DEPARTMENT of PSYCHOLOGY

# Virtually Immersed in Language:

# a DTI Study on Rapid Neural Plasticity and Virtual Learning Environments in L2 acquisition

Inari Tuulia Kuusela

Master's Thesis (30 hp)

Fall 2023

Supervisors: Johan Mårtensson and Nicola Spotorno

## Acknowledgments

First, I would like to extend my deepest gratitude for my wonderful supervisors Johan Mårtensson and Nicola Spotorno who lent me their expertise, encouragement and patience for my numerous questions and confusion, and helped me turn my mere novice enthusiasm for neuroimaging into skill and knowledge. I enjoyed this project immensely and was lucky to have you both on board.

I would also like to wholeheartedly thank my family and friends, who have supported me throughout my entire education journey. You made me laugh when I got too serious, stood with me in times of entropy, and very importantly, politely smiled and nodded as I yet again went on a lengthy tangent about existence, brains, bodies and the psyche. Your efforts did not go unnoticed.

I truly appreciate you all for the kindness, support and love I have received

# Abstract

*Objective* Second language learning continues to be a field of growing interest both in academia and everyday life. As such, a more comprehensive understanding of the best learning methods remains an important area of focus. This study aimed to investigate single-session vocabulary acquisition and neural underpinnings of language learning through Diffusion Tensor Imaging (DTI), in relation to immersive language learning contexts.

*Method* A total of 46 native Swedish speakers were compared before and after learning Mandarin Chinese in three groups: Virtual reality learning group (VR), Video game learning group (VE) and Controls. Potential changes in white matter organization were assessed using fractional anisotropy (FA) and an immediate and delayed vocabulary test provided a link between neural changes and learning outcomes.

*Results* A significant increase in FA was found for the learning groups together (VR+VE) compared to the control group ($t(31) = 3.89$, $p < .001$) in the posterior part of the corpus callosum, known as the splenium. However, comparison of the learning groups separately revealed no significant differences in FA scores between the gaming modalities. This finding was further reflected in the vocabulary assessment: immediate and delayed recall indicated no significant differences between the gaming conditions ($F(1,30) = 3.58$, $p = 0.068$).

*Conclusion* These findings contribute to our understanding of the rapid neural reorganization underlying language learning. The results highlight the effect of language learning on white matter plasticity but do not support more immersive 3D virtual reality outperforming a traditional 2D game environment in this learning method.


Keywords: *language learning, diffusion tensor imaging, fractional anisotropy, virtual reality, virtual environment*

# Virtually Immersed in Language: a DTI Study on Rapid Neural Plasticity and Virtual Learning Environments in L2 Acquisition

Our first interaction with language starts before we are even born, and our natural ability to pick up tones, grammar, vocabulary and the signature use of our inherited language ensures the learning of communication. As defined by a famous linguist Noam Chomsky, language is an innate cognitive capacity in humans, characterized by universal underlying principles and structures that enable the creative generation and understanding of an infinite range of meaningful utterances (Chomsky, 2002). Language, as a complex cognitive and social phenomena, enables us to convey messages and comprehend ideas, thoughts and emotions of others and our own. It could be further suggested that rather than a mere collection of meaningful symbols and structure, language can be considered a large-scale, embodied experience.

Second language (L2) acquisition is in high demand in our growingly connected world. Learning another language provides not just enjoyment, but career and academic advancements, enriched travel experiences, cognitive benefits as well as enabling communication and connection with a wider range of people from different cultures and backgrounds. However, as we age, our brains have pruned an efficient, yet largely less flexible learning system making a new language acquisition a more effortful task (Van Hell, 2023). Therefore a comprehensive understanding of the most effective methods provide the basis for real-life applications, and in the field of L2 learning the potential benefits of immersive and enriched environments have garnered significant attention. Drawing inspiration from the natural processes of first language acquisition, researchers have explored the notion that second language learning could be enhanced through experiences and environments that mimic the multifaceted nature of real-life immersive contexts

(Li et al., 2014; Parmaxi, 2023). This study aims to further illuminate this notion and investigate language learning in immersive environments that resemble a natural learning experience.

The way of teaching a new language has stayed relatively unchanged for a long time. A classroom environment combined with text-based vocabulary, grammar and translation exercises are still the most common practices at learning institutes. Conversely, L2 acquisition is known to be the most comprehensive and fast in immersive real-life contexts: study abroad experiences (Stein et al., 2012), and other, either real-life or simulation of it, communicative contexts and interpersonal interactions have been found to yield better learning outcomes (Young et al., 2012). However, offering such possibilities can be costly and challenging to arrange both for the individuals and educational institutes. Moreover, the level of immersion needed for beneficial outcomes is not yet comprehensively understood. Alternative learning methods, such as digital learning spaces have been recognised as advantageous on numerous domains, subjective and objective, regarding learning. The gamified learning environments have been assessed as more motivating, engaging and fun (Bedwell et al., 2012; Young et al., 2012), generating improvement over general language skills, grammar and particularly in vocabulary (Zhonggen, 2018), while also providing relatively accessible immersive conditions. Virtual environments also make possible varied levels of immersion, from a less immersive conventional video game to the highly immersive virtual reality. Different types of immersion can allow us to explore how we simulate experiences in our brains: does the brain respond similarly when physically moving ourselves around an environment versus moving an avatar in the same environment in a 2D form. It could be that despite the differences in subjective immersiveness, the brain does not differentiate between these states on a neural level (Calvo-Merino et al., 2005), since it needs to simulate any movement it sees to be able to understand it. If this is the case, virtual reality and traditional computer games might produce similar neural alterations and learning outcomes, making the more expensive VR environment unnecessary for equally proficient learning.

Presently, the current body of knowledge comparing different learning environments focuses mainly on either the cognitive and behavioral outcomes or the neural changes, particularly in the context of immersive learning interventions (Connolly et al., 2012; Parmaxi, 2023). In a real-life setting, Stein et al (2012) found significant increases in grey matter correlated with language proficiency among participants studying German in Germany, while the potential benefits of virtual reality have mostly been explored for linguistic and cognitive outcomes (Lan et al., 2015; Mroz, 2015). Nevertheless, examining neurocognitive aspects of immersive learning environments remain relatively scarce, despite their promising advantages. The prior work by Mårtensson et al. (2011;2012) discovered a broader understanding by combining both approaches. Their investigation into intensive language learning unveiled unexpected and non-uniform hippocampal volume changes which were linked to measures of educational struggle and proficiency and thereby revealing an intriguing interaction between brain plasticity and study situation. Importantly, neither approach in isolation would have been sufficient to uncover this effect. Furthermore, most previous research has focused on relatively long time frames, from weeks to months, for their language tasks. Studies by Sagi et al (2012) and Hofstatter et al. (2016) show preliminary results for very brief study sessions (1-2 hours) reflecting adaptations on the neural level, yet much remains to be discovered about the timescale of neuroplasticity regarding short-term language learning. Building on these prior findings, this study aims to bridge a gap in current research by focusing on the rapid neurocognitive aspects underlying L2 learning and relate it to the immersive environments of virtual learning.

Various brain imaging methods have been used to explore the structural and functional neural aspects of language learning, most commonly (f)MRI and EEG. As a relatively recent method, Diffusion Tensor Imaging (DTI) is less employed for cognitive and behavioral studies, while it is well established in clinical research (Tae et al., 2018). DTI is an analysis strategy based on diffusion weighted imaging, a magnetic resonance (MRI) technique that measures the diffusion of water molecules in biological tissues, such as the brain (Assaf & Pasternak, 2008). Specifically, DTI can help infer information about white-matter connectivity and organization by leveraging the differential diffusion properties of water in more (e.g., white matter tracts) or less (e.g., cerebral spinal fluid) restricted spaces. DTI represents water displacement as a tensor, which describes the magnitude and direction of water diffusion in three-dimensional space (Assaf & Pasternak, 2008). This information can be used to identify and quantify changes in white matter structure and connectivity, making DTI a valuable tool for investigating individual variability in white matter pathways. For language research it has mostly been utilized for anatomical changes in language disorders, despite the unique perspective it can provide on how the brain adapts and reorganizes itself during the language learning process.

Several different measures within DTI can be chosen to represent the changes in white matter, but Fractional Anisotropy (FA) serves as the most commonly used metric for measuring diffusion in highly anisotropic environments such as white matter tracts. FA assumes values between 0 and 1, calculated from the DTI tensor (Assaf & Pasternak, 2008). It reveals changes in the white matter and examining changes in FA can reveal the neural plasticity associated with learning processes - with higher FA in tract values potentially reflecting a higher level of coherence in the tracts. Identifying specific tracts showing significant FA differences allows a deeper understanding of the effectiveness of interventions and additionally, can serve as biomarkers for monitoring individual differences in learning abilities and potential predictors for learning outcomes.

Central brain areas and networks linked to language learning have been identified in numerous studies, demonstrating the multimodal involvement needed for language processing. Traditionally the aspects of language are considered have a left-brain emphasis, with many key areas such as the left inferior frontal gyrus (IFG), commonly referred to as Broca's area (Broca, 1861), involved in syntax and speech production and the superior temporal gyrus (STG), known as Wernicke's area (Wernicke, 1874), involved in comprehension and auditory information. In contrast, the more tonal aspects of language including prosody, intonation, and emotional expression in speech have been associated with areas in the right hemisphere (Price, 2010). DTI studies have further identified the involvement of specific white matter tracts, such as the Arcuate Fasciculus (AF) connecting Broca's and Wernicke's areas (Catani & Mesulam, 2008) , the Superior Longitudinal Fasciculus (SLF) integrating information from frontal, parietal, and temporal regions (Kamali et al., 2014), the Uncinate Fasciculus (UF) connecting the anterior temporal lobe, including the amygdala and the hippocampus, with the frontal cortex (Papagno, 2011) and the Inferior Fronto-Occipital Fasciculus (IFOF) connecting the frontal and occipital lobes (Friederici & Gierhan, 2013). Additionally, the corpus callosum (CC) further facilitates communication and information transfer between the language-related regions (Catani & Mesulam, 2008). Overall, the extensive engagement and integration of multiple brain regions including auditory, visual, motor, and emotion-linked areas, suggest the importance of multimodal information in language processing. These findings provide a neural basis for the concept that language learning incorporating multiple modalities in an enriched learning environment can support memory encoding and improve comprehensive learning outcomes.

The overarching objective of the present study is to enhance our understanding of optimal L2 learning contexts by investigating the impact of short-term learning and immersiveness in virtual learning modalities on white matter adaptations. The study aims to examine neural changes by comparing two levels of immersion: conventional virtual environment (VE) and the more immersive virtual reality (VR). Furthermore, the study will assess vocabulary acquisition as an outcome measure to expand on the knowledge of the learning environment influence. The first objective is to compare combined learning groups (VR+VE) with a control group after a short-term language learning task, hypothesizing that the learning groups will exhibit increased fractional anisotropy (FA) in one or multiple white matter tracts compared to the controls. The second objective is to compare the learning groups to each other, with the expectation the higher immersion levels of the VR group will reflect in higher FA in the white matter tracts than the VE group. Additionally, the study will evaluate the acquired language using both an immediate and a delayed vocabulary test, providing a link between neural changes and learning outcomes. An enhanced vocabulary recall is expected from the more immersive learning environment at both timepoints. This study will contribute to the existing language learning literature by illuminating rapid neural adaptations underlying language learning, and guiding the development of targeted interventions, optimizing learning processes and educational practices.

# Methods

## Participants

Twenty-eight female and nineteen male native Swedish speakers were enrolled through a digital participant recruitment platform Accindi (http://accindi.se), between the ages of 19 and 40 (M = 25.6, SD = 5.34). The 47 participants had no prior experience in studying Mandarin Chinese and all reported being right-handed. Participants were given an oral and written description of the study including any potential psychological or physical concerns that might arise during the process as well as their right to withdraw consent at any time without giving a reason. Informed consent was obtained from each participant, and the study was conducted in accordance with the principles outlined in the Declaration of Helsinki as well as an ethical approval was obtained from the Lund University ethics committee. Furthermore, participant data was encoded and stored in an anonymous manner to ensure confidentiality and privacy.

## Materials

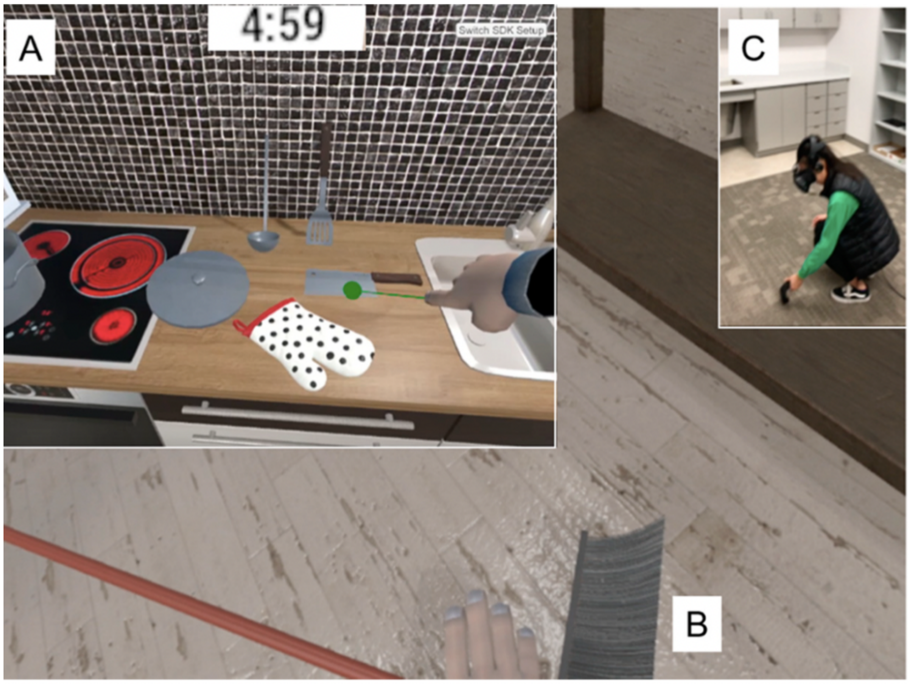
The VR and VE gaming environments utilized in this study were acquired in collaboration with Pennsylvania State University and are based on the earlier work by Lan et al. (2015) investigating the impact of different learning contexts on language learning, yet were further developed for the present study at the Lund Humanities Laboratory.

### *Virtual Reality*

The VR game environment simulates a kitchen as a realistic, physical space. Within the kitchen setting, participants encounter 30 kitchen-related items that they can freely explore and interact with. By physically moving around the virtual kitchen and manipulating the objects, participants engage in a multimodal learning experience. Upon pointing at an item, the corresponding Mandarin name is audibly presented, allowing the participant to associate the words with the objects they interacted with. Figure 1 demonstrates an example of how the environment is set up. To ensure participants' interaction with all items, a red arrow prompts participants to engage any objects they might have missed before the session ends. This environment promotes enriched learning in a life-like manner, with multiple sensory inputs, including tactile, vision, and auditory.

**Figure 1**

*The Game Environment*




*Note*. An illustration of the 3D kitchen environment (Lan et al., 2015). The participant was able to point objects (A), interact with them (B) while physically moving around (C).

### *Virtual Environment*

The VE game environment mirrored the VR condition with the same interactive kitchen setting. However, in this condition, the environment was presented as a 2D game on a laptop and the participants had the possibility to navigate the kitchen area and interact with the objects using a mouse and keyboard.

### *The vocabulary recall task*

The language proficiency test was developed using the PsychoPy software (Peirce et al., 2019) and conducted on a laptop computer at the Humanities laboratories. The test aims to assess the participants' knowledge of the 30 Mandarin Chinese kitchen-related words that they had been exposed to during their learning sessions.

During the test, and in line with multisensory learning, participants are presented with auditory words and simultaneously shown three images on the screen. The images consist of both VR/VE rendered objects that are representative of the items they encounter during the learning sessions. The participant is required to select the image that corresponds to the auditory word they have just heard, repeating this process for all 30 words.

**Procedure**

This thesis has been refined in a context of a larger-scale study, in which all the measures are to be integrated answering additional and more extensive research questions not in the scope of this thesis. All the steps are described in the following, but including the more in-depth descriptions only relevant for this study, also illustrated in Figure 2.

First, participants were given a thorough overview of the study, its purpose and their right to withdraw at any point, after which they signed informed consent and filled out a questionnaire considering demographic questions such as age, education, handedness as well as language and gaming background. Before the first scanning took place, participants were initially administered a battery of behavioral and cognitive measures encompassing working memory, attention, inhibitory control, language aptitude, and background assessments. These measures included the llama B and D tasks (Meara, 2005), N-back (Kirchner, 1958), Stroop tests (Stroop, 1935), and Raven's Progressive Matrices (Raven, 2000) that served to establish the participants' cognitive profiles and provide a baseline for subsequent analyses. Following the completion of the behavioral assessments, participants underwent brain scans to acquire high-resolution Diffusion

MRI scans, conducted using state-of-the-art imaging equipment, the Siemens 7 tesla scan located at Lund University laboratory.

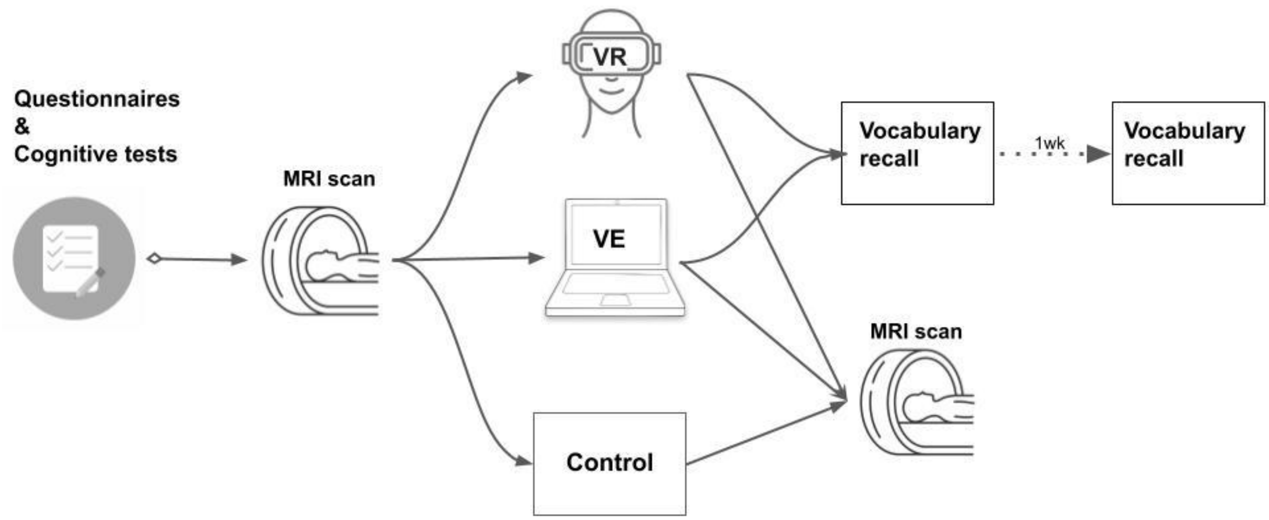
Participants were then randomly assigned to one of three groups based on the study design with the overarching objective to investigate enriched language learning environments. The first group, referred to as the virtual reality (VR) group, engaged in the learning sessions within a highly immersive virtual reality environment at the Lund Humanities laboratory. This environment allowed participants to actively explore and interact in a simulated physical space with virtual kitchen objects and milieu, facilitating an enriched, embodied learning experience. The second group did their language training in a Virtual Environment (VE), at which the learning context was a conventional computer-based game. The group interacted with the virtual kitchen environment and engaged in the learning activities in a similar manner as the VR group. The learning sessions for both the VR and VE groups were conducted over the duration of approximately 30 minutes. During this time, participants familiarized themselves with the kitchen objects and learned their corresponding Mandarin names. To establish a control condition, a third group of participants were designated to watch a nature documentary during the study period, providing a neutral comparison for the effects specific to the learning interventions.

Following the completion of the learning sessions, participants underwent a second brain scan to capture potential changes in the brain resulting from the learning experience. Furthermore, to evaluate the immediate impact of the interventions in another modality, participants completed a vocabulary recall test after the learning sessions. This test assessed their ability to recall the 30 Mandarin names of the kitchen objects they had been exposed to during the learning phase.

Approximately one week later (M = 6.78, SD = 3.01), participants returned for a delayed vocabulary test. This follow-up assessment allowed for an examination of the retention and recall of the learned vocabulary over an extended period, providing a basis for insights into the retention of the acquired language knowledge.

**Figure 2**

*The Study Procedure*




*Note*. A flow chart displaying the steps in the study procedure

## Imaging

### Imaging protocol

The DTI imaging was performed using a 7 Tesla Philips Achieva MRI scanner. A single-shot echo-planar imaging sequence was used to acquire 30 diffusion-weighted imaging volumes (repetition time: 8010 msec; echo time: 65 msec; resolution: 2x2x2 $mm^3$; b values range: 0,100 and 1000 sec/$mm^2$ distributed over 3, 6 and 30 directions). The Phase Encoding Direction was set to Anterior-Posterior, enabling efficient encoding of spatial information during image acquisition. These specifications ensured high-quality DTI images with detailed information about white matter tracts, enabling accurate analysis and interpretation of the data.

**Preprocessing**

Various artifacts can affect the quality of imaging data, which, in turn, can influence the reliability and validity of the following analyses and interpretations. To ensure optimal data quality for white matter fiber tract dissection in the subsequent steps, all images underwent meticulous manual inspection and subsequent automated preprocessing using FSL and MRtrix3 (Jenkinson et al., 2012; Tournier et al., 2019). One participant was excluded from the analysis at this stage due to inadequate imaging quality. The preprocessing pipeline encompassed essential steps including denoising procedures and corrections for image distortions arising from motion artifacts and eddy current-induced distortions.

**Tractography**

The whole brain white matter tracts were segmented using Tractseg (version 2.5; Wasserthal et al., 2018). TractSeg is an automated segmentation program for identifying and delineating major white matter bundles in diffusion MRI images. It utilizes a combination of deep learning and tractography techniques, drawing from previous research and manual dissection models. To standardize the brains for analysis, the pre and post scans were aligned to the Montreal Neurological Institute atlas (MNI) (Fonov et al., 2011), which serves as a common spatial reference. Subsequently, a pipeline was built to identify the start and end points of each tract, generate bundle-specific orientation maps, and combine this information to create 72 tract bundles for each individual scan.

In line with typical DTI imaging at 7T, some scans exhibited signal loss in regions adjacent to air-filled spaces such as sinuses and ears, as well as in the cerebellum, where interrupted signal is commonly observed. This signal loss resulted in incomplete tracts for 13 participants, with 6 individuals having one incomplete bundle and 7 individuals having multiple incompleted bundles. The majority of the incomplete bundles were located in the cerebellum area, with some also found in the frontal and temporal lobes. However, considering the reasonable extent of the data loss, no exclusion of participants was deemed necessary.

## Analysis

The present study investigated potential group differences in behavioral measures and fractional anisotropy (FA) in white matter tracts. Prior to statistical analysis, the normality of scores and homogeneity were carefully examined via visual and numerical approaches using Q-Q plots, Levene's test and Mauchly's test of sphericity, ensuring that all assumptions for proceeding with parametric tests were satisfied. In order to determine similarity of groups at baseline an analysis of the behavioral measures and comparison of the groups for differences in means of scores for each measure was carried out in SPSS (version 29). Analysis of Variance (ANOVA) was performed on the results from behavioral and cognitive measures tasks for all three groups. Furthermore, to investigate whether the more or less enriched gaming conditions made a difference in learned kitchen vocabulary, a repeated measures ANOVA was performed with the learning modality (VE and VR) as the between-subjects factor and time of testing (immediately and delayed word recall) as the within-subjects factor.

The tractography analysis was conducted in Tractseg (Wasserthal et al., 2018; Yeatman et al., 2012), which employs a multiple permutation-based test to compare and correct for multiple comparisons across each white matter bundle. First, the FA scores of pre and post measurements were subtracted to obtain a difference score for each individual after which the FA analysis on the difference scores was performed on thirteen bundles, accounting for multiple comparisons (FWE correction).

As a final step to enhance the robustness of the possible findings, a signal smoothing procedure was implemented. Tractseg allows for segmentation of each tract in the bundle into 100 segments, enabling more precise examination of different parts of the tract. However, smaller segments are more susceptible to false positives and local inhomogeneity of the signal. Therefore, the tracts were analyzed using larger segmentations (i.e., only 10 segments for tract) to increase the confidence in the interpretation of the findings.

# Results

## Cognitive Data and Vocabulary Recall

A total of 46 participants were considered in the final analysis: 17 in the VR group, 15 in the VE group and 14 in the controls. No significant differences were found in the ANOVA for the cognitive and language abilities among the groups. The detailed ANOVA results for each of these measures obtained from the analysis can be found in Appendix A.

Regarding the vocabulary assessment, a significant main effect of the time of testing was observed. However, interaction of the temporal factor on the learning modality did not yield significant differences in learning outcomes $F(1,30) = 0.20$, $p = 0.656$. Furthermore, the main effect between the VR and VE conditions on learned kitchen vocabulary did not show statistically significant differences between the learning modalities $F(1,30) = 3.58$, $p = 0.068$. The detailed ANOVA results are presented in Table 1.

**Table 1**

*Repeated measures ANOVA of Vocabulary Recall*

Within Subjects Effects

| | Sum of Squares | df | Mean Square | F | p | η² |
|---|---|---|---|---|---|---|
| Time of testing | 0.011 | 1 | 0.011 | 9.48 | .004 | .018 |
| Time of testing ✻ learning condition | 2.33e-4 | 1 | 2.33e-4 | 0.20 | .656 | .000 |
| Residual | 0.034 | 30 | 0.001 | | | |

Between Subjects Effects

| | Sum of Squares | df | Mean Square | F | p | η² |
|---|---|---|---|---|---|---|
| learning condition | 0.059 | 1 | 0.059 | 3.58 | .068 | .098 |
| Residual | 0.492 | 30 | 0.016 | | | |

*Note.* Results of the immediate and delayed vocabulary recall assessment in VR and VE

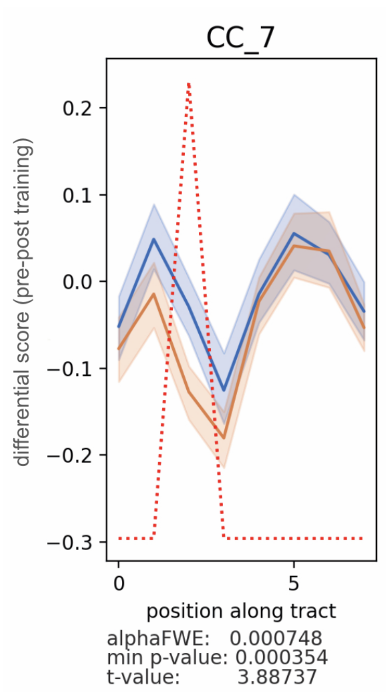
## DTI data

### *Learning versus Control*

When comparing the learning conditions (VR+VE) to the control group, a significant difference $t(31) = 3.89$, $p < .001$ in fractional anisotropy (FA) was observed in one specific white matter bundle CC_7, namely the splenium, in the posterior part of the corpus callosum. Higher FA values were found in the post-measurements of the learning groups leading to negative differential scores. Figure 3 visualizes the significant result between the groups in FA and its position in the bundle, while Figure 4 provides an anatomical depiction of the tract in the brain and the FA scores along its trajectory. This finding lends support to the initial hypothesis of white matter alterations associated with the learning conditions compared to the neutral control group.

**Figure 3**

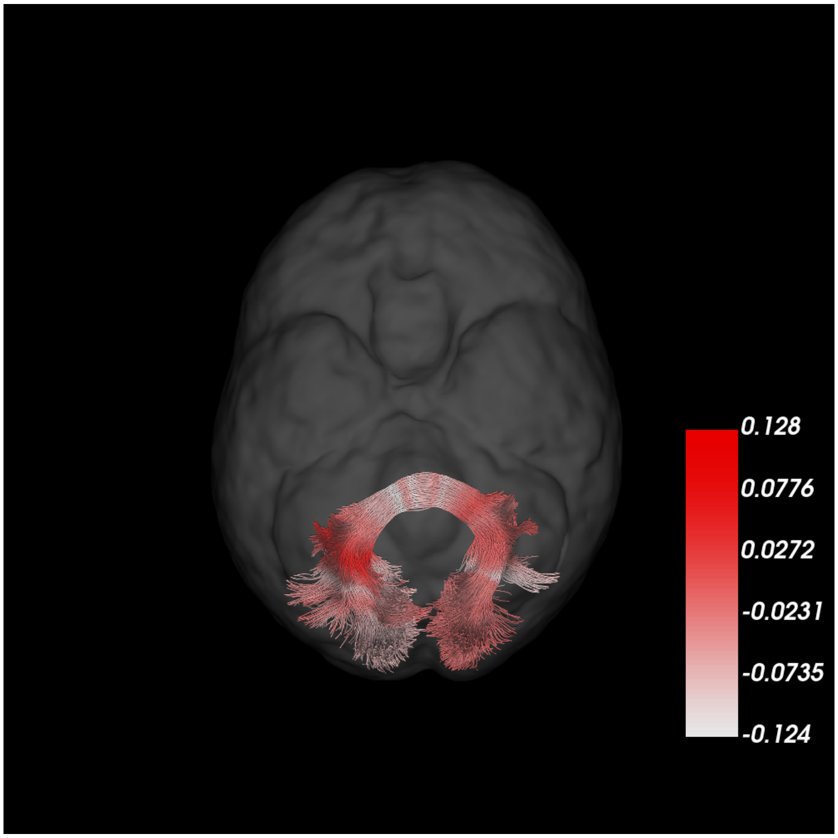
*Comparison of Pre-Post FA Values between Learning Groups (VR+VE) and Control Group*




*Note*. The blue line represents the control group, while the red line represents the learning groups (VR+VE), the shaded area marking the 95% confidence interval. AlphaFWE is the alpha value adjusted for multiple comparisons (multiple positions per bundle and multiple bundles). The minimum p-value is calculated for each bundle (minimum over all 10 positions) and bundles with a min p-value < alphaFWE indicate significant results while the red dotted line denotes the positions within the bundle where p-value < alphaFWE.

**Figure 4**

*The Splenium of Corpus Callosum*




*Note*. Inferior view of the tract bundle splenium of corpus callosum. The coloring showcases the pre-post difference scores of FA along the tract.

***VR versus VE***

The second hypothesis suggested higher FA in the post measurements for the more enriched environment VR in comparison to the less immersive VE. The analysis revealed no tracts had a significant difference in the differential FA scores between the two language game conditions, thus refuting the second hypothesis. A full display of all the DTI results can be discovered in Appendix B.

## Discussion

This study investigated short-term language learning and the impact of virtual language learning environments on white matter alterations in the brain. Alongside brain imaging analyses, learning outcomes were assessed using a vocabulary test administered at two different time points. The results indicated microstructural alterations as increased FA in the posterior part of the corpus callosum, commonly referred to as the splenium. However, contrary to the initial hypothesis, the results did not support the notion that the more immersive virtual reality-based learning outperformed the computer game based learning, as indicated by both the statistically insignificant FA analysis and vocabulary test outcomes. These findings contribute to the fields of second language acquisition as well as rapid learning-related neuroplasticity and virtual environment learning.

To ensure methodological rigor and minimize confounding factors, this study accounted for potential background influences by implementing a comprehensive set of cognitive and behavioral measures. These measures served as a baseline assessment, enabling the identification of any pre-existing differences among the participant groups. Specifically, assessments encompassed domains such as working memory, attention, inhibitory control, language aptitude, and a self-rated language background questionnaire. Importantly, the results indicated no significant differences between the groups, suggesting that they started the study on an equal footing enhancing the reliability and robustness of the subsequent comparisons made in the analysis.

The first research question aimed to investigate the effects of a single, short-form language study session on microstructural organization by comparing the combined learning groups (VR+VE) to the no-study control group, with the expectation of observing significant alterations in white matter among the former. The FA analysis revealed a notable change in FA in the splenium of the corpus callosum. Within the context of language learning, the corpus callosum plays a crucial role in integrating linguistic information and coordinating language-related processes (Catani & Mesulam, 2008). While the precise function of the splenium remains only partially understood, the current knowledge highlights its intricate connections with the hippocampi and suggests its involvement in facilitating the transfer of visual and linguistic information between the cerebral hemispheres (Blaauw & Meiners, 2020), as well as the interplay between prosody and syntax during speech comprehension (Sammler et al., 2010). Overall, the splenium of corpus callosum is linked to transmitting information between several multimodal features in language processes. Hence, this finding underscores the effect of language learning altering the brain and an enhanced level of relaying and integrating of linguistic information in the process of language learning, aligning with the prior findings in the field.

The second hypothesis aimed to investigate potential differences in FA between the different learning modalities, expecting the more immersive VR environment to produce superior neural adaptations in comparison to the computer game. Contrary to this assumption, no significant differences among the participants were observed in the FA analysis. It is plausible that while our subjective experience may differ when moving and interacting within virtual environments compared to a two-dimensional screen, our brains do not necessarily differentiate between these modalities in terms of the neural responses elicited: previous studies have demonstrated similar patterns of neural activation when merely imagining a movement as compared to physically executing it (Calvo-Merino et al., 2005; Jeannerod, 1995; Kilteni et al., 2018; Pearson, 2019). Therefore, from a learning perspective, it is possible that our brains respond in a similar manner regardless of whether the learning environment is presented in a 2D, or a more real-life equivalent, 3D format.

The differences in learning outcomes between the gaming groups were also investigated through a vocabulary test administered immediately after the learning session and repeated approximately one week later. Overall participants recalled fewer words during the second assessment, reflecting a natural decay in memory over time. However, no significant differences in vocabulary recall were observed either between the VR and VE modalities or at either time point. These findings, coupled with the results of the FA analysis, collectively suggest that the immersive VR environment does not confer an advantage over the more conventional VE in terms of language acquisition at both neural and cognitive levels.

However, it is important to acknowledge certain methodological constraints within this study. Although the study focused specifically on rapid neural adaptations, the designated timeframe employed may not have been sufficient to reliably capture discernible differences between the levels of immersiveness. Future research on virtual language contexts would benefit from additional study sessions in order to shed light on the specific temporal dynamics of emerging distinction. Furthermore, the study employed solely a passive, non-study group as controls. While the alignment of the results with prior established research is noteworthy, it is possible the microstructural changes between the learning groups and controls could be attributed more to the nature of engagement in an active gaming task rather than a distinct language-specific effect. By implementing a control group engaging in another activity or a group undertaking a more conventional yet less immersive language learning method, could provide a more robust overview about the rapid effects of learning in the brain. Lastly, another consideration is this study consisted exclusively of adults. Children and adults are suggested to employ brains differently during language learning processes (Ojima et al., 2005; Olesen et al., 2006; Van Hell, 2023), yet studies focusing on children's neural underpinnings of L2 acquisition are not widely conducted. Exploring the interplay of age-related factors in L2 learning has the potential to deepen our understanding in the field. Overall, future research should consider exploring varied short-form timeframes and amount of sessions, different types of control group configurations and age cohorts.

## Conclusion

The present study investigated rapid changes in white matter plasticity in the brain as a result of language learning. The results showed support for language rapidly altering our brains, highlighting the effect of learning on white matter organization. However, virtual reality does not appear to be an advantage in this particular learning method. Future research into temporal dynamics and virtual language learning context might consider employing varied parameters for time, age and type of learning activity to expand on the knowledge gained from this study.

Examining language from a neurocognitive perspective is a valuable approach in the field. By identifying specific tracts associated with these processes and the timeframe related with these changes, we gain a deeper understanding of the brain's structural adaptations during learning and identify brain regions and networks that are involved in specific language-related processes. It is also important to be cautious in making causal claims about the relationship between white matter changes and learning, as there are likely to be many factors that contribute to both. Overall, investigating neurocognitive changes benefits our understanding of the neural basis of learning, enhances theories of neuroplasticity, and has practical implications for educational interventions, cognitive rehabilitation, and personalized learning approaches.

## References

Assaf, Y., & Pasternak, O. (2008). Diffusion tensor imaging (DTI)-based white matter mapping in brain research: A review. *Journal of Molecular Neuroscience*, *34*, 51–61.

Bedwell, W. L., Pavlas, D., Heyne, K., Lazzara, E. H., & Salas, E. (2012). Toward a Taxonomy Linking Game Attributes to Learning: An Empirical Study. *Simulation & Gaming*, *43*(6), 729–760. https://doi.org/10.1177/1046878112439444

Blaauw, J., & Meiners, L. C. (2020). The splenium of the corpus callosum: Embryology, anatomy, function and imaging with pathophysiological hypothesis. *Neuroradiology*, *62*(5), 563–585. https://doi.org/10.1007/s00234-019-02357-z

Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé, suivies d’une observation d’aphémie (perte de la parole). *Bulletin et Memoires de La Societe Anatomique de Paris*, *6*, 330–357.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action Observation and Acquired Motor Skills: An fMRI Study with Expert Dancers. *Cerebral Cortex*, *15*(8), 1243–1249. https://doi.org/10.1093/cercor/bhi007

Catani, M., & Mesulam, M. (2008). The arcuate fasciculus and the disconnection theme in language and aphasia: History and current state. *Cortex*, *44*(8), 953–961. https://doi.org/10.1016/j.cortex.2008.04.002

Chomsky, N. (2002). *On nature and language*. Cambridge University Press.

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, *59*(2), 661–686. https://doi.org/10.1016/j.compedu.2012.03.004

Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., & Collins, D. L. (2011). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, *54*(1), 313–327. https://doi.org/10.1016/j.neuroimage.2010.07.033

Friederici, A. D., & Gierhan, S. M. (2013). The language network. *Current Opinion in Neurobiology*, *23*(2), 250–254. https://doi.org/10.1016/j.conb.2012.10.002

Hofstetter, S., Friedmann, N., & Assaf, Y. (2017). Rapid language-related plasticity: Microstructural changes in the cortex after a short session of new word learning. *Brain Structure and Function*, *222*, 1231–1241.

Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia*, *33*(11), 1419–1432.

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Kamali, A., Flanders, A. E., Brody, J., Hunter, J. V., & Hasan, K. M. (2014). Tracing superior longitudinal fasciculus connectivity in the human brain using high resolution diffusion tensor tractography. *Brain Structure and Function*, *219*, 269–281.

Kilteni, K., Andersson, B. J., Houborg, C., & Ehrsson, H. H. (2018). Motor imagery involves predicting the sensory consequences of the imagined movement. *Nature Communications*, *9*(1), 1617. https://doi.org/10.1038/s41467-018-03989-0

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352–358. https://doi.org/10.1037/h0043688

Lan, Y.-J., Fang, S.-Y., Legault, J., & Li, P. (2015). Second language acquisition of Mandarin Chinese vocabulary: Context of learning effects. *Educational Technology Research and Development*, *63*(5), 671–690. https://doi.org/10.1007/s11423-015-9380-y

Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex*, *58*, 301–324. https://doi.org/10.1016/j.cortex.2014.05.001

Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., & Lövdén, M. (2012). Growth of language-related brain areas after foreign language learning. *NeuroImage*, *63*(1), 240–244. https://doi.org/10.1016/j.neuroimage.2012.06.043

Mårtensson, J., & Lövdén, M. (2011). Do Intensive Studies of a Foreign Language Improve Associative Memory Performance? *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00012

Meara, P. (2005). LLAMA language aptitude tests: The manual. *Swansea: Lognostics*.

Mroz, A. (2015). The development of second language critical thinking in a virtual language learning environment: A process-oriented mixed-method study. *CALICO Journal*, *32*(3), 528–553. https://doi.org/10.1558/cj.v32i3.26386

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP Study of Second Language Learning after Childhood: Effects of Proficiency. *Journal of Cognitive Neuroscience*, *17*(8), 1212–1228. https://doi.org/10.1162/0898929055002436

Olesen, P. J., Macoveanu, J., Tegner, J., & Klingberg, T. (2006). Brain Activity Related to Working Memory and Distraction in Children and Adults. *Cerebral Cortex*, *17*(5), 1047–1054. https://doi.org/10.1093/cercor/bhl014

Papagno, C. (2011). Naming and the role of the uncinate fasciculus in language function. *Current Neurology and Neuroscience Reports*, *11*, 553–559.

Parmaxi, A. (2023). Virtual reality in language learning: A systematic review and implications for research and practice. *Interactive Learning Environments*, *31*(1), 172–184. https://doi.org/10.1080/10494820.2020.1765392

Pearson, J. (2019). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience*, *20*(10), 624–634. https://doi.org/10.1038/s41583-019-0202-9

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Price, C. J. (2010). The anatomy of language: A review of 100 fMRI studies published in 2009.

*Annals of the New York Academy of Sciences*, *1191*(1), 62–88.

Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology*, *41*(1), 1–48.

Sammler, D., Kotz, S. A., Eckstein, K., Ott, D. V. M., & Friederici, A. D. (2010). Prosody meets syntax: The role of the corpus callosum. *Brain*, *133*(9), 2643–2655. https://doi.org/10.1093/brain/awq231

Stein, M., Federspiel, A., Koenig, T., Wirth, M., Strik, W., Wiest, R., Brandeis, D., & Dierks, T. (2012). Structural plasticity in the language system related to increased second language proficiency. *Cortex*, *48*(4), 458–465. https://doi.org/10.1016/j.cortex.2010.10.007

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Tae, W.-S., Ham, B.-J., Pyun, S.-B., Kang, S.-H., & Kim, B.-J. (2018). Current Clinical Applications of Diffusion-Tensor Imaging in Neurological Disorders. *Journal of Clinical Neurology*, *14*(2), 129. https://doi.org/10.3988/jcn.2018.14.2.129

Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H., & Connelly, A. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, *202*, 116137. https://doi.org/10.1016/j.neuroimage.2019.116137

Van Hell, J. G. (2023). The Neurocognitive Underpinnings of Second Language Processing: Knowledge Gains from the Past and Future Outlook. *Language Learning*, lang.12601. https://doi.org/10.1111/lang.12601

Wasserthal, J., Neher, P., & Maier-Hein, K. H. (2018). TractSeg—Fast and accurate white matter tract segmentation. *NeuroImage*, *183*, 239–253. https://doi.org/10.1016/j.neuroimage.2018.07.070

Wernicke, C. (1874). *Der aphasische Symptomencomplex: Eine psychologische Studie auf anatomischer Basis*. Cohn & Weigert.

Yeatman, J. D., Dougherty, R. F., Myall, N. J., Wandell, B. A., & Feldman, H. M. (2012). Tract Profiles of White Matter Properties: Automating Fiber-Tract Quantification. *PLoS ONE*, *7*(11), e49790. https://doi.org/10.1371/journal.pone.0049790

Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., Simeoni, Z., Tran, M., & Yukhymenko, M. (2012). Our Princess Is in Another Castle: A Review of Trends in Serious Gaming for Education. *Review of Educational Research*, *82*(1), 61–89. https://doi.org/10.3102/0034654312436980

Zhonggen, Y. (2018). Differences in serious game-aided and traditional English vocabulary acquisition. *Computers & Education*, *127*, 214–232. https://doi.org/10.1016/j.compedu.2018.07.014

# Appendix A

**Table 1**

*Univariate ANOVA Results for Cognitive Measures by Group*

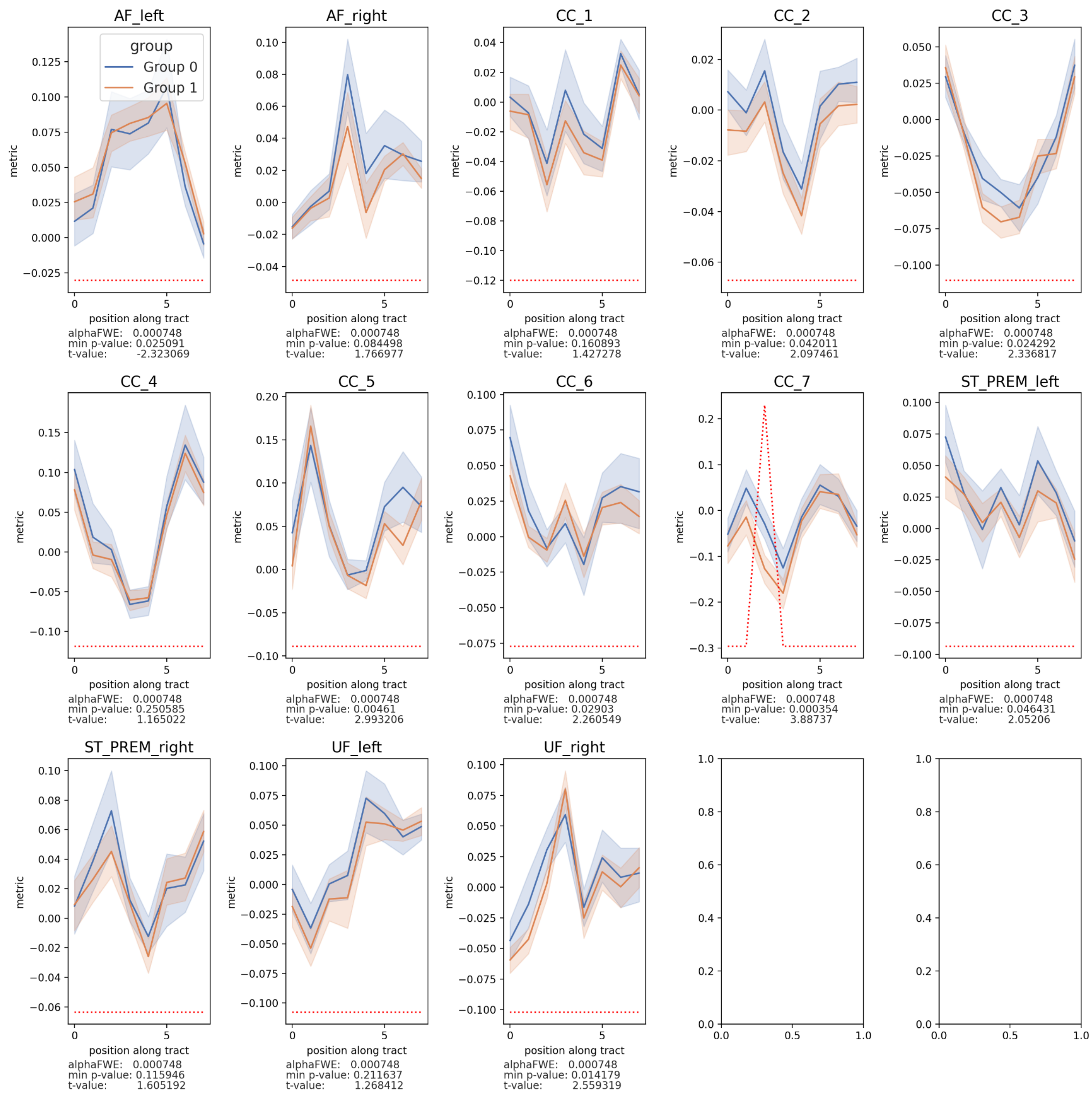
| | F | df | p |
|---|---|---|---|
| N_back | 1.87 | 2 | .173 |
| SART | 0.25 | 2 | .781 |
| Ravens | 1.89 | 2 | .168 |
| Stroop | 0.10 | 2 | .904 |
| llamaB | 0.33 | 2 | .722 |
| llamaD | 1.89 | 2 | .169 |

*Note.* Overview of all the cognitive measures conducted at the start of the study across all three groups

# Appendix B

## Figure B1

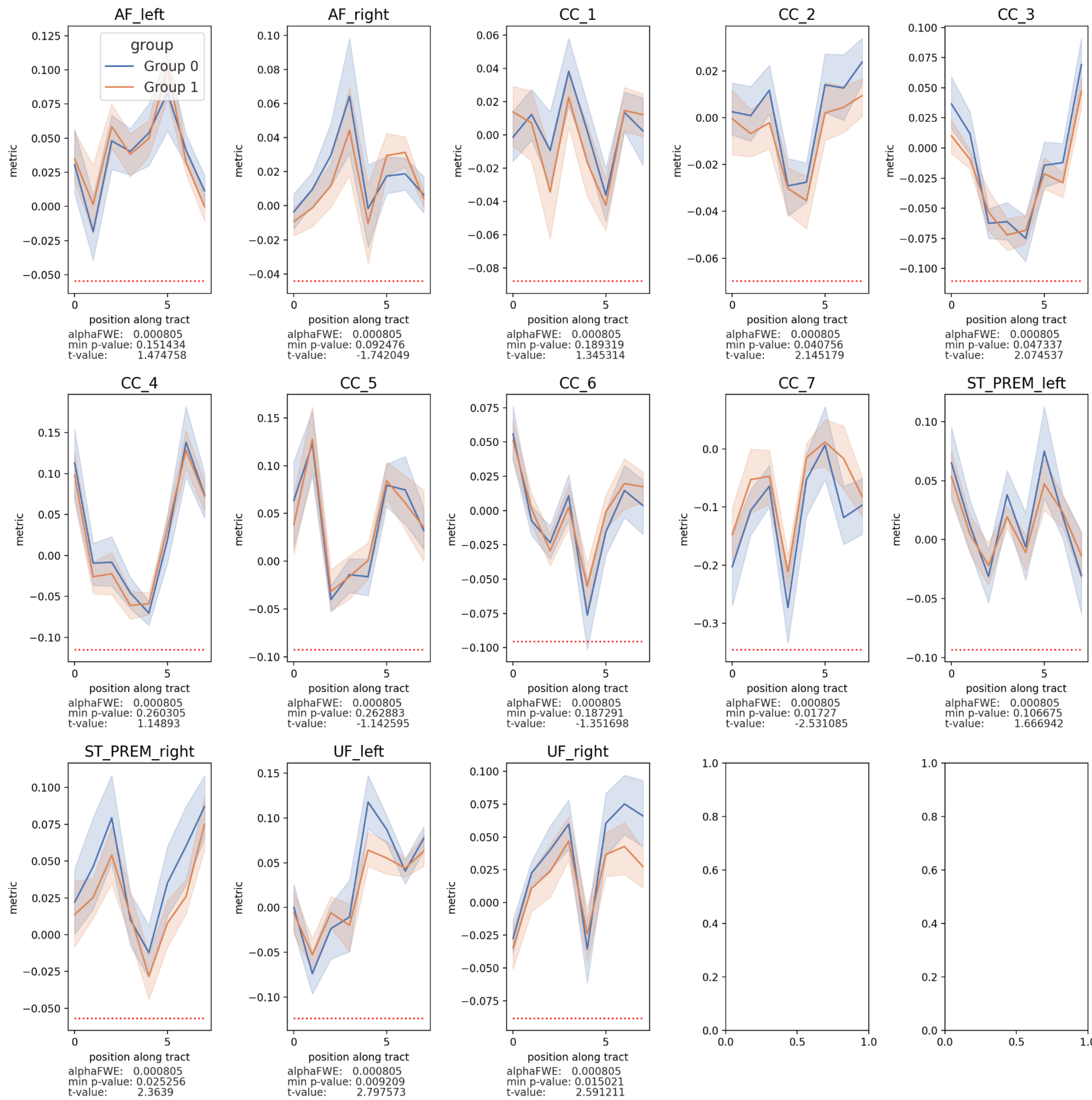
*The statistical FA results of VR+VE versus Control group*




*Note.* Overview of all the pre-post FA results, Group 0 (blue line) represents Controls and Group 1 (red line) VR+VE

**Figure B2**

*Statistical FA results of VR versus VE*




*Note.* Overview of all pre-post FA results, Group 0 (blue line) represents VE and Group 1 (red line) VR

**Abbreviations**

AF (arcuate fasciculus)

ST_PREM (striatal-premotor connections),

UF (uncinate fasciculus)

Parts of corpus callosum:

CC_1 (Rostrum)

CC_2 (Genu)

CC_3 (Rostral body (Premotor))

CC_4 (Anterior midbody (Primary Motor))

CC_5 (Posterior midbody (Primary Somatosensory))

CC_6 (Isthmus)

CC_7 (Splenium)