Effects of Visual Stimuli on Memory Performance in Virtual Environments

Shreya Shankar*   Mark Steelman†

ABSTRACT

One of the most typical memorization techniques in education is the use of flash cards, which is a way of associating terms with their meanings. In our experiment we plan to compare different associative memorization techniques of varying levels of presence within a virtual environment in order to assess the improvement of memory performance with more interactive memorization techniques that correspond to higher presence fidelity. Within the same virtual environment, we will compare meaning association with terms and varying interaction levels. Our hypothesis was that it will become easier to memorize terms as the user progresses through higher levels of presence and interactivity. We collected both quantitative and qualitative data on the memory performance of 24 subjects and found that there was not any consistent pattern indicating that added visual stimulus necessarily correlates with improved memory performance. However, we made some interesting observations about how the visual stimulus affected the participants performance, with results varying by individuals.

1 INTRODUCTION

Virtual reality has the potential to significantly improve the learning experience by immersing students in their education, and as part of understanding how this can be done it is important to understand the ways in which virtual reality can improve and enhance the memorization process of educational content. So far, there are few educational tools that provide an interactive and engaging experience, and most of them are in 2D platforms which inherently limits their potential interactivity and induced presence. The motivation behind this project is to create a learning platform for educational institutions that acts as a tool to make educational experiences more enjoyable while increasing memory performance. For almost every subject in education, teachers teach the same content in the same medium year after year and we believe that it would be more effective to translate this content into a virtual environment. There are many unique affordances of VR that provide the option for experiential and more natural learning processes at lower costs. To go further, there is a lot more that could be built on top of such a platform such as integrating online multiplayer capabilities so that you could have an unlimited number of students all in the same virtual classroom from anywhere in the world. We believe that gathering evidence for the baseline capabilities of VR for enhancing memory performance is an essential starting point for eventually building a robust VR educational platform. The purpose of this research is to understand whether virtual reality environments are more effective for memorization to be used as a learning tool in educational settings. The results of this research could drastically influence the field of virtual reality and its capabilities as a memorization tool within learning platforms.

2 LITERATURE REVIEW

There have been many research projects done that explore the effects of using Virtual Environments (VE's) as a learning tool. In regards to education, the range of the uses of virtual environments have varied greatly. For example, there have been projects that have explored the effects of interactivity in virtual environments within the education space [5]. In regards to this specific research project, it relied on qualitative methods to form conclusions. Important findings included realizing that virtual environments allowed participants to remain more focused because they felt immersed and present in the virtual world as compared to the real world (where participants felt that they could be easily distracted by the world around them). However, another factor for us to consider when designing our study is the use of text, especially because we are testing the effects of virtual environments on memorization. From this study, the researchers concluded that text was harder to read in a virtual application and that a subject with lots of text to read would be most suitable for 2D interfaces, while subjects that require simulations and 3D representations would be a better fit for virtual environments. As our study targets developing memorization tools for concepts in virtual reality, we feel that a virtual environment will still be effective in improving memorization techniques. Though we started with the idea of using 3D models in virtual environments as an improved memorization tool, we have found projects that provide interesting insight into other techniques that could potentially be more effective. Ragan, Sowndararajan, Kopper, and Bowman [8] discuss the values and limitations of supplementing conceptual information with spatial information in educational virtual environments. Although this study focuses on the effects of immersion and we are interested in the effects of presence, it was still useful to understand different ways in which memorization could be tested. Participants in this study memorized procedures in a virtual environment and then attempted to recall those procedures. This study was helpful for us to detail ways in which memorization within virtual environments could potentially be tested. Other studies focused on memorization in the context of spatial awareness within virtual environments, specifically how virtual environments aid spatial memory recollection [1]. The conclusions from this study indicated that spatial and object memory resulting from interactions with the virtual environment were enhanced only for aspects of the environment directly involved in the interaction, specifically, the spatial layout through which participants were required to navigate. An important aspect to consider with memory testing is time allocation for recollection, as it plays a huge factor in the results. In the Brooks, Attree, Rose, Clifford, Leadbetter study [1], participants were given five minutes to draw their recollection of the layout of the virtual reality rooms. This study also indicated that enhanced memory recollection only occurred in the participants that were active and navigated the virtual environment with the joystick, and had negligible effects on passive participants who observed the actions of the active participants.

An essential component to the formation of our research topic is that VR technology has the power to facilitate learning. Studies have shown that the application of VR technology in education enriches teaching and learning in the current education model. Virtual learning environments (VLEs) improve students abilities to analyze problems and explore new concepts. When immersive, interactive, and imaginational aspects are also integrated, the virtual learning space allows learners to model, act, and express anything they want as long as the environment provides the tools to do so [7]. This study defines four principle components of VLEs: Knowledge Space, Communication Community, Active Action, and Facil-

*e-mail: shreya.shankar@duke.edu
†e-mail:mark.steelman@duke.edu

ity Toolkit. These four principles will be helpful in guiding us in the direction of creating a holistic teaching and learning tool that follows the guidelines of virtual learning environments. Furthermore, this study provides insight into interaction enhancing active learning, which is consistent with other studies that have been conducted within the virtual environment space. From these studies, strong conclusions can be drawn that interaction in virtual environments is necessary for learning in VLEs to be successful. Interaction can come in a variety of forms, including more passive interaction such as the inclusion of additional sensory inputs/outputs that have been shown in other studies outside of virtual reality to improve memory as a result of the emotional and other psychological states that other senses trigger. Dinh, Walker, Song, Kobayashi, and Hodges [4] investigated the effects of tactile, olfactory, and audio sensory cues on the sense of presence and memory in a virtual environment, finding a positive correlation between increased sensory modalities and the users sense of presence and memory for objects in the VE. Interestingly, while increasing the fidelity of the three aforementioned senses increased memory and presence, increasing visual detail did not correspond with increased memory or presence. There have been previous predictions of such observations (such as Ivan Sutherlands) about visual realism and that it would not necessarily contribute to increased presence, so we will probably not focus too heavily on the visual quality of our experimental design as long as it is minimally convincing. The study used a series of VEs representing typical areas of an office building, and used presence and memory questionnaires to collect data about the users presence and sense of spatial memory. These questionnaires will be useful to look at when designing our own data collection methods.

Another important aspect crucial to the success of our study is how information within a virtual environment will be presented. As we would like to explore VEs within any educational context, the aim is to be able to use this virtual environment to help facilitate learning of abstract or complex symbolic concepts (potentially related to the space). Bowman, Hodges, Allison, Wineman [2] conducted a study to compare learning between traditional lectures and classroom material enhanced with the use of a virtual environment in the context of a zoo exhibit. The participants were given tests on material relating to specific information about the design of the zoo exhibit habitat. Students who had their learning augmented by the virtual environment performed better on the test, suggesting that students were able to draw on their experiences in VEs in educational settings where virtual environments were used to augment learning. From studies like this one, it is intuitive that virtual environments are better at creating associations between spatial and abstract information, and they add a strong experiential component to educational settings. An important factor in this study was the time after the study in which the test was given - students were unaware that they were participating in an experiment and that they would be tested, and the test was given five days after students were presented the content so that they could internalize the material as they would in a classroom setting, and so that too short of a time post the experiment wouldnt serve as a confounding variable. We want to make sure that we are testing for learning and memory in a way thats consistent with traditional paper-and-pencil measures, and one useful assessment that was found to meet this criteria is the Virtual Reality Cognitive Performance Assessment Test [10]. The VRCPAT focuses on refined analysis of neurocognitive testing that assesses the users recalling targets in a virtual city environment. In this study, the VRCPAT was compared to standard neuropsychological measures as a way of validating the VRCPAT as an appropriate method for measuring neuropsychological effects. The results of the study showed some conflicting results. On the one hand, the VRCPAT correlated significantly with the traditional neuropsychological learning composite as well as the traditional neuropsychological memory composite. However, there were no significant correlations between the VRCPAT and the neuropsychology test composites for executive functions, attention, processing speed, or verbal fluency. For the sake of our project, since there was a significant correlation with learning and memory we may want to draw from the VRCPAT in the design of our memory/learning assessment.

Intuitively it makes a lot of sense that VR would be an excellent medium for a learning environment due to its engaging and interactive nature, but there is still limited knowledge about precisely which features of VR provide the most opportunity for enhanced learning. The best approach to education is also very dependent on what is being taught, so in using the features of VR that can enhance education we must also find the ways to adapt that usage to different educational contexts. All of this is indicative of the complex nature of education and the even more complex task of combining the medium of VR with complex education. Salzman, Dede, Loftin, and Chen [9] have tried to identify, use, and evaluate VRs affordances to facilitate the learning of complex and abstract concepts. They present a general model describing how these affordances along with certain factors in learning work together. The major lessons that they came away with were that various features in VR increased motivation for learning, different features are appropriate for different concepts (there is no general rule), there are various trade-offs to be made, and individual learner characteristics - such as gender, domain experience, spatial ability, computer experience, motion sickness history, and immersive tendencies - are important. It is also important to also remember that the educational benefit of a VE is emphasized not in the novelty of the environment, but in its ability to teach relationships between different pieces of information. Wickens [11] proposes that the educational benefit of a VE comes from the learner being able to draw relationships between existing pieces of information and abstract material. These relationships, seen by students in various forms, will aid students in recollecting the information they experience within virtual environments. Experiencing information in different forms and having the ability to explore and connect the relationships between pieces of information should give the student more opportunities to later retrieve the information.

As designers of the virtual environment that will be used for this study, we recognize that the success of the study also depends on the usability of the virtual environment. Wickens [11] argues that the naturalness of the interface must be increased to reduce the cognitive effort required by the user to navigate and interpret the VE. By aspiring to make the interface as natural as possible, we avoid the possibility of participants experiencing the cognitive overload of processing new material while learning how to navigate a new environment in unfamiliar ways. As Wickens points out, we want to mitigate any distractions from learning due to usability issues of the interface. Unfortunately it is possible that some of these potential distractions and added complexities are inherent to the use of virtual reality for cognitive measurement, as suggested by meta-analysis on various studies assessing cognitive processes using virtual reality and analogous classical or computerized assessment tools of the same process [6]. In this meta-analysis the results of a random effects model indicate a moderately increased overall task difficulty in VR, demonstrating a poorer cognitive performance in VR compared to traditional assessment. However, their findings do not directly correlate with any objectively defined factors of VR that would be the cause of such cognitive performance, but rather theorize that the causes are factors such as stressors, distractions, and complex stimuli that add more information for the user to process while still completing assessment tasks. As we have already discussed, it is possible to mitigate these effects through good design and by controlling certain factors, but we should still be wary that we may not have full control over these effects.

Since the motivation for our research is to work towards a widely applicable educational tool to use VR to enhance any and all relevant topics in education, it is important that we consider the attributes of VLEs in relationship to educational theory and pedagogical practice. Bricken [3] discusses several key aspects of VR that have inherent ties to educational theory, including the experiential nature of VR, the natural interaction with information that is afforded, the possibility of shared experiences, and the flexibility of VR to be tailored to individuals. In education there are almost always multiple paths to achieving learning outcomes, but usually certain paths are more efficient and effective than others, and VR can easily assist in providing access to those better paths when otherwise unavailable. For instance, there are many situations in which abstract methods are used to describe or teach inherently experiential or physically interactive concepts like architectural design where there are usually insufficient resources to learn this through non-abstract methods. Many educational institutions struggle to provide sufficient individual attention to students due to the lack of teachers or resources, but since VR inherently collects individual data that can be analyzed (especially higher quality VR), the barrier to individual tailorship can be significantly reduced. However, despite the many affordances of VR, there are still various challenges such as information accuracy (sometimes difficult to translate content accurately into VR), fear of overburdening the teachers, lack of historical perspective, and other issues that have yet to even be discovered.

3 Hypothesis

Our overall hypothesis was that an increase in the level of immersion would allow for better performance in the memory recall task. Our hypothesis was motivated by our desire to understand how virtual reality could be used to enhance memory performance in educational environments. This hypothesis was influenced by the fact that the experiential nature of virtual reality and the natural interaction with information that is afforded provide a solid foundation for using VR as an educational tool (Bricken). This hypothesis was also formed because we knew that virtual reality has the capacity to allow users to draw relationships between existing pieces of information and abstract material (Wickens). These relationships that will take form in the virtual environment and be seen in various levels of immersivity will aid students in recollecting the information they experience within virtual environments. Therefore, experiencing the information in various forms and having the ability to explore and connect the relationships between pieces of information should give the user opportunities to later retrieve the information. In this study, we chose to investigate the effects on short term memory only, since we realize that memorization evaluation research is limited to effectively testing working memory. Therefore we recognize that working memory has space and duration limitations when dealing with new, unfamiliar information because it can only hold 5 to 9 elements for about 20 seconds and process only two to four elements at a time. Processed information in working memory is only stored in long-term memory if it is passed through the sensory stage. By extension, information will not be processed or stored if it is not engaging, meaningful, and emotionally compelling enough to pass the sensory memory stage. Even with this understanding, not all educational materials are designed to enable learners to store and retrieve all essential information, causing significant losses to education. In the present study, we investigate the use of VR to ignite learners brains with visual stimuli. By focusing on the visual stimuli, we believe that the information will pass the sensory memory stage and be processed in working memory, therefore making it more likely to be stored in long-term memory. Additionally, being in a virtual environment will cause an increase in brain activity that will then allow the learner to focus on and internalize the environment, and make the learning experience incredibly realistic, natural, and memorable.

Although memory can be influenced by a variety of senses within virtual environments, in this study we chose to focus on visual stimuli. We hypothesized that increasing levels of visual stimuli would facilitate increased memory recall amongst learners. By using abstract terms and attributes associated with each term, we were able to reduce the influence of external or existing knowledge. We also chose to randomize the order in which the participants completed the recall questionnaire so as to negate any effects of spatial memory.

4 Experimental Design

In this experiment, participants were asked to complete three memorization tasks. Each memorization task consisted of eight made-up characters, with each character having four attributes: hometown, age, favorite food, and occupation. The names of the characters were fictional and borderline fantasy, as were the hometowns to negate any bias participants might have had relating to real names or places. The characters and attributes used are shown in the appendix. Participants performed three memorization tasks, each having a different level of visual stimulus and each with eight different characters. We determined that the characters in each task would be different to account for any improvements in memory that may have occurred from repetition. The three levels of visual stimulus were simple text flashcards (with no visual stimulus), 2-Dimensional picture cards, and 3-Dimensional models. These levels served as the within-subject independent variable. As we were testing memory improvement, we chose to further mitigate any confounding effects due to memorization fatigue or novelty effects by creating six participant groups. These groups were formed by generating a combination of all possible orderings of the visual stimulus, creating a total of six groups, each with three trials. The orderings of the six groups are listed below:

1	Text, Picture, Model
2	Text, Model, Picture
3	Picture, Model, Text
4	Picture, Text, Model
5	Model, Text, Picture
6	Model, Picture, Text

Participants were randomly placed into one of the six groups. Each participant was given five minutes within each virtual environment, and then proceeded to complete a memory recall questionnaire after each level. The memory recall questionnaire was designed as a fill in the blank sheet for each attribute, with the names of the characters given for each trial. In the trials where 2D or 3D visual stimuli were used, flat pictures or pictures of isometric views of the models were also provided in the questionnaires. This was decided upon since we were interested in the capability of virtual reality to create relationships between abstract information and interactive content, not just text. If only text (ie. the names of the characters) was provided in the questionnaire for each level of visual stimuli, we feared that the participants would spend more time relating attributes to the name rather than the visual provided. We also consciously chose not to place a time limit on the recall questionnaire so that the participants could completely sift through their memory without being anxious about time pressure. Since this experiment focused on working memory, we predicted that the time span for participants to complete the recall questionnaires wouldnt be longer than a few minutes since working memory operates under a limited time, as described in sections above. A sample of the memory recall questionnaire can be seen in figureS 10, 11, and 12 in the appendix. The dependent variable was the number of errors in the memory recall phase. An error was counted as each attribute the participant failed to remember. Each trial was scored out of 32 points (eight characters, four attributes per character, one point per correct attribute) and the resulting score was calculated as a percentage of the total correct points out of 32.

4.1 Participants

24 voluntary, unpaid participants were recruited in total, out of which 14 were male and 12 were female. The mean age of the participants was 20 years old. From the background survey the participants completed prior to the experiment, 16 of the participants had previous experience in a virtual environment using a head-mounted device. The participants came from a variety of different educational backgrounds and levels ranging from political science to electrical engineering as well as undergraduates to pHD candidates.

4.2 Experimental Procedure

Before the start of the experiment, the participant completed a background questionnaire that provided information on prior experience in virtual reality. The participants then performed a introductory memory test with three characters, one for each level of visual stimuli. Within the virtual environment, the participant was placed in the center of a waist-level blue table arrangement, with a table placed on either side of the participant and one directly in front of the participant. Each of the objects was placed on the table, and the rest of the environment was left empty to assure that the participants would be focused solely on the characters that they were centered around. The introductory test was used as the learning stage as well, allowing participants to become familiar with using the controllers and motions to pick up each of the objects. They were then given two minutes in the introductory scene to memorize the four attributes for each of the three characters. The participants then filled out the memory recall questionnaire in the real world so that they were able to get a feel for what the questionnaire would be like. This practice run was done in order to make sure that there were no confounding variables of novelty of VR or uncertainty of what the trials would look like when the real trials were conducted. After the practice run, participants completed the three trials in the order decided by the group they were placed in. Before starting the experiment, participants were informed of the objective of this study. Participants were told which visual stimulus they would be seeing before each trial, and were reminded that the questionnaire would include the 2D pictures or isometric model views if they were about to complete a trial with either of the visual stimuli. We also requested that participants stay focused for the minutes they were performing the memorization task. Participants were allowed to pick up all the objects and rotate them as they wished, but were asked to place the objects back on the table when they had completed interacting with it. Besides the motion necessary to pick up objects, no other virtual navigation or locomotion was allowed. At the end of each trial the participants completed a presence questionnaire to provide us with well rounded results. At the end of the three trials, after the experiment was complete, the participants were also asked to complete a subjective questionnaire and a usability questionnaire.

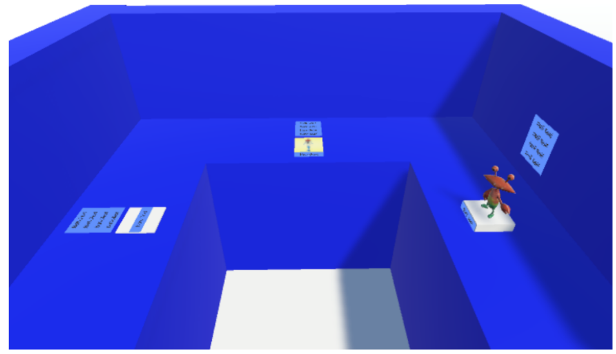

Figure 1: Introduction Scene

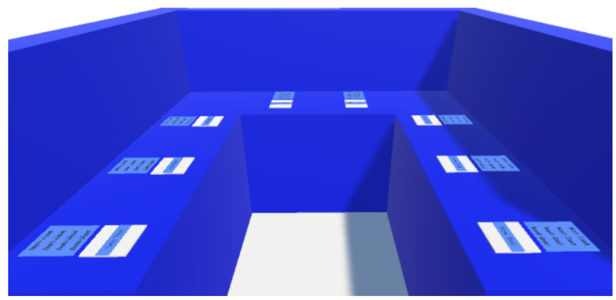

Figure 2: Text (No Visual Stimulus)

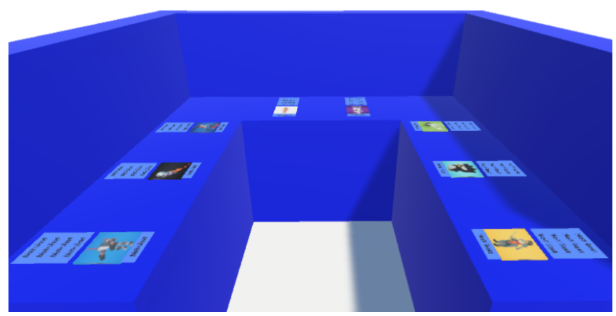

Figure 3: Picture Visual Stimulus

Figure 4: Model Visual Stimulus

4.3 Apparatus

For our experiment, we used the HTC Vive (Head Mounted Display) headset and controllers that provided participants with the ability to be immersed in our virtual environment with high visual fidelity and the ability interact with the objects by grabbing them with the trigger button on the controller. No audio input was used. The environment and virtual reality application was created using Unity and the VRTK Unity Plugin, with 3D models taken from Poly by Google. The pictures of the characters were just 2D screenshots of the 3D models. All of the experiments were run in the same testing room with a 5ft by 4ft testing area. The participants completed each of their questionnaires in the real world on a laptop, doing the recall questionnaire through an excel spreadsheet and all other questionnaires through qualtrics.

5 Results

We performed an analysis of qualitative and quantitative results from the recall questionnaires and subjective questionnaires that participants were asked to complete. The following sections summarize the data and information collected from the 24 participants that completed the experiment.

Figure 5: HTC Vive Testing Apparatus

5.1 Quantitative Results

As previously mentioned, we had 6 different groups of subjects for all of the orderings of the 3 types of trials, and each subject completed all three trials for one of the 6 orderings. We calculated the average percent score on the recall questionnaire for each of the 6 groups across all 3 trials, which can be seen in Figure 6. These values are the primary performance measure that we used to assess the subjects ability to memorize the character objects within each of the trials. Per our hypothesis, we expected to see higher scores for the trials with added visual stimulus, especially the trials with 3D models. Also included is a table of the standard deviations for all of these average group scores.

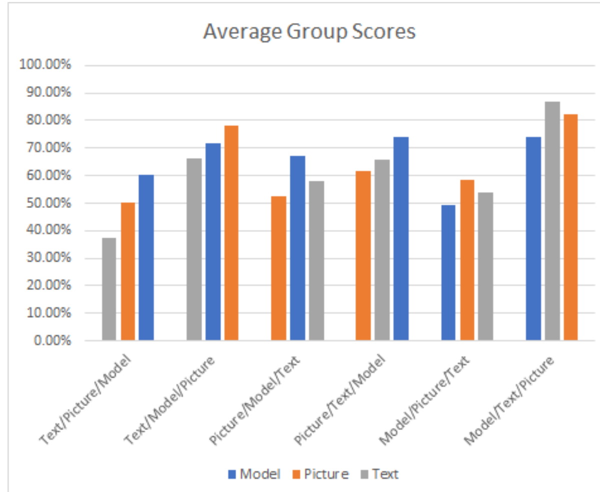

Figure 6: Average group scores for all three trials

Using these average score values for each trail for each group, we also calculated the change in score between trials to see how subjects would improve their memory performance going from one trial to another (Figure 7). In general, we expected to see scores improve in consecutive trials due to familiarity with the procedure, VR, and adjusted memorization strategy. Similarly, we calculated the change in score between trial types (Text, Picture, or Model) to see what kinds the percentage difference would be going from one visualization type to another. This was done in order to determine whether the participants were influenced more by any kind of visual stimuli or by the 3D models in particular.

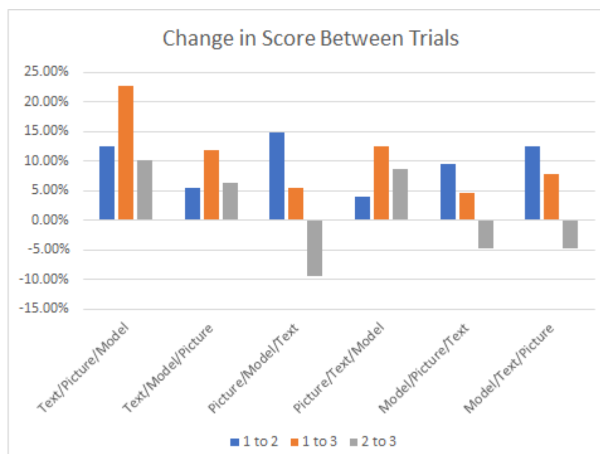

Figure 7: Average change in percent score between trials

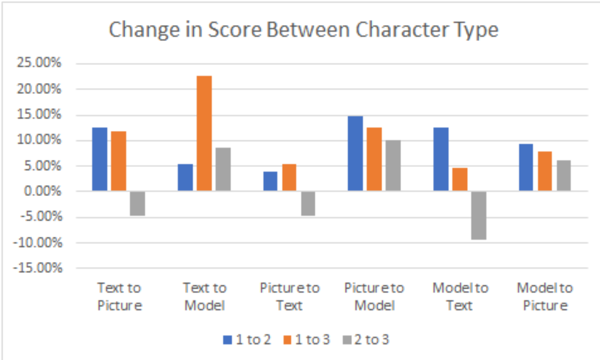

Figure 8: Average change in percent score between the different types of visual stimulus

Since the goal of our study is to see how added visual stimulus can potentially improve memory performance over basic text information, analyzing the score differences between the different trials is essential to evaluating our hypothesis.

Standard Deviations	Model	Picture	Text
Text/Picture/Model	17.49%	10.52%	17.38%
Text/Model/Picture	19.33%	19.43%	17.12%
Picture/Model/Text	26.69%	23.45%	28.47%
Picture/Text/Model	12.34%	23.52%	6.44%
Model/Picture/Text	25.31%	17.00%	28.11%
Model/Text/Picture	22.15%	18.29%	12.60%

Table 1: Standard deviations for the average group scores.

5.2 Qualitative Results

The participants were given subjective questionnaires so that the information we received would be as holistic as possible. It was useful because it gave participants to explain their feelings towards each trial and counterbalance any discrepancies in quantitative results due to novelty, fatigue, or any other factors. In the overall subjective questionnaire, participants were asked to rate their opinions on each of the three types of visual stimuli on a five point Likert scale. The scale ranged from 1: very bad to 5: very good. The questionnaire also prompted the participants to pick their top choice from the three visual stimuli. Finally, the questionnaire gave room for additional comments that the participants may have had about any part of the experiment.

Text	2.25
2D Picture	3.458
3D Model	4.083

Table 2: Average participant opinions for each trial (based on 5-point Likert scale)

Out of the 24 participants, only one participant chose the 2D picture technique as the tool most preferred, and the rest 23 said that they would choose the 3D model visual stimulus as their preferred memorization tool. This outcome was consistent with our hypothesis that allowing learners to interact with material and develop a relationship with abstract content allows for more a more engaging exchange of information and deeper understanding. The subjective questionnaire was helpful in determining the participants overall feelings towards the experiment and the possibility of virtual reality becoming accepted and widely used as a memorization platform for educational usage.

6 Discussion

All of the tested components of immersion and visual stimuli contributed to differences in performance. The results are interpreted and their implications for the future of VR in education are discussed in the following section.

6.1 Interpreting the Results

The quantitative results unfortunately did not show any completely consistent patterns, although there were a number of interesting observations to be made that give us a partial understanding of how the visual stimulus affected memory performance. For the average group scores, there is not really a clear trial type where the memory performance was higher across the board. The total average score for each trial type was as follows: 61.33% for text, 63.80% for pictures, and 66.15% for models. While this pattern is somewhat in the direction of supporting our hypothesis, the difference between the trial types is still pretty small and this total average does not reflect the diversity of our data.

In general, our results were also fairly inconsistent because there were fairly large standard deviations for a number of the trials across the different groups. As can be seen in Figure 9, there are multiple instances of a standard deviation above 20%, which indicates that the results of the recall questionnaires were largely dependent on the individuals and their personal ability to memorize the character attributes. The variability in these scores indicates that there are probably a lot of factors explaining the results that we have not properly controlled for. In future versions of this experiment, we should design to control these additional variables that were determined in this experimental phase. Even though there are various pieces of evidence that support our hypothesis, they do not do so strongly and it is hard to explain the specifics of the data gathered.

6.2 Memorization Over Time

One significant pattern that can be seen is that for all groups, the average score for the first of the three trials is the lowest score. This indicates that the subjects were still adjusting to the task of memorizing the characters attributes in the first trial. This result is consistent with the participants subjective feelings towards the study, the majority commented on the difficulty of the memorization task decreasing after each trial due to familiarity. The introductory trial may not have been sufficient enough in preparing the subjects for the level of difficulty associated with the full 5 minute trials, so it would be interesting to see if this pattern would be different if instead we had given each of the participants full practice trials for each of the trial types instead of just one introductory trial. However, this pattern of doing better in subsequent trials was not true in all cases when going from the second to the third trial, and in three of the groups the scores actually went down from the second to the third trial. This inconsistency made it somewhat difficult to interpret our results, but by looking at the change in score between the trials we were able to see what was going on a little bit more.

Another way to evaluate these jumps in scores was to calculate the difference in score between each visualization type instead of just the trial numbers, see figure 8. This graph is very similar to the previously discussed graph, but makes it a little bit easier to visualize what is happening between different visualization types. The largest increase in performance occurred going from the text trial when the text trial was first to the model trial when the model trial was third. However, the largest increase in back to back trials (1 to 2 or 2 to 3) occured going from the picture trial to the model trial.

While this is a good sign that the model trials have better performance, there were still significant increases when going from the model trial to the text trial from trial 1 to 2. This is somewhat contradictory to our hypothesis, although considering our observation that the scores always increase from the first trial, its possible that this had to do more with the fact that the subjects tend to improve after the first trial no matter what.

6.3 Distinguishing Between Visual Stimuli

An important realization that was drawn was the importance on what visual stimuli contributed to the biggest difference in memory performance for each of the participants. In the majority of the groups, any difference in scores between a trial and one of the trials that followed it was positive, meaning the subjects generally performed better as they progressed through the trials. However, there were three instances where the scores went down, and all of them were going from the second trial to the third trial. The largest of these drops occurred in going from the model trial to the text trial (-9.37%), which is the kind of trend we expected to see according to our hypothesis. However, the other two drops were going from text to picture and vice versa, both of equal magnitude (-4.69%), which indicates that there may not be a significant difference between the text and picture object types.

Qualitatively however, we received feedback from the participants that the use of any kind of visual stimulus (not necessarily just the models) was helpful in performing the memorization task. However, participants were clear in expressing that they appreciated the models because it allowed them to humanize the characters, interact, and develop a relationship with a human-esque object rather than a static picture. The ability for participants to do so provided them with the tools to form relationships with the content. As we hypothesized, this was an aspect of virtual reality that we felt was crucial to its success in educational environments. This shows us that even if the quantitative results are not in line with participant opinions, there is room for VR to become accepted in the learning space as participants enjoyed interacting with material in order to facilitate the learning process.

6.4 Spatial vs. Non-spatial Memorization

Although our results are not completely consistent with our hypothesis, there are some alignments. However, our experiment also gives rise to interesting discussions about spatial versus non-spatial memorization. We purposely did not test spatial memorization in this experiment, because we were only trying to evaluate the ability to create relationships with the items themselves. Spatial memorization is an entirely different concept that should be researched under different conditions, and it would be interesting to test whether VR serves better as a memorization tool using spatial or non-spatial memorization techniques.

6.5 Biases and Memorization Techniques

Our experiment was designed with great effort to remove almost all bias possible. However, since our characters and hometowns were fictitious and cartoon-esque, it was difficult to completely remove all bias. We deliberately attempted to create attributes for each character that didn't have a relation to the character's figurine. However, as mentioned above, we saw that each participant had a unique way of relating the characters to the attributes, whether it was through the figure or through external knowledge such as relating attributes to a person they knew. We discovered from this that it is impossible to remove every bias possible, and we recognize that if this were to be used in educational settings, the learners would develop their own tool in conjunction with this tool that would aid them to memorize any material given.

Many subjects noted that in certain instances the visual stimulus would have a natural correlation with the information they were trying to memorize, and when this was the case, the visual stimulus was more effective in their memorization process. For instance, a subject described that one of the characters whose occupation was a singer had a model that looked like it could be holding a microphone, so they were able to associate that occupation with that character more easily. This kind of association is what one would expect to make a visual stimulus more effective, although we were trying to avoid having these factors to try and see if having unrelated visual stimulus would still improve memory.

Additionally, we saw that many participants decided to focus on memorizing attributes that were not fictitious (the age, favorite food, and occupation). This could be the case because it was easier to create a story around or relation with an item that a participant has previous knowledge of. This was found from looking through the recall questionnaire responses for each participant and seeing that most got their points from the real-life attributes rather than the fictitious ones. This is consistent with the opinions of the participants; a majority stated that trying to learn a new word as well as memorize it for each character was too difficult in the time given. Perhaps using real places would have been better to facilitate memorization within the virtual environment since there would be less cognitive load on the participants.

6.6 Presence

Although we gave participants a presence questionnaire after each type of trial, we did not see much value in the results it displayed. Participants were asked to respond to their feelings of "being there" in the virtual environment on a seven-point Likert scale from 1: not at all to 7: very much. The average of the participants' feelings of being there was 4.9, which tells us that most participants felt like they were at some level in a different environment. The effects of presence on memorization within VR must be further researched, as our environment was consistent across all trials and no further data was collected regarding presence.

7 Conclusions and Future Work

There have been many experiments and research done in the past that test the capability of VR as an educational tool, but little evidence has been found that it can facilitate better memorization. Because the actions being performed in VR are still able to be performed in the real world, there is currently little evidence that VR will allow learners to better understand the material. Although it gives the learner the ability to develop relationships with the material, this can still be done in the real world. Clearly, there is much future work that must be done. Our experiment did not give us any clear conclusions that confirmed our hypothesis. Further research is required to determine the capability of VR as a memorization tool, focused on aspects that were potentially lacking during this experiment such as an increase in presence (potentially creating a room rather than a skybox), as well as developing methods to overcome certain confounding variables such as memorization fatigue. Additionally, the number of participants we tested that were unfamiliar with VR or with the procedure (even after the introductory scene) was high. In future iterations, it's important to note that there is a learning curve and a fair amount of time should be attributed to allowing participants to understand the environment (probably through a longer introductory scene). In our experiment, participants were also frequently transitioning in and out of the virtual environment. This made it so that their presence was broken after each five minute chunks, potentially causing some distractions or discomfort. In a future version of this experiment, it could be worthwhile to test the effects of having the participants perform the memory recall task within the virtual environment. However, this poses complexities if this type of tool were to be applied to learning environments, because most activities or material that is learned is not used or tested within a virtual environment.

In the future, there is potential to add even more interactivity to the models such as the ones in our experiment. By doing so, future research can be done on higher levels of interactivity and the effects on memory capacity. It would be especially interesting to see if using visual stimulus and interaction that correlates with the meaning of the information being absorbed would help, as our subjective feedback indicated that this helped certain subjects. Our approach of adding visual stimulus in somewhat abstract and not very meaningful way was somewhat limiting the potential of how visual stimulus like 3D models could be used to enhance the memorization process.

As mentioned above, this experiment specifically researched on VR's capability with working memory. It would be interesting to test other types of memory using distractor tools to determine whether VR is capable of allowing participants to retain information for a longer period of time and store it in their long-term memory capacity.

References

[1] B. BM, A. EA, R. FD, C. BR, and L. AG. The specificity of memory enhancement during interaction with a virtual environment. *US National Library for Medicine*, 7(1):65–78, Jan. 1999.

[2] D. A. Bowman, L. F. Hodges, D. Allison, and J. Wineman. The educational value of an information-rich virtual environment. *Presence*, 8(3):317–331, June 1999.

[3] M. Bricken. Virtual reality learning environments: Potentials and challenges. *SIGGRAPH Comput. Graph.*, 25(3):178–184, July 1991.

[4] H. Q. Dinh, N. Walker, C. Song, A. Kobayashi, and L. F. Hodges. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. *IEEE*, pages 222–228, 1999.

[5] M. Hussein and C. Natterdal. The benefits of virtual reality in education. *Department of Computer Science and Engineering*, 1(2):1–15, June 2015.

[6] A. Negu, S.-A. Matu, F. A. Sava, and D. David. Task difficulty of virtual reality-based assessment tools compared to classical paper-and-pencil or computerized measures: A meta-analytic approach. *Computers in Human Behavior*, 54:414 – 424, 2016.

[7] Z. Pan, A. D. Cheok, H. Yanga, J. Zhua, and J. Shia. Virtual reality and mixed reality for virtual learning environments. *Elsevier Ltd*, 30(1):20–28, Jan. 2005.

[8] E. D. Ragan, A. Sowndararajan, R. Kopper, and D. A. Bowman. The effects of higher levels of immersion on procedure memorization performance and implications for educational virtual environments. *MIT Press Journals*, 19(6):527–543, Dec. 2010.

[9] M. C. Salzman, C. Dede, R. B. Loftin, and J. Chen. A model for understanding how virtual reality aids complex conceptual learning. *Presence: Teleoperators and Virtual Environments*, 8(3):293–316, Jun 1999.

[10] P. TD and R. AA. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis. *National Center for Biotechnology Information*, 39(3):250–261, Jul 2008.

[11] C. D. Wickens. Virtual reality and education. In *[Proceedings] 1992 IEEE International Conference on Systems, Man, and Cybernetics*, pages 842–847 vol.1, Oct 1992.

A Questionnaires

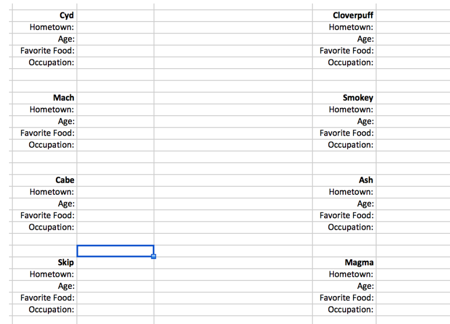

Figure 9: Text Recall Questionnaire

Figure 10: Picture Recall Questionnaire

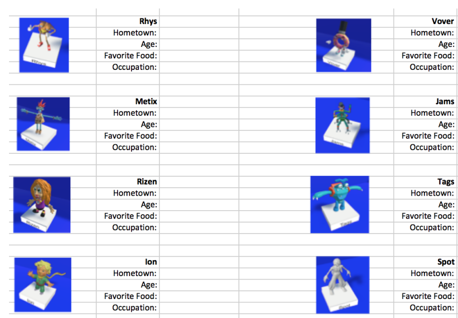

Figure 11: Model Recall Questionnaire

Duke UNIVERSITY

Please rate your opinion of the flashcard memorization technique:

1 - Very bad

2

3

4

5 - Very good

Please rate your opinion of the 2D picture memorization technique:

1 - Very bad

2

3

4

5 - Very good

Please rate your opinion of the 3D model memorization technique:

1 - Very bad

2

3

4

5 - Very good

If you had to choose a memorization technique, would you choose:

Flashcards (only text)

Flashcards (with 2D pictures)

3D Models

Please give any additional comments you may have:

Figure 12: Subjective Questionnaire