dHealth 2025
M. Baumgartner et al. (Eds.)

doi:10.3233/SHTI250148

A Preliminary Evaluation of a VR Application for Associative Memory and Cognitive Rehabilitation

Prince AMPOFO[a,1], Andrea VITALI[a], Francesca MORGANTI[a], Caterina RIZZI[a] and Daniele REGAZZONI[a]
[a] *University of Bergamo, Bergamo, Italy*

Abstract. Background: Associative memory is essential for episodic memory formation, and training in this domain has been shown to enhance memory in both clinical and non-clinical populations. However, the application of VR-based training in this memory domain remains underexplored. **Objectives:** This study aimed to develop and evaluate a VR-based intervention to improve associative memory, focusing on verbal-visual and visual-auditory stimuli associations. **Methods:** Five healthy younger adults (mean age = 24.6) completed 8 trials of object-name and object-sound matching tasks in a virtual environment. **Results:** Significant performance improvements were observed across trials, with object-name matching showing higher recognition accuracy and faster response times than object-sound matching. **Conclusion:** Both tasks demonstrated increased accuracy with reduced response times with training. This study underscores the importance of tailoring VR-based cognitive training to specific associative tasks, offering promising applications in memory rehabilitation and cognitive enhancement.

Keywords. Virtual Reality, Memory Training, Paired-Associate Learning

1. Introduction

Digital health technologies, which include telemedicine, wearable devices and virtual reality (VR) are revolutionising healthcare by enhancing patient care, improving outcomes, and personalising treatments. These technologies are addressing modern healthcare challenges such as accessibility, efficiency and targeted intervention [1]. Among these technologies, VR has emerged as a powerful tool in digital health, creating immersive 3D environments for medical training, rehabilitation, and cognitive interventions. VR seamlessly bridges the gap between theory and practice through its ability to simulate realistic scenarios [2].

One critical cognitive domain where VR has demonstrated significant potential is in enhancing associative memory. Associative memory, a critical component of episodic memory, enables an individual to learn and retain relationships between unrelated items [3]. This fundamental cognitive function plays a vital role in daily life, enabling individuals to associate objects with their names, sounds, shapes, locations, and other perceptual properties, thereby creating rich and meaningful experiences. Furthermore,

[1] Corresponding Author: Prince Ampofo, University of Bergamo, Bergamo, Italy,
E-Mail: prince.ampofo@unibg.it

associative memory contributes to the formation of semantic memory, which is essential for cognitive tasks such as recognition, language acquisition, and navigation within environments [4]. Despite its importance, associative memory is more fragile than item memory, making targeted training crucial for both clinical and non-clinical populations across all age groups. Associative memory experiments often involve learning and retrieving verbal-visual, verbal-auditory, or visual-spatial stimulus pairs using computerised or manual exercises. These approaches have proven effective in enhancing memory function among healthy individuals [5] and dementia patients [6].

The advent of digital health solutions like VR technology has further enhanced cognitive training by enabling the creation of ecologically valid environments, interactive experiences, and immersive scenarios that closely mimic real-world conditions, which are difficult to achieve through traditional or non-immersive digital methods [7]. VR-based associative memory studies have also demonstrated effectiveness in detecting memory deficits [8] and improving memory performance [9], offering a more personalised and scalable approach to cognitive health within the digital health framework. While VR systems may have higher initial costs than conventional digital tools, their ability to provide interactive, real-world training environments justifies their use for optimising cognitive outcomes.

The present study aimed to design and develop a VR-based associative training intervention and evaluate performance across multiple trials, aligning with the broader goals of digital health by delivering tailored interventions that improve outcomes and patient engagement. Given the limited research on VR-based training that integrates visual-verbal and audio-visual learning pairs in a multi-trial framework, this study focused on these specific memory constructs. Previous studies have reported that associative recognition often favours stimuli pairs that include verbal stimuli over auditory stimuli. For example, Hocking and Price (2009) investigated audiovisual integration between verbal and non-verbal stimuli using spoken and written object names, and pictures of objects paired with naturally occurring object sounds. Their findings revealed significantly better performance and faster response times for pairings that include verbal stimuli, and reduced structural processing demands compared to pairs of pictures and object sounds [10].

Based on the literature, it was hypothesized that object-name learning would have higher learning accuracy and faster response time than object-sound learning.

2. Methods

2.1. Participants

Six students from the University of Bergamo (age range: 21 – 26 years) were recruited voluntarily using flyers and posters for this study. Inclusion criteria included healthy individuals without physical, psychological, neurological, or audiovisual impairments. The mean Mini-Mental State Examination (MMSE) [11] score of participants was 29.4, indicating normal cognitive functioning and absence of dementia. Before the protocol began, all participants provided informed consent receiving a detailed explanation of the main goals and procedures of the study. They were also informed they could withdraw at any time if they experienced any discomfort while completing the task. One participant reported that the VR headset was too heavy after the second trial and withdrew from the study. The study protocol was approved by the University of Bergamo Research Ethics Committee (Project number: 2024_07_03) and was conducted under the research ethical standards outlined in the Declaration of Helsinki (2013 revision).

2.2. Materials and Technology

The study materials included 3D objects from www.turbosquid.com and melodies from www.pixabay.com. A single 3D toy model was downloaded, duplicated and customised into ten toys with distinct colours, enhancing object naming [12]. The melodies were trimmed to 10 seconds in length, and ten were randomly selected among 30 melodies. Familiar melodies were either reversed or excluded. The verbal stimuli comprised ten nonword consonant-vowel-consonant (CVC) trigrams, sampled from Ampofo et al., (2023).

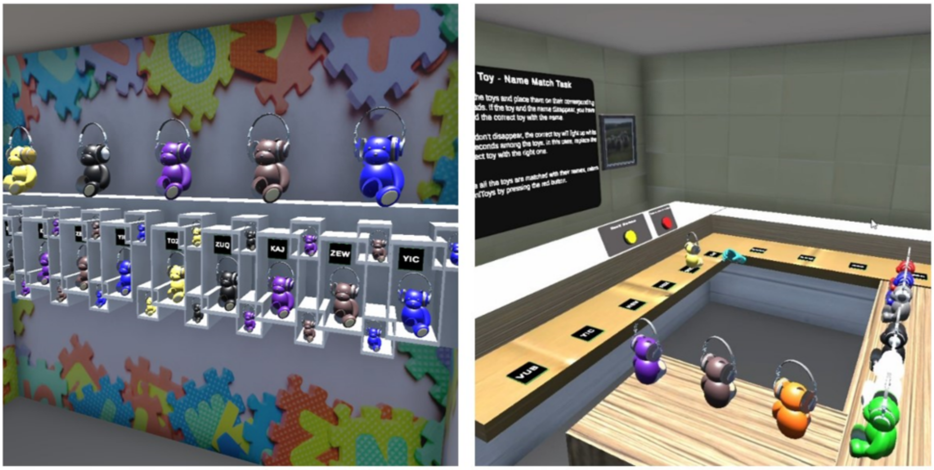
The environment was built using Unity software (2021.3.36f1) and deployed via Meta Quest 3 headset. It features a 4K+ Infinite Display (2064 x 2208 per eye), 120Hz refresh rate, built-in headphones for audio immersion, and hand/controller tracking. The headset was connected via Airlink to an HP Omen laptop (16 GB RAM, AMD Ryzen 7, NVIDIA RTX 4060).

2.3. Study Design and Procedure

This study employed a within-subject experimental design where participants completed object-name and object-sound matching tasks, earning one point for each correct match. The study consisted of two sessions (4 repeated trials per session) over two days in a week.

During the VR protocol activities, participants entered the virtual environment, which consisted of encoding and training rooms. In the encoding room (entry point), they observed toys with name tags and mentally associated each toy with its name. Afterwards, they entered the training room to complete an object-name matching task, where they were instructed to place the toys onto their corresponding name tags. Correct matches made the toy and tag disappear, while incorrect matches caused the correct toy to illuminate white among the toys, requiring participants to retry.

Next, participants returned to the encoding room to associate toys with their unique melodies by grabbing and holding each toy for ten seconds. Grabbing a toy triggered its sound, enabling sound-to-toy association. They then moved to the training room for the object-sound matching task, where they pressed buttons to play the melodies and placed the toys on pads corresponding to the buttons. Accurate matches made the toy and pad disappear, while incorrect matches illuminated the correct toy for retry.
Figure 1 depicts the encoding room and the object-name matching training room.



(A) (B)

Figure 1. (A) Virtual encoding room where participants explore toys and their names. (B) Incorrectly matching the yellow toy triggers the correct toy (pink) to illuminate white during the object-name matching task.

3. Results

Performance indexes of interest included learning over trials and response time. Using JASP (version 0.19.2), a repeated measures ANOVA was conducted to compare these indexes across all trials.

Learning over trials reflects cumulative performance improvement across trials. Dependent variables included the number of correct matches in object-name and object-sound matching tasks per trial. There was a significant effect of matching type, $F(1, 4) = 9.529$, $p = 0.037$, $\eta^2p = 0.704$, with object-name matching progressively out-performing object-sound in all trials, except in Trial 2, as indicated in Figure 2(A). Also, a significant effect of trial was found, $F(7, 28) = 5.687$, $p < 0.001$, $\eta^2p = 0.59$, with no significant interaction effect between the matching type and trial, $F(7, 28) = 1.33$, $p = 0.27$, $\eta^2p = 0.25$. Post hoc comparisons revealed a significant mean difference of 0.900 ($p = 0.037$), favouring object-name matching relative to the object-sound matching task. Also, pairwise comparison showed no significant differences in trials (all p-values > 0.05), indicating consistent performance across trials with no significant differences.

Response time refers to the time spent completing a task. Therefore, the time spent completing the object-name and object-sound matching tasks were used as dependent variables. The analysis revealed a significant effect in response time, $F(1, 4) = 88.528$, $p < 0.001$, $\eta^2p = 0.957$, with object-name response time consistently lesser and decreasing across trials compared to object-sound response times which had a similar decreasing pattern but higher response times, see Figure 2(B). Also, there was a significant effect of trial, $F(1, 4) = 9.225$, $p < 0.001$, $\eta^2p = 0.698$, showing a successive reduction in response times for both tasks across trials. No significant interaction effect was observed between response time and trial, $F(1, 4) = 0.502$, $p = 0.825$, $\eta^2p = 0.112$. Post hoc comparison showed faster response times for object-name matching than object-sound matching tasks ($p < .001$), but no significant differences were observed between trial pairs after Bonferroni correction (all p-values > 0.05).

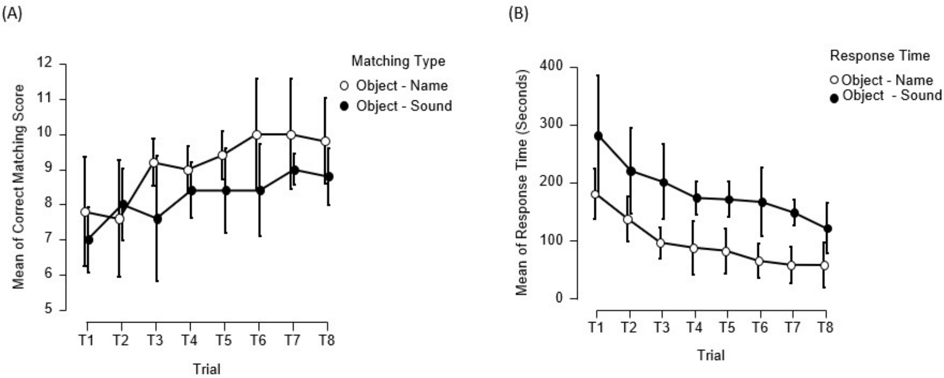


Figure 2. Mean of correct object-name and object-sound matching scores (A) and response time (B).

4. Discussion

This study designed, developed, and evaluated associative recognition memory using virtual reality (VR), comparing performance in object-name and object-sound matching tasks across multiple trials. Results indicated a progressive improvement in performance accuracy across trials and a decrease in response time, suggesting evidence of learning through repeated exposure. Notably, participants demonstrated significantly higher accuracy and faster response times in the object-name matching task compared to the object-sound task, supporting the hypothesis of the study. These findings are consistent with prior research indicating that verbal-visual pairings have higher learning accuracy and faster response time than audio-visual pairings.

The observed result can be attributed to the distinct processing mechanism underlying verbal and auditory stimuli. Phonological processing, which activates the temporal region, facilitates verbal stimulus associations by accelerating processing and reducing reliance on structural object processing, leading to faster and more accurate performance [10]. In contrast, auditory stimuli processing engages more complex cognitive processes, including auditory working memory and temporal integration, which demand greater effort and repeated exposure to establish robust connections [14]. The integration of auditory and visual stimuli further necessitates maintaining temporal coherence, which imposes a higher cognitive load and longer processing time.

Moreover, visual-verbal stimuli are processed instantly upon presentation whereas audiovisual recognition requires full stimulus duration for processing [12]. This difference accounts for the faster response time observed with the visual-verbal stimuli.

Furthermore, these findings could be attributed to using colours in distinguishing objects. Colour has been found to enhance object naming by facilitating semantic processing and reducing cognitive demands, reflected in the faster response time[12]. The addition of colour may have supported the processing of the object-name matching task relative to the object-sound matching task.

The findings of this study have practical implications for memory training and cognitive rehabilitation using VR. Associative memory training in VR could take advantage of repeated exposure techniques to improve memory performance. However, it is important to note that object-sound learning tasks may require longer training periods and additional support to achieve comparable results to object-name learning tasks.

The small sample size and lack of a control group limit the generalization of the findings of this study. Future research should address these by increasing the sample size, incorporating a control group, and including diverse populations such as older adults and people with memory impairments. Comparative analyses with other cognitive training methods and longitudinal studies are needed to assess effectiveness across cognitive domains and long-term impact as a rehabilitation tool.

5. Conclusion

VR integration in digital health has effectively enhanced associative memory, yet multi-trial training of verbal-visual and audio-visual associations remains understudied. This study investigated performance in these memory constructs and observed improved accuracy and reduced response time across trials. This study highlights the potential of VR for targeted cognitive training and its role in scalable digital health solutions.

References

[1] Y. K. Chan, Y. M. Tang, and L. Teng, "A comparative analysis of digital health usage intentions towards the adoption of virtual reality in telerehabilitation," *Int J Med Inform*, vol. 174, (2023).

[2] A. Singh Jat and T.-M. Grønli, "Virtual Reality in Digital Health: A Literature Review," in *Proceedings of the 7th International Conference on Digital Technologies in Education, Science and Industry (DTESI 2022)*, Almaty, Oct. 2022.

[3] N. Ribeiro, C. Sagnier, V. Quaglino, Y. Gounden, and E. Loup-Escande, "Effect of a short rest period on associative and relational memory performance: A Virtual Reality study," *International Journal of Virtual Reality*, vol. 20, no. 1, (2020), pp. 21–32, doi: 10.20870/ijvr.2020.20.1.3186.

[4] E. Tulving, "Episodic and semantic memory," in *Organization of Memory*, A. Tulving and W. Danaldson, Eds., Cambridge: Academic Press, 1972, pp. 381–403.

[5] S. A. Savage, O. Piguet, and J. R. Hodges, "Giving words new life: Generalization of word retraining outcomes in semantic dementia," *Journal of Alzheimer's Disease*, vol. 40, no. 2, (2014), 309–317.

[6] T. E. Gaston, S. Nair, J. B. Allendorfer, R. C. Martin, J. F. Beattie, and J. P. Szaflarski, "Memory response and neuroimaging correlates of a novel cognitive rehabilitation program for memory problems in epilepsy: A pilot study," *Restor Neurol Neurosci*, vol. 37, no. 5, (2019), pp. 457–468.

[7] V. Catania, F. Rundo, S. Panerai, and R. Ferri, "Virtual Reality for the Rehabilitation of Acquired Cognitive Disorders: A Narrative Review,", *Multidisciplinary Digital Publishing Institute (MDPI)*. (2024), doi: 10.3390/bioengineering11010035.

[8] J. M. Fernandez Montenegro and V. Argyriou, "Cognitive evaluation for the diagnosis of Alzheimer's disease based on Turing Test and Virtual Environments," *Physiol Behav*, vol. 173, (2017), pp. 42–51.

[9] S. Serino *et al.*, "A novel virtual reality-based training protocol for the enhancement of the 'mental frame syncing' in individuals with Alzheimer's disease: A development-of-concept trial," *Front Aging Neurosci*, vol. 9, no. JUL, Jul. (2017), doi: 10.3389/fnagi.2017.00240.

[10] J. Hocking and C. J. Price, "Dissociating verbal and nonverbal audiovisual object processing," *Brain Lang*, vol. 108, no. 2, (2009), pp. 89–96.

[11] M. F. Folstein, S. E. Folstein, and P. R. Mchugh, "Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician," *J. gsychiaf. Res*, vol. 12, (1975), pp. 189–198.

[12] J. Hocking and C. J. Price, "The influence of colour and sound on neuronal activation during visual object naming," *Brain Res*, vol. 1241, (2008), pp. 92–102.

[13] P. Ampofo *et al.*, "Cross-cultural evaluation of learning and memory using a consonant-vowel-consonant trigram list," *Journal of the International Neuropsychological Society*, vol. 29, no. 10, (2023).

[14] N. McLachlan and S. Wilson, "The Central Role of Recognition in Auditory Perception: A Neurobiological Model," *Psychol Rev*, vol. 117, no. 1, (2010), 175–196.