Comparing Cognitive Training Tests: Paper-Based, Computer-Based, and Virtual Reality Stroop Tests

LORENZO MONDRAGON, CASSIDY ROBERTS, and ISABELLA SETTJE

This study investigates the effectiveness of cognitive training across three distinct learning environments: Virtual Reality (VR), traditional paper-based, and computer-based tests. With advancements in education and cognitive science, understanding the potential benefits of VR becomes crucial. Our objective is to assess the impact of these different learning modalities on cognitive performance, particularly in problem-solving abilities. Utilizing the Stroop Test as a cognitive assessment tool within VR, computer-based, and paper-based settings, we aim to compare their efficacy. Leveraging the immersive capabilities of VR and the convenience of computer-based testing, alongside the traditional paper-based approach, we seek to elucidate the potential advantages of immersive technologies in enhancing cognitive outcomes. Previous research has shown that VR's immersive nature can significantly improve engagement and cognitive test scores by simulating realistic and interactive scenarios. By addressing a gap in comparative analyses between immersive virtual environments and traditional learning settings, this study contributes insights into the future of educational technology and cognitive enhancement strategies.

Additional Key Words and Phrases: Mixed Reality, Cognitive Test, Virtual Reality, Stroop color and word test

ACM Reference Format:

Lorenzo Mondragon, Cassidy Roberts, and Isabella Settje. 2024. Comparing Cognitive Training Tests: Paper-Based, Computer-Based, and Virtual Reality Stroop Tests. 1, 1 (May 2024), 12 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

+

1 INTRODUCTION

In this project, we are going to study cognitive training and compare its effectiveness with three learning environments, Virtual Reality, paper-based and a computer-based setting. Education is something that has been going on for ages and with technology ever changing, there is an opportunity to improve learning across the board. The goal for the project is to shed light on the potential VR experiences and enhance the cognitive training outcome. This project is significant because it is addressing the challenges in education, psychology and cognitive science. Research is constantly being done by educators on how to optimize the experience and make the learning the most effective. By looking at the effectiveness of cognitive training in all three settings, we hope to learn more about how to optimize learning and how to make the cognitive tests be the most successful. With VR being relatively new, cognitive training will represent immersive learning education technology. VR can offer a new dimension of experiential learning that can do more than a traditional, physical classroom could do. In a study called "VIRTUAL REALITY (VR) AND THE KID'S BRAIN: EMERGING EVIDENCE FOR BRAIN COGNITIVE FUNCTION REHAB AND NEURO PAIN MECHANISM OF VR"

Authors' address: Lorenzo Mondragon, lorenzo.mondragon@colostate.edu; Cassidy Roberts, cassidy.roberts@colostate.edu; Isabella Settje, isabella.settje.edu.

Manuscript submitted to ACM

written by Henry Xiang, virtual reality has been proven to help rehabilitate children with traumatic brain injuries. [31] We want to explore this further because cognitive function in children is important. This has also been shown to have a similar effect in older adults with mild cognitive impairment [28]. Using a Stroop test in virtual reality to improve cognitive skills would be the goal. We want to highlight and explore the immersive nature that Virtual Reality has on learners. If they are more immersed, will that make them more engaged leading to a higher score on cognitive tests?

The Stroop Color and Word Test (SCWT) has long been a fundamental tool in neuropsychological assessment, serving to gauge the ability to inhibit cognitive interference. Originating in 1935, the SCWT presents individuals with tasks that challenge their cognitive processes by requiring them to name colors while inhibiting the automatic response of reading color names. This is known as the Stroop effect, and has been extensively studied and applied not only in evaluating attention, processing speed, cognitive flexibility, and working memory. A systematic review by Scarpina and Tagini (2017) examined the various scoring methods used in the SCWT, highlighting the need for a comprehensive scoring approach that considers both speed and accuracy of responses across different test conditions [23]. Building upon this research, our study aims to compare the efficacy of paper-based, computer-based, and virtual reality Stroop tests in enhancing cognitive performance and problem-solving abilities. Through this investigation, we seek to contribute to the ongoing discourse on cognitive training and educational technology.

There has been previous research that has highlighted the growing technology and effectiveness of immersive virtual reality for learning. In a specific study, Buttussi and Chittarao [7] dove deep into the efficacy of virtual reality when doing aviation safety training. Stroop tests improve cognitive functions because "In the context of the Stroop task,the dorsolateral prefrontal cortex appears to be involved in the relevant executive functions, particularly in maintaining response set (to name the color, not to read the word), whereas the anterior cingulate cortex plays a central role in selecting the appropriate response and evaluating its accuracy" [14] They were able to simulate realistic scenarios and it helped improve the "students" decision-making skills. In this specific study, they found that VR training was highly effective because it was able to simulate dangerous scenarios for the students without putting them at risk. This improved hazard recognition and safe behavior. This VR training was proven to be more effective because they were able to immerse the students in more, allowing them to learn more. Another article discussed using virtual reality as a tool for enhancing cognitive functioning in the elderly population, Bauer and Andringa [4] highlighted the number of elderly people experiencing cognitive decline and the dire need to mitigate/prevent this phenomena. Immersive VR is a promising approach because it can provide personalized training in safe environments, offer feedback, gather psychological data, enhance motivation and enrich engagement. The article further discusses the specific features of immersive VR because it is flexible and it has potential to alleviate professionals from certain aspects of their jobs. Bauer and Andringa emphasize the importance of technological factors in creating immersive VR experiences and vow for the need for more research in this field, which is what we aim to do. Another interesting fact about our experiment is that each one of our participants is under 30. According to Sophie Melissa Clare Davison, Catherine Reeprose and Sylvia Terbeck, " Younger adults completed significantly more parking simulator levels ($p<0.001$), placed significantly more objects ($p<0.001$), and located significantly more items than older adults ($p<0.01$), demonstrating higher levels of performance" [10]. Going forward, we can use this knowledge to interpret results and make an accurate conclusion.

Our team aimed to implement a novel approach by combining elements of traditional cognitive assessments with immersive virtual reality technology. Unlike previous studies that have primarily focused on either traditional methods or VR or a combination of the two, we seek to compare the effectiveness of cognitive training in all three settings simultaneously. By doing so, we hope to provide a comprehensive understanding of how these different learning environments impact cognitive outcomes. According to Stefan Gradl, in the article "The Stroop Room: A Virtual Reality-Enhanced Stroop Test" he talks about how " A Stroop Test was designed for this virtual classroom with the purpose of evaluating complex attention performance. They showed that a virtual 2D Stroop Test has the same real-world Stroop interference" [12]. This is necessary because it is showing us how education could potentially be changed by using virtual reality. Starting from the bare bones cognitive test(s). This project is a human computer interaction project as it explores how the interaction of VR technology can enhance cognitive training outcomes compared to traditional methods.

Our experiment will have three sides that will address the traditional paper-based learning cognitive assessment, the VR cognitive assessment and then a computer-based cognitive assessment. First, before our participants were given any of the test, they were asked to provide their previous experience with using VR and being a participant in cognitive training tests. After the Stroop test is done, we also administered a post experiment questionnaire to gauge their preferences of the three tests, as well as their predicted performance.

2 RELATED WORKS

2.1 Insights into Stroop Performance

In the pursuit of understanding cognitive processes, neuro-scientists have leveraged advanced imaging techniques to probe the neural correlates of attentions tasks, exemplified by the Stroop color word test. Through the application of positron emission tomography (PET), researchers embarked on a systematic exploration of brain activity during task engagement [10]. The evolution of cognitive assessments from traditional to modern platforms reflects significant scientific and technological progress. Historically, tools like the Stroop Color and Word Test (SCWT) have been foundational in understanding cognitive control and attention processes. Advances in neuro-imaging, such as the use of PET scans, have revealed complex brain activity patterns associated with these tasks, challenging prior assumptions and setting the stage for further innovations.

In the initial endeavor, participants underwent PET scanning while performing the Stroop test, yielding unexpected findings regarding regional cerebral blood flow (rCBF) dynamics. Contrary to established literature, activation's were observed in right orbito-frontal and bilateral parietal structures, accompanied by significant temporal fluctuations in rCBF [5]. Subsequent investigations sought to corroborate these observations, leading to the identification of focal activation's in the right anterior cingulate and right frontal polar cortex. Correlation analyses unveiled a distributed network of brain regions implicated in attention processing, illuminating the intricate interplay between anterior and posterior cortical areas.

2.2 Virtual Reality Applications

The advent of virtual reality (VR) has revolutionized the application of the Stroop Test, leading to the development of the "Stroop Room." This VR-based adaptation extends the original test's demands while ensuring accessibility and acceptability across diverse populations [12]. Designed as an open-source platform, the Stroop Room serves as a laboratory-quality stressor, facilitating large-scale in-field stress research without the constraints of traditional laboratory settings. Evaluation studies demonstrate its efficacy in provoking cognitive stress, eliciting physiological responses such as increased heart rate, skin conductance level, and salivary cortisol and alpha-amylase concentrations. Comparative analyses with traditional Stroop Test methodologies reveal improvements in stress induction metrics, alongside strong user engagement and flow-induction as indicated by questionnaire evaluations. These findings underscore the potential of immersive VR environments in stress research and cognitive assessment, offering a novel approach to understanding and addressing stress-related phenomena [27]. Furthermore, in a study conducted by Coban et al. their results show the meta-analysis suggests that the impact of immersive virtual reality on learning outcomes was relatively modest. While there were some significant effects observed in certain contexts, the general trend leaned towards a small effect size [8]. This is relevant as it shows VR might not be better in an education setting, at least not full time. In addition, This article provides an overview of the use of virtual reality (VR) in assessing and treating anxiety and related disorders, such as posttraumatic stress disorder (PTSD) and obsessive-compulsive disorder (OCD). Despite the rapid development of VR technology, its integration into clinical practice remains limited. The review explores the effectiveness of VR in therapy, including virtual reality exposure therapy, and discusses its alignment with existing theoretical models. In addition, in an experiment conducted by Meyerbröker and Morina where they use VR in assessment and treatment of anxiety and related disorders, they examine patient acceptance of VR and therapist reluctance to use it. It focuses on factors such as therapeutic alliance [16]. Future directions for VR in clinical practice, including its potential for use with younger populations, are also discussed. It is important to highlight participants reluctance to VR, as I know some of the people we reached out to to be potential participants showed disinterest due to the fact of VR and it causing motion sickness.

2.3 Computer-based Applications

The very first Stroop Test was created and conducted prior to the invention of the computer. Since then, there have been several variations of a computerized version of the original Stroop test. Pilli et al.'s study on a computerized Stroop test for evaluating psychotropic drug effects provides insights into the efficacy of computer-based cognitive assessments, relevant to comparing cognitive training tests across paper-based, computer-based, and virtual reality formats. The study concluded that the computerized Stroop test could be a valuable tool for assessing attention function and drug effects in both healthy individuals and patients [19]. In the same realm as Pilli et al.'s study, Basu examines the valuable implications of the computerized version of the Stroop test on children aged five to 13 years. This provides our study with more of an educational context as we are ultimately aiming to examine the effects of different learning environments [3]. They found that the results from the paper test and the computer test didn't always match up, and stressed the effectiveness of such adaptations in educational settings. These findings are important when comparing the two like we did in our experiment, however the participants we had all have significant experience with computers. Moreover, Coelli et al. demonstrated the feasibility and usability of a completely computerized system for the administration of the Stroop model in a clinical environment, providing further support for the integration of computerized cognitive assessments into various contexts [9]."

2.4 Comparative Analysis

A comparative examination of stress perception in virtual reality (VR) versus real-world environments sheds light on intriguing differences in subjective experiences. While stressful tasks in VR are perceived as less stressful than their real-world counterparts, the inclusion of additional stress factors, such as head movement, can alter this perception [24]. This suggests the value of transferring stressful tasks from traditional office environments to VR platforms for more nuanced understanding and intervention strategies. Recent investigations into the Stroop effect and Reverse Stroop phenomenon within the VR environment further elucidate these dynamics. By implementing interactive applications with instruction-based tasks, participants demonstrated varied responses influenced by subjective difficulty and cognitive processes. These insights highlight the complexities inherent in VR-based research methodologies and underscore the need for comprehensive approaches to stress assessment and intervention. Additionally, Barbosa et al. (2010) provides evidence of the effectiveness of computerized Stroop tests in inducing cardiovascular reactivity [2]. This could emphasize the potential benefits of computer-based cognitive training methods over paper-based ones. Moreover, such computerized tests offer advantages in terms of accuracy, control over stimuli presentation, and potential applications in clinical settings. While VR and computer-based methods offer innovative features like immersive environments and precise data tracking, traditional methods are valued for their accessibility and established reliability. Comparative studies often reveal that while technological methods can enhance certain cognitive processes, they also introduce new challenges such as technological barriers and the need for specialized equipment.

2.5 Expanding the Horizons

Virtual Reality (VR) has transformed cognitive testing by introducing immersive environments that extend beyond the capabilities of traditional paper-based tests. The development of VR platforms like the "Stroop Room" has not only replicated traditional tests in a virtual setting but has also enhanced the scope of data collection by eliciting stronger physiological and cognitive stress responses. The use of VR in cognitive training has been shown to improve engagement and memory retention, offering substantial benefits for educational psychology and therapeutic interventions. Despite its benefits, VR's effectiveness varies across different learning and therapeutic contexts, suggesting a need for tailored VR solutions to maximize outcomes.

3 METHODOLOGY

3.1 Participants

Participants were recruited by word of mouth, the only qualifier for participants was that they must be over the age of 18 and did not have color blindness. A total of 15 participants were run through the experiment. The demographics of the participants were collected using a google form that consisted of the following questions:

- **Q 1:** *H*ow old are you?
- **Q 2:** *W*hat is your gender?
- **Q 3:** *W*hat is your major?
- **Q 4:** *W*hat is your experience level with VR?
- **Q 5:** *H*ave you previously participated in cognitive assessments or experiments?
- **Q 6:** *D*o you have any visual impairments that may affect your ability to see colors or text clearly?

All the participants were enrolled students with a variety of majors. All of the participants had a varying degree of a VR headset experience before the experiment.

3.2 Materials

3.2.1 Paper-Based. For the paper-based learning assessment, we used a printed out sheet of paper that contains 16 words with varying color names with different colored texts. An iPhone was used to record a voice memo of the participant.

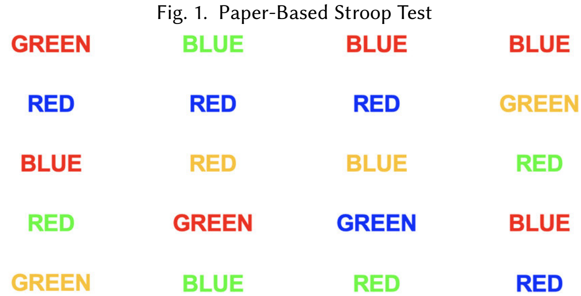
3.2.2 Computer-Based Stroop Test. Our second source was to conduct a Stroop test on a computer monitor. We used a standard mouse and a keyboard. This test was downloaded from the PsyToolkit experiment library and run locally on a PC. Participants will engage with a series of color names displayed in in-congruent colors, for example the word "green" displayed in yellow-colored text. To start, the online version of the Stroop test gave the flexibility and control over the potential conditions. From there, we can properly gather data and understand the stimuli going on at the same time. The control that the computer gave means that we can repeat experiment easy, have a consistent environment and get accurate results. Another reason we chose to do this is because it served as a “buffer” between the physical paper and the virtual reality headset. With that being said, the participants had something technical, but familiar at the same time. The computer interface served as a middle man for this experiment.

3.2.3 Virtual Reality. With the design environment for the VR Stroop Test in mind, A Meta Quest 2 headset was chosen for this project. Participants were equipped with this Meta Quest 2 headset to perform the experiment. They wore the VR headset and entered an immersive environment where color stimuli were presented in a three-dimensional space.

3.3 Procedure

3.3.1 Paper-based. Participants were asked to sit down with a desk in front of them while we flipped over the paper that contained names of colors with different text colors. This paper was derived from a previous pilot study on interactive multi sensory environments [11]. This paper contained 16 different words, split up into four rows of four words. Participants were asked to read the color of the text of the word out loud reading each row from left to right, starting with the first. Participants were recorded via iPhone voice memo and we manually tracked the accuracy and got the time results from it as well.

Fig. 1. Paper-Based Stroop Test



3.3.2 Computer-based. For the computer-based learning environment, participants were sat in front of a computer and were given instructions on how to use the Stroop Test. The reason that it is important is because "This suggests that executive function, as measured by the Stroop test, declines with age and that the decline is more pronounced in people with a low level of education. By doing this in an educational setting, we are balancing this disconnect. We will be using the schools computer. [29]"Stroop Test User Instructions: In this task, you will see color names (red, green, blue, yellow) in different text colors. You need to response to the print color. For example, if you see green displayed in red-colored text. You need to print the color (red), and press the associated button ("r"). The other buttons used in this study are "g", "b", and "y", for green, blue, and yellow. After each response, there would be a new word displayed. Participants were asked to identify 40 words in total. Responses were automatically timed in the application to measure reaction times. We recorded the participant responses, including the Status (1=correct, 2=wrong, 3=timeout).

3.3.3 Virtual Reality. Participants were equipped with the Meta Quest 2 headset and the VR functionality was enabled for this portion of the experiment. After adjusting the headset to fit comfortably and ensuring proper focus, participants entered the virtual reality environment designed for the VR Stroop Test. Participants were instructed to read the color of the word out loud and not the word. During the task, color names (red, green, blue, yellow) were displayed in various "print" colors. Participants were required to respond to the print color by saying the answer out loud. "The Stroop Color and Word Test (SCWT) is a neuropsychological test extensively used to assess the ability to inhibit cognitive interference that occurs when the processing of a specific stimulus feature impedes the simultaneous processing of a second stimulus attribute, well-known as the Stroop Effect" [23] After each response, a new word would be presented and this repeated about 40 times. We recorded the participant responses, including the accuracy for identifying the ink color and the time taken for each response.

3.4 Design

3.4.1 Paper-based. The Paper-based Stroop Test was modeled after the the original Stroop Test named after John Ridley Stroop. It typically consists of a list of color words printed in in-congruent ink colors, where participants are required to name the ink color while ignoring the word's meaning. This traditional format serves as a fundamental tool for assessing cognitive control and attentional processes [6]. The Stroop test is also very necessary because "The Stroop Color and Word Test (SCWT) is a neuro-psychological test extensively used to assess the ability to inhibit cognitive interference that occurs when the processing of a specific stimulus feature impedes the simultaneous processing of a second stimulus attribute, well-known as the Stroop Effect" [23]

3.4.2 Computer-Based. The design of the computer-based Stroop Test was also modeled after the original Stroop Test. It adheres closely to the fundamental principles of the task, where participants are required to name the text color of the words while ignoring the actual word meaning. This classic version of the Stroop Test aims to demonstrate the interference between automatic processing of word meaning and the effortful processing of ink color, as originally observed by Stroop in 1935. The task provides a standardized method for assessing cognitive control, attention, and processing speed.

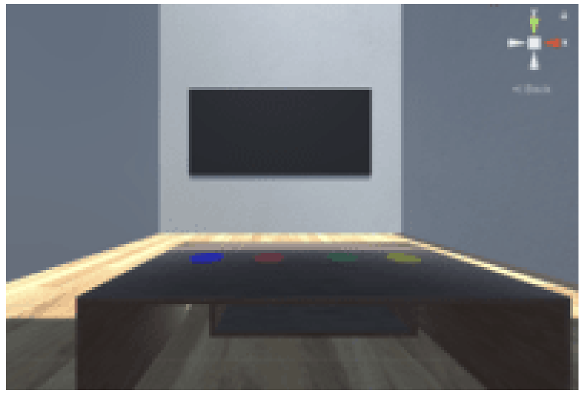
3.4.3 Virtual Reality. The development of the mock-up drew inspiration from the work of Smith and Johnson (2022), who conducted an Emotional Virtual Reality Stroop Task as part of their immersive cognitive test [15]. Within this environment, there was a TV screen-like element displaying words on the opposite wall where the participant was standing, in front of a table with buttons associated with colors for participants to click when presented with the word. The mock-up drawing provided a visual representation of this setup, as well as the YouTube video. The VR Stroop Tests aimed to involve participants physically, providing a more experiential approach to cognitive training.

4 RESULTS

In this section, we present the findings from three different implementations of the Stroop test: a traditional paper-based format, a computer-based interface, and a Virtual Reality (VR) setting. Each test aimed to explore the cognitive load and response accuracy in differing modalities, offering insights into how user interfaces can impact cognitive performance. Initial analyses focus on the comparative results of accuracy rates across these modalities, followed by a detailed examination of the results from each test. The integration of VR technology, in particular, provides a novel perspective on interface design and its cognitive implications, which are detailed in the subsequent sections.

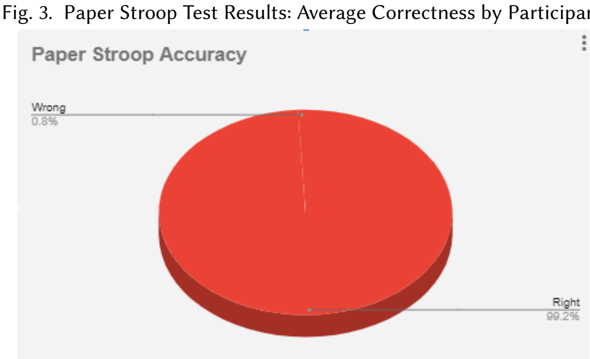
4.0.1 Paper Test. The implementation of the paper-based Stroop test with our 15 participants resulted in incredibly high accuracy. The test consisted of 16 words per participant, which means that 240 responses were recorded. Out of

Fig. 2. Virtual Reality Environment: Stroop Test



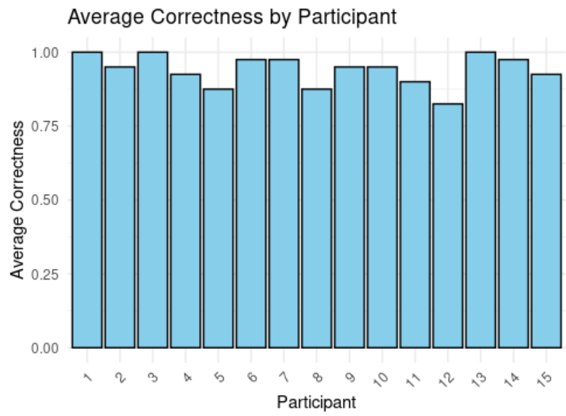
the 240 responses, 238 were correct meaning the success rate was 99.17. The high accuracy means a lot because of the nature of the Stroop test. The Stroop test is designed to challenge cognitive flexibility. The interesting thing about the two errors is that both of then occurred on the 14th word. One participant made a mistake, said the next word, then went back and realized he made a mistake [21]. The other participant confidently said the wrong word. A lot of factors could have attributed to this which will be discussed further in the conclusion. Below is a graph that is visually showing how high the accuracy rate was.

Fig. 3. Paper Stroop Test Results: Average Correctness by Participant



4.0.2 Computer-based Test. In the Computer Stroop test, a total of 15 participants were evaluated for their ability to correctly identify colors as they were presented on a computer screen by using corresponding first letter key of the name of the color. Out of all three tests, this test had the lowest level of accuracy. Despite being the lowest accuracy rate, the results indicate a high level of accuracy across the board, with the average percentage of correct responses being 94 percent. Figure 4 is a bar plot showing the average correctness for each of the 15 participants. The majority of participants (10 out of 15) achieved a correctness score above 90 percent, suggesting consistent performance across tasks. Examining the individual performance of each participant further illuminates the nuances within the dataset. While some participants achieved perfect or near-perfect scores, others displayed slightly lower levels of accuracy, contributing to the overall variability in the results.

Fig. 4. Computer-based Test Results: Average Correctness by Participant



The results show that while some people did better than others, overall, everyone did pretty well on the computer test. This suggests that people are good at handling tasks like this on a computer. It's like they're adaptable and can quickly make decisions, even in a virtual setting.

The computer-based Stroop test is helpful for understanding how people think and make decisions. It gives us clues about how flexible people's thinking is, how well they can focus, and how good they are at controlling their responses when using a computer. The Stroop test is interesting and in the article "The Stroop Effect On the Internet", they mention that "The Stroop task was developed to study interference effects on attention" and found that "he frequency of incorrect responses was 1.2 (188 errors on 15360 trials), and there was no significant difference in t

he frequency of errors between the type of stimuli or Stroop format." [13] This is interesting because we found similar results when going through our data.

It is worth highlighting the responses that we received from our Post Experiment Questionnaire form that we required each individual to fill out. Out of the 15 participants, 3 participants (20 percent) indicated they preferred the computer-based Stroop test the most out of the three distributed tests. They were also asked to explain their reasoning and all three participants indicated that it was the easiest test for them.

4.0.3 Virtual Reality Test. In the Virtual Reality (VR) Stroop test, a total of 15 participants were evaluated for their ability to correctly identify colors under immersive conditions. The results indicate a high level of accuracy across the board, with the average percentage of correct responses being approximately 97.83 percent. The standard deviation of performance was relatively low at 2.29 percent, reflecting a consistent performance among participants.

Most participants achieved perfect or near-perfect scores, with 73.33 percent scoring a full 100 percent accuracy, and the remaining scores not dropping below 95 percent. This high consistency suggests that the VR environment did not negatively impact the cognitive ability to perform the Stroop test when compared to traditional methods. The minimum score recorded was 95 percent, and the maximum was a perfect score of 100 percent. The majority of participants (median performance) identified 39 out of 40 colors correctly, underscoring the potential of VR as a stable platform for cognitive testing [18].

These findings highlight the efficacy of VR interfaces in maintaining cognitive performance, which could be particularly relevant in designing applications that require high cognitive engagement in immersive environments [25].

5 DISCUSSIONS

While the paper-based test offers a traditional and effective approach to cognitive training, the computer-based tests introduce the benefits of digital precision and repeatability. On the other hand, the VR environment, as evidenced by the data, did not impair cognitive performance, potentially due to its immersive and engaging nature [16], which might help sustain or even enhance cognitive focus and processing.

Despite these findings, our study is not without limitations. The sequence in which the tests were administered might have influenced the performance due to participants' increasing familiarity with the task format. Furthermore, conducting all tests in the same room might have introduced distractions that could have affected the outcomes [1]. Future studies should consider these factors to refine the testing protocol and further validate our findings

Future work includes having another aspect and measuring heart rate and be able to record stress levels throughout the process. In the article, "Same Same but Different: Exploring the Effect Stroop Color Word Test In Virtual Reality" written by Romina Poguntke, Markus Wirth and Stefan Gradl, they conclude that "During the ... trial, participants showed slightly higher HR values in the VR condition and also for the head movement condition, compared to performing the Stroop test in front of a desktop screen" [20]. This could be very interesting to explore to see how stress can effect the brain's cognitive ability and what the most "calming" interface is for the user.

5.0.1 Limitations. During the experiment, we stayed consistent and had the participants do the paper test, computer test then VR test. This could have caused an issue because as the tests progressed, the participants could have become more familiar with the Stroop tests. This familiarity may have lead to better results the second time around because participants have become accustomed to the task's format, instructions, or requirements. By the time they got to the VR section, they had two other "practice tests". Another thing worth mentioning is that we conducted the three tests in the same room at the same time [11]. This resulted in everyone saying different colors out loud which could have effected performance [26]. The resolution of the HMD might have also had an effect on the amount of correct/incorrect answers. [22]

6 CONCLUSION AND FUTURE WORK

Our investigation has provided valuable insights into the comparative effectiveness of cognitive training across three distinct platforms: paper-based, computer-based, and virtual reality (VR). Our findings suggest that each platform offers unique benefits that can enhance cognitive performance [17], with VR demonstrating a particularly notable capability for maintaining high levels of accuracy and engagement due to its immersive nature [30].

REFERENCES

[1] Jeremy N. Bailenson, Andrew C. Beall, Jack Loomis, Jim Blascovich, and Matthew Turk. 2004. Transformed Social Interaction: Decoupling Representation from Behavior and Form in Collaborative Virtual Environments. *Presence: Teleoperators and Virtual Environments* 13, 4 (08 2004), 428–441. https://doi.org/10.1162/1054746041944803 arXiv:https://direct.mit.edu/pvar/article-pdf/13/4/428/1624137/1054746041944803.pdf

[2] Daniel Fernandes Barbosa, Francisco José A Prada, Maria Fátima Glanner, Otávio de Toledo Nóbrega, and Cláudio Córdova. 2010. Resposta cardiovascular ao Stroop: comparação entre teste computadorizado e verbal. *Arq. Bras. Cardiol.* 94, 4 (April 2010), 507–511.

[3] Sandhya Basu. 2023. Examining the reliability and validity of computerized stroop test in children aged 5-13 years: a preliminary study. *Qual. Quant.* 57, 1 (Feb. 2023), 645–653.

[4] Anna Cornelia Maria Bauer and Gerda Andringa. 2020. The potential of immersive virtual reality for cognitive training in elderly. *Gerontology* 66, 6 (Sept. 2020), 614–623.

[5] C.J. Bench, C.D. Frith, P.M. Grasby, K.J. Friston, E. Paulesu, R.S.J. Frackowiak, and R.J. Dolan. 1993. Investigations of the functional anatomy of attention using the stroop test. *Neuropsychologia* 31, 9 (1993), 907–922. https://doi.org/10.1016/0028-3932(93)90147-R

[6] Jack Block. 2005. The Stroop effect: its relation to personality. *Personality and Individual Differences* 38, 3 (2005), 735–746. https://doi.org/10.1016/j.paid.2004.05.027

[7] Fabio Buttussi and Luca Chittaro. 2018. Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. *IEEE Transactions on Visualization and Computer Graphics* 24, 2 (2018), 1063–1076. https://doi.org/10.1109/TVCG.2017.2653117

[8] Murat Coban, Yusuf Islam Bolat, and Idris Goksu. 2022. The potential of immersive virtual reality to enhance learning: A meta-analysis. *Educational Research Review* 36 (2022), 100452. https://doi.org/10.1016/j.edurev.2022.100452

[9] Stefania Coelli, Giulia Tacchino, Elisa Rossetti, Mario Veniero, Luigi Pugnetti, Francesca Baglio, and Anna Maria Bianchi. 2016. Assessment of the usability of a computerized Stroop Test for clinical application. In *2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI)*. 1–5. https://doi.org/10.1109/RTSI.2016.7740597

[10] Sophie Melissa Clare Davison, Catherine Deeprose, and Sylvia Terbeck. 2018. A comparison of immersive virtual reality with traditional neuropsychological measures in the assessment of executive functions. *Acta Neuropsychiatrica* 30, 2 (2018), 79–89.

[11] Mathyas Giudici, Eleonora Beccaluva, Mattia Gianotti, Jessica Barbieri, Giacomo Caslini, and Franca Garzotto. 2022. A Pilot Study on Interactive Multisensory Environments for Neuropsychological Assessment. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces* (Frascati, Rome, Italy) *(AVI 2022)*. Association for Computing Machinery, New York, NY, USA, Article 17, 5 pages. https://doi.org/10.1145/3531073.3531169

[12] Stefan Gradl, Markus Wirth, Nico Mächtlinger, Romina Poguntke, Andrea Wonner, Nicolas Rohleder, and Bjoern M. Eskofier. 2019. The Stroop Room: A Virtual Reality-Enhanced Stroop Test. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 28, 12 pages. https://doi.org/10.1145/3359996.3364247

[13] Clas Linnman, Per Carlbring, Åsa Åhman, Håkan Andersson, and Gerhard Andersson. 2006. The Stroop effect on the internet. *Computers in Human Behavior* 22, 3 (2006), 448–455. https://doi.org/10.1016/j.chb.2004.09.010

[14] Colin M MacLeod. 1980. The Stroop phenomenon: the reversal of interference produced by color-word mismatches. *Journal of Experimental Psychology: Human Perception and Performance* 6, 2 (1980), 228–238. https://doi.org/10.1037/0096-1523.6.2.228

[15] Deniz Mevlevioğlu, Sabin Tabirca, and David Murphy. 2021. Emotional Virtual Reality Stroop Task: Pilot Design. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology* (Osaka, Japan) *(VRST '21)*. Association for Computing Machinery, New York, NY, USA, Article 66, 3 pages. https://doi.org/10.1145/3489849.3489952

[16] Katharina Meyerbröker and Nexhmedin Morina. 2021. The use of virtual reality in assessment and treatment of anxiety and related disorders. *Clinical Psychology & Psychotherapy* 28, 3 (2021), 466–476. https://doi.org/10.1002/cpp.2623 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpp.2623

[17] Roxana Moreno and Richard Mayer. 2007. Interactive multimodal learning environments. *Educ. Psychol. Rev.* 19, 3 (Sept. 2007), 309–326.

[18] Thomas D. Parsons and Albert A. Rizzo. 2008. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry* 39, 3 (2008), 250–261. https://doi.org/10.1016/j.jbtep.2007.07.007

[19] Raveendranadh Pilli, Mur Naidu, Usha Rani Pingali, J C Shobha, and A Praveen Reddy. 2013. A computerized stroop test for the evaluation of psychotropic drugs in healthy participants. *Indian J. Psychol. Med.* 35, 2 (April 2013), 180–189.

[20] Romina Poguntke, Markus Wirth, and Stefan Gradl. 2019. Same Same but Different: Exploring the Effects of the Stroop Color Word Test in Virtual Reality. In *17th IFIP Conference on Human-Computer Interaction (INTERACT)*. Paphos, Cyprus, 699–708. https://doi.org/10.1007/978-3-030-29384-0_42 hal-02544574.

[21] Giuseppe Riva, Fabrizia Mantovani, and Andrea Gaggioli. 2004. Presence and rehabilitation: toward second-generation virtual reality applications in neuropsychology. *J. Neuroeng. Rehabil.* 1, 1 (Dec. 2004), 9.

[22] Ben D Sawyer, Benjamin Wolfe, Jonathan Dobres, Nadine Chahine, Bruce Mehler, and Bryan Reimer. 2020. Glanceable, legible typography over complex backgrounds. *Ergonomics* 63, 7 (2020), 864–883.

[23] Federica Scarpina and Sofia Tagini. 2017. The Stroop Color and Word Test. *Front. Psychol.* 8 (April 2017), 557.

[24] Shahnaz Shahrbanian, Xiaoli Ma, Najaf Aghaei, Korner-Bitensky N, and Maureen Simmonds. 2012. Use of virtual reality (immersive vs. non immersive) for pain management in children and adults: A systematic review of evidence from randomized controlled trials. *European Journal of Experimental Biology* 2 (01 2012), 1408–22.

[25] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1535 (Dec. 2009), 3549–3557.

[26] Kay M. Stanney, Ronald R. Mourant, and Robert S. Kennedy. 1998. Human Factors Issues in Virtual Environments: A Review of the Literature. *Presence: Teleoperators and Virtual Environments* 7, 4 (08 1998), 327–351. https://doi.org/10.1162/105474698565767 arXiv:https://direct.mit.edu/pvar/article-pdf/7/4/327/1623006/105474698565767.pdf

[27] Aleksei Tepljakov. 2020. Investigation of the Stroop Effect in a Virtual Reality Environment. (2020).

[28] Ngeemasara Thapa, Hye Jin Park, Ja-Gyeong Yang, Haeun Son, Minwoo Jang, Jihyeon Lee, Seung Wan Kang, Kyung Won Park, and Hyuntae Park. 2020. The effect of a virtual reality-based intervention program on cognition in older adults with mild cognitive impairment: A randomized control trial. *J. Clin. Med.* 9, 5 (April 2020), 1283.

[29] Wim Van der Elst, Martin P J Van Boxtel, Gerard J P Van Breukelen, and Jelle Jolles. 2006. The Stroop color-word test: influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment* 13, 1 (March 2006), 62–79.

[30] J.D. Westwood. 2011. *Medicine Meets Virtual Reality 18: NextMed*. IOS Press. https://books.google.com/books?id=2YZtot_CN_YC

[31] Henry Xiang. 2023. VIRTUAL REALITY (VR) AND THE KID'S BRAIN: EMERGING EVIDENCE FOR BRAIN COGNITIVE FUNCTION REHAB AND NEURO PAIN MECHANISM OF VR. *IBRO Neuroscience Reports* 15 (2023), S910–S911. https://doi.org/10.1016/j.ibneur.2023.08.1912 IBRO 11th World Congress of Neuroscience Supplement.