

MOTION AND MEMORY IN VR: THE INFLUENCE OF VR CONTROL METHOD ON MEMORIZATION OF FOREIGN LANGUAGE ORTHOGRAPHY

Noel H. Vincent, Zhen Liang, and Yosuke Sasao

Graduate School of Human and Environmental Studies, Kyoto University,
Japan

ABSTRACT

A small-scale investigation of 22 first-time VR users was conducted to determine the extent to which different control methods ("Touch Mode" and "Point Mode") influenced participants' ability to memorize the educational content (Hebrew orthography) in an original VR game. In Touch Mode, participants selected user interface (UI) elements by physically moving their hands (and upper bodies) toward targets in virtual space, whereas in Point Mode, participants could manipulate a virtual laser pointer with their wrists to select UI targets from a stationary position. The Point Mode control method required less time to learn on average than the Touch Mode with high-medium effect size (Cohen's $d = 0.76$). Ultimately, the difference in post-test scores between Touch Mode users and Point Mode users was not statistically significant ($p = .294$). Thus, the increased bodily motion necessitated by Touch Mode did not have a significant influence on memory function as initially hypothesized.

KEYWORDS

Virtual Reality (VR), Educational Software, Foreign Language Education, Neuropsychology, Interaction Design

1. INTRODUCTION

Following the introduction of affordable consumer VR headsets like the Oculus Rift and HTC Vive in the latter half of the previous decade, researchers have been exploring the educational applications of immersive VR hardware and software. With the changing of Facebook's name to Meta and its announcements regarding "the metaverse," there is renewed enthusiasm for all things VR-related. However, some researchers understandably doubt the practical benefits of educational VR and wonder how it differs from other computer-based learning methods. Like desktop-based applications, VR applications can utilize real-time 3D graphics and stereo sound, but they can also take advantage of motion-based controls and stereoscopic vision. However, there is ongoing debate as to whether these novel features of VR contribute to learning in a meaningful way.

1.1. Literature Review

Regarding the topic of motion control, an investigation by Fuhrman et al. found that participants in their study were able to better retain Finnish words that they had learned by pantomiming with

3D digital props in VR (e.g., saying the word for “hat” while grasping and donning a 3D hat)[1]. This type of learning activity recalls the Total Physical Response (TPR) method first proposed by James J. Asher in the late 1960s, in which language instructors act out words or phrases while performing an associated action (e.g., jogging in place while saying “run”)[2].

This avenue of research suggests that pantomiming words or phrases may aid in memorizing foreign vocabulary, but the exact mechanism facilitating this remains unclear. One possibility is that the benefits derived from these pantomiming activities are due to the neurophysiological effects induced by cardiovascular activity. There is a wide body of research suggesting that cardiovascular activity prior to study helps learners to better absorb study content [3][4][5], especially when the intervening time duration is short [6]. This line of thinking could be referred to as the “kinetic hypothesis” (i.e., that any benefit derived from pantomiming activities is explained by the physiological effects of bodily motion). Alternatively, one could postulate a “semantic hypothesis,” in which the benefits of gestural learning are the result of meaningful associations formed between the motor function and language processing centers of the human brain [7][8][9].

Thus, a key question is whether the findings reported by Fuhrman et al. and modern adaptations of TPR [10] are best explained by excitation of the cardiovascular system, or by a process of semantically mapping words onto physical actions across various brain regions. Disentangling this question is of key importance, as it could inform whether content designers take a more kinetic or semantic approach to motion-controlled VR learning software.

Some evidence against the kinetic theory can be inferred from the research of Hartfill et al. [11], who developed an educational adaption of the popular VR game, *Beat Saber*, called *Word Saber*. In this rhythm game, users use virtual lightsabers to slash at blocks which approach from the horizon in sync with music. At higher levels of play, motion becomes quite intense, and many people use the game as a form of cardiovascular exercise. The researcher’s game functioned similarly, except that players had to slash at blocks displaying foreign language vocabulary. While participants reported enjoying *Word Saber*, the VR activity was found to be less efficient for learning vocabulary than conventional flashcard study.

The VR game developed for the present study is similar to the *Word Saber* game developed by Harfill et al. in terms of functionality, with users having to select the correct answer using VR motion controls. However, this game incorporates two control methods (*Point Mode* and *Touch Mode*)—the latter of which requires substantially more physical movement than the former. Thus, if the score of the memory post-test were substantially higher for the group using *Touch Mode* controls, then it can be inferred that cardiovascular excitation made some contribution to learning, whereas if the scores for users of this control method are similar or lower, it would indicate that cardiovascular excitation did not play a significant role, or that the physical exertion required to play this VR game was of insufficient quantity or quality to elicit the effects reported in the neurophysiology studies.

1.2. Research Questions

The primary questions this research sought to investigate are as follows:

1. Which of the two VR control methods included in this study (*Touch Mode* or *Point Mode*) is most readily learned by first-time VR users?
2. To what extent does the control method used during VR play influence users’ ability to recall educational content from their play session?

3. What influence, if any, does increased bodily motion have on users' ability to recall educational content from their play session?

2. METHOD

The present study involved the development and testing of a custom VR learning game designed in Unity software with the aid of the VRChat software development kit (SDK). The software consists of a learning activity intended to help users memorize the characters (glyphs) of a foreign language (in this case, Hebrew). The completed VR software was tested by 26 students of Kyoto University, who used the software for a period of 40 minutes before completing a post-test to measure their short-term recall of the phonetic readings of 27 Hebrew letters. The Hebrew alphabet was selected for use in this experiment because it is unfamiliar to most Japanese and contains relatively few characters. Additionally, its letters do not change shape (e.g., by forming ligatures) or have context-dependent positioning (as in Korean writing).

2.1. Participant Demographics

Students applied to join the experiment via a Kyoto University job posting website. The advertisement offered a 1000-yen gift card as compensation for participation in the experiment. From this applicant pool, only students with zero prior experience using immersive VR were selected to participate, to ensure that prior familiarity with motion controllers did not bias research outcomes. For purposes of recruitment, "immersive VR" was defined as virtual reality hardware including both a stereoscopic headset and motion-tracked hand controllers. In total, 26 students of Kyoto University participated in the study. The majority were undergraduates ($n = 20$), while the remainder were MA ($n = 4$) and PhD ($n = 2$) students. All participants had some familiarity with English, and none had prior experience learning the target language, Hebrew. The majority of participants were native speakers of Japanese ($n = 24$). Two participants were native speakers of Mandarin Chinese. The gameplay activity log was unable to be recovered for 4 participants (due to loss of network connection), resulting in a final sample size of 22.

2.2. Software Development Considerations

The VR software was purpose-built for this research using the Unity game engine and VRChat's software development kit (SDK). Unity can be used to create and publish games to a wide variety of platforms, including Windows, MacOS, web browsers, and VR headsets (such as the Meta Quest 2 used in this experiment). Implementing a VR movement system and UI from scratch is challenging, so VRChat's SDK for Unity was used to create an interactive game that runs within VRChat. The resulting software is a sort of game within a game (Meta, indeed).

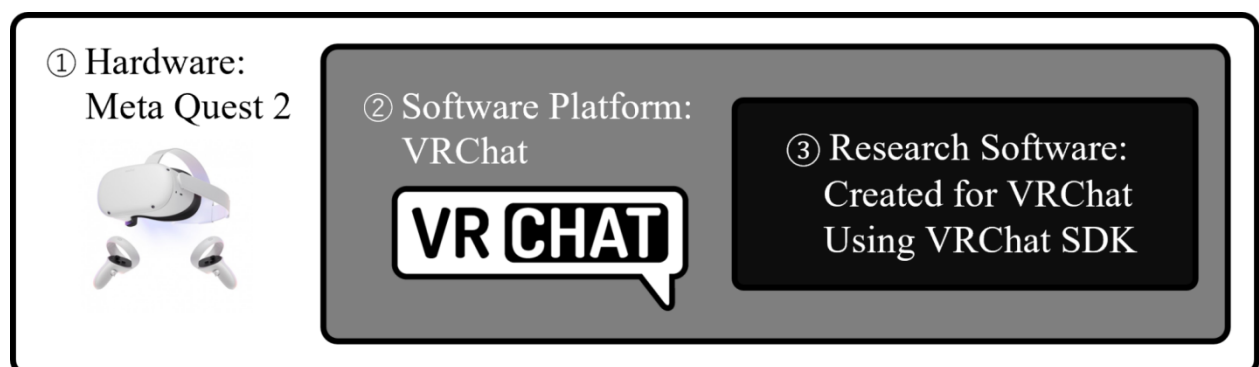


Figure 1. A diagram of the platform and hardware on which the experimental software runs

In the Unity engine, game logic is typically defined using scripts written in the programming language C#, which is similar in many respects to Java. However, importing the VRChat SDK into Unity enables use of a proprietary scripting language called Udon. Udon can be programmed visually, by connecting nodes (representing events, functions, data types, etc.), or by using a written variant called Udon#. The VRChat SDK also includes “prefabs” (prefabricated components) for basic game functionality, such that simple environments can be created without the need for any programming. Finished content can be uploaded to the VRChat platform, where avatar movement and network communication is handled by VRChat—greatly reducing the work required to create interactive content for VR.

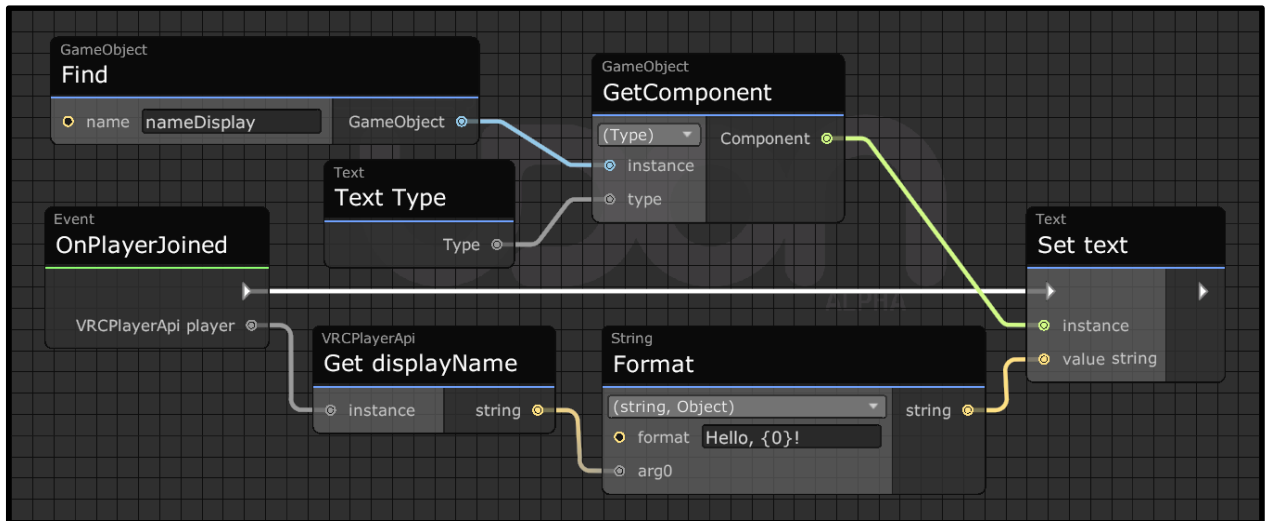


Figure 2. An example of a node graph created with the Udon visual programming language

Despite its ease of use, however, there are some drawbacks to using the VRChat SDK for VR content development. This is because VRChat deliberately disables certain native capabilities of the Unity engine that might lead to unauthorized use of its game platform, including functions or modules that allow for communication with external websites. Additionally, VRChat does not currently allow for the storage of persistent save data in uploaded worlds. This made the collection of participants' in-game play activity somewhat challenging.

As a workaround to these limitations of the VRChat SDK, player statistics (such as play duration, answer accuracy rates, etc.) were compiled in a visual log that could be captured via screenshot after participants finished their play sessions. This method of data collection requires an experimenter to manually put on each headset and perform a screenshot command. Unfortunately, there were three instances in which the network connection to a participant's headset was lost. This caused VRChat to eject the user from the game and resulted in a complete loss of play data. Additionally, one participant's data had to be discarded because the log contained too many events and the overflowed text could not be captured in a screenshot.

2.3. Description of User Interface and Control Methods

The VR game software incorporates two control methods: *Point Mode* and *Touch Mode*. *Point Mode* makes use of the industry standard interaction method [12] for user interfaces in VR and is available by default in the VRChat SDK. *Touch Mode*, by contrast, is a less commonly used interaction method that was implemented with custom code for purposes of this research.

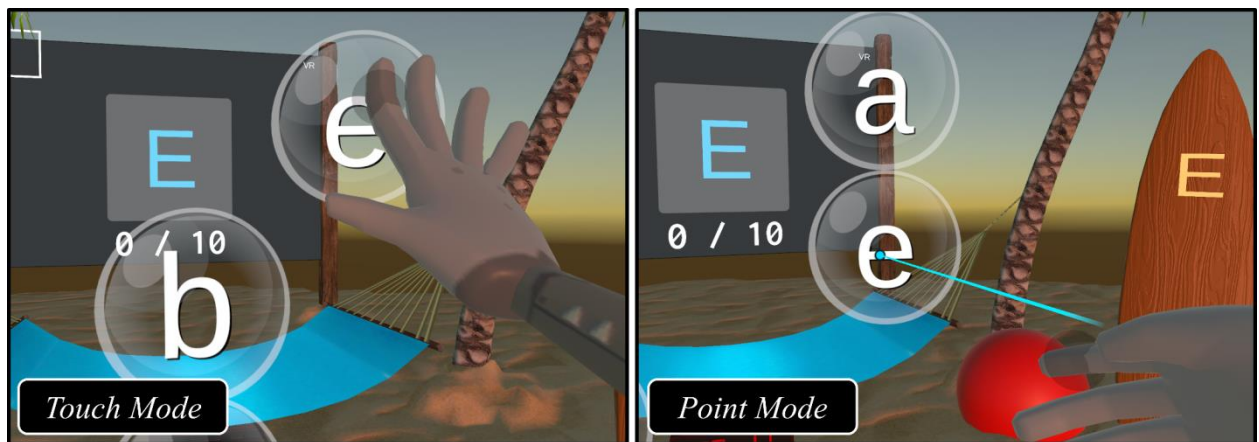


Figure 3. Images of the game's tutorial when used in *Touch Mode* (left) and *Point Mode* (right)

When *Point Mode* is active, the user can select UI elements by aiming a virtual laser beam at the target and squeezing the main trigger on the hand controller. This works ambidextrously. The visibility of the laser beam is context-dependent—appearing only when the controller is oriented toward an interactable object. This interaction method is employed in most VR applications, especially when interacting with menu screens containing numerous and or densely-packed UI elements. This interaction method is analogous to moving and clicking a mouse cursor, but in 3D space. Similar to using a mouse, it requires minimal bodily movement (rotating one's wrist is generally enough to reach all UI elements).

When *Touch Mode* is active, the user can select UI elements by directly reaching out toward a target in virtual space. This also works ambidextrously. In contrast with *Point Mode*, *Touch Mode* requires a considerable amount of bodily motion, such as leaning forward, extending one's arms, or even walking towards a distant UI target. The *Touch Mode* as implemented in the present VR software does not offer any contextual clues (such as the laser beam visible in *Point Mode*) regarding which UI elements can be selected by touching.

While currently not as prevalent as the point and squeeze method, Meta is currently experimenting with touch-based UI interactions. In its developer guidelines documentation, it recommends that touchable UI elements should respond to hand proximity with visual or tactile feedback (e.g., having the UI target grow in size or making the controller vibrate). The inclusion of this recommendation in Meta's developer guidelines indicates that these contextual clues are important for helping users determine which UI elements are interactable via touch [12].

2.4. Description of Gameplay

Participants are given 40 minutes (measured by an in-game timer) to learn the phonetic readings of the 27 letters of the Hebrew alphabet through a series of matching activities within the VR game. To avoid overwhelming participants, the 27 letters are divided into 5 “challenges” consisting of 5 to 9 letters. Completion of one challenge unlocks the subsequent one, until all challenges are unlocked and the player can access a comprehensive quiz of all 27 letters. After all challenges have been unlocked, the player can repeat previous challenges until 40 minutes has elapsed.

Participants begin the VR game in one of the two interface modes (*Touch Mode* or *Point Mode*) as selected in advance by the researcher. This mode remains active for the first 30 minutes of the

experiment, and then automatically switches to the other mode for the remaining 10 minutes. In this way, all participants gain some experience with both modes.

All interactable elements in the game's UI were made to resemble bubbles, so that once users learned to interact with one UI element, they would easily recognize subsequently encountered bubbles as interactable UI targets. The menu consists of a series of bubbles arranged into rows, which when selected, reveal subdirectories of additional bubbles (see Figure 4).

Within the software's code, quiz content is stored as question-answer pairs (e.g., 1 : "one", 2 : "two", 3 : "three", etc.), in which the first component is the foreign language symbol and the second is its phonetic representation in latin characters (e.g., א : "a").

During a challenge, a prompt is displayed on a screen in the background of the VR environment, accompanied by an audio clip indicating how the prompt is pronounced (see Figure 4). A prompt can consist of either the foreign language symbol, or its phonetic representation. Along with the prompt, 5 bubble graphics appear at randomized locations in the foreground and each contain an answer choice. If the prompt displays the foreign language symbol component of the question-answer pair, then the bubbles will display the phonetic equivalent (and vice versa).

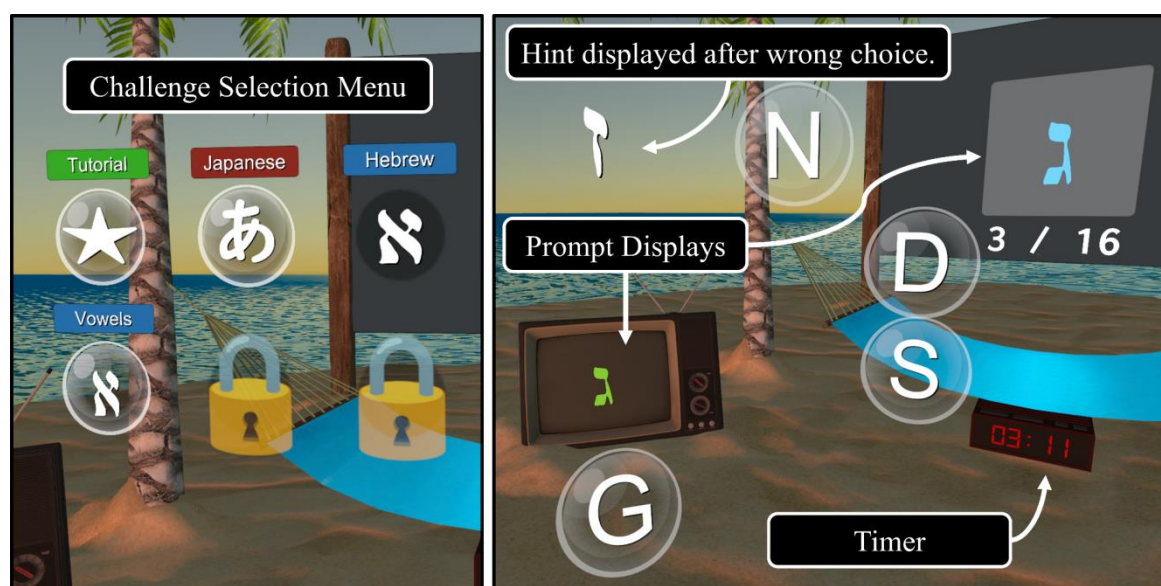


Figure 4. The menu UI for selecting challenges (left) and the Hebrew letter “gimel” displayed as a prompt with 4 of 5 answer bubbles still available for selection (right)

When the selected answer corresponds to the prompt, one point is awarded to the player and the game updates to display a new prompt and answer choices. If the selected answer is incorrect, a point is deducted and a “hint” is displayed consisting of the corollary to that choice (e.g., if the answer bubble initially displayed “a”, it would display “N” and vice versa), along with a pronunciation cue. This repeats until the player selects the correct answer.

To complete a challenge, participants must achieve a certain point threshold by selecting correct answers. The threshold is automatically calculated to be two times the number of question items (e.g., if there are 5 question items, the point threshold will be 10). Selecting an incorrect answer results in a deduction of 1 point, unless the point value is already 0.

The presentation of prompts in a challenge alternates between the set of foreign language symbols and the set of phonetic equivalents. Because the player is assumed to be unfamiliar with the foreign symbol component of the question-answer pair, the software is designed to begin a new challenge by first displaying the complete set of foreign language symbols as prompts (with the phonetic equivalents displayed on bubbles). In this way, the player can first look at the foreign symbol as a prompt, hear the audio cue and then select the bubble whose phonetic equivalent seems to correspond to the audio prompt. In the inverse case, the player would have no choice but to select answers at random. The presentation order within these sets is randomized. For instance, given a collection of 5 question-answer pairs, the presentation of prompts might be ordered as shown in Table 1.

Table 1. An example of the presentation order of prompts displayed during a challenge

Cycle Count	Display Type	Example Ordering
Set 1	(Symbols)	2,3,1,5,4
Set 2	(Phonetic)	one, three, four, two, five
Set 3	(Symbols)	3,2,5,4,1

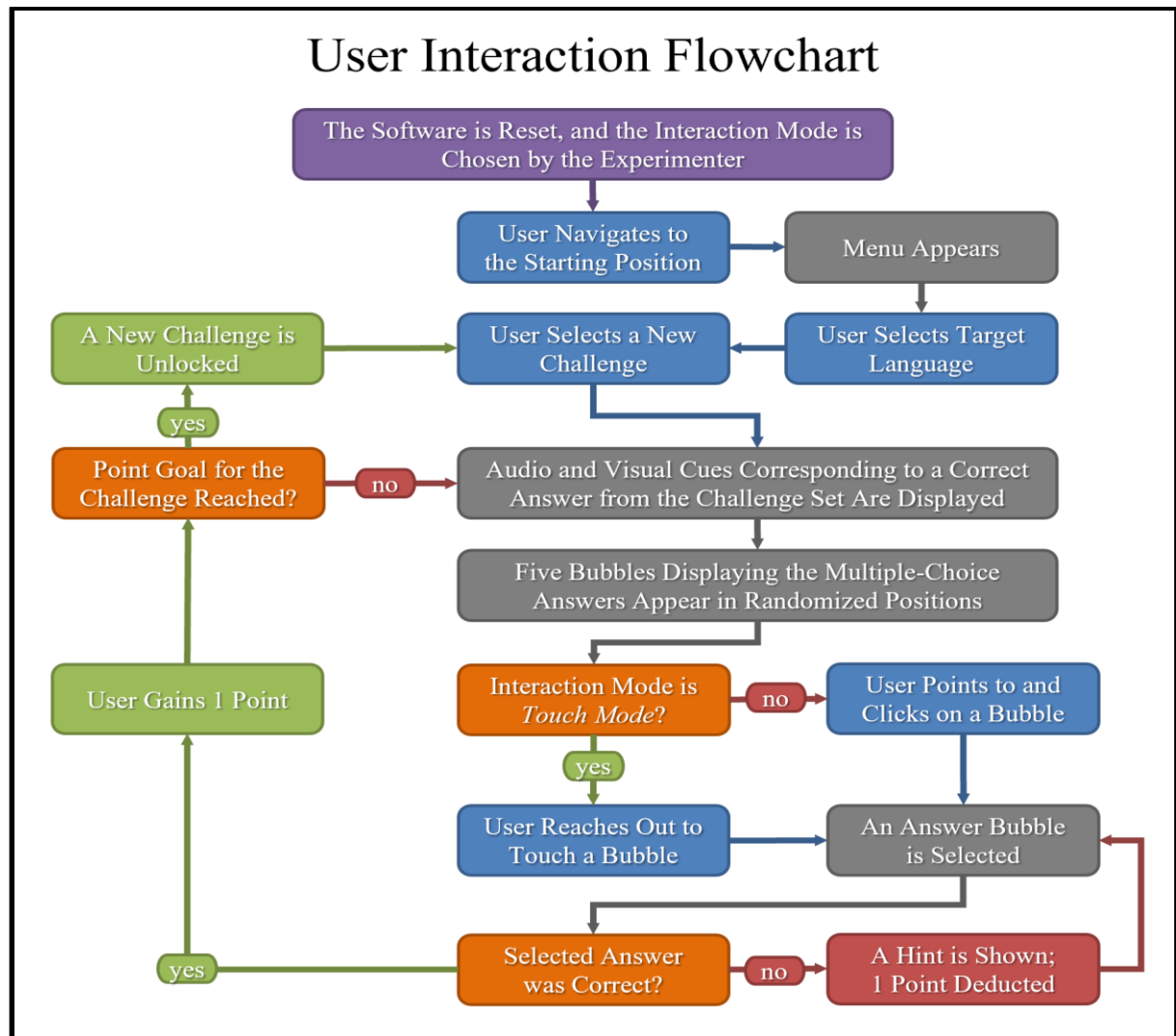


Figure 5. A diagram representing the flow of game logic within the experimental software

2.5. Description of VR Hardware

This investigation made use of Meta Quest 2 VR headsets (initially released under the name Oculus Quest 2 prior to Facebook Inc.'s rebranding as Meta Platforms, Inc.). These devices are ideal for an experimental setup with multiple simultaneous VR users because they feature an "inside-out" tracking system that eliminates the need for external tracking sensors. Additionally, the Quest 2 is a standalone device that can function independently of a desktop computer. Not only is this more cost-effective than purchasing computers and VR hardware separately, but the standalone nature of the device means that participants can experience VR without the encumbrance of cables. One shortcoming of standalone devices, however, is that they can quickly deplete their internal batteries. This issue can be mitigated with the use of portable USB battery packs.

2.6. Experimental Setup and Procedure

The experiment was conducted in eight, 60-minute sessions over the course of two days. There were as many as 4 participants per experimental session. Each participant was assigned to one of 4 VR stations consisting of a desk, a chair, a Meta Quest 2 VR headset with hand controllers, charging accessories, and a portable battery for emergency charging. These corridor-shaped stations were positioned adjacent to one another across the length of the testing location. The VR stations were made sufficiently wide to ensure that participants would not accidentally collide with one another when reaching out their arms. The boundaries between stations were programmed into the VR headsets using the Quest 2's Guardian system.

Before each experimental session, the VR headsets were (re)initialized by checking the virtual boundaries for each station, repositioning the game avatar in the starting position and orientation, resetting volume controls and selecting the control method (*Touch Mode* or *Point Mode*) that participants would use at the beginning of their play sessions. All participants present for a given time slot began with the same control method, and the initial control mode was alternated for each new group of participants. After initialization, headsets and controllers were connected to charging adapters and sanitized using disinfecting wipes.

Upon arrival at the testing location, participants were signed consent forms and received a debriefing in Japanese regarding use of the VR controllers and what to expect during their play session.

In the debriefing, participants were told the following:

- 1) The game consists of learning letters of the Hebrew alphabet.
- 2) There will be a quiz following the play session.
- 3) This experiment will test what controls are intuitive for first-time VR users.
- 4) Specific instructions for controlling the game will not be given.
- 5) Please attempt to learn the controls on your own.
- 6) Do not touch the controllers until instructed.
- 7) You will need to use the left controller stick to move your character.
- 8) During gameplay, do not touch any of the circular buttons.
- 9) You may try squeezing the triggers.

From a seated position, participants were asked to put on their VR headsets and adjust the headset straps. Participants were asked to shake their heads laterally to see whether the headset was sufficiently fastened to their head. Excessively loosened or tightened straps were adjusted by the experimenter.

With headsets properly adjusted, the experimenter placed a pair of VR controllers in each participants' hands, reminding participants to refrain from touching any inputs (buttons, triggers, etc.). Controllers in hand, participants were then told to stand up and begin moving the stick on their left controller to navigate their avatar to the starting point within the VR game (indicated by a rotating 3D arrow).

Participants were then monitored to see whether they could successfully learn the game controls. If participants were still appearing to struggle after approximately 3 minutes, the experimenter offered hints until the participant was able to figure out the control method.

Once all participants learned to control the menu, they were monitored to ensure that they did not collide with other players or otherwise injure themselves. Additionally, the experimenter took note of interesting behavior exhibited by participants. After 30 minutes, the software would automatically switch from one control method to the other (e.g., from *Touch Mode* to *Point Mode*). As before, participants were given hints after struggling for a sufficiently long time, or after asking for assistance.

After 40 minutes, the game automatically ends whatever challenges the participants may have been engaged in and displays a message instructing participants to remove their headset. To ensure data capture, the experimenter put on each headset to take screenshots of the gameplay log (made visible via a debug menu). These screenshots are saved to the local storage of the VR device in a pictures directory for later retrieval.

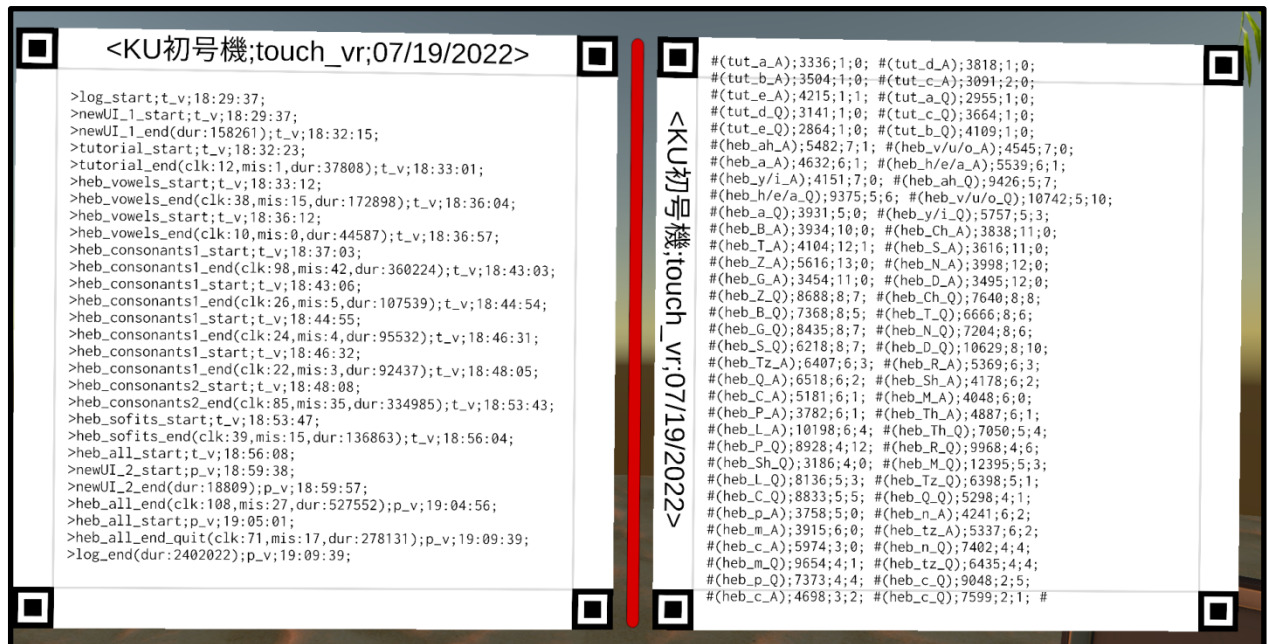


Figure 6. A typical data log captured from the VR software featuring game events (left) and individual letter statistics (right)

After capturing data from all VR devices, the experimenter directed participants via QR code to a Google Form containing a post-test related to what they learned during their VR session. The post-test was formatted as a multiple-choice quiz in which participants had to match Hebrew letters to their appropriate phonetic equivalents, as they practiced in the VR software. Participants were given until the end of their 60-minute time slot to submit their quizzes. Once they had

submitted answers to their post-tests, participants were compensated for their time with 1000-yen gift certificates and allowed to leave the testing location.

Log data collected during the experiment was retrieved from the VR headsets' local storage via USB cable. Then, using a Python script, image files were grouped by participant name using the image timestamp and device ID. The sorted images were then viewed by the researchers and key statistics were manually recorded in a spreadsheet.

3. RESULTS

The following subsections present several comparisons between the group of users who began their VR session using *Touch Mode* and those who began their session using *Point Mode*. All data were analyzed by independent sample *t*-tests to examine the difference between the two modes. Additionally, Cohen's *d* metric is provided as a means of estimating effect size whereby ($d = 0.2$) is considered a small effect, ($d = 0.5$) medium, and ($d = 0.8$) large [13].

3.1. Relationship Between Control Method and Post-Test Scores

The maximum score for the post-test was 27 points (corresponding to the number of letters in Hebrew). The mean score for participants who began the experiment with *Touch Mode* was 16 points, whereas the mean for those who started with *Point Mode* was 2.63 points higher at 18.63 points. This difference, however, is not statistically significant ($t(10) = -1.08$, $p = .294$, $d = 0.459$).

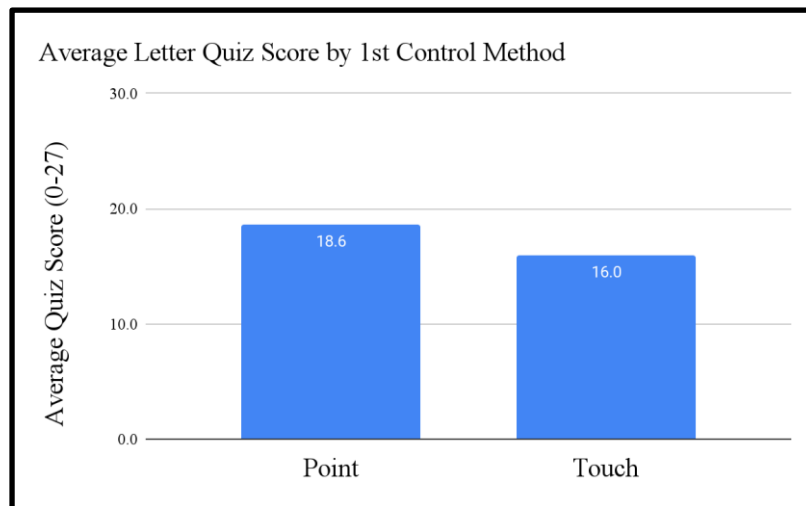


Figure 7. Mean score as grouped by starting control method

3.2. Time Required to Learn a Control Method

The game software automatically logs the duration between certain events that occur during the course of gameplay, such as the time between when a player begins and ends a challenge. The following data concerns two crucial time durations:

1. *Time to Learn 1st Control Method:*
(time from when a participant first touches the left controller stick until the participant selects a UI target in the challenge select menu)

2. *Time to Learn 2nd Control Method:*

(time from when the software switches to the second control mode until the user interacts with a UI element)

The mean time required to learn the game controls when the first control method was *Touch Mode* was 260.7 seconds versus a mean time of 152.1 seconds when the first method was *Point Mode* (a difference of 108.6 seconds). ($t(10) = 1.77, p = .091, d = 0.76$)

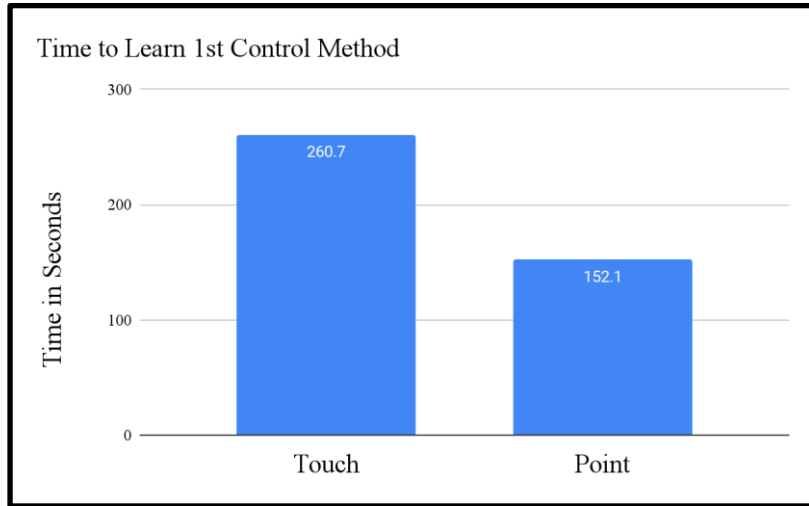


Figure 8. Time required to learn the initial control method

The mean time required to learn the game controls when the second control mode was *Touch Mode* was 74.1 seconds versus a mean time of 55.1 seconds when the second mode was *Point Mode* (a difference of 19.0 seconds). ($t(10) = 0.54, p = .592, d = 0.23$)

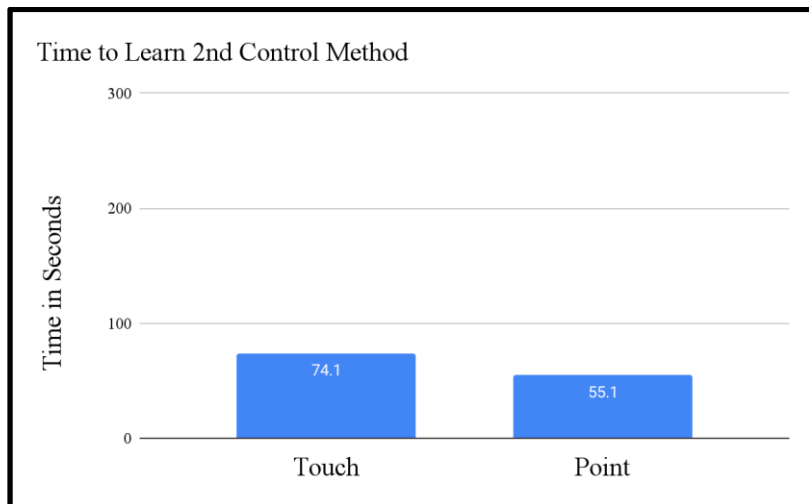


Figure 9. Time required to begin using the second control method

4. DISCUSSION

The following subsection discusses investigative findings as they relate to the research questions initially stated in the *Introduction* section.

1. Which of the two VR control methods included in this study (*Touch Mode* or *Point Mode*) is most readily learned by first-time VR users?

The present study found that *Point Mode* was able to be learned more quickly on average than *Touch Mode*. This constitutes the most statistically significant finding of the present study, with a p value of .091 and a Cohen's d value of 0.76 (high-medium effect size).

Initially, it was hypothesized that *Touch Mode* might require less time to learn than *Point Mode*, since selection during *Touch Mode* does not require the use of any buttons. In reality, it took nearly twice as long on average for participants to discover how to select objects using *Touch Mode*. There are several possible explanations for this result.

First, given that these participants had no prior experience with immersive VR hardware, their intuitions regarding how to use the controllers were likely informed by use of conventional gamepads. Thus, the notion of physically extending one's hands to interact with menus may not have initially occurred to participants. While some participants are likely to have had experience with some variety of motion-controlled gamepad (such as the "Wiimote" packaged with the Nintendo Wii game console), these types of devices typically only measure rotation, as opposed to position in 3D space. In the case of the Wiimote specifically, its motion controls were commonly utilized for aiming a cursor on a television screen. Thus, prior experience with this category of motion controller may have helped some participants quickly intuit the control method in *Point Mode*.

A second explanation concerns contextual feedback. During *Point Mode*, a virtual laser beam would appear when the avatar's hand was oriented toward a selectable object. The conspicuous appearance and disappearance of this laser beam likely helped participants to infer that selecting objects was related to hand rotation. Conversely, the *Touch Mode* control method, as implemented in the experimental software, does not provide any contextual feedback to users (based on proximity to the UI target, etc.).

Despite the relatively large difference in time required to learn *Touch Mode* versus *Point Mode* at the beginning of a VR session, there was relatively little difference between the average times required to learn the second control method ($t(10) = 0.54$, $p = .592$, $d = 0.23$). It seems, therefore, that additional control methods can be learned with comparably little effort once the user has grown acclimated to moving within the VR environment. Moreover, commercial VR games are likely to display explicit instructions to the user on how to control the software—rendering the initial "intuitiveness" of the control method largely irrelevant.

In light of these findings, therefore, it is difficult to recommend one method of control over the other based solely upon the amount of time required for a new user to independently deduce its functionality. Instead, developers of VR content should select control schemes based upon other criteria particular to the application.

2) To what extent does the control method used during VR play influence users' ability to recall educational content from their play session?

The average difference in post-test scores between users beginning with *Touch Mode* and users beginning with *Point Mode* was insignificant ($p = .294$) with a small effect size ($d = 0.459$). Thus, the influence of these control methods on participants' ability to memorize the educational content was minimal. Nonetheless, users who began in *Touch Mode* scored 2.63 lower on average. It is uncertain whether a larger-scale experiment would yield a similar result. However,

at least one explanation for this lower average can be inferred: *Touch Mode* took significantly longer to learn than *Point Mode*. Consequently, the average amount of time that participants who began with *Touch Mode* could dedicate toward memorizing the educational content was reduced.

3) *What influence, if any, does increased bodily motion have on users' ability to recall educational content from their play session?*

The control method that required the most physical movement to use was *Touch Mode*. As mentioned above, the average difference in post-test scores between *Touch Mode* users and *Point Mode* users was below the threshold of statistical significance, but with a slightly lower average score for the *Touch Mode* users. It can be reasonably inferred that the slight underperformance of *Touch Mode* users was related to having less practice time overall, since the *Touch Mode* controls required more time to learn on average than *Point Mode* controls. These results echo the findings of Hartfill et al., which found that while VR motion controls did not negatively impact the *efficacy* of language learning tasks, they resulted in lower *time efficiency* than traditional flashcard study. While the present study does not have the benefit of comparing against a flashcard study control group, it is likely that such a comparison would have yielded similar results.

While motion-controlled VR applications may be less time efficient for short-term study, this may not hold true over longer timespans. The experiment described in Fuhrman et al., which involved teaching foreign language vocabulary through object manipulation in VR, included both an immediate and delayed post-test, as well as a control group that watched VR objects without any interaction. Though the recall accuracy of vocabulary items was similar across these groups on the immediate post-test, the object-manipulation group achieved a higher score on the (1 week) delayed post-test than the control group. This research suggests that learning activities incorporating VR motion control may promote long term recall.

5. CONCLUSION

5.1. Summary

Though initially hypothesized that the incorporation of bodily motion into a conventional learning activity might have a positive influence on participants' ability to memorize in-game content, the insignificant difference in average post-test scores between the *Touch Mode* users and *Point Mode* users indicate that the inclusion of "motion for motion's sake" confers little, if any, educational benefit. Indeed, the data presented here lends credence to the idea that the benefits reported by investigations into gesture-based learning methods are not merely the product of random kinetic motion.

5.2. Future Directions

Advancements in technology can often outpace ideas regarding how to use them. Early motion pictures, for instance, were just uncut stageplays filmed from a fixed position, and it was not until filmmakers began splicing film reels and varying camera positions that the medium's true potential began to be realized.

The same can be said of virtual reality. When the prototype version of the Oculus Rift VR headset was first made available, hand controllers were not yet available and developers focussed on adapting existing computer games for use with a VR headset. Then, as hand controllers became a staple of consumer-level VR technology, design conventions shifted toward

incorporating this control scheme—resulting in new types of gameplay that were previously difficult or impossible using conventional gamepads.

With institutional interest in VR technologies steadily gaining momentum in recent years, the question for researchers is no longer “if” but “how” virtual reality can be used for educational purposes. This presents an opportunity for educational content creators to devise new ways of learning, as well as opportunities for repurposing previously devised methods in this new medium—particularly those incorporating bodily motion or navigation of 3D spaces.

The present research demonstrates some challenges with adapting conventional learning methods for a virtual reality context, and suggests some interesting directions for future research. If the incorporation of bodily motion in and of itself does not contribute meaningfully to learning outcomes, further investigation is required to see how motion-controls can best be utilized for educational purposes. Toward this end, gesture-based learning applications—where the motion of users is semantically related to the educational content, rather than merely a feature of gameplay or menu navigation—warrant further exploration.

REFERENCES

- [1] Fuhrman, O., Eckerling, A., Friedmann, N., Tarrasch, R., & Raz, G. (2019). The moving learner: Object manipulation in VR improves vocabulary learning. doi:10.31234/osf.io/wgzxu
- [2] Asher, J. J. (1969). The Total Physical Response Approach to Second Language Learning. *The Modern Language Journal*, 53(1), 3. doi:10.2307/322091
- [3] Roig, M., Nordbrandt, S., Geertsen, S. S., & Nielsen, J. B. (2013). The effects of cardiovascular exercise on human memory: A review with meta-analysis. *Neuroscience & Biobehavioral Reviews*, 37(8), 1645-1666. doi:10.1016/j.neubiorev.2013.06.012
- [4] Statton, M. A., Encarnacion, M., Celnik, P., & Bastian, A. J. (2015). A Single Bout of Moderate Aerobic Exercise Improves Motor Skill Acquisition. *Plos One*, 10(10). doi:10.1371/journal.pone.0141393
- [5] Winter, B., Breitenstein, C., Mooren, F. C., Voelker, K., Fobker, M., Lechtermann, A., . . . Knecht, S. (2007). High impact running improves learning. *Neurobiology of Learning and Memory*, 87(4), 597-609. doi:10.1016/j.nlm.2006.11.003
- [6] Roig, M., Thomas, R., Mang, C. S., Snow, N. J., Ostadan, F., Boyd, L. A., & Lundbye-Jensen, J. (2016). Time-Dependent Effects of Cardiovascular Exercise on Memory. *Exercise and Sport Sciences Reviews*, 44(2), 81-88. doi:10.1249/jes.0000000000000078
- [7] MacSweeney, M., Campbell, R., Woll, B., Giampietro, V., David, A. S., McGuire, P. K., . . . Brammer, M. J. (2004). Dissociating linguistic and nonlinguistic gestural communication in the brain. *NeuroImage*, 22(4), 1605-1618. doi:10.1016/j.neuroimage.2004.03.015
- [8] Schippers, M. B., Gazzola, V., Goebel, R., & Keysers, C. (2009). Playing Charades in the fMRI: Are Mirror and/or Mentalizing Areas Involved in Gestural Communication? *PLoS ONE*, 4(8). doi:10.1371/journal.pone.0006801
- [9] Schippers, M. B., Roebroek, A., Renken, R., Nanetti, L., & Keysers, C. (2010). Mapping the information flow from one brain to another during gestural communication. *Proceedings of the National Academy of Sciences*, 107(20), 9388-9393. doi:10.1073/pnas.1001791107
- [10] Wang, F., Hwang, W., Li, Y., Chen, P., & Manabe, K. (2019). Collaborative kinesthetic EFL learning with collaborative total physical response. *Computer Assisted Language Learning*, 32(7), 745-783. doi:10.1080/09588221.2018.1540432
- [11] Hartfill, J., Gabel, J., Neves-Coelho, D., Vogel, D., Räthel, F., Tiede, S., . . . Steinicke, F. (2020). Word saber. *Proceedings of the Conference on Mensch Und Computer*. doi:10.1145/3404983.3405517
- [12] Best Practices | Oculus Developers. Developer.oculus.com. (2021). Retrieved 11 September 2022, from <https://developer.oculus.com/resources/hands-design-bp/>.
- [13] Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.