An Immersive Memory Palace: Supporting the Method of Loci with Virtual Reality

Full Paper

Jan-Paul Huttner  
University of Brunswick – Institute of Technology, Brunswick, Germany  
j-p.huttner@tu-braunschweig.de

Susanne Robra-Bissantz  
University of Brunswick – Institute of Technology, Brunswick, Germany  
s.robra-bissantz@tu-braunschweig.de

Abstract

We investigated the impact of an increased level of immersion on the effectiveness of the method of loci (also memory palace). This is an old memorization strategy which relies on the mental association between the learning content and spatial cues in an environment like an apartment (memory palace). An experiment was conducted in which the participants were assigned to one of two groups. Subjects were instructed to use the method of loci to memorize and recall five lists of words. These lists consisted of eleven words each. The first group was told to use a virtual environment, presented on a laptop, as a template for their memory palace. The second group received an immersive head-mounted-display to explore the memory palace in virtual reality (VR). It was hypothesized and partially shown, that the VR group outperforms the laptop group in terms of accuracy and actual use of the instructed method.

Keywords

Method of Loci, Memory Palace, Virtual Reality, Design Science, Immersion.

Introduction

This design science approach addresses especially researchers in the domain of virtual reality interfaces for education. Following the Design Science Research Methodology defined by Peffers et al. (2007), the upcoming introduction will outline the theoretical foundation. Afterwards, the problem and research question will be explained. Finally, the resulting artefact and its evaluation will be described and discussed.

The basic idea of mnemonic methods, e.g. the method of loci, is to build a mental association between the learning content and an entity or object. For instance, an image, a catchy phrase or an abbreviation could be such an entity. This procedure makes it easier for the user to remember the content (Yates 1999). Even a mental map of a spatial environment could serve as an association anchor for several items that have to be remembered (Putnam 2015). This is possible, as it requires less effort to memorize these kind of entities or objects than plain text. Nevertheless, certain design principles should be respected when choosing such an object to avoid the effect of cognitive overload (Mayer and Moreno 2003).

These mnemonic techniques can certainly be applied to a problem like managing today’s information oversupply (Fassbender and Heiden 2006). However, especially the application of those methods in students’ curriculum may foster their chance to successfully perform in college (Hartwig and Dunlosky 2012; McCabe 2011). In 2015, Putnam found a positive correlation between the use of mnemonics and students’ motivation to educate themselves, however, being able to easily retrieve information from the memory is the main goal of mnemonic strategies (Putnam 2015). The role of mnemonics in the overall learning process can therefore be found in the first level of Bloom’s revised taxonomy of learning domains. According to this well-established model, the first level of learning is defined as “remembering” followed by “understanding”, “applying”, “analyzing”, “evaluating”, and “creating” (Krathwohl 2002).
The idea to integrate mnemonic methods into the students’ curriculum was already suggested at the beginning of the 90s. Levin and Levin (1990) created a set of picture compositions, called mnemonomies, to illustrate the hierarchical relationships within a botany classification system. They conducted an experiment and could prove that it is substantially easier to memorize them instead of the abstract botanical terms. However, mnemonic methods were not fully integrated in the teaching community (Putnam 2015).

The artefact in this study is used to evaluate the mnemonic technique called method of loci (MOL or memory palace). The main principle of the classical MOL is to navigate in mind through a familiar architecture or environment with spatial characteristics. On this tour, the user mentally places the to-be-remembered items in certain places (lat. locus for place, pl. loci). For instance, a simple object like a chair, a cupboard or a table can serve as an adequate locus. Hence, if the user wants to memorize an apple, s/he establishes an association by thinking of how the apple, for example, rolls over the table (locus). In order to recall these items, s/he has to visit these loci again. The mental association between the loci and the items facilitates the process of bringing them back to mind.

In 2002, Maguire et al. analyzed the underlying process for the success of the MOL. They found out, that people who have a superior memory tend to apply a learning strategy that is based on spatial aspects. Moreover, they investigated the participants’ brain activity using neuropsychological measurement instruments and brain imaging. Results indicated, that the subjects had a higher engagement in brain regions that are associated and crucial to the human spatial memory. Hence, the effectiveness and longevity of the MOL suggests a natural human tendency to use spatial context to memorize and recall information (Maguire et al. 2002).

**Related Work & Research Approach**

The approach of using information systems to enhance mnemonic strategies was already promoted in the late 90s. Storkerson and Wong (1997) suggested the application of the MOL to multimedia and hypermedia due to a higher intelligibility of communications in a spatial context. Three years later, Hedman and Bäckström (2000) built a 3D virtual environment (VE) for standard personal computers. It was designed to implement the characteristics of a memory palace. The goal was to enhance the students learning performance in a class of philosophy. The design followed the idea of a traditional museum. Note that the virtual memory palace (virtual world) in this and in the following related work was built to serve the user as a template for her/his mental memory palace. However, an experiment with two groups (five students each) did not indicate any effect to prove the superiority of the virtual memory palace over an uninstructed learning technique. As a result, authors suggested a more sophisticated interaction design. Later in 2006, Fassbender and Heiden (2006) used a similar approach. They hypothesized that the exploration of a virtual memory palace is easier for the user than stressing the own mind to apply the MOL only in a mental environment. Their prototype allowed the users to self-select the locations (loci) and to put the learning content in context with them. Fassbender and Heiden evaluated their virtual memory palace by conducting an experiment including 15 participants. They stated, that especially the participants’ long-term memory performance improved. In 2012, six years later, Legge et al. (2012) performed a study that involved a total of 142 participants to evaluate the potential that lies in the application of the virtual MOL. Participants were randomly assigned to one of three groups: a control group (which was not instructed to use a specific learning method), a group using the traditional MOL and a group that was instructed to use the virtual MOL. The task was to memorize ten lists of eleven words each. Authors did not find significant differences between the recall performance of the traditional and virtual MOL groups. However, the control group was outperformed and it could be shown that one does not necessarily need a familiar and richly detailed environment. Additionally, authors pointed out that the practice time, which is necessary to apply the MOL, can be reduced to five minutes. Earlier studies reported on training phases that endured between two hours and two days (Brehmer et al. 2007; Roediger 1980).

In 2016, Jund et al. conducted an experiment to explore the impact of the users’ frame of reference on memorization in virtual environments. To do so, they chose the MOL as memorization technique and found out that an egocentric frame outperforms an allocentric frame to offer the user spatial cues. Moreover, the authors strongly suggest the use of virtual reality (VR) for the application of the MOL. Due to the rich spatial cues, they identify an immersive VR environment as the perfect technology for this use.
case (Jund et al. 2016). Witmer and Singer describe immersive presence (also immersion) as the perception of being in a certain location although one is physically in another one (Witmer and Singer 1998). It could be shown that immersive presence is a crucial factor to perform successfully in virtual environments. It reduces the cognitive burden (Agarwal and Karahanna 2000) and supports cognitive processing (Ragan et al. 2012), which fosters the task performance (Witmer and Singer 1998). Moreover, Liu et al. (2014), Bredl et al. (2012) and Dede (2009) found that the perceived enjoyment, learning and engagement are also positively influenced by the level of immersion. Furthermore, the user’s memory recall performance concerning virtual objects and the spatial layout are fostered (Lin et al. 2002; Mania and Chalmers 2001). In 2010, Ragan et al. (2010) performed a study to analyze the influence of different levels of immersive presence on procedural memory. Participants had to fulfil procedural tasks in a virtual environment like sorting or rearranging simple objects on a table. The virtual world was presented using different devices to generate different levels of immersion. Later, they had to repeat this task in the real world. The study was motivated by the idea, that a higher level of immersion results in more spatial cues, which improve the users’ spatial perception of the environment. However, they found a positive correlation between an increasing level of immersion and the memory performance. That implicates, that a higher immersion facilitates the association between a spatial cue and an information. This is the main principle of the MOL. Therefore, a head-mounted display (HMD), which generates a higher level of immersion (Ragan et al. 2010), should improve the performance of the virtual MOL (vMOL) (Jund et al. 2016) in terms of learning success and actual use of the MOL. The learning success should increase as the process of building an association between a *locus* and an information becomes easier. The actual use of the vMOL should increase as the users’ cognitive burden is lowered due to better spatial cues. Hence, using the virtual reality environment as a memory palace should become easier.

Thus, the evaluation of these two measures, the learning success and actual use, will be the central aspect in this investigation. Learning success is measured as accuracy of the recall ability. Legge et al. (2012) measured accuracy with two scores. These scores are the proportion of words each participant is able to recall after memorizing a list of eleven words. The first one, the strict score, measures how many words the participant was able to remember, considering the correct position. For instance, if a list consists of the words *fish, spoon, table* and the participant enters the words *table, spoon, fish* then the strict score would be 0.33 as only the word *spoon* is correct and in the right position. The second score is calculated in a more lenient way. The lenient score also measures the percentage of the words that could be recalled, but disregards the position. Following the example above, the lenient score would be 1.0 or 100% as every word in the list could be recalled. The words in the lists were not chosen randomly. Legge et al. (2012) confirmed in their study, that high imaginable words (words that have a high score of concreteness, e.g. “tree”) are easier to remember than words with a low concreteness (e.g. *wisdom*). However, they did not find any evidence for an advantage of the MOL protocol for words of high concreteness. Hence, they varied the concreteness of the words but the growing learning performance of the MOL groups was not significantly better than in the control group. In this study, the variance of the concreteness will not be part of the research approach. Therefore, the words that were used here, are all of a high concreteness to facilitate the application of the MOL. The source of those words is a list of Kanske and Kotz (2010). They conducted a study with a total of 64 participants to rate, amongst other word properties, the concreteness of approximately 1000 German words. However, the word lists in this investigation were created in a way to ensure an equal level of concreteness. Moreover, all terms are words that are commonly used (e.g. *fork, table or apple*). This way, the experimental design ensured that participants most likely knew the meaning of these words and could easily build a mental representation of those objects.

Finally, Legge et al. (2012) used a questionnaire to determine the actual use of the vMOL. It was measured as the compliance rate. This rate indicates whether and how strong each participant was compliant to the instructed vMOL. It is defined by the number of lists a participant was able and willing to memorize and recall by applying of the MOL. A participant was regarded as compliant if s/he used the MOL for at least 50% of the lists (Legge et al. 2012).

The previous work in the domain of the MOL and virtual worlds as well as previous findings about the effect of immersive presence on several dependent variables like memory performance lead to the research question, outlined by Huttner and Robra-Bissantz (2016): Is the application of the vMOL in a high immersive device more effective in terms of accuracy and actual use than in a setting with a computer screen? If this is the case, future research and practical concepts use should focus on artifacts
that are built for HMDs or other devices that foster a higher immersive experience. The following three hypotheses will be investigated to answer the research question:

**Accuracy**  
\( H_{1a} \): Using a HMD to apply the vMOL will result in a higher strict score than applying the vMOL with a computer screen.  
\( H_{1b} \): Using a HMD to apply the vMOL will result in a higher lenient score than applying the vMOL with a computer screen.

**Actual Use**  
\( H_{2} \): Using a HMD to apply the vMOL will result in a higher compliance rate than applying the vMOL with a computer screen.

**Experimental Design**

An experiment was performed to analyze this research approach. The experimental design, including its procedure and the measurement of the learning success, is closely aligned to the study performed by Legge et al. in 2012 as its design fits the research question in this study (Legge et al. 2012).

**Participants**

A total of 78 unpaid students took part in the experiment. They were aged 17 to 29 (mean = 23.79, 21 females). Participants were required to have German as their mother language as the study involved the reading and writing of German words. The students were invited to the experiment and could subscribe anonymously. About 55% of the participants reported to have at least heard of the MOL in advance, but none of them uses it regularly or knows exactly how it works. Note that the participants, who had at least a little knowledge about the MOL, were almost evenly distributed between the experimental groups.

**Technology**

As mentioned above, the level of immersion should have a positive influence on the vMOL effectiveness. In order to increase the user’s immersive presence, participants of the virtual reality group were equipped with a HMD to explore the virtual reality. The applied HMD is a system that needs a smartphone that serves as the display. As the smartphones display would be too close to the users eyes to successfully focus the image, this type of HMD has two lenses that generate a stereoscopic effect. This means, the virtual world is presented on a split screen on the smartphone while the two lenses enable the user to perceive the two views as one. This effect gives the user the illusion of spatial depth. Hence, the user optically perceives the virtual environment in three dimensions.

The virtual environment was developed with the game engine Unity 3D (version 5.5.0) and is closely aligned to the virtual apartment that was used by Legge et al. (2012). Figure 1 illustrates the bird and the first-person view of the virtual apartment. Every room has several loci (e.g. plants, tables, chairs, cupboards or images) that can be used to establish a mental association between them and the learning content. Hence, this apartment was the virtual environment thought to serve every participant as the template for their memory palace.
In order to ensure a smooth and fluent virtual experience, the smartphones’ hardware performance was good enough (Google Nexus 6P, last generation of Google’s high end smartphone). The same applies to the type of laptop (computer screen) that was used for the other set of participants. The detailed procedure of the experiment will be explained in the following section.

**Procedure**

Participants were randomly assigned to one of two groups, the computer screen and the virtual reality group. All participants performed under the same conditions except the medium that was used to present the virtual environment. Figure 2 shows the sequence of the essential elements in this experiment. First, every participant was given a few minutes to become confident in the navigation in both devices (laptop/computer screen first, then HMD). The practice in both media was thought to ensure that the practice environment and medium did not affect the participants' performance in the later phase in any way. However, a simple, virtual practice room was implemented for the computer screen and the HMD device. This virtual room was empty and therefore did not contain any loci. Participants were told to get familiar with the navigation in the virtual environment. The integrated keyboard and an external mouse were used for navigating in the computer screen environment. The user was able to walk by using the arrow keys while the mouse was used to look around. In the virtual reality environment, the user’s physical head movement was tracked by the smartphone in the HMD and then translated directly into the virtual environment. The gamepad was used for walking. After a few minutes, when the particular participant confirmed to be confident in the navigation s/he was asked to remember and reproduce a practice list of words. The list was presented on a computer screen and contained eleven words. This practice task was given to ensure that each participant understood the procedure correctly which was important later in the recall phase. All lists that were presented followed the same protocol used by Legge et al. (2012). Hence, each word in a list was shown solely for 5000 milliseconds in the center on a white background.
An inter-stimulus break of 150 milliseconds divided the presentation of each word and its successor. The participants’ task was to remember the list and reproduce it by entering the words in a single web based form field, one after another and if possible in the same order. The test person entered the remembered word and pressed enter to confirm the input. If s/he could not recall the word at a certain position, there was a skip button implemented. By clicking this button, the form field was marked as empty at this position and the participant could enter the next word. This recall phase was limited to 120 seconds, so if a participant needed more than 120 seconds, the form automatically saved every word s/he entered until this point. Participants did not receive any information about how many or which words were already entered or how much time they had to enter the lists. After completing the practice list task, the practice phase was finished and participants were introduced to the virtual environment. As described in the related work, this virtual environment should serve them as a template for their memory palace. Depending on the group (computer screen or VR), they either received a computer screen or a HMD. Participants had a maximum of five minutes to explore the virtual apartment. After that, they received a written instruction on how to use the MOL. This instruction was the same one that was used by Legge et al. and is based on the description of Yates (Legge et al. 2012; Yates 1999, see Appendix A). It was translated into German, as the test persons were predominantly German students. After participants were informed how to apply the vMOL, they were presented five lists of eleven words each. The procedure was the same as in the practice phase, but this time s/he had to remember the words by building an association between them and the loci in the virtual apartment. Finally, every participant was given a questionnaire to report, amongst others, the demographic information and her/his compliance rate.

**Data Analysis**

The data analysis was conducted with the open source software R (version 1.0.136). The two-sampled Welch’ t-test and an exact chi-square test (Fisher-Yates test) were used to analyze the differences of means and the differences of the distribution of compliance rates between the two groups. The Welch’ t-test was preferred against the students t-test due to its robustness against unequal sample sizes (Ruxton 2006; Welch 1947). Similarly to Legge et al. (2012), an alpha level below 0.1 was considered as a trend effect (p<0.1). Also, the analysis was conducted on all participants and again for the “compliant only” subset. Therefore, results will be shown for these two groups to show whether the learning success (accuracy) differs between them. It is expected, that the “compliant only” subset will show better results due to the hypothesized superiority of the MOL.
Results

In the following the data is statistically analyzed and interpreted to illustrate findings and relevant questions for future research in the domain of the vMOL.

Accuracy

The analysis of the accuracy involved the measurement of the strict and lenient scores of all participants. As described in the experimental design, these scores were measured using lists of highly concrete words. To determine the average scores for each group (computer screen and VR) and participant, each score was calculated as the mean over the five lists. So, these means were considered as the overall strict or lenient score for a particular participant. For instance, if a participant reached the strict scores 0.30, 0.70, 0.30, 0.30 and 0.40 for the lists one to five, his overall strict score would be 0.40. These means of strict and lenient scores were then used to compare the accuracy of both groups. Table 1 gives an overview of the means of the strict and lenient scores.

<table>
<thead>
<tr>
<th></th>
<th>Strict Score</th>
<th>Lenient Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Computer screen</td>
<td>Virtual Reality</td>
</tr>
<tr>
<td>All participants</td>
<td>0.4327</td>
<td>0.4880</td>
</tr>
<tr>
<td>Compliant only</td>
<td>0.4293</td>
<td>0.4923</td>
</tr>
<tr>
<td>Difference between compliance</td>
<td>-0.0034</td>
<td>0.0043</td>
</tr>
</tbody>
</table>

Table 1: Analysis of accuracy means of groups computer screen and virtual reality

The strict score ranges between 42% and 50%, while on average the virtual reality group achieved better results than the computer screen group (approximately 5 to 6%). However, t-tests did not show a significant difference of means, neither for all participants, nor for the compliant only subset of the groups (all participants: df=75,84; p-value=0,2353 and compliant only: df=62,26; p-value=0,2051). The lenient score lies between 64% and 73%. Again, the virtual reality group performed better than the computer screen group on average (approximately 5 to 7%). T-tests showed a trend effect for the set of all participants on a level of p<0.1 (all participants: df=70,59; p-value=0,094 and compliant only: df=53,15; p-value=0,2092). Except for the strict score in the computer screen group, the difference of the scores between the compliance levels implicates a small improvement on average. Hence, hypotheses h1a cannot be confirmed. Hypotheses h1b cannot be confirmed either as only the all participants analysis showed a trend effect (p<0,1).

Compliance Rate

As mentioned earlier, the compliance rate is the number of lists each participant memorized and recalled by the use of the MOL. It was measured by self-assessment while participants were considered compliant if they applied the MOL for at least 50% (three lists). Results are shown in Table 2.

<table>
<thead>
<tr>
<th>Number of compliant lists</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Σ</th>
<th>Compliance rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer screen</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>10</td>
<td>14</td>
<td>7</td>
<td>40</td>
<td>77.5%</td>
</tr>
<tr>
<td>Virtual Reality</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>6</td>
<td>20</td>
<td>38</td>
<td>97.4%</td>
</tr>
</tbody>
</table>

Table 2: Frequency distribution of compliant lists

Analysis shows that the participants of the virtual reality group were significantly more often compliant compared to the computer screen group (computer screen: 31/40, VR: 37/38). The exact chi-squared test resulted in a strongly significant difference (p<0.01) between the two groups (p-value=0.0021). Therefore, the hypothesis h2 can be confirmed.
Discussion and Conclusion

The approach in this study was to evaluate the possible improvement of the vMOL. Multiple studies were found that promote the positive effects of immersive presence on important factors like memory performance (Bredl et al. 2012; Dede 2009; Ragan et al. 2010). Consequently, two groups were instructed to apply the vMOL, while one group was told to use a device (HMD) that generates a higher level of immersion. Two central performance indicators were tested: the accuracy and the actual use or compliance rate. It was hypothesized, that the group with the HMD achieved better results in the accuracy and compliance rate. The analysis of the data does not fully support the expected effects. However, it was shown, that the learning success varies between the computer screen and the virtual reality group. All the average scores (strict and lenient) in the virtual reality group are approximately 5 to 7% higher than those in the desktop group. This indicates a trend promoting the advantage of the virtual reality MOL. A possible reason for the accuracy difference might be that the level of immersion was not high enough in the virtual reality group to show significant effects. In this study, the level of immersive presence was theoretically increased by the use of a HMD. Other aspects, that help to improve the immersion should also be considered. For instance, some participants commented that the feeling of being entirely in the virtual world would increase with a sound feedback like footsteps or other noises that are typical for an apartment. According to Dede (2009), sensory cues are a key factor for immersive presence. Therefore, auditory cues would be a reasonable feature to increase the level of immersion (Dinh et al. 1999). Of course, the difference in accuracy levels might also be caused by the fact that participants in the VR group were significantly more compliant. As explained in the introduction, the MOL improves the memorization and recall performance. Therefore, the factors accuracy and compliance rate are most probably not independent from each other. However, the exact correlation between the level of immersion, the compliance rate and the accuracy should be investigated in future studies. That includes the measurement of immersion and its influence on those factors.

Another limitation is given by the homogeneous group of participants (students only) and a majority of male participants, and of course, an increased number of compliant participants could help to find more significant results in this area.

It is certainly necessary to invest further research in the application of the virtual reality based MOL, but results in this study indicate a relevant and unused potential lying in this approach. For future research, we suggest the improvement of the immersive experience by integrating multi-sensory cues and to investigate the its statistical correlation to the factors compliance and accuracy. Nevertheless, Legge et al. (2012) emphasized that the compliance rate is especially important for researchers as the number of compliant participants of an experiment is crucial for reliable research results. Therefore, the application of the vMOL with a high immersive virtual reality is encouraged at this point.
Appendix A

The instructions to use the MOL (Legge et al. 2012)

The Method of Loci has been proven to significantly increase the effectiveness of memory. Below is a description of the Method of Loci, paraphrased from The Art of Memory by Yates, the established historical text on the Method of Loci. In this method, memory is established from places and images. If we wish to remember an object, we must first imagine that object as an image, and then place it in a location. If we wish to remember a list of objects, then we must make a path out the many locations. The easiest way would be to imagine a familiar environment and place the imagined objects inside it. Then, you can pick up the objects as you imagine navigating the environment, thereby remembering the object list in order.

REFERENCES

Hedman, A., and Bäckström, P. 2000. “Rediscovering the Art of Memory in Computer Based Learning—An Example Application.”