https://doi.org/10.2352/EI.2023.35.10.HVEI-220
© 2023, Society for Imaging Science and Technology

Comparison of AR and VR memory palace quality in second-language vocabulary acquisition

Nicko R. Caluya, Xiaoyang Tian, and Damon M. Chandler; Ritsumeikan University, Kusatsu City, Japan

Abstract

The method of loci (memory palace technique) is a learning strategy that uses visualizations of spatial environments to enhance memory. One particularly popular use of the method of loci is for language learning, in which the method can help long-term memory of vocabulary by allowing users to associate location and other spatial information with particular words/concepts, thus making use of spatial memory to assist memory typically associated with language. Augmented reality (AR) and virtual reality (VR) have been shown to potentially provide even better memory enhancement due to their superior visualization abilities. However, a direct comparison of the two techniques in terms of language-learning enhancement has not yet been investigated. In this paper, we present the results of a study designed to compare AR and VR when using the method of loci for learning vocabulary from a second language.

Situating Spatial Memorization on Immersive Environments

As more experiences in extended reality (XR; e.g., virtual and augmented reality (VR, AR)) become more immersive and interactive, many studies over the past years have explored if these innovations can improve skills and further learning.

Additionally, AR and VR reported to have advantages compared to traditional learning media like videos or audio. Providing learners with an immersive and interactive learning environment, boosting interest and motivation, and lowering mental effort.

An interesting investigation is the use of VR and AR as a memorization tool and environment, as the inclusion of interactive 3DCG content can potentially affect the way the user remembers various stimuli. In the case of spatial memory, spatial relationships between the user and these objects rely on the quality of the environment that both are situated in.

XR Technologies for Memory Palaces

We would like to further explore the idea of memorizing new words based on their attachments to 3DCG objects. With spatial memory interacting with semantic memory, placements of objects in the AR/VR world can provide an additional memorization anchor for learning new vocabulary. Our proposed technique is inspired by popular routines such as the method of loci (MoL; e.g., memory palace), with the additional benefit of an enhanced environment tailored for the experience of memorization.

Memory palace, also known as the method of loci, utilizes physical locations to build meaningful connections to other items (especially information) and remember the information. This technique has great value for real-life use[1]. The technique involves 2 key steps. First is the memorization of a space, either imaginary or real, to find key spots in the space. Next, connect these locations of importance to things to be remembered. This way, a list of items can be remembered in a certain order. The technique is popular for general memory training and also popular for language learning, especially vocabulary.

In a study comparing virtual and conventional MoLs [3], they reported very little difference between them, thus making virtual MoL a viable environment for memorization experiments. Handheld, mobile AR apps have also utilized for MoL. Users were allowed to choose the stimuli to designate on each locus, whether as online or in-phone images as AR content [4] or as actual sheets of paper with QR codes to register the 3DCG content [2]. Another study highlighted the advantages of situating participants on a VR-based memory palace versus image-only counterparts, as they were tasked to retrieve and recall information from scholarly articles [8].

XR Technologies for Language Learning

Weerasinghe et al. [7]. proposed an AR prototype ”ARigato” and conducted an experiment in which users learned phrases from 2 different topics. The unique focus of the study is on the amount of instructions given to the participants and the semantic associations the participants were allowed to build. The researchers also developed an experiential learning cycle based on Kolb’s learning cycle theory.

The closest related work to this user study is the work by Santos et al. [5], where they compared AR (i.e., fiducial marker-based) and non-AR (i.e., app-only) presentations of a vocabulary learning material in Filipino and in German languages. The intention of their study is more towards situated learning of vocabulary, which is a little different from memorization. Additionally, the locations of the AR markers and the objects to be memorized are fixed, and both conditions were displayed on a handheld mobile tablet device (e.g., they used an iPad). The results showed that presenting the vocabulary learning scenario in a context-rich environment of the real world has its merits over the non-AR approach, where multimedia (e.g., sound and image) are displayed alongside the word to be learned. Similar to their work, the idea of virtual/augmented memory palaces also rely on the context in which the participants are situated in. However, in the case of this study, we give the participants more selections in terms of stimuli or objects to associate the words with.

Comparing AR and VR as Learning Environments

The main differences that we focus on this user study are (1) the level of realism of the environment, (2) the involvement of the user through presence, and (3) the possibility of skill transfer from training to the real world.

First, we expect users to notice immediately the difference between the inclusion of the real world in AR and the elimination of realistic stimuli in VR. Given this difference, users may employ contrasting memorization strategies being affected by realism. The real world may provide more visual landmarks for the loci, while the modifiable world in VR can offer an environment with reduced real world distractions.

Second, presence is commonly evaluated when comparing AR and VR experiences. Details like sense of embodiment, quality of interactions, and immersion may contribute toward the overall experiences. In this user study, we used the Igroup Presence Questionnaire [6] to guide us in determining differences in presence factors like realism or involvement.

Last, we envision that memorizing a new language can only be fully realized outside of the training environment, which is usually in the real world. As such, we hypothesize that the AR condition can provide better skill transfer, as it can easily map spatial relationships from the real world experienced in AR with the actual real world where the skill is put to the test.

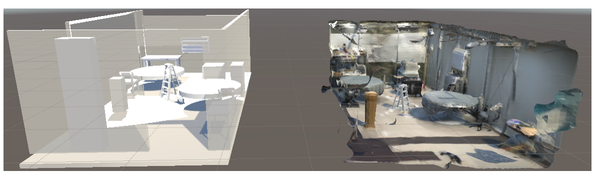

Figure 1. *Rendered perspective views of the same room in VR (left) and AR (right)*

Experiment Task

We conducted a preliminary user study to verify if the differences between VR and AR that we have just mentioned contribute toward better performance in spatial memorization. We developed an experiment environment where subjects perform a language learning task in 2 target conditions: VR and AR, and a real-world control setup using paper. We instructed participants to learn a small vocabulary (15 words) from a target language (Filipino: Tagalog) in the environment within 15 minutes. In each environment, the same vocabulary was given to the participants, and they have the same level of interactivity and exploration. This study is focused on three factors for comparison: presence, environmental realism, and skill transfer.

Experiment Details

We recruited 12 university students for this user study, with ages ranging from 22 to 26 years old (M = 23.1, SD = 1.3). We asked them to answer a pre-experiment survey, disclosing any vision problems or potential cybersickness. Even when 9 participants declared having vision problems and 6 participants had previous encounters with cybersickness, all participants were able to finish all the tasks in the user study, as we adjusted the experiences for corrected-to-normal vision.

Only 2 participants reported that they have experienced AR/VR in an educational setting. Using a 5-point Likert scale, participants on average felt more comfortable using AR headsets (M = 3.3, SD = 1.2) over VR (M = 2.7, SD = 0.9). All participants know at least 3 languages, with one participant knowing 5.

As the user study involved the comparison of the same task in different environments, we conducted a within-subjects study that counterbalanced the order of three conditions (VR, AR, and Paper/Control). The control condition is only a typical memorization scenario where pairs of words are laid out on a table for them to look at without any other interactions (Fig. 2). Participants experienced this control condition in a different experiment area, separate from the AR and VR conditions. The differences between the AR and VR environment setups are shown in Table 1.

Table 1. Experiment Conditions: AR vs. VR

	AR	**VR**
Device	HoloLens 2	Vive Pro
Room View	Actual Room	3D Mesh of Room
Real World	Visible	Invisible
Controls	Vive Controller	Hand (Gestures)
User's Body	Visible	No View

It is also important to note that the visual fidelity of the room environment in the VR condition has been reduced to simple 3D objects with minimal, white textures, and a simple lighting setup (see Fig. 1 and 3). This is to provide a stark contrast between the real world and virtual view. The positions, sizes, and orientations of all the objects in both conditions are the same.

To evaluate their memorization, each participant will be asked to pair the newly-learned words with (1) their English counterparts, and (2) the objects they are associated with, in (3) the specific location where the participant placed them.

On the post-experiment survey, aside from IPQ, we also asked our participants what memorization strategy they used to pair up the Filipino word with the English word.

Preliminary Results

The environment was tested by one participant each, the results that were taken from their tests are shown in Figure 5.

Levene's test did not detect a significant difference between the levels of variance across the test scores ($F(2,15) = 1.248$, $p \approx 0.3152$ for *Condition*). However, a Shapiro–Wilk test showed that the results were not normally distributed ($p < 0.05$). Thus we carried out a Friedman's test for all these observations with Bonferroni correction for the adjustment of p values.

Figure 2. *Participant's view of the control (paper) condition*

Considering the training scores, we did not find any statistically significant difference among the three *Condition* levels (AR vs. VR vs. Paper) from the results of the Friedman's test, $\chi^2(2) = 5.1579, p \approx 0.08$. A post-hoc test using Fisher's Least Significant Difference (LSD), however, showed that only the test scores after doing the Paper condition were significantly higher than the VR condition.

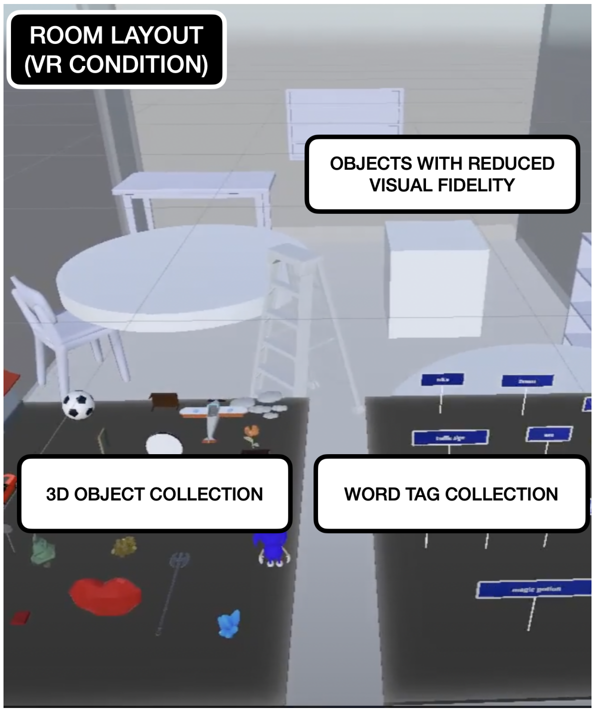

Figure 3. *User's view of the two boards (word tags and 3D objects) within the experiment space in the VR condition.*

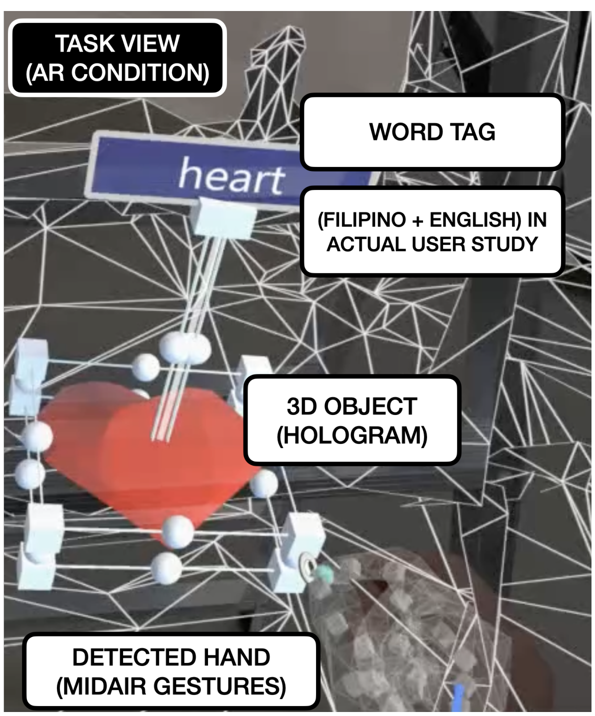

Figure 4. *Interaction components of the task, shown in the AR condition.*

According to the results for scores, the control paper condition still had the highest scores. We can attribute this to the fact that the skill transfer task is performed in the real world, and no augmentations or any CG is involved when immersed in this experience with no devices.

However, the average score of the AR condition is higher than that of VR. In this case, the advantages of viewing the real world shows. Again, as the skill transfer task is performed in the real world, there is a sense of realism and direct learning from the actual environment.

Data obtained from participants regarding their memorization strategies (Fig. 6) follow similar properties of the test scores, with Levene's test not detecting a significant difference between the levels of variance, but Shapiro-Wilk test revealing that the distribution is not normal.

Thus, we can only perform Friedman's test according to *Condition* and *Strategy*. We found a statistically significant difference in *Strategy* ($\chi^2(3) = 10.184, p \approx 0.02$), but not in *Condition* ($\chi^2(2) = 0.195, p \approx 0.907$). Using pairwise Wilcoxon tests, the prior vocabulary option was chosen more than the other three strategies.

Additionally, we also ran an ANOVA on rank-transformed results and found possible interactions between *Condition* and *Strategy* ($p \approx 0.195$). As shown on the graph, the "prior vocabulary" answer was at the highest for the paper condition, however, the "semantic" answer was the lowest in the paper condition as well.

With this stark contrast of the participants' selections of prior vocabulary and semantics, this implies not only possible interactions between the memorization strategy and the way the memorization task was delivered. This also highlights that participants adjusted their strategy according to the learning environment that they were situated in. The strategies were leaning more towards the relationship between the target word to learn and prior knowledge (i.e., from a word they knew before, or a visual stimuli that they have already encountered).

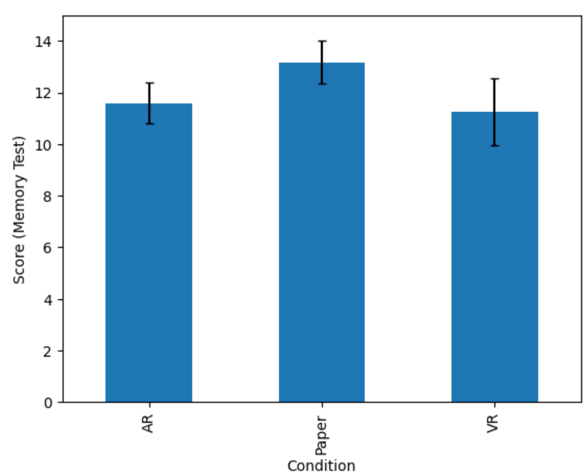

Figure 5. *Average scores according to* ***Condition***

This also implies that order and position may not have any significant impact as cues for the participants to memorize a new set of vocabulary. In future work, it may be interesting to look at adding more words and exploring various sizes of the memorization space (i.e., from a tabletop to a building).

Performing the same normality tests for the IPQ scores, we found that even if the scores have approximately equal variance (Levene: $F(1,94) = 1.098$, $p \approx 0.2989$), their distribution was not normal (Shapiro-Wilk: $p < 0.05$). Thus, we also had to perform non-parametric tests. We found a statistically significant difference in *Condition* ($\chi^2(1) = 8.333, p \approx 0.04$). Using pairwise Wilcoxon tests, we verified that only involvement had a statistically significant difference between the VR and AR ratings of presence.

The results from the presence questionnaire provide us with insights regarding the presentation of the experience. Even when VR rating from the IPQ appear to be higher than AR, the lack of statistical significance suggests that they offer similar levels of presence. However, the case of higher involvement ratings may also imply that participants had a higher sense of awareness that they are experiencing the virtual world with the virtual objects they are interacting with, instead of reconciling virtual objects versus in the real world in the AR condition. For future iterations of this study, the difference in AR and VR can be limited further to only just one aspect, e.g., the gestures or interactions, the visibility of the real world, or the realistic details of the real world. This way, we can identify the affordance that contributes to these presence ratings.

Limitations and Future Work

The context in which a memory palace strategy will be more effective might lie on how participants utilize the immersive environment to enhance the skill of memorizing.

The results of the survey on their memorization strategies reveal future work for the study. First, the high dependence on prior vocabulary on the paper condition implies that participants had to rely on the readily available stimuli. Since we introduced virtual objects as an additional stimulus while they memorized new words, it was evident that participant relied less on prior vocabulary and prioritized the semantic association of the new word with the objects. As investigated by Weerasinghe et al.[7], the adaptive guidance approach provided additional cues by creating virtual objects as an abstract conceptualization of homophonic word (combinations) from prior vocabulary.

However, to fully realize the episodic and spatial properties of an immersive virtual memory palace, the association of the assisting stimuli to the loci still needs to be explored further. In Nevermind [4], the memorization strategies primarily relied on the location and sequence of words because of the high amount of ideas to memorize, the fixed virtual object associated with the idea, and the words were set in English.

As this is still an exploratory study, there are limitations to the work, especially in the control design of the environment. As both AR and VR did not outperform the paper control condition, it may be necessary to think about the ways that the virtual/augmented task lends itself to skill transfer. Even with the promise of enhanced presence, tactile interactions may also play a part in the memorization process. The hand or controller gestures do not necessarily translate to the paper placement task in the immediate recall task. As the AR and VR devices also serve as visual implements, we have to take into consideration the way the environment is being rendered to the user to create a more immersive environment.

Conclusions

The research was able to accomplish the creation of an AR and VR learning environment for a foreign language, Filipino. The AR and VR environment was able to create a memory palace that lets the player interact with the objects around them. The environments are meant to help the users in vocabulary acquisition of another language. The preliminary experiment that we conducted aimed to know which environment is better when it comes to language learning. The main results show that the paper-based task outperformed both AR and VR conditions, and that AR and

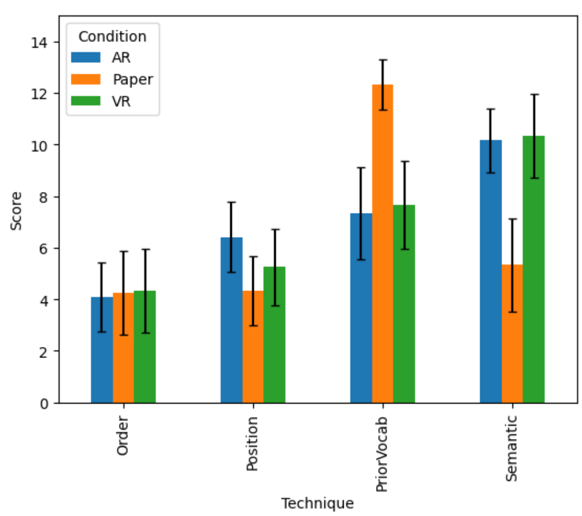

Figure 6. *Frequency of memorization strategies for all words*

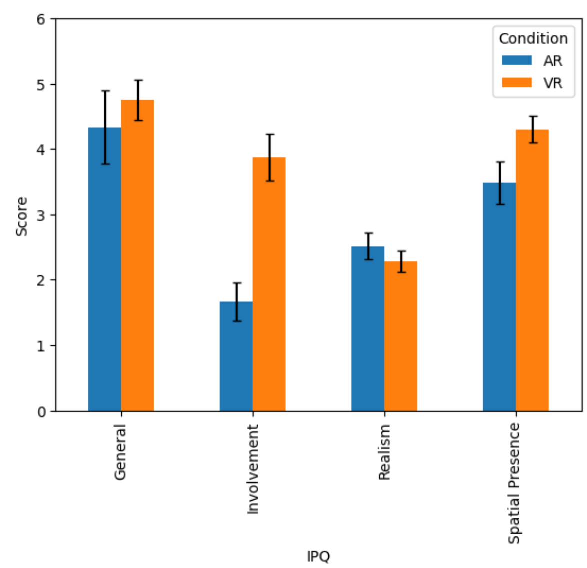

Figure 7. *Post-Experiment Igroup Presence Questionnaire (**IPQ**) results*

VR scores did not have statistically significant difference to determine which one was better. The higher involvement ratings in VR on the presence questionnaire suggest the importance of awareness between virtual and real worlds when participants are immersed in this experience. Furthermore, due to the interference of different factors like memorization strategies, further investigation is needed to validate our findings.

References

[1] Sauvik Das et al. "The memory palace: Exploring visual-spatial paths for strong, memorable, infrequent authentication". In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. 2019, pp. 1109–1121.

[2] Alexia Larchen Costuchen, Stephen Darling, and Clare Uytman. "Augmented reality and visuospatial bootstrapping for second-language vocabulary recall". In: *Innovation in Language Learning and Teaching* 15.4 (2021), pp. 352–363.

[3] Eric L.G. Legge et al. "Building a Memory Palace in Minutes: Equivalent Memory Performance Using Virtual Versus Conventional Environments with the Method of Loci". In: *Acta Psychologica* 141.3 (2012), pp. 380–390.

[4] Oscar Rosello, Marc Exposito, and Pattie Maes. "Nevermind: Using augmented reality for memorization". In: *Adjunct Proceedings of the 29th Annual ACM Symposium on User Interface Software and Technology*. 2016, pp. 215–216.

[5] Marc Ericson C. Santos et al. "Augmented reality as multimedia: the case for situated vocabulary learning". In: *Research and Practice in Technology Enhanced Learning* 11.4 (2016), pp. 1–23.

[6] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. "The Experience of Presence: Factor Analytic Insights". In: *Presence: Teleoperators and Virtual Environments* 10.3 (June 2001), pp. 266–281.

[7] Maheshya Weerasinghe et al. "Arigatō: Effects of Adaptive Guidance on Engagement and Performance in Augmented Reality Learning Environments". In: *IEEE Transactions on Visualization and Computer Graphics* 28.11 (2022), pp. 3737–3747.

[8] Fumeng Yang et al. "A virtual reality memory palace variant aids knowledge retrieval from scholarly articles". In: *IEEE Transactions on Visualization and Computer Graphics* 27.12 (2020), pp. 4359–4373.

Author Biographies

Nicko R. Caluya *received his B.S. in Computer Science from Ateneo de Manila University (2013) and his M.Eng. and D.Eng. in Information Science at Nara Institute of Science and Technology (2018, 2022). He is currently an Assistant Professor in the Visual Information Engineering Laboratory at Ritsumeikan University. His research interests are augmented/virtual reality, human-computer interaction, and educational technology.*

Xiaoyang Tian *is currently finishing his Bachelor of Engineering degree at Ritsumeikan University. His research interests are augmented and virtual reality.*

Damon M. Chandler *received the B.S. in Biomedical Engineering from The Johns Hopkins University (1998); and the M.Eng., M.S., and Ph.D. in Electrical Engineering from Cornell University (2000, 2004, 2005). From 2005-2006, he was a postdoc in the Department of Psychology at Cornell. From 2006-2015, he was on the faculty at Oklahoma State University. From 2017-2021, he was an Associate Professor at Shizuoka University. He is now a Professor at Ritsumeikan University, leading the Visual Information Engineering Laboratory. His research focuses on modeling properties of human vision. He is as an Associate Editor for the IEEE Journal of Perceptual Imaging and Journal of Electronic Imaging.*