Natural Vs Unnatural: VR Hand Tracking vs Ray Interactors for Learning Retention

Ria Mukherji Southern
Methodist University
6425 Boaz Lane, Dallas, TX75205
rmukherji@smu.edu

Abstract

This Study investigates whether "natural" hand tracking or "unnatural" ray based interaction methods in Virtual Reality (VR) lead to better learning retention outcomes. While the VR industry commonly assumes natural interactions enhance user experience, prior research has shown conflicting results about effectiveness for learning. We implement a comparative study using a 3D object manipulation learning task with identical content presented through both interaction modalities. Through measuring task performance, cognitive workload (NASA-TLX). presence (IPQ), and knowledge retention through post-task assessments, this research provides insights into how interaction design affects cognitive processing in educational VR applications. Results contribute to understanding the relationship between interaction naturalness and information retention, with implications for designing effective immersive learning environments (/).

1. Introduction

The growth of VR in training, education, and simulation applications has raised questions about optimal interactive design for knowledge acquisition and retention. This research explores a counterintuitive finding in VR research - that less natural interaction methods sometimes outperform more natural ones in learning context.

1.1. Objectives

To measure and compare learning retention between users using hand tracking versus ray-based interaction (unnatural) in identical VR learning environments. Our primary research question examines whether the cognitive processing demands of different interaction methods impact how well users retain information.

1.2. Motivation / Novelty

As VR becomes increasingly embedded in educational contexts, understanding how interface design influences

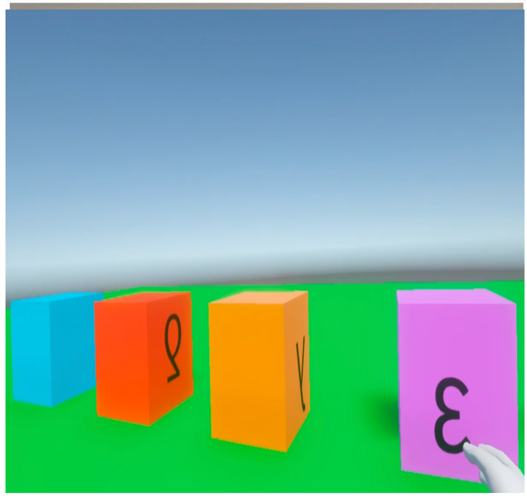

Figure I. Example of a modular memory game used in the study for short term recall and procedural memory in VR.

teaming outcomes is critical. Studies have shown that VR training can lead to a 75% retention rate compared to just 10% with traditional methods (2). However, specific impacts of interaction modality on retention remain underexplored. Studying different interaction methods in this contributes insights into designing intuitive VR systems that support cognitive retention informing both academic and industry VR development.

1.3. Background or Context

Research on VR interaction methods reveals several important considerations to consider when conducting this study:

I. Cognitive Load: Hand tracking reduces the mismatch between proprioception and visual feedback, which may improve spatial memory and retention (3). However, the impression of hand tracking may increase cognitive demands of complex manipulation tasks (4).

2. Leaming retention: Some studies indicate that "unnatural" interactions may reduce cognitive load by simplifying interactions, allowing users to focus on content rather than mechanics (1). A PwC study found VR can result in 14% faster learning time overall (2).

3. Interaction Trade-offs: Recent evaluations of interaction methods show that while hand tracking enhances embodiment and presence, ray interactors excel at precision and consistency, particularly for remote

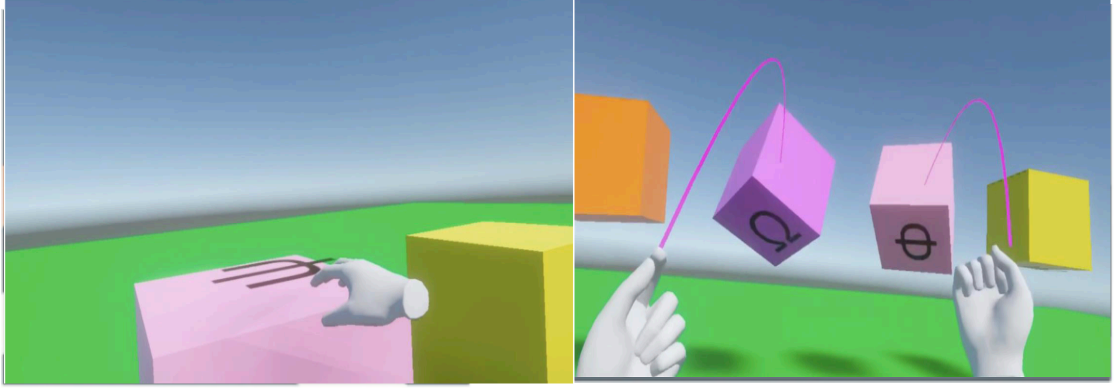

Figure 2. Comparison of interaction modalities used in the study. Left: Hand tracking allows direct manipulation of virtual objects. Right Ray-based interaction uses laser pointers to select to manipulate objects. Both were tested in identical learning environments.

interactions (4)(5).

4. Affect and performance: Research indicates that interaction method influences user affect (emotional state) with hand tracking making users feel more positive but less excited and less in control than when using controller-based methods.

2. Methods

2.1. Concept and user experience

We will build a VR learning environment in Unity where users explore six color-coded cubes, each labeled with a Greek letter on only one face. The environment supports two interaction modes:

I. Controller-based hand simulation (natural interaction using animated hand models)
2. Ray interactors using Unity XR Interaction Toolkit (unnatural interaction)

All users will hold controllers. The learning task is identical in both modes to allow a fair comparison of memory retention. To make sure the data is balanced users will alternate on whether they start the experiment using hand tracking, or the ray interactor version. After exploring the cubes in VR for 1.5 minutes, users complete a color-letter recall quiz, followed by a 2-minute distraction period and then a NASA-TLX and IPQ survey.

2.2. Hardware and Software

We will use Meta Quest headsets with standard VR controllers. In the "natural" condition, users interact with cubes via animated hand models that respond to trigger inputs, simulating hand movement. In the "unnatural" condition, users interact using ray pointers. Unity will be used for development, with the XR Interaction Toolkit supporting both interaction types. Post-task evaluation includes a color-letter recall quiz, NASA-TLX for cognitive load, and IPQ for presence.

2.3. Development plan

This project will be executed in four phases, guiding participants from baseline testing through immersive interaction and final evaluation. The overall objective is to assess whether interaction modality, either naturalistic hand-based manipulation or ray-based pointing, affects short-term memory retention within a virtual learning environment.

In the first phase, participants are introduced to the VR environment and told the premise of the experiment. After an explanation of what they will have to remember and how to operate in the VR environment, they are randomly assigned to one of two interaction conditions. In the "natural" condition, participants use animated hand models controlled through trigger-based input on standard VR controllers. This simulates grasping and object manipulation without relying on true optical hand tracking or the Meta SDK. In the "unnatural" condition, participants use a ray interactor to select and inspect objects from a distance. In both conditions, participants hold controllers throughout the experience to maintain consistency.

During the second phase, participants are given 1.5 minutes to explore the virtual environment containing six color-coded cubes, each labeled with a unique Greek letter. The Greek letter appears only on one face of the cube, requiring participants to rotate the objects using their assigned interaction method. This encourages spatial engagement and simulates a realistic recall-based learning scenario.

In the third phase, a brief distraction period (2 minutes) is introduced to reduce short-term memory priming. Immediately afterward, participants complete a computer-based multiple-choice quiz asking them to match each cube's color to its corresponding Greek letter. This assesses their ability to recall symbolic associations from the exploration phase.

The final phase involves completion of two standardized instruments: the NASA Task Load Index (NASA-TLX) to measure perceived cognitive workload, and the Igroup Presence Questionnaire (IPQ) to assess perceived immersion and spatial presence. These instruments allow for comparison not only of memory performance, but also of user experience across interaction modalities.

2.4. Assessment Plan

The study will use both objective and subjective evaluation methods to assess the relationship between interaction modality and memory retention. See section 3.Results for actual questions. The primary outcome variable is retention accuracy, measured by the proportion of correct letter-color matches on a post-task multiple-choice quiz. Task completion time and error rates will be recorded to evaluate performance efficiency and precision.

Subjective measures include the NASA-TLX, which captures mental demand, physical effort, temporal pressure, perceived performance, frustration, and exertion on a continuous 0-100 cale. Presence will be measured using the IPQ, which includes scaled statements on spatial presence, involvement, realism, and overall immersion for the final analysis.

2.5. Code repository

All project code, including Unity implementation of both interaction methods, will be maintained in a version controlled repository. The repository will be included documentation for replication of the experiment once completed.

3.Results

3.1 Memory Task Accuracy

To evaluate how well participants retained symbolic associations from the cube-matching task, we compared the number of correct answers (out of 6) for each interaction modality. (Actual data excel sheets and full graph analysis .ipynb can be found in the repository)

Ray Interactor:
Mean = 5.33, Median = 6.0, Mode = 6,
Standard Deviation = 0.98
Participants demonstrated high memory accuracy with the ray interactor. Most scored a perfect 6, and the low standard deviation indicates consistent performance across the group.

The bar graph clearly shows that participants using the ray interactor achieved higher and more consistent accuracy than those using the hand interactor. This suggests the ray-based interaction supported better memory retention for object-symbol pairings during the task.

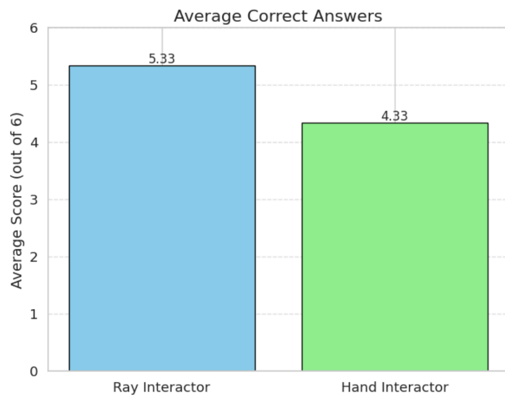

Figure 3. Results of Memory test

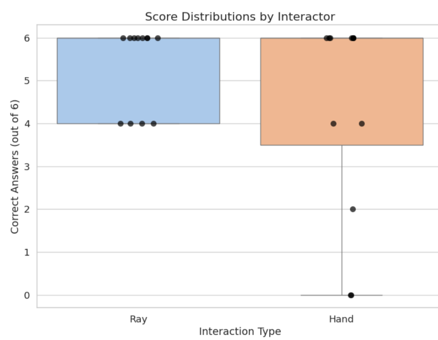

Figure 4. Score Distribution for Memory test

Results are based on 13 individuals most of which were between the ages of 20-25. Future testing would include a larger sample size distributed across more age ranges

Full graph and statistical analysis is provided in the ipynb within the associated repository.

Hand Interactor:
Mean = 4.33, Median = 6.0, Mode = 6,
Standard Deviation = 2.39
Although the median and mode matched the ray condition, the lower mean and higher variability suggest more participants made errors with the hand interactor.

3.2 IPQ (Igroup Presence Questionnaire)

Participants completed six IPQ questions for both interaction types. Below is a comparison of how participants rated presence, realism, and engagement using the Hand and Ray interaction methods.

Q1 – "I felt like I was really inside the virtual environment."
Hand Interactor: Mean = 8.17, Median = 8.0, Mode = 8.0, SD = 0.58
Ray Interactor: Mean = 6.83, Median = 7.0, Mode = 6.0, SD = 2.17
Participants reported a noticeably stronger sense of immersion using the hand interactor, with lower variability.

Q2 – "How real did the virtual world seem to you?"
Hand Interactor: Mean = 6.25, Median = 6.5, Mode = 7.0, SD = 2.01
Ray Interactor: Mean = 6.25, Median = 6.0, Mode = 5.0, SD = 1.76
Perceived realism was equal on average, though the hand interaction had slightly higher ratings at the median and mode.

Q3 – "I had a sense of acting in the virtual space, rather than operating something from outside."
Hand Interactor: Mean = 6.75, Median = 7.0, Mode = 7.0, SD = 2.80
Ray Interactor: Mean = 6.58, Median = 6.0, Mode = 7.0, SD = 2.57
Both interaction types scored similarly, with a slight edge toward the hand condition in perceived agency.

Q4 – "How aware were you of the real-world surroundings while navigating?"
Hand Interactor: Mean = 4.33, Median = 4.0, Mode = 3.0, SD = 2.06
Ray Interactor: Mean = 4.67, Median = 5.0, Mode = 3.0, SD = 2.25
A lower score indicates less awareness of the real world. Both scores were moderate, with no major differences between conditions.

3.3 NASA TLX (Task Load Index)

NASA TLX questions measured subjective cognitive load, stress, and performance across both interaction types.
Q1 – "How mentally demanding was the task?"
Hand Interactor: Mean = 4.17, Median = 4.0, Mode = 3.0, SD = 2.56
Ray Interactor: Mean = 4.08, Median = 4.0, Mode = 3.0, SD = 2.84
Both interaction methods yielded similar cognitive load ratings.

Q2 – "How physically demanding was the task?"
Hand Interactor: Mean = 4.25, Median = 4.0, Mode = 3.0, SD = 2.60
Ray Interactor: Mean = 2.42, Median = 1.5, Mode = 1.0, SD = 2.23
The hand interactor was rated as more physically demanding, suggesting users experienced greater exertion.

Q3 – "How hurried or rushed did you feel?"
Hand Interactor: Mean = 3.83, Median = 4.0, Mode = 5.0, SD = 2.89
Ray Interactor: Mean = 2.75, Median = 3.0, Mode = 1.0, SD = 2.19
Users felt slightly more hurried during the hand interaction.

Q4 – "How successful do you feel you were at completing the task?"
Hand Interactor: Mean = 7.0, Median = 7.0, Mode = 7.0, SD = 2.24
Ray Interactor: Mean = 7.58, Median = 8.0, Mode = 9.0, SD = 2.15
Participants rated their task performance higher with the ray interactor.

Q5 – “How hard did you have to work to accomplish the task?”
Hand Interactor: Mean = 5.17, Median = 5.0, Mode = 5.0, SD = 2.59
Ray Interactor: Mean = 3.67, Median = 3.5, Mode = 2.0, SD = 2.22
Users reported more effort was needed in the hand condition.

Q6 – “How irritated, stressed, or annoyed were you while trying to complete the task?”
Hand Interactor: Mean = 3.17, Median = 3.0, Mode = 1.0, SD = 2.64
Ray Interactor: Mean = 2.75, Median = 2.5, Mode = 1.0, SD = 2.45
Both conditions yielded low stress scores, though the hand interaction had slightly more variability.

4. Summary and Future Work

This study contributes to the field of VR educational design by comparing the effectiveness of hand tracking and ray-based interaction methods on memory retention. While both modalities offered valuable user experiences, results showed that the ray interactor led to higher and more consistent accuracy in symbolic recall, whereas the hand interactor produced stronger feelings of presence and embodiment.

These findings highlight the trade-offs between cognitive performance and immersive experience in VR environments. Future work could explore hybrid interaction approaches that dynamically combine hand and ray inputs to leverage the strengths of both. Additionally, long-term retention studies and broader participant samples would help validate the impact of these interaction styles on sustained learning outcomes.

5. References

[1] “Shaping the Future of VR Hand Interactions: Lessons Learned from Modern Methods,” *arXiv*, Apr. 2025. Available: https://arxiv.org/html/2504.00337v1

[2] “Implementing Virtual Reality in Training for Enhanced Learning,” *Vorecol*, 2023. Available: https://vorecol.com/blogs/blog-implementing-virtual-reality-in-training-for-enhanced-learning-9147

[3] “The Effects of Hand Tracking on User Performance: An Experiment Study of an Object Selection Based Memory Game,” *International Symposium on Mixed and Augmented Reality (ISMAR)*, 2023. Available: https://research.gold.ac.uk/id/eprint/33107/1/The_Effects_of_Hand_Tracking_on_User_Performance_an_experimentstudy_of_an_object_selection_based_memory_game_ISMAR_.pdf

[4] “Comparing and Combining Virtual Hand and Virtual Ray Pointer Interactions in Immersive Analytics,” *IEEE Transactions on Visualization and Computer Graphics*, 2022. Available: https://vvise.iat.sfu.ca/user/data/papers/stcinteraction.pdf

[5] K. Pfeuffer, B. Mayer, D. Mardanbegi, and H. Gellersen, “Gaze + Pinch Interaction in Virtual Reality,” in *Proc. of Symposium on Spatial User Interaction (SUI '17)*, Brighton, United Kingdom, Oct. 2017, pp. 1–10. Available: https://3dvar.com/Pfeuffer2017Gaze.pdf

[6] “Detecting Affect and Cognitive Load from Free-Hand Gestures in VR,” *arXiv*, Sep. 2024. Available: https://arxiv.org/html/2409.12921v1

[7] “Design of a Virtual Reality Serious Game for Experiencing the Colors of Cultural Heritage,” *Nature Scientific Reports*, vol. 14, no. 1, Oct. 2024. Available: https://www.nature.com/articles/s40494-024-01477-x