A Correlational Analysis of Confidence and Error Rates of Virtual Reality Data

Introduction

Technology is advancing at an ever-increasing speed. One of the most exciting areas of this development is Virtual Reality. Virtual reality has made impressive leaps in sophistication over the past decade and shows no sign of slowing down. The technology has been used in video games, field trips, and as quarterback training in college and professional sports (Bailenson, 2018). The greatest advantage that Virtual Reality (VR) has to offer is the ability to place the user in an environment that is almost indistinguishable from reality, this gap will narrow as the technology is refined. These experiments are the Turing test of our time. This technology has widespread humanitarian benefits by showing the users the effects of ocean acidification, hunger and poverty, the split-second decision making of high intensity situations, and visiting the wonders of the world all by eliciting emotional reactions that are becoming more and more indistinguishable from reality. Using this technology to bring awareness, foster empathy, and safely train individuals is a revolutionary application of this technology. Science must work to understand these effects and how they are obtained in order to raise the efficacy and ethicality of this new technology. Several studies have already been conducted on the immersive qualities of this technology. One of the most notable uses of this technology has been the ability to treat clinical patients with phobias. Studies have been conducted in a harmless virtual environment to treat arachnophobia, fear of heights, public speaking, and agoraphobia (Bailenson, 2018). A study by Nick Yee (2014) demonstrated the after-effect of using virtual reality in a real-life simulation. In this study Yee and colleagues would place individuals into avatars, virtual renderings of themselves, and have them act out scenarios with others. One of the interesting findings from this study was that shorter individuals placed inside taller avatars of themselves exhibited higher self-esteem, higher confidence, and more erect posture days after experiencing this experiment (Yee, 2014).

Studies like these display the potential for clinical application to help these users with lasting effects into the future, but how can this technology help users with past issues of trauma and memory. In one study from Jeremy Bailenson’s lab at Stanford, researchers were able to work with soldiers suffering from PTSD. Their goal was not to replicate a first-person shooter video game that satisfies a “revenge fantasy” but rather to evoke states of high anxiety (Bailenson, 2018). A clinician will set the scene of this virtual experience by altering lighting, by changing sounds like ambient street noise to explosions and gunfire, all the way to implementing physical sensation in the therapy (clients are placed on a platform that can vibrate and sway to simulate physical experience). The goal is to make this sensation as close to indistinguishable from reality as possible. This is done by triggering sensory mechanisms and cognitive mechanisms within the participant who then faces these high anxiety situations (Bailenson, 2018). In theory, this is because the new experience encodes over the traumatic memories in the participant, lowering their anxiety levels. This ability to retroactively influence the past has opened the door for hundreds of variations on this type of treatment. The question now becomes; how can this technology effect non-traumatic memories?

A study recently done by Rubo et al. (2020), sought to discover the effects that VR may have on our memory systems. This study used a source recall test between three modalities: Virtual reality, reality, and computer monitor. By having participants recall what modality they were exposed to a stimulus in, researchers could investigate how our memory systems distinguish and store VR data. This is application is grounded in work done by Hoffman et al. (2001) who defined source monitoring as the process individuals make decisions where memories obtained from a range of unique sources and/or origin (Virtual vs. Real) are separated and categorized (Hoffman et al., 2001). The definition for source monitoring is the main theoretical focus of the article I chose to analyze.

Their Study

In the article by Rubo et al. (2020), researchers sought to investigate how our memory systems distinguished between reality, copycat manifestations within a virtual reality environment, and a two-dimensional computer monitor. The Rubo et al. (2020) study wanted to identify if virtual reality experiences were viewed as being more real compared to a computer monitor.

The study discusses a unique approach to how we examine and experience our virtual environments to determine how convincing this new artificial world is. We retain memory traces after an experience, these memory traces are thought to be accompanied by “source tags” that provide individuals with the ability to recall content as well as the memory’s source (Rubo et al., 2020). For instance, if someone saw an elephant in VR but did not, or could not, define the source tag well enough, the recall may attribute that elephant experience to have occurred in physical space a real physical environment. The article goes on to suggest that when there is a deficit in the quality of the source tag or the source tag is not defined well enough, the memory system could rely on a heuristic recall that infers information about a memory's details (Rubo et al., 2020).

To test their participants Rubo et al. (2020) used toy replicas of animals, 30 different kinds, across a within-subjects design using a memory confusion paradigm to influence the source tags of memory. This was to determine if VR were categorized as more real than stimulus from a computer monitor. Copies of these animals very rendered in high fidelity virtual reality software that was also used in the computer monitor condition. Each modality had three potential conditions: interaction with one, both, and not at all. The interaction group could hold the animal in hand and manipulate however they desired during the trial. The virtual reality interaction used a handheld controller to manipulate the animal, and the computer condition allowed for the use of directional keys to manipulate the animals.

The procedure for testing these three modalities involved assigning participants to one of the modalities and one of the interactions. The participants categorized into no interaction, interaction with one, and interaction with both. Each category displayed all 30 animals in one of the three possible modalities at random. In the interaction category, participants could handle the objects placed in front of them. The real interaction meant the user could feel and grasp the stimulus. The virtual reality could grab and move the object but could not feel the stimulus, only the controller. The computer monitor condition only allowed for directional keyboard interaction. Participants where rested and then re-presented shown animals at random and asked to identify which modality they had experienced the stimuli in. Participants were asked to give confidence ratings using a visual analogue scale that ranged from "Not sure at all" to "Very sure" of their decision.

The study revealed that it was easiest for participants to misattribute virtual reality and monitor modalities (25.64% error rate). Researchers do make a point to identify that misattribution rates were higher for the virtual reality-real modality comparison compared to the monitor-real modality comparisons (approx. 6% higher), suggesting that there is a potential misinterpretation of memory source tags in the participants' memory recall systems.

When looking at confidence ratings of decisions, researchers found that for 43.98% of choices, participants gave the highest confidence rating. The average was .91 on a scale of 0 to 1 (most confident), where a score greater than .91 was deemed a confident answer. Researchers found a negative correlation between confidence ratings and error rates which suggests that confidence ratings reflected memory system performance. Confidence was higher in virtual reality-monitor error. Real life modality gave the highest confidence to accuracy ratio, the highest error rates and widest range of confidence ratings were in the virtual reality condition.

The study did not demonstrate any significant influence of interaction on memory error rates.

Methods

I was curious as to see how the participants' confidence in their decisions may have had an influence on the source recall error rates of these participants. I also wanted to see if it was possible to tease out data that revealed how the brain was storing virtual reality data. Was the brain assimilating VR data into one of the other categories in their memory system or was the brain generating a defined virtual reality category? While the findings demonstrated the individuals could often recall when they had seen something in three-dimensional space, there were clear signs that some confusion begin to appear when they were exposed to stimulus in virtual space. Participants had little issue recalling if they had seen stimulus from a computer monitor, but distinguishing between the real and virtual conditions proved more difficult.

The confidence values were given before the participants solidified their decision on source modality. The confidence rating was meant to capture the level of assurance the participants had in their recollection and choice regarding the modality the stimulus had been presented to them.

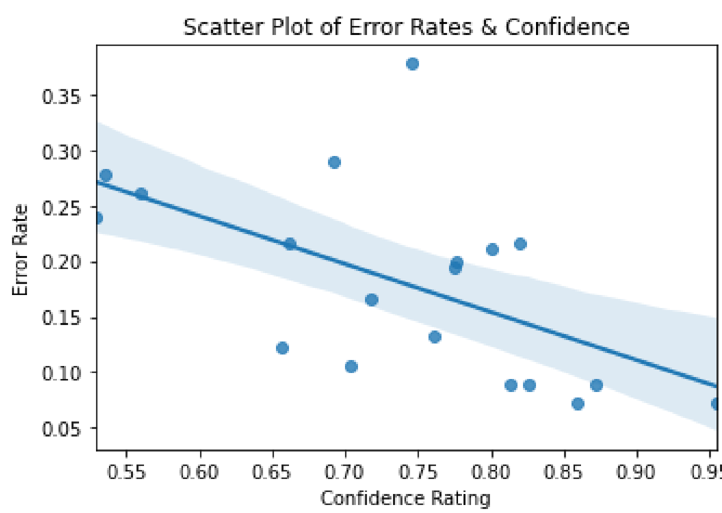
The study being discussed investigated whether or not the virtual reality technology was adequate to influence the ability of a participant's short term memory systems distinguished and processed different modalities.

From the data presented in this study, I expected to find a relationship between the confidence ratings of participants and their error rates in determining the presentation source of the stimulus. I also expected to see that participants were most confident in determining a source shown to them in physical space as well as on a computer monitor. I expected to find that the mean confidence scores to be lower in determining the virtual reality category.

In order to determine if a potential relationship existed between these two categories I had to calculate the error rates for each participant. This was done by subtracting how often they gave an incorrect determination for source modality from the total number of trials, then dividing that number by the total amount of trials. Once the data had been processed, I ran a correlation analysis for the two categories.

Results

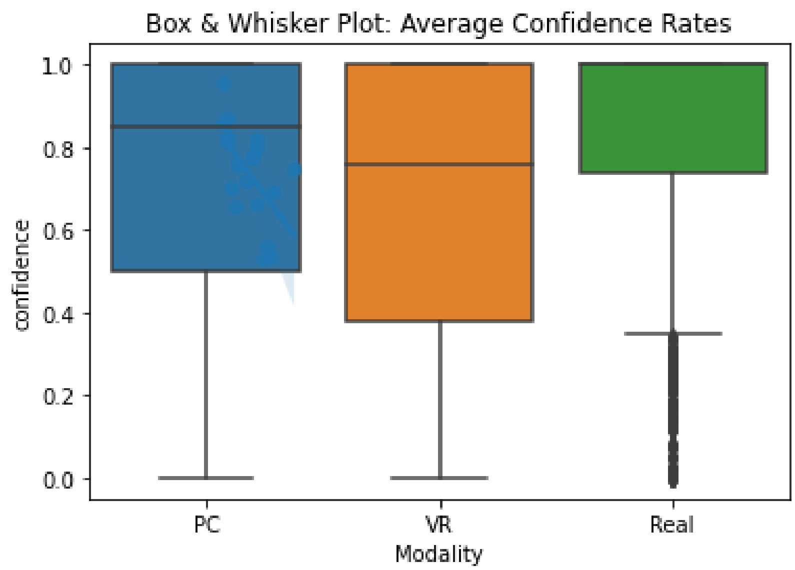
The results for the correlation analysis revealed a negative correlation between the two categories. The data clearly shows that as participant confidence increased the error rates tended to decrease. This reveals that seems to be a notable relationship between error rates and confidence ratings of participants, however, we cannot discern how this relationship manifests across the three modalities separately due to this being a general correlation.



	Error_Rate	Confidence
Error_Rate	1.000	-0.578142
Confidence	-0.578142	1

Participant data was also plotted to a box and whiskers graph. By presenting the data in this manner, I was able to determine where error rates were highest based on the confidence correlations from the scatterplot above.

In the Box plot we can see that the PC and Real modalities both show high mean confidence ratings. We can see that the PC modality had a much wider range in confidence ratings than the Real condition, which stayed in the high 0.8 - 1.0 range. Participants could easily determine the source modality of these two conditions when asked. However, the VR condition displays the lowest of the three conditions for the mean confidence rating. The VR condition also has the widest range of confidence ratings. These two pieces of information tell us that participants were least confident when determining if a stimulus was presented to them in this modality. These findings demonstrate that participants had more difficulty in recollecting source modality in this category which implies the brain is grouping this experience with the PC condition or the Reality condition rather than storing this modality in a separate conditional category. More research will need to be done in order to fully understand how the brain is storing virtual reality data.



Discussion

While these findings are derived from a small participant pool, I believe that they still show promising potential. Technology will only become more immersive as time moves on. I believe that the technology and budget of this study may have had a negative impact on the study. This study was a proof of concept with room to improve.

If we were to attempt to replicate this study, I would suggest investing in haptic feedback glove which act as a controller in virtual space. While the head mounted display that responds to head rotations does the majority of heavy lifting in both terms of immersion and distraction if the HUD is uncomfortable. A study by Cadet et al. (2020) found that immersion, defined as the objective ability of the HUD to create sensorial stimulation and head tracking, was one of the key factors for VR and the sense of presence, otherwise known as the *feeling* of being in a place.

A higher sense of presence was correlated with higher rates of immersion. The findings of the study revealed that the virtual fidelity, or graphic quality, did not influence the immersion factor of the VR system. The key impact on the sense of presence and immersion were more dependent on the type of device that was used (Cadet et al., 2020). A heavy HUD system or handheld controllers, rather than one's hand, can all drastically influence how effectively one is placed in the virtual reality environment and can potentially influence how the mind encodes its experiences. I believe it is the finer details that prevent this study from demonstrating the effectiveness of VR on memory encoding and source tags. The immersion quality of this study is hindered by the use of a handheld virtual reality controller, which equates to a held stick that interacts via button input. Using haptic gloves would remove this most glaring immersion breaking quality.

Another study by Dietrichkeit et al. (2020) placed participants into a virtual environment and exposed them to common place interactions and tasks to simulate a real-world scenario. In order to test learning and confidence they did not inform participants on what information would be useful in their interaction or what they should pay attention to as they went about completing tasks and providing confidence ratings in their abilities of doing so successfully (Dietrichkeit et al., 2020). I reference this study because they had rated their task system as moderately difficult. If replication of the Rubo et al. (2020) study was to be done, offering a more diverse set of exposures without informing the participants on what they are to distinguish between alongside higher fidelity interactive equipment (more comfortable and naturally controlled) may prove a more fruitful design when testing how the brain distinguishes and encodes experiences in Virtual Reality.

One other direction would be to place EEG recording devices on participants while they undergo this experiment. The immersion would not be hindered as the participants must also wear a head mounted display; in effect the distraction of the EEG would be rationalized as the HUD to prevent any confounding variable. These memory studies and immersion could be particularly effective in treatments or therapies for individuals diagnosed with dementia. Overall, there are numerous studies that could be done on memory and immersion in the realm of VR.

In sum, the use of Virtual Reality in research is expanding faster with each passing year. As the fidelity of the programming and the user interface becomes more intuitive the ability of our memory system to appropriately tag and distinguish virtual experience from physical will only shrink. Researchers are already looking to see the wider applications of this capability, but it is also important to research potential down sides. This type of research may bring to light any deleterious effects implementing this disruptive and powerful technology may have; not just on our traumas and entertainment but how our brains perceive our very reality.

References

Bailenson, J. (2018). *Experience on Demand*. W.W. Norton & Company, Inc.

Cadet, L. B., & Chainay, H. (2020). Memory of virtual experiences: Role of immersion, emotion and sense of presence. *International Journal of Human-Computer Studies, 144* https://doi.org/10.1016/j.ijhcs.2020.102506.

Dietrichkeit, M. (2020). Using virtual reality to explore differences in memory biases and cognitive insight in people with psychosis and healthy controls. *Psychiatry Research, 285*

Hoffman, H. G., Garcia-Palacios, A., Thomas, A. K., & Schmidt, A. (2001). Virtual reality monitoring: phenomenal characteristics of real, virtual, and false memories. *Cyberpsychology & behavior: the impact of the Internet, multimedia and virtual reality on behavior and society*, *4*(5), 565–572. https://doi.org/10.1089/109493101753235151

Rubo, M., Munsch, S., & Messerli, N. (2020, October 20). The Human Source Memory System struggles to distinguish Virtual Reality and Reality. https://doi.org/10.31234/osf.io/n24cw

Yee, N. (2014). *The Proteus Paradox: how online games and virtual worlds change us- and how they don't*. Yale University Press.