visual feed of the virtual environment. The audio was delivered through headphones connected to the HMD. Participants started in the SteamVR Home scene [Ste17] to check proper fitting and visual clarity of the HMD. After the initial hardware check, the first experimental condition was loaded. Participants were asked to take a few moments (10 seconds) upon entering the virtual world to look around and orientate themselves. After orientation, the participants were instructed by the experimenter to approach the speakers and remember the delivered words best that they could. Orientation and listening of both word lists took 2-3 minutes per condition. When finished listening to both speakers the participants were asked to remove the Vive HMD. Participants were then given a set of maths questions on a sheet of paper to work through for one minute, to encourage forgetting and reduce the use of short term memory strategies like repetition (Word list 1) when subjects approached them. A difference in methodology used in this study was that the distraction task and the free recall occurred in the real world and not the virtual, acting as a control for any context reinstatement effects. 2) it allowed a more direct replication of previous distraction and recall task.

Four sets of 10 words were required for this study, 40 words in total. In line with previous experiments [PTT’16], the MRC Psycholinguistic Database was used [Psy97] to generate the words. Each word was one syllable, five letters long, and range in word frequency from 20 to 103 per million. 226 words were generated and then 40 of these randomly selected for this study. A subjective check was made to make sure that the words within each list were not too similar or held a strong implicit association (Word list used can be found here [Wat21]). Two virtual environments were developed within the Epic Unreal Engine 4 [Eag04] (Figure 1) for the conditions of this study: A single room and a two-room environment. The overall virtual area used for each condition was the same, and the distance between the speakers were 2m for both conditions (Figure 2). However, the aesthetics used within each room were different. For the 2R condition, a dividing wall and doorway connected the rooms that opened automatically once the first word list had been listened to. In both conditions, there were two speakers that when approached by the participant, voiced ten randomly selected words with a delay of one second between each. Subjects were able to physically walk within the virtual world between word lists and through rooms to reflect the same navigation method used in previous work.

To adapt this methodology to the VR platform, a human male was recorded speaking the word lists. These voice files were delivered to subjects through virtual speakers. Speakers would automatically deliver word lists (A one second gap was used between each word) when subjects approached them. A difference in methodology used in this study was that the distraction task and the free recall occurred in the real world and not the virtual, acting as a third room for recall. This approach was chosen as 1) previous work ( [Smi82] - experiment 3) suggest that recalling a word list in the same room as the encoding, will elicit a context reinstatement effect, whereby the environmental cues exposed to during encoding aid recall. However, context reinstatement is significantly reduced when using multiple rooms. Therefore recalling in the same room for both conditions would give a context reinstatement advantage to the single room condition and is not the subject of this study. Using a different context (room) for recall for both conditions will control for any context reinstatement effects. 2) it allowed a more direct replication of previous distraction and recall task.
Table 1: Participant Demographics

<table>
<thead>
<tr>
<th>Demographic Information</th>
<th>Age 18-24</th>
<th>25-34</th>
<th>35-44</th>
<th>45-54</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>21</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td>Female</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used VR before?</td>
<td>Yes</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Post experiment survey

<table>
<thead>
<tr>
<th>Survey Questions</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did the environment with the doorway feel like two separate rooms?</td>
<td>27</td>
<td>2</td>
</tr>
<tr>
<td>Did you use any memory strategies to help remember the words?</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

For a video walkthrough of the virtual conditions: https://youtu.be/M9whJPyB6E0

4. Results

In post experiment questionnaires (Table 2), participants were asked if they used a memory strategy when encoding the word lists. A memory strategy was defined as a technique that would aid long term recall. For example, creating a story from the presented words during encoding. Strategies that aided short term recall were categorised as not using a memory strategy, as the distractor task between encoding and recall should minimise any benefit to recall from such strategies. On analysis, no significant effect of memory strategy use was found.

![Figure 3: Mean recall between the two-room and one-room conditions. The similar value reflects no significant difference between conditions.](https://example.com/figure3)

Of the 29 participants recruited for this study: 15 participants presented increased recall in the 2R condition, 12 participants presented an increase in recall in the 1R condition, and 2 participants saw no change in recall between conditions. Mean recall for the 2R condition 4.59/20 (SD = 1.88). Mean recall for the 1R condition was 4.45/20 (SD = 1.96) (Figure 3). Recall data for speaker 2 and 3 (Figure 2 for speaker layout) were not normally distributed as assessed by Shapiro–Wilk’s test (p = < 0.05). Therefore, a Wilcoxon signed-rank test was run to analyse the degree of recall difference between the 2R and 1R conditions. Statistically, there was no significant median difference in recall between conditions, Z = 0.503, p = .615, Hedges g = 0.08 90% CI [-0.37 – 0.53].

To test for recall differences between individual word lists, a Friedman test was run. Word list recall was similar across word lists 1-3 (Mdn = 2.00) with an increase at word list 4 (Mdn = 3.00). These differences were not statistically significant, 2(3) = 6.536, p = 0.88.

5. Discussion

Following on from real world studies that suggest segmenting information between multiple rooms can aid recall [PTT16], [Smi82], we investigated if the same effect could be observed within immersive virtual environments. Even though participants perceived the virtual two room condition as two different rooms, no significant difference in recall was observed when segmenting information between multiple immersive virtual rooms.

Improving recall of word lists by segmenting the encoding between different locations has been replicated multiple times in real world methodologies. There are instances when replication did not reach significance [Smi84]. However, there was still some difference in recall between using a single room and multiple rooms for encoding. Meta analysis of environmental context work suggest that segmenting information between contexts (rooms) to learn word lists is a reliable effect with a moderate effect size (d = 0.45). This, combined with event cognition work that has observed enhanced recall from segmenting word lists between event models (through narrative, locations and other media), suggests that it is a replicable effect across media platforms for both within and between subject experimental designs.

This study closely followed past methodologies from event cognition work [PTT16], applied to a VR platform. Interestingly, in comparing this study to previous results, not only do we observe a non-significant difference between conditions but average recall across this study was lower. The two room condition observed a 9% (2 words) drop in comparative recall and the single room condition 5% (1 word) drop in comparative recall. Additionally, the effect of memory strategies was also not significant. The use of internal strategies should supersede effects of location on memory and show greater recall independent of how the information is segmented between rooms [Smi85]. This suggests more interference in the encoding or recall process and potentially higher difficulty of applying memory strategies. Our current study may have limited statistical power due to the sample size (N = 29). Previous work [PTT16] observed a moderate effect size d = 0.629 for improving recall by separating word lists between locations. Analysis of power based on this work [Wes15] estimates approximately
55 participants are required to obtain the statistical power of 0.80. Therefore, this current study may not be sensitive enough to observe the same effect. In the presented study we observe a smaller effect size ($d=0.07$). Given that overall recall is lower than previous work, memory strategies did not appear to significantly increase recall, and a smaller effect size observed, these results suggest that an element of the methodology or the VR platform interferes with the use of event models to structure and aid the recall of word lists.

One methodological explanation for no effect of separating information between rooms, is that previous work [PTT16] had subjects recall words in the same room as they listened to the word lists. In this current study, the subjects removed their HMD to perform the distraction and recall tasks. This approach was used to make sure that there was a different environment for recall so no context segmentation would give advantage to the single room condition. This approach should not interfere with using multiple locations to encode word lists. However, some recent work suggests that there may be a more fundamental reduction in recall when recalling words in the real world that were memorised in VR [LL21]. Subjects a set of memorisation tasks at a virtual desk or real world desk. They were then called back 24 hours later to recall and recognise this information in the same or opposite reality. Those that learnt information in the VR context, recalled significantly less in the real world, compared to all other conditions. However, more work is required to establish the extent of this effect in larger groups and different interval times before final recall. It does suggest though that a simple replication of an environment may not engender the same cognitive response as the real world. Based on this work, developing a unique recall environment within VR would minimise the effects of context reinstatement while also avoiding any mediating influence of removing the headset on the final recall.

The navigation and interaction method for this current study were chosen to be as natural as possible to reflect real world methodologies being replicated. Each room was distinct in colour, shape, and décor. The culminations of these attributes was to simulate the belief that subjects were passing into separate rooms and minimise overlap in environmental cues that could cause competition at recall. Participants overwhelmingly felt as though they were passing between two separate rooms in the 2R condition. This suggests that the broader, global details were sufficient for event model creation. However, some local details overlapped between conditions. The speaker design and voice used to deliver words lists were the same at each information point. From an event cognition perspective, event models can both help and hinder recall of information [Rad12]. When experienced information is clearly segmented between separate event models, then there is potential for better recall. If event models share information, then interference occurs and recall is reduced. Environmental context work [SV01] suggests a key driver for larger effect sizes is variation in the delivery of information, for example a different experimenter delivers the word lists. It is plausible then that using the same speakers and voice for the word lists was enough to generate interference from competing environmental cues within the conditions. Potentially local information variation is as important as global room differences for this study.

Relating more broadly to environmental context dependent work is the phenomena of context reinstatement, where returning to a context exposed during encoding can aid recall. For example, enhanced memory of information when recalling in the same room that he information was learnt. Although distinct to multiple context segmentation of information, they share fundamental drivers, the linking of encoded information to a location. Context reinstatement has had difficulties with replication in VR. Walti [WWW19] could not replicate enhanced recall from context reinstatement across simple background contexts (pc screen with colours and words), visual richness (using landscapes as contexts) and immersion levels (HMD vs none). Throughout these trials, subjects were sitting down and had limited interaction with the locations. Another well-studied phenomenon when walking through doorways, is a reduction in recall and recognition of information shared between the rooms. For example, knowledge of an item that has been picked up or placed down in a previous room. This is known as the location updating effect. The interference is caused by competition between the multiple event models assigned to each room. Although considered a reliable effect, this was not replicable within VR, using previously established methodologies [MNP21]. This study used passive navigation (movement through rooms was automated) and each room was purposefully identical. And yet Shin [SMOK21] was able to evidence of context reinstatement for free recall of collected objects within two VR contrasting worlds. A key difference in Shin’s work, was developing an understanding of the worlds through exploration, objectives, and distinct interactions. The rationale being that if an individual has experience and a developed schema of a location, they will have a deeper mental representation to build cues for recall. After familiarisation with the locals, the subject’s task was to search for objects and rate their usefulness before a surprise recall test. This approach would most likely use more tools to aid memory than just context reinstatement. It may encourage a method of loci style memory strategy on the subjects as they use local environmental constructs to associate object names to. This has shown to be effective in VR [KPV19] even when the subjects are not explicitly taught the method of loci technique [HQRB19]. Additionally, it has been evidenced that a simple search for objects within a VR room without the intent to memorise objects is more effective for recognition tests and object location knowledge compared to being shown the objects in a room with explicit instructions to memorise them [HDV20]. Shin’s work therefore shows that familiarisation and schema development of an environment can strengthen cues used for context reinstatement, or aid the efficacy of loci based memory strategies, without the need of an individual to explicitly apply the strategy. However, regarding more broadly the use of environmental cues to help structure and recall information, it may suggest that if an individual is not able to create an internally understood context for each location then there is a chance that many locations can be merged into a higher level context on recall. For example, putting on a VR headset and entering a new reality overrides the individual contexts of each virtual locations. Although the virtual environments used in our study are realistic in style, the colours, shapes, lighting, and composition are synthetic and in high contrast to the laboratory that the participants start in. This may be amplified by the difference in field of view that the HTC Vive HMD provides. The HTC Vive has a field of view of 110 degrees. Although wide, it is less than natural vision...
(approximately 180 degrees), with the borders of the lenses viewable in the periphery of the user. This will give the effect of wearing big goggles. If subjects are not given sufficient time to familiarise themselves to the VR environment and the technology, then environmental cues of individual rooms may be superseded but the general context of the VR platform. Therefore, a longer familiarisation time within VR could help the establishment of environmental cues and would control for technology experience variation within the sample [LDC16].

6. Limitations and Future Work

Although the design of separating information between two virtual room locations did not enhance memory recall, this has identified future work that would help to clarify if the effect was superseded or too weak to be pragmatic as a design approach to information delivery on VR platforms.

It is unclear if the act of removing the HMD for interferes with real world recall. Repeating the same work with free recall tasks inside of the VR world would help address this interesting question. Future work could also consider replicating real world segmentation of information between rooms to compare directly against the VR platform.

A familiarisation period should be established to make sure that subjects are sufficiently acclimatised to the VR world and control for potential population variation of experience with VR platforms.

A larger sample size is suggested to increase the sensitivity of this methodology to help identify confounding effects or the potential that the effect of multiple rooms aiding recall of word lists is weaker within VR platforms.

The strength of the VR platform is that the virtual space can be adapted or kept consistent between methodologies. The difference between locations, and in particular local foci of attention (speakers aesthetics) could be exaggerated for more distinct cues to associated delivered information.

7. Conclusions

We investigated if separating word lists between two virtual rooms could help recall compared to a single virtual location. Results indicate that there is no significant benefit to recall when separating spoken information between virtual environments. However, this study also found no significant impact of memory strategy when used by participants, and recall rates were lower than previous work performed in the real world. We have discussed potential improvements to the methodology used that may interfere with enhancement with recall when segmenting word lists between rooms as observed in real world studies. VR is a great tool for information delivery, but applications that value retention of spoken information may need to account for a possible reduced recall when using VR that could mediate information assimilation compared to real world counterparts.

References


[HQRB19] Huttner J.-P., Qian Z., Rohra-Bissantz S.: A virtual memory palace and the user’s awareness of the method of loci. 5


