dHealth 2026
G. Schreier et al. (Eds.)

doi:10.3233/SHTI260067

Cross-Cultural Evaluation of Associative Memory Intervention Using Virtual Reality

Prince AMPOFO[a,1], Andrea VITALI[a], Francesca MORGANTI[a], Daniele REGAZZONI[a], Delali A. DOVIE[b], Joseph OSAFO[b] and Caterina RIZZI[a]
[a] *University of Bergamo, Bergamo, Italy*
[b] *University of Ghana, Accra, Ghana*

Abstract. Background: Associative memory is essential for daily functioning; however, VR interventions in this domain are limited. Objectives: The present study examined associative memory in different modalities and investigated the effects of culture on performance and response time in a virtual environment. Methods: Twenty-two healthy younger adults from Italy and Ghana (ages 18 – 26) completed eight training trial sessions on object-name matching and object-sound matching tasks in a virtual environment. Results: Performance improved progressively across trials in both tasks, with greater accuracy and faster response times in the object-name task than in the object-sound task. Cross-cultural comparisons revealed no significant performance differences between the Ghanaians and Italians. However, the Italians completed both tasks significantly faster than the Ghanaians. Conclusion: The findings of the present study indicate that associative memory is culturally invariant, whereas response-time differences may reflect cultural influences on processing strategies. Overall, this study supports the use of VR as a culturally adaptable platform for associative memory assessment and training.

Keywords. Virtual Reality, Associative Memory, Cross-cultural Study, Pair-associate learning

1. Introduction

The advent of digital health technologies has transformed cognitive assessment and interventions. Virtual reality (VR) has emerged as an effective cognitive training tool for creating interactive, realistic, and repetitive tasks in a controlled immersive environment, capabilities often limited in conventional 2D or paper-and-pen methods [1],[2]. Integrating VR-based interventions into digital health offers an opportunity to improve memory, promote independent living, and reduce healthcare system burdens [3].

Associative memory, a core component of episodic memory responsible for forming and retrieving links between multiple elements of an experience (e.g., a face and a name or an object and its corresponding sound), is particularly susceptible to age-related decline and is therefore a critical target for cognitive training [4], [5]. It enables individuals to connect perceptual information and navigate their environment to create meaningful experiences [6]. Associative memory involves integrating information across multiple sensory and cognitive systems, requiring encoding and binding of different stimuli into a coherent memory trace. According to Baddeley's Working Memory Model, these processes involve the visuospatial sketchpad for visual-spatial processing, the

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

phonological loop for verbal stimuli, the central executive for attention and control, and the episodic buffer, which integrates inputs into a unified memory trace [7]. While verbal stimuli are processed efficiently in the phonological loop, nonverbal auditory stimuli, which lack linguistic structure, place a greater demand on the central executive and episodic buffer for attention and cross-modal integration, thereby reducing performance [7]. VR-based associative memory training has been effective in improving memory in healthy individuals [5], [8] and in patients with dementia [2]. A pilot study comparing object–name and object–sound matching in VR reported better performance in the object–name condition, supporting the Working Memory Model [5].

Culture shapes cognitive processes, including associative memory, by influencing semantic networks, object familiarity, exposure to everyday sounds, and strategies for forming associations[9]. However, cultural effects are not universal across all binding processes [10]. Studies explicitly exploring the effect of culture on VR-based associative memory training are limited. This study compared performance in visual–verbal and visual–auditory tasks and examined cultural effects on performance and response time in young adults. Studying this population allows associative learning to be examined in a cognitively stable group without age-related confounds. A custom VR application with unique stimuli and an immediate correction strategy was developed to support the task.

2. Methods

2.1. Participants

Twenty-two students (12 Ghanaians, 10 Italians; aged 18-26) were conveniently recruited from the University of Ghana and the University of Bergamo via flyers and posters. An a priori power analysis using G*Power (α = .05, power = .90, effect size = 0.30), based on a preliminary study showing a larger effect (η^2p = .704), indicated a minimum sample size of 14 participants (7 per group) to detect an interaction effect. Only healthy young adults without physical, neurological, psychological, or sensory impairments were included; anyone with conditions affecting VR use or cognitive function was excluded. All participants scored above 29 on the Mini-Mental State Examination (MMSE) [11], which indicates normal cognitive function. Written Informed consent was obtained, and the study was approved by the Research Ethics Committees of the University of Bergamo (Project number: 2024_07_03) and the University of Ghana (Project number: ECH 282/24-25) and conducted in accordance with the standards outlined in the Declaration of Helsinki (2013 revision).

2.2. Materials and Technology

Materials included ten melodies downloaded from www.pixabay.com, 3D toy objects from www.turbosquid.com, and ten verbal stimuli, comprising nonword consonant-vowel-consonant (CVC) trigrams sampled from Ampofo et al. (2023). The melodies were randomly selected from a pool of 30 and trimmed to 10 seconds each, while a single 3D toy model was duplicated into ten colour-differentiated versions. The virtual environment was developed using Unity (2021.3.36f1) and deployed on a Meta Quest 3 headset via Airlink, connected to an HP Omen laptop (AMD Ryzen 7, 16 GB RAM, NVIDIA RTX 4060).

2.3. Study Design and Procedure

The study employed a mixed (within–between) design in which Ghanaian and Italian participants completed eight trials of object–name and object–sound matching tasks, with each correct match scoring one point. The virtual environment included two areas: an encoding room and a training room. The task was autonomously controlled by participants, allowing interaction with stimuli at their own pace, which increased ecological validity while maintaining accurate measurement of behavioural performance. During the VR task, participants first entered the encoding room to learn and associate the 3D toys with their names, as shown in Fig. 1(a), then moved to the training room to perform an object-name matching task by placing each toy on its respective name (see Fig. 1b). Correctly matched caused the toy and name to disappear, while incorrect matches caused the correct toy to light up white for retry. After matching all pairs, participants returned to the encoding room to learn toy–melody associations, with each toy emitting a unique 10-second melody when held (Fig. 1(c); at this stage, only the toys remained visible. Following the toy-sound encoding, participants performed an object-sound matching task in the training room by pressing buttons to play melodies and placing the toys on the corresponding pads, as shown in Fig. 1(d). Similarly, accurate matches made the toy and pad disappear, whereas incorrect matches triggered the correct toy to illuminate white for a retry.

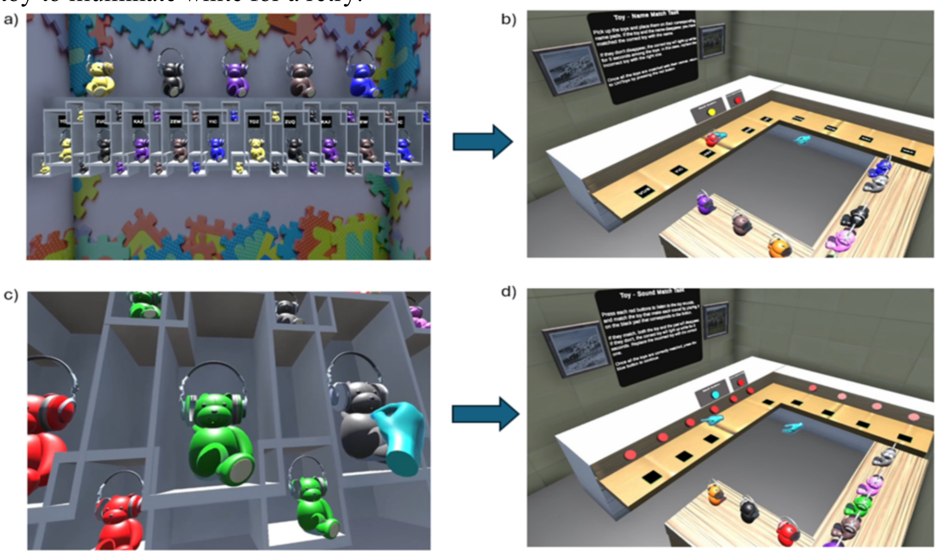

Figure 1. (a) Object–name encoding task. (b) Object–name matching task. (c) Object–sound encoding task. (d) Object–sound matching task.

3. Results

Two separate 2(Nationality; Ghana, Italy) × 2(Training Type; object-name, object-sound) × 8(Trial; trial 1-8) mixed design repeated measures ANOVA was conducted to analyse training performance and response times over trials. Training type and trial were within-subject factors, and nationality was a between-subject factor. All analyses were computed with JASP (version 0.95.04). Mauchly's test indicated that sphericity was violated for the within-subject factors and their interactions ($p < .05$); therefore, Greenhouse–Geisser corrections were applied to all affected effects.

Learning across trials was compared between the two training types, with the number of accurate matches as the dependent variable. Results showed a significant main effect of Training Type, $F(1, 20) = 15.06$, $p < .001$, $\eta^2p = 0.429$ and Trial, $F(3.53, 70.55) = 56.02$, $p < .001$, $\eta^2p = .737$, as shown in Fig. 2(a). Nationality had no significant effect, $F(1, 20) = 1.60$, $p = .221$, $\eta^2p = .074$, indicating comparable performance between the Italian and Ghanaian groups (See Fig. 3(a) and 3(c). The Training Type × Nationality interaction was not significant, $F(1, 20) = 3.58$, $p = .73$, $\eta^2p = 0.152$, nor were Trial × Nationality interaction, $F(3.53, 70.55) = 1.81$, $p = .144$, $\eta^2p = 0.083$, Training Type × Trial interaction, $F(3.98, 79.62) = 1.80$, $p = .137$, $\eta^2p = 0.083$ and Training Type × Trial × Nationality interaction, $F(3.98, 79.62) = 1.11$, $p = .357$, $\eta^2p = .053$. Pairwise comparisons showed higher performance in the object–name than in the object–sound matching task ($p < .001$). Post hoc analyses indicated a strong early learning effect, with performance in Trial 1 significantly lower than in all subsequent trials (all $p < .05$). Performance increased across Trials 2–4, whereas no significant differences were observed among Trials 5–8 (all $p > .05$).

Response times between object-name and object-sound matching tasks were compared, with task completion time across trials as the dependent variable. As depicted in Fig.2(b), the analyses revealed a significant main effect of Response Time, $F(1, 20) = 199.61$, $p < .001$, $\eta^2p = .909$, and main effect of Trial, $F(2.27, 45.33) = 57.42$, $p < .001$, $\eta^2p = .742$, indicating significant differences in completion times between the matching tasks and intertrial time changes across trials. Also, a significant effect of Nationality was observed, $F(1, 20) = 12.26$, $p = .002$, $\eta^2p = .380$. There were significant effects of Response Time × Nationality interaction, $F(1, 20) = 4.40$, $p = .049$, $\eta^2p = .180$ and Response Time × Trial interaction, $F(3.19, 63.70) = 2.73$, $p = .048$, $\eta^2p = .120$. However, significant effects were not observed in the Trial × Nationality interaction, $F(2.27, 45.33) = 1.51$, $p = .230$, $\eta^2p = .070$ and Response Time × Trial × Nationality interaction, $F(3.19, 63.70) = 1.72$, $p = .168$, $\eta^2p = .079$. Post hoc analysis revealed that Response times were faster in the object–name than in the object–sound task ($p < .001$) (see Fig.2(b)) and faster in the Italian than in the Ghanaian group ($p = .002$), with Italians responding faster across both tasks (all $p < .05$), as shown in Fig. 3(b) and 3(d). Response times decreased significantly across early and mid-trials (all $p < .05$) and plateaued in later trials (all $p > .05$). Trial-by-trial comparison revealed faster response times for the object-name task across all trial levels (all $p < .05$). Both groups showed a comparable learning curve; however, the Italians consistently completed the tasks faster.

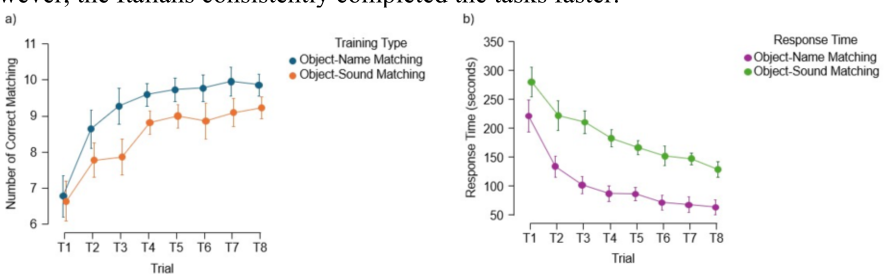

Figure 2. Means and 95% confidence intervals for matching performance (a) and response times (b) for object–name and object–sound tasks

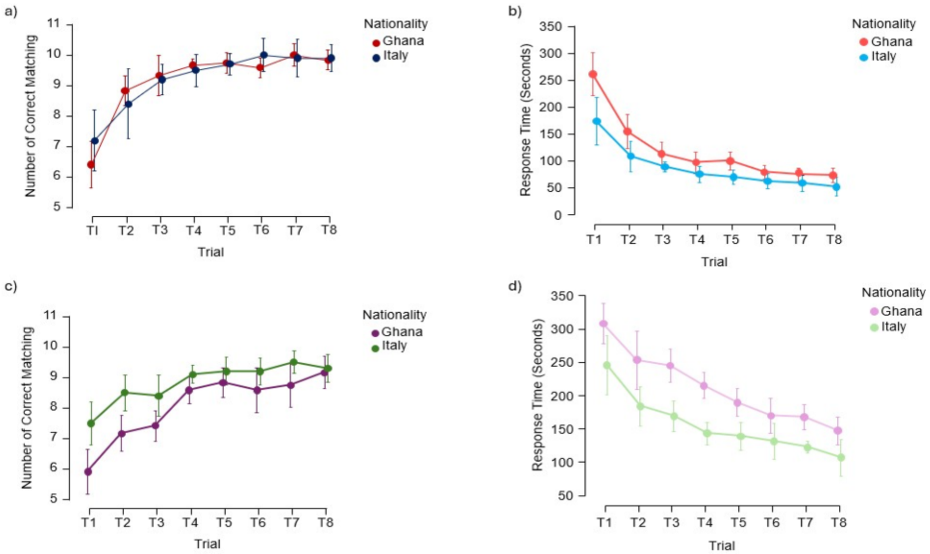

Figure 3. Means and 95% confidence intervals by nationality for object–name matching performance (a) and response times (b), and for object–sound matching performance (c) and response times (d).

4. Discussion

The study compared performance in visual-verbal (object-name) and visual-auditory (object-sound) associative tasks and examined cultural effects in a virtual environment. Results revealed progressive performance improvement and faster response times across trials in both tasks, with higher accuracy and quicker response times in the object-name task relative to the object-sound task. These findings are consistent with a previous study [5] and can be explained using the Baddeley Working Memory Model [7]. According to the model, verbal stimuli are efficiently encoded in the phonological loop, enabling rapid rehearsal and consolidation, whereas nonverbal auditory stimuli require greater involvement of the central executive and the episodic buffer for attention and cross-modal integration, potentially leading to reduced associative performance [7].

Cross-cultural comparisons of object–name and object–sound tasks revealed similar accuracy among Ghanaian and Italian participants; however, Italians performed slightly better on the object–sound task, although this difference was not statistically significant. Moreover, the Italians completed both tasks significantly faster than Ghanaians. Thus, while accuracy did not differ, Ghanaians required more time per trial. These results partially align with Cheung et al. (2024), who found no cultural effect in a colour–shape binding task, and concluded that feature integration may be culturally universal [10]. The absence of a cultural effect on accuracy may stem from the unfamiliar auditory and non-semantic verbal stimuli, providing no group advantage. The observed difference in response time may reflect attentional and visual processing patterns: individuals from Western, individualistic cultures, such as Italy, often show more focused attention and efficient search strategies, whereas those from collectivist cultures, such as Ghana, tend to adopt broader attention, requiring more processing time [13].

Several limitations should be noted, though the study provides evidence of cross-cultural cognitive training using VR: the lack of control groups, the inclusion of only two cultural groups, and the exclusive focus on young adults limit generalisation and attribution of the observed outcome solely to the VR training. Future studies should include control conditions and include more diverse cultural and age groups.

5. Conclusion

The findings of this study reveal modality-specific differences in associative memory, with superior performance on verbal relative to nonverbal auditory tasks, consistent with working memory theory. The lack of cultural disparities in performance and the presence of differences in response time indicate that associative memory mechanisms are independent of culture, whereas processing speed may vary across cultures. Overall, the findings support VR as a culturally adaptable tool for cognitive assessment and training.

References

[1] A. S. Rizzo and S. T. Koenig, "Is clinical virtual reality ready for primetime?," *Neuropsychology*, vol. 31, no. 8, pp. 877–899, Nov. 2017, doi: 10.1037/neu0000405.

[2] 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.

[3] D. Lupton, *Digital Health*. Milton Park, Abingdon, Oxon ; New York, NY: Routledge, 2018. |: Routledge, 2017. doi: 10.4324/9781315648835.

[4] A. R. Bender, M. Naveh-Benjamin, K. Amann, and N. Raz, "The role of stimulus complexity and salience in memory for face-name associations in healthy adults: Friend or foe?," *Psychol Aging*, vol. 32, no. 5, pp. 489–505, Aug. 2017, doi: 10.1037/pag0000185.

[5] P. Ampofo, A. Vitali, F. Morganti, C. Rizzi, and D. Regazzoni, "A Preliminary Evaluation of a VR Application for Associative Memory and Cognitive Rehabilitation," *Stud Health Technol Inform*, vol. 324, pp. 1–6, 2025, doi: 10.3233/SHTI250148.

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

[7] A. Baddeley, "Working memory: Theories, models, and controversies," *Annu Rev Psychol*, vol. 63, pp. 1–29, 2012, doi: 10.1146/annurev-psych-120710-100422.

[8] 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, pp. 21–32, Mar. 2020, doi: 10.20870/ijvr.2020.20.1.3186.

[9] H. F. Chua, J. Boland, and R. E. Nisbett, "Cultural variation in eye movements during scene perception," *Proceedings of the National Academy of Sciences*, vol. 102, no. 35, pp. 12626–12631, 2005, doi: 10.1073/pnas.0506162102.

[10] H. W. Cheung, N. Geeraert, and V. M. Loaiza, "Evidence Against Effects of Cultural Group and Prior Knowledge on Feature Binding in Working Memory," *J Cogn*, vol. 7, no. 1, 2024, doi: 10.5334/joc.390.

[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, pp. 189–198, 1975, doi: 10.1016/0022-3956(75)90026-6.

[12] 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, doi: 10.1017/S1355617723000656.

[13] A. Alotaibi, G. Underwood, and A. D. Smith, "Cultural differences in attention: Eye movement evidence from a comparative visual search task," *Conscious Cogn*, vol. 55, pp. 254–265, Oct. 2017, doi: 10.1016/j.concog.2017.09.002.