

Enhancing Cognitive Screening for Alzheimer's Disease: Integrating Virtual Reality and AI-Driven Speech

Andreas Sarantopoulos¹, Aristomenis Alevizopoulos², Julia Vasiliades³,
Georgios Alevizopoulos⁴, Antonia Stergiou⁵, Iakovos Kritikos⁶

¹Medical School, European University Cyprus, Nicosia, Cyprus

²School of Medicine, National and Kapodistrian University of Athens, Athens, Greece

³College of Engineering, Northeastern University, Boston, USA

⁴Psychiatric Clinic, Agioi Anargyroi General Oncological Hospital of Kifisia, Athens, Greece

⁵Hertfordshire University Partnership NHS Foundation Trust, Colchester, UK

⁶Global Business School for Health, University College, London, UK

Email: andreas.sarantopoulos01@gmail.com, alevizopoulosaristomenis@gmail.com, vasiliades.j@northeastern.edu, galevs@uoa.gr, antonia.stergiou@nhs.net, iakovos.kritikos.24@ucl.ac.uk

How to cite this paper: Sarantopoulos, A., Alevizopoulos, A., Vasiliades, J., Alevizopoulos, G., Stergiou, A. and Kritikos, I. (2025) Enhancing Cognitive Screening for Alzheimer's Disease: Integrating Virtual Reality and AI-Driven Speech. *Advances in Artificial Intelligence and Robotics Research*, 1, 23-40.

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Recent advances in virtual reality (VR) technology and AI-driven speech detection have opened new avenues for cognitive screening. This study integrates the Mini-Mental State Examination (MMSE) within a VR environment, coupled with AI-based speech detection, to enhance diagnostic accuracy and patient engagement in detecting early signs of Alzheimer's disease. Eighteen participants diagnosed with early symptoms of Alzheimer's disease were assessed using both traditional and VR-based MMSE methods. The VR system included immersive environments, motion tracking, and electrophysiological sensors to provide a comprehensive and interactive assessment. AI-based speech metrics, including speech latency, frequency of speech errors, and response completeness, were automatically extracted and compared with traditional MMSE scoring. Statistical analysis using intraclass correlation coefficients (ICCs) demonstrated excellent agreement between traditional and VR-based MMSE scores. The results indicate that the VR-based approach, combined with AI-driven speech detection, offers reliable and accurate cognitive assessments, supporting its potential for early detection of Alzheimer's disease. This innovative approach promises to transform cognitive screening by providing a more engaging and standardized testing environment.

Keywords

Virtual Reality (VR), AI-Driven Speech Detection, Cognitive Screening,

1. Introduction

VR technology provides an immersive and interactive platform that can simulate real-world environments, making cognitive assessments more engaging for patients [1]-[3]. By creating a controlled and consistent testing environment, VR can reduce external distractions and provide a standardized setting for administering the MMSE. This approach not only improves the reliability of the test results but also makes the experience more enjoyable for patients, potentially increasing their willingness to participate in regular cognitive screenings.

AI-driven speech detection further enhances this integration by providing real-time transcription and analysis of patient responses. Recent studies have demonstrated the effectiveness of AI [4]-[7] models in predicting the progression of cognitive impairment to Alzheimer's disease with high accuracy. By analyzing speech patterns, language structure, and other linguistic features, AI can offer valuable insights into a patient's cognitive state. This automated analysis can complement traditional cognitive tests, providing a more comprehensive assessment of cognitive function.

Combining VR and AI technologies in cognitive screening represents a significant advancement in the early detection of Alzheimer's disease [5] [8]. This innovative approach aims to improve diagnostic accuracy, enhance patient engagement, and provide a more holistic understanding of cognitive health. As research in this field continues to evolve, the integration of these technologies holds promise for transforming cognitive assessments and improving outcomes for individuals at risk of Alzheimer's disease.

Exposure therapy is a well-established psychological treatment used to manage anxiety disorders [9]-[11], phobias [12], and posttraumatic stress disorder (PTSD) [13] [14]. As well as Covid-19 [15]. Traditionally, it involves exposing patients to anxiety-provoking stimuli in real-world settings or through guided imagination, helping them gradually face and diminish their fears. However, recent technological advancements have introduced innovative approaches that enhance the effectiveness and engagement of exposure therapy. One such technology is VR, which creates immersive, three-dimensional environments that simulate real-life scenarios. This allows patients to confront their fears in a controlled and safe setting, offering significant advantages over traditional methods by providing a highly customizable and repeatable therapeutic experience.

Additionally, integrating motion tracking technology enables precise monitoring of a patient's physical responses during therapy sessions [16]. This real-time feedback allows therapists to adjust treatments to better suit the patient's needs and track their progress more accurately. Electrophysiological sensors, such as electroencephalography (EEG) and Heart Rate Variability (HRV) monitors [17] [18],

further enhance this approach by measuring neural and cardiovascular activity. These sensors provide objective data on the patient's arousal and emotional state, enabling a comprehensive assessment and facilitating more effective, personalized interventions [19]. The combination of VR, motion tracking [20] [21], and electrophysiological sensors offers a multidisciplinary approach that promises to improve exposure therapy outcomes and expand its applicability to various mental health disorders.

Prior research has explored the use of virtual reality (VR) and artificial intelligence (AI) individually in cognitive assessment and Alzheimer's detection. For example, VR has been employed to simulate realistic environments for evaluating memory, spatial navigation, and executive function, providing an immersive alternative to traditional cognitive tests. Similarly, AI—particularly natural language processing (NLP) and acoustic analysis—has shown promise in analyzing speech for early signs of cognitive decline, often using datasets of spoken responses from standard cognitive tests. However, most of these studies treat VR and AI as separate tools, and few have attempted to integrate them into a unified framework. Moreover, existing studies often rely on scripted tasks or retrospective datasets, limiting ecological validity and real-time interactivity. In contrast, our study uniquely combines a VR-administered MMSE with real-time AI-driven speech detection within an interactive, immersive clinical scenario. This integration allows not only for dynamic, context-sensitive cognitive testing but also for the automated extraction of nuanced speech metrics—such as latency, error frequency, and completeness—that are evaluated alongside traditional MMSE scores. By bridging these technologies, our work addresses limitations in engagement, scalability, and standardization found in previous approaches, offering a more holistic and modernized cognitive screening solution.

We hypothesize that integrating the Mini-Mental State Examination (MMSE) within a virtual reality (VR) environment, coupled with AI-driven speech detection, will enhance the diagnostic accuracy and patient engagement in cognitive screening for early signs of Alzheimer's disease. Specifically, we expect that the VR-based MMSE will show high agreement with traditional MMSE scoring methods, as measured by intraclass correlation coefficients (ICCs). Additionally, we anticipate that the AI-based speech metrics, including speech latency, frequency of speech errors, and response completeness, will provide valuable insights into cognitive function, further supporting the reliability and validity of this innovative approach. This hypothesis is grounded in the potential of VR to create immersive and controlled testing environments and the capability of AI to analyze speech patterns with high precision.

One of the key contributions of this study is the development of a fully integrated VR-based platform for administering the Mini-Mental State Examination (MMSE), which aims to enhance both the ecological validity and user engagement of cognitive assessments. Traditional paper-based MMSE tests, while widely used, may lack contextual immersion and interactive features that support real-world

applicability. By embedding the MMSE within a virtual environment, our system offers a more naturalistic and immersive experience, potentially reducing test anxiety and improving participant compliance. This innovation not only modernizes cognitive screening practices but also facilitates remote and repeatable testing, which is particularly valuable in monitoring early signs of Alzheimer’s disease in aging populations.

2. Materials and Methods

2.1. Hardware Configuration

The system was powered by a high-performance desktop setup, which included an NVIDIA GeForce GTX 1070 graphics card, an AMD Ryzen 7 2700X processor, and 16 GB of G.Skill TridentZ DDR4 RAM. The setup featured HDMI 1.3 video output along with multiple USB ports (three USB 3.0 and one USB 2.0) to support external peripherals. To facilitate an immersive VR experience, we utilized an Oculus Rift headset, complemented by a Plantronics Blackwire audio system for high-quality sound input and output [22].

2.2. Software Framework

The VR environment was built using a combination of industry-standard software tools. The system ran on Windows 11, configured with the necessary Oculus Rift drivers. Unity 3D served as the core development platform, integrating various hardware components such as sensors, controllers, and tracking systems. To design 3D models, animations, and textures, we employed Blender 3D for asset creation and Adobe Photoshop for image enhancements. The OVR Plugin enabled seamless integration of the Oculus Rift headset with Unity. For the speech-processing component, we incorporated the Speechmatics API, enabling real-time verbal interactions between the system and the user. Additionally, ChatGPT was utilized to analyze and process dialogue-based responses. Data extraction and analysis were conducted using Python within the JupyterLab environment, while a PowerShell script was implemented to synchronize and manage automation processes efficiently.

2.3. System Workflow and Automation Loop

The proposed system involves several interconnected components to provide real-time interaction. A general visualization of the workflow is depicted in **Figure 1**. The entire system operates in a loop, executing every 15 seconds. During each cycle, the files `dialogn.txt` and `answer.txt` are overwritten with new data, ensuring that the system remains up-to-date with the latest dialogue. A timer is implemented to trigger the loop every 15 seconds using a PowerShell script.

2.4. Speech Recognition and Natural Language Processing

Using Flow by Speechmatics, a speech recognition API, real-time audio dialogue is captured and transcribed using automatic speech recognition (ASR). The

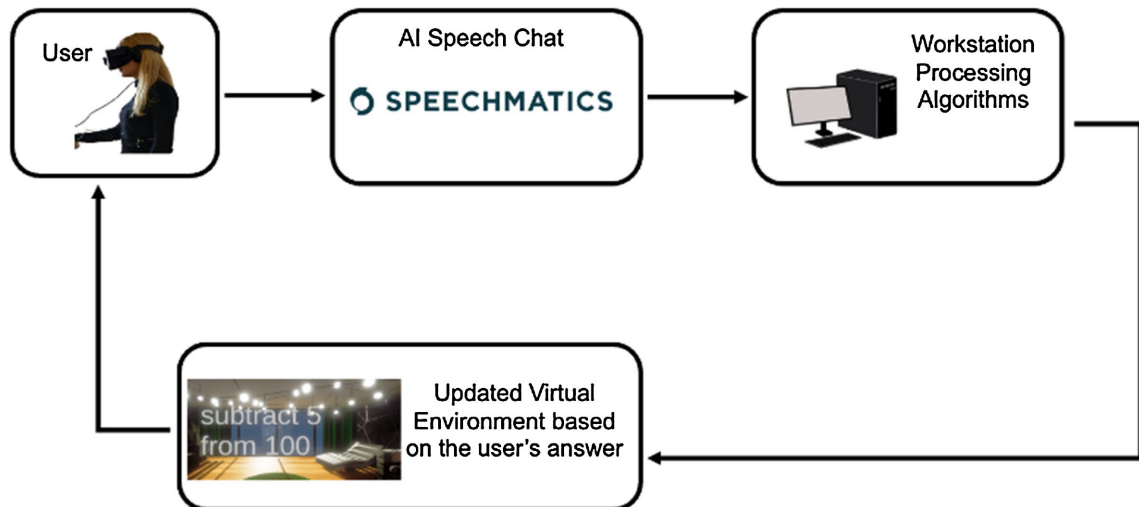


Figure 1. Illustrating the system architecture of the proposed VR-MMSE diagnostic platform. It includes the VR-based interaction layer, speech recognition and natural language processing modules, real-time scoring engine, and data storage for clinician access. The figure has been updated to reflect all key components and data flows.

transcribed data is saved to `dialogn.txt`. A Python script reads the latest content, parses user intent, and submits it to the ChatGPT API. The API's structured response is written to `answer.txt`, formatted and timestamped for downstream processing.

2.5. Virtual Reality Interaction Pipeline

Within Unity, a C++ script continuously reads the contents of `answer.txt` and parses the AI response to determine the required interaction. This may include displaying text, triggering animations, or validating object presence. Conditional logic ensures that the virtual world updates in real time, accurately reflecting user inputs and AI interpretations.

2.6. Simulation Design

Initial Simulation: This phase begins by asking the participant for their name to establish familiarity. It then assesses basic orientation (person, place, time) and recognition of five virtual objects. Following this, users complete simple mathematical calculations to assess problem-solving skills and short-term memory.

Object Verification Simulation: Participants are presented with questions about virtual objects within the environment and are asked to verbally identify them. The system matches their responses with the presence and identity of displayed 3D models (Figure 2).

Mathematical Calculation Simulation: Participants perform a verbal serial subtraction task (e.g., subtracting 5 from 100 repeatedly). Their responses are captured and logged (Figure 3). This simulation is used to evaluate attention, working memory, and numerical fluency. An overview of the processing function of this system is visually represented in Figure 4.

Participants: From a pool of 18 applicants, participants were selected and

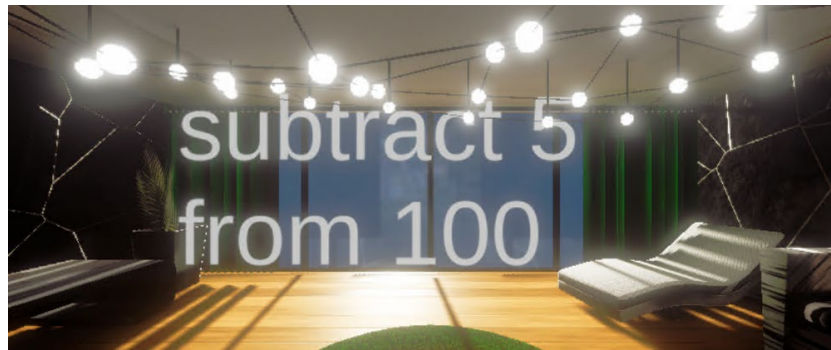


Figure 2. Simulation 1; the users are required to answer verbally about the object that they see.

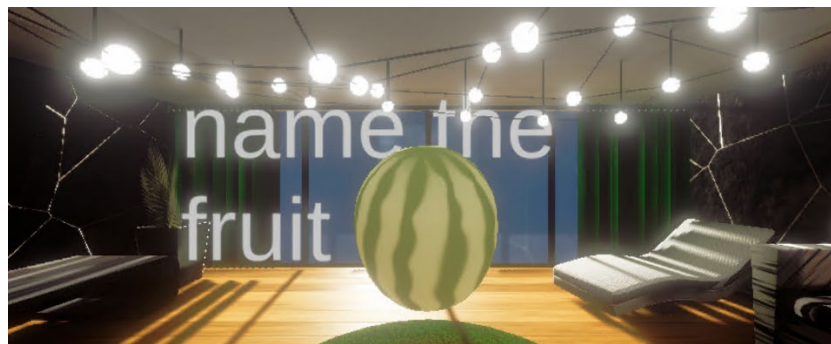


Figure 3. Simulation 2; the users are required to answer verbally about the result that occurs every time they subtract 5 from the result of the immediately previous subtraction, starting from the initial calculation 100 minus 5.

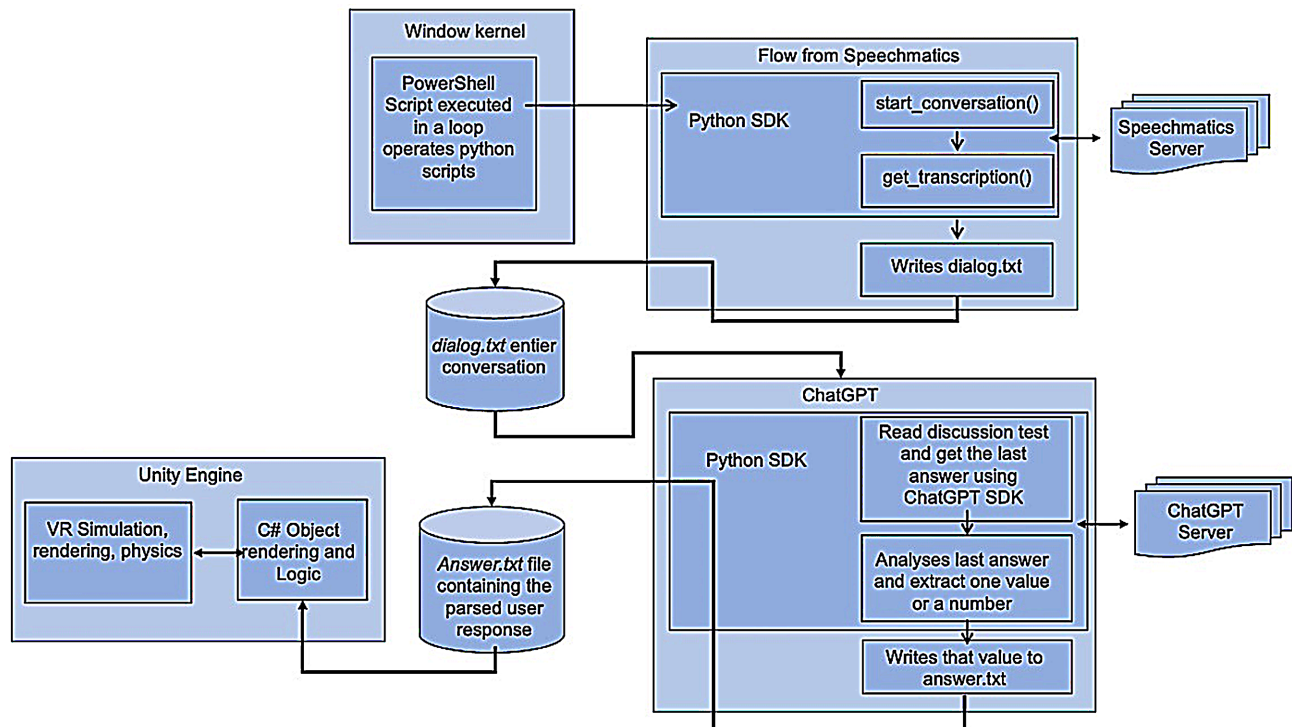


Figure 4. The flow of the system. The Windows kernel and PowerShell scripts execute the Python screens. The first one is a Python script that operates the Speechmatics SDK, and the second one is the ChatGPT SDK. Data are stored in the txt files dialog and answer, respectively, and those data are updated every 15 seconds.

diagnosed with early symptoms of Alzheimer’s disease based on the criteria of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5). Personal interviews were conducted by an experienced psychiatrist who was trained in the application of the Structured Clinical Interview for DSM Disorders (SCID) by the author who validated the SCID in Greek populations [23]. All 18 participants provided written informed consent prior to their involvement in the study. Inclusion Criteria: Individuals aged 65-85 with suspected mild cognitive impairment, fluent in the test language, with normal or corrected-to-normal vision suitable for VR headsets. Exclusion Criteria: Severe visual, hearing, or motor impairments that prevent VR usage; history of motion sickness or other conditions exacerbated by VR; acute psychiatric conditions.

Procedure: The experiment was conducted at the National and Kapodistrian University of Athens. The study was approved by the Ethics Committee of the Clinic with protocol number 35/2025, Agioi Anargyroi General Oncological Hospital of Kifisia, Psychiatric Clinic. Each of the 18 participants participated in a one-session simulation. The total duration of each session was 15 minutes, divided as follows: 5 minutes for a screening test and placement of the equipment, 2 minutes for the initial simulation, 4 minutes for the VR simulation, and 4 minutes for recovery from the VR simulation and a quick screening test to check for any complications during the VR simulation. In more detail, each participant was given specific instructions about the procedure of the simulation, the duration, and the differences between simulations, as well as instructions. The participants were informed about the study goals and were asked to sit comfortably on a chair to avoid hand movements, preventing potential movement artifacts during the simulation. Then, the appropriate equipment was placed on each participant [24]. At all times, a clinician had control of the simulation to intervene if something unexpected happened.

- **Briefing (5 minutes):** Participants are given a brief explanation of VR equipment and the study purpose. They are fitted with the VR headset and oriented to the virtual environment.
- **Virtual Environment Familiarization (2 minutes):** The participant navigates or looks around the virtual clinic to reduce novelty and anxiety.
- **MMSE Administration (10-15 minutes):**
 - **Orientation Questions:** Displayed as floating text prompts; participant answers verbally. AI records speech for analysis.
 - **Registration:** The VR environment cues participants to repeat named objects displayed as 3D models (e.g., an apple, pen, table).
 - **Attention:** A visual prompt instructs participants to subtract 7 from 100 repeatedly or spell “WORLD” backward.
 - **Recall:** 3D icons of the objects appear if needed; participant must recall their names verbally. Naming virtual objects (e.g., a pencil, a watch) shown in 3D.
- **Post-Test Debrief (5 minutes):** Assess for any discomfort or side effects from

VR. Gather qualitative feedback on the experience.

Assessment of Confounding Variables: Potential confounding factors, including participants' prior experience with virtual reality (VR), VR-induced anxiety, and motion sickness, were carefully assessed and monitored. Participants completed a pre-study questionnaire documenting any previous exposure to VR technologies. Immediately following the VR assessment, anxiety levels were evaluated through brief self-report scales, and structured post-session interviews were conducted to identify and record any motion sickness or discomfort. Recognizing that anxiety, motion sickness, or unfamiliarity with VR could potentially influence cognitive performance by increasing cognitive load or reducing engagement, these factors were explicitly considered in our analysis and interpretation of the cognitive screening results.

Ethical Considerations: Participant confidentiality and data security were strictly maintained throughout the study. All participants provided written informed consent prior to participation. Speech recordings and related participant data were encrypted and securely stored within protected digital databases, accessible solely to authorized research personnel. Data handling procedures adhered to the General Data Protection Regulation (GDPR) guidelines. Additionally, ethical approval was granted by the Institutional Review Board (IRB), ensuring compliance with ethical standards as established by the Declaration of Helsinki.

3. Results

The data analysis for the 18 participants involved in the VR-based cognitive screening simulations was conducted using a combination of AI-based speech metrics, MMSE scoring comparisons, and statistical analysis to evaluate the effectiveness and accuracy of the VR-based approach (**Figure 5**). The data collected encompassed AI-extracted speech features, traditional and AI-generated MMSE scores, and demographic covariates, ensuring a comprehensive assessment of cognitive performance [25] [26].

The AI-based speech metrics included speech latency, frequency of speech errors, and response completeness, all automatically extracted from participants' verbal responses during the VR simulations. Speech latency, a continuous variable, was measured as the time taken by participants to respond to prompts. The frequency of speech errors was recorded as count data, representing the number of incorrect or incomplete responses per participant. Response completeness was evaluated on an ordinal scale, categorizing responses as incomplete, partially complete, or fully complete. These speech-derived parameters were analyzed to explore their relationship with MMSE performance and cognitive impairment severity [27] [28].

For MMSE scoring, traditional clinician-administered MMSE scores were compared with AI-generated MMSE results derived from natural language processing (NLP) algorithms. The traditional MMSE was manually scored by experienced clinicians, whereas the AI-based approach utilized predefined linguistic and

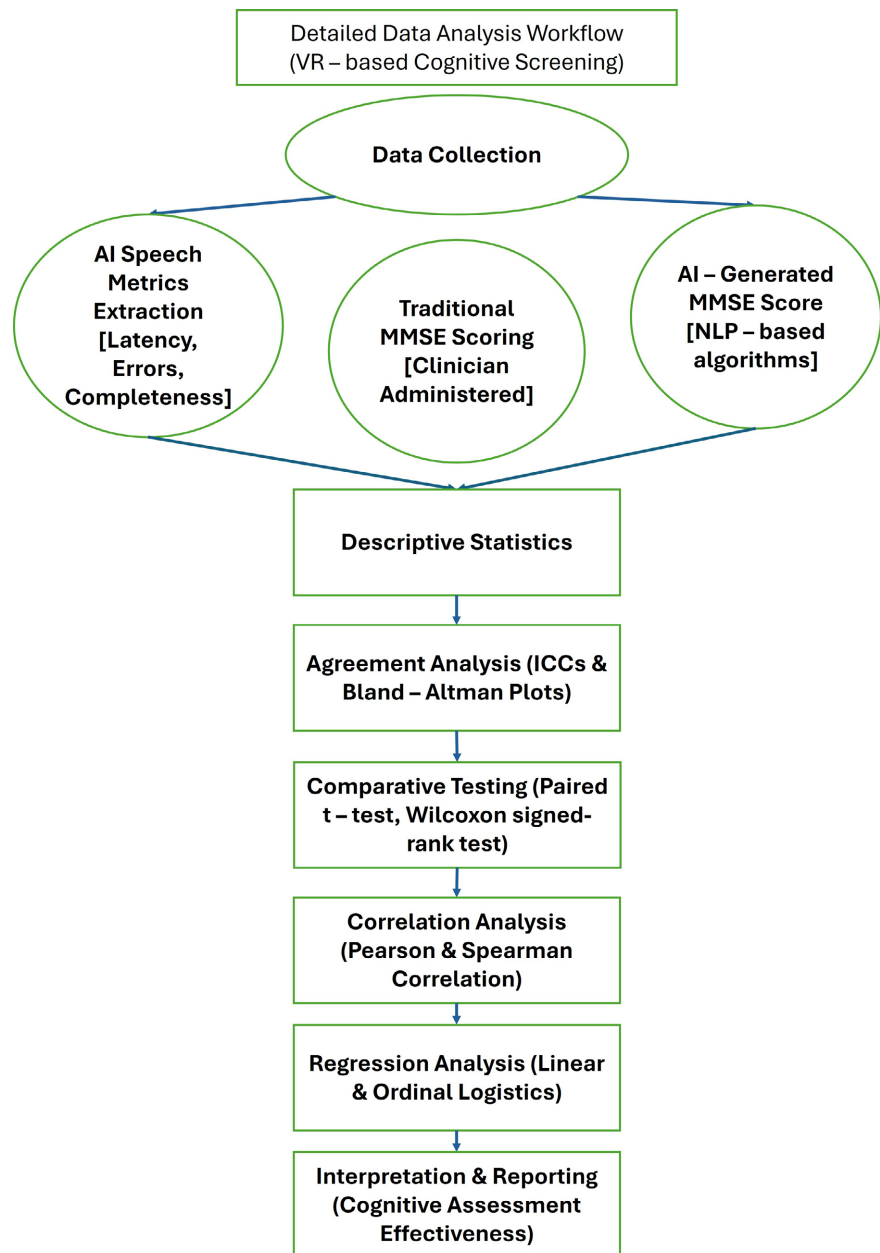


Figure 5. Data analyses workflow.

cognitive criteria to assign scores. The goal of this comparison was to determine the reliability and consistency of AI-driven scoring in relation to the established clinical standard [29].

To assess the level of agreement between traditional clinician-administered MMSE scores and AI-generated MMSE scores, Intraclass Correlation Coefficients (ICCs) were calculated using a two-way mixed-effects model with absolute agreement. The resulting ICC was 0.83 (95% CI: 0.65 - 0.93), indicating excellent agreement between the two scoring methods. According to established guidelines, ICC values above 0.75 reflect strong consistency, supporting the reliability of the AI-based scoring system. Additionally, a Bland-Altman plot was generated to visually

examine the level of agreement and detect any systematic bias between the two methods. The mean difference between traditional and AI scores was 0.44 points, with 95% limits of agreement ranging from -2.3 to $+3.2$. While most differences fell within the limits, a slight trend toward underestimation of scores by the AI system at higher MMSE values was observed, suggesting the need for minor calibration in future AI models.

To evaluate whether there were significant differences between the traditional MMSE scores and AI-generated scores, data distribution was first assessed using the Shapiro-Wilk test. Results indicated that the score differences were normally distributed ($p = 0.21$). Accordingly, a paired t-test was conducted, revealing no significant difference between the two scoring approaches ($t(17) = 1.29, p = 0.21$). The mean traditional MMSE score was 23.8 ($SD = 2.8$), while the mean AI-based score was 23.4 ($SD = 3.1$). The small mean difference of 0.44 points was not statistically or clinically significant.

A two-way mixed-effects model with absolute agreement was chosen for the ICC calculation because it assumes fixed raters (clinician and AI) and accounts for both rater and subject variability, which is appropriate for assessing agreement between a specific set of measurement methods [30]; an ICC value of 0.83 indicates excellent agreement based on their guideline, which defines values above 0.75 as good and above 0.90 as excellent.

To evaluate the predictive value of AI-extracted speech features on cognitive performance, a multiple linear regression analysis was conducted with traditional MMSE scores as the dependent variable. Independent variables included speech latency (in seconds), number of speech errors, response completeness (percentage), and participant age, sex, and education level. The regression model was statistically significant, $F(6, 11) = 5.21, p < 0.01$, indicating that the combination of speech and demographic variables significantly predicted traditional MMSE scores. The model explained approximately 72% of the variance in MMSE scores ($R^2 = 0.72$). Among the predictors:

- Speech latency had a significant negative association with MMSE scores ($\beta = -0.58, p = 0.02$), suggesting that increased response time was associated with lower cognitive performance.
- Response completeness positively predicted MMSE scores ($\beta = 0.41, p = 0.03$), indicating that more complete verbal responses were linked to higher cognitive functioning.
- Speech error frequency was negatively related to MMSE scores, though this trend did not reach statistical significance ($\beta = -0.27, p = 0.09$).
- Age, sex, and education level were included as control variables; however, their contributions were not statistically significant in the presence of speech-based features.

These results support the hypothesis that real-time speech features extracted through AI algorithms can reliably predict cognitive performance, aligning closely with clinical assessments. The findings highlight the potential of integrating speech

detection into VR-based cognitive screening tools for early detection of Alzheimer's-related impairment.

4. Discussion

The findings from this study underscore the potential of integrating VR and AI technologies for cognitive screening [26] [31], particularly in the early detection of Alzheimer's disease. The high level of agreement between traditional and VR-based MMSE scores indicates that the VR environment does not compromise the validity of cognitive assessments. The AI-driven speech metrics provide valuable insights into participants' verbal responses, enhancing the overall evaluation process. This innovative approach offers a more engaging and standardized method for cognitive screening, which could lead to improved patient compliance and more accurate assessments.

This study notably contributes to the field by introducing a sophisticated, yet practical application of immersive virtual reality combined with advanced AI-driven speech detection. The integration of VR and AI technologies addresses critical challenges in traditional cognitive assessments, such as environmental variability and subjective scoring biases. By employing automated speech metrics, including speech latency and error frequency, this system achieves a higher degree of precision and consistency. Moreover, the interactive nature of VR significantly enhances patient engagement, making cognitive screening less burdensome and potentially encouraging more frequent monitoring. These factors combined highlight the approach's promise for broader clinical application, particularly in routine screening and early intervention programs for Alzheimer's disease.

This study distinguishes itself from prior work by combining VR immersion with real-time AI-driven speech detection to administer and evaluate the MMSE. While previous approaches have used VR or AI independently, our integration allows not only automated scoring but also contextual understanding of verbal responses within an immersive environment.

Additionally, VR platforms enable the simulation of engaging, naturalistic cognitive challenges while maintaining a controlled experimental environment, thereby increasing ecological validity in cognitive assessments. This immersive approach allows for the replication of real-world scenarios, providing a more accurate evaluation of an individual's cognitive abilities in everyday contexts [32]. Secondly, AI-driven speech detection has emerged as a powerful tool in detecting cognitive decline. Recent studies have demonstrated that automatic speech detection can serve as objective assessment tools for individuals with cognitive impairments, offering a non-invasive and efficient method for early detection. By analyzing speech patterns, AI can identify subtle linguistic markers associated with Alzheimer's progression, facilitating timely intervention strategies [33].

While the findings are encouraging, certain limitations must be considered:

- **Small Sample Limitations:** The relatively small sample size ($n = 18$) and the cross-sectional nature of this pilot study limit the generalizability of our find-

ings. While the high ICC value (0.83) between VR-MMSE and traditional MMSE is promising, future work will require larger and more diverse samples with longitudinal follow-up to assess test-retest reliability and sensitivity to cognitive decline progression over time. This small sample size of 18 participants limits the generalizability of the findings. Future research should involve larger and more diverse populations to validate these results further. The primary objective of this research is to evaluate the technological usability and feasibility of integrating VR and AI-driven speech detection in cognitive screening, rather than to establish the system's diagnostic accuracy at this stage. This proof-of-concept study aims to assess the practicality, user experience, and potential benefits of immersive cognitive assessments, setting the groundwork for future large-scale validation studies. The small sample size (N=18) reflects the study's exploratory nature, focusing on initial usability insights rather than broad generalizability. Additionally, given the complexity of VR-based cognitive testing, our study prioritizes in-depth qualitative observations and system refinement over statistical significance in diagnostic outcomes. Resource constraints, including access to specialized AI models, VR equipment, and eligible participants, further influenced the sample size. Future research will expand participant diversity, incorporate a power analysis for sample size determination, and validate the system across different demographics and clinical stages to enhance its diagnostic reliability and generalizability.

- **VR limitations:** reliance on VR technology could create accessibility barriers for individuals with severe visual, auditory, or motor impairments [34]. Additionally, VR usage may induce motion sickness or other discomforts, potentially affecting participant performance; these effects require careful monitoring and mitigation strategies.
- **Model Training Limitations:** In this study, we did not train the speech detection model or the LLM ourselves. Instead, we integrated pre-trained models via their respective APIs through libraries like Chatbot, leveraging their existing capabilities. This approach ensures robustness and reliability while allowing us to focus on effective implementation and analysis rather than model training.
- **Ethical Biases Limitations:** AI models can inherently exhibit biases due to the nature of their training data, algorithmic design, or implementation context. In this study, we acknowledge the potential biases that may arise, particularly in speech detection and language processing, where model accuracy can vary based on demographic factors, accents, and contextual nuances. To mitigate these concerns, we ensured that the AI models used were industry-standard, pre-trained systems that have undergone extensive testing. Additionally, we implemented diverse test cases to assess performance across different scenarios and carefully analyzed outputs for any significant discrepancies. While this study is exploratory in nature, future implementations involving direct user interactions would require a more comprehensive ethical framework to address

bias-related concerns proactively.

- While this study primarily compares VR-MMSE with the traditional MMSE, we acknowledge the importance of evaluating its performance against other modern diagnostic methods. AI-driven diagnostic tools, such as speech detection systems and machine learning models trained on large datasets, have shown promise in cognitive assessment. These approaches leverage NLP and acoustic analysis to detect early signs of cognitive decline. Compared to these methods, VR-MMSE offers an interactive, immersive environment that engages patients in real-time tasks, potentially improving test sensitivity and patient experience. However, further research is needed to directly compare VR-MMSE with AI-based diagnostic models to assess their respective strengths, limitations, and applicability across different patient populations.
- Lastly, this study's cross-sectional design precludes analysis of longitudinal changes and the cognitive assessments' long-term efficacy.

The AI components leveraged pretrained APIs for semantic parsing and Speechmatics for speech-to-text transcription. While these services offer high accuracy and scalability, their use raises questions about adaptability and bias. The models were not fine-tuned on clinical or dementia-specific datasets, and thus their responses may reflect limitations when applied to nuanced cognitive impairments. Future iterations of this system will involve custom model training and benchmarking against clinically validated NLP frameworks to ensure precision and fairness in diagnostic inference.

Potential confounding variables, including VR-induced anxiety, participant familiarity with VR technology, and motion sickness, were systematically monitored throughout the study. These factors could influence cognitive performance by increasing cognitive load or reducing overall comfort during the assessments. While efforts were made to mitigate their impact, such as providing pre-assessment VR acclimatization and ensuring breaks when necessary, their effects cannot be entirely ruled out. Participants with prior VR experience may have adapted more quickly, whereas those unfamiliar with the technology might have experienced heightened anxiety or disorientation. Future research should further investigate these confounding variables by incorporating larger sample sizes, control groups, and alternative VR calibration methods to enhance the reliability and validity of VR-based cognitive screening tools.

While the current study benchmarks VR-MMSE against the traditional MMSE, a comparison with other AI-based cognitive assessment tools would offer further context. Preliminary results suggest that VR-MMSE may provide advantages in user engagement and ecological validity, yet systematic comparisons with tools like Cognetivity, Altoida, or AI-driven speech classifiers will be necessary to confirm its diagnostic value and scalability.

Future Research Future research should focus on expanding the sample size and including participants from various demographic backgrounds to enhance the generalizability of the findings. Longitudinal studies are needed to assess the long-

term efficacy and reliability of VR-based cognitive screening. Additionally, exploring the integration of other AI-driven tools and technologies, such as machine learning algorithms for predictive modeling, could further enhance the accuracy and utility of cognitive assessments. Investigating the potential for VR-based interventions to improve cognitive function and slow the progression of Alzheimer's disease is another promising area for future research.

In future clinical applications, the platform is envisioned as a semi-autonomous screening tool that could be used with minimal clinician supervision, especially in primary care or community settings, while allowing for clinician oversight when needed. Although the current study focused on Alzheimer's detection, the system's modular design and use of generalized AI-driven speech metrics make it adaptable for other neurocognitive disorders such as Parkinson's disease or mild cognitive impairment. While formal usability testing was not conducted, participant feedback and in-session monitoring indicated that users were able to navigate and understand the VR environment intuitively, with minimal instruction or technical issues.

5. Conclusion

In conclusion, this proof-of-concept study demonstrates that a VR-based administration of the MMSE, combined with AI-driven speech detection, is both feasible and comparable in accuracy to traditional methods. The immersive and standardized VR environment, coupled with automated speech metrics, enhances cognitive screening by improving patient engagement and minimizing scoring variability. We recommend future studies with larger, more diverse cohorts to validate these findings, explore longitudinal outcomes, and assess integration into clinical workflows. The system's adaptability and scalability highlight its potential for broader use in early Alzheimer's detection and cognitive monitoring programs.

Ethics Approval

This study was conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments. The study was approved by the Ethics Committee of the Clinic with protocol number 35/2025, Agioi Anargyroi General Oncological Hospital of Kifisia, Psychiatric Clinic. Each of the 18 participants participated in a one-session simulation.

Consent for Participation and Publication

Informed consent for participation and publication of anonymized data was obtained from all individual participants included in the study.

Availability of Data and Materials

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

IK (Iakovos Kritikos) conceptualized the study, designed the VR-based MMSE environment, collected and analyzed data, and drafted the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Baus, O. and Bouchard, S. (2014) Moving from Virtual Reality Exposure-Based Therapy to Augmented Reality Exposure-Based Therapy: A Review. *Frontiers in Human Neuroscience*, **8**, Article ID: 112. <https://doi.org/10.3389/fnhum.2014.00112>
- [2] Society of Clinical Psychology (n.d.) What Is Exposure Therapy? American Psychological Association: Division 12. <https://www.apa.org/ptsd-guideline/patients-and-families/exposure-therapy>
- [3] Caravas, P., Kritikos, J., Alevizopoulos, G. and Koutsouris, D. (2021) Participant Modeling: The Use of a Guided Master in the Modern World of Virtual Reality Exposure Therapy Targeting Fear of Heights. In: Perego, P., TaheriNejad, N. and Caon, M., Eds., *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Springer International Publishing, 161-174. https://doi.org/10.1007/978-3-030-76066-3_13
- [4] McCarten, J.R., Rottunda, S.J. and Kuskowski, M.A. (2004) Change in the Mini-Mental State Exam in Alzheimer's Disease over 2 Years: The Experience of a Dementia Clinic. *Journal of Alzheimer's Disease*, **6**, 11-15. <https://doi.org/10.3233/jad-2004-6102>
- [5] Lee, J., Lee, H., Yoo, H.B., Choi, J., Jung, H., Yoon, E.J., et al. (2019) Efficacy of Cistostazol Administration in Alzheimer's Disease Patients with White Matter Lesions: A Positron-Emission Tomography Study. *Neurotherapeutics*, **16**, 394-403. <https://doi.org/10.1007/s13311-018-00708-x>
- [6] König, A., et al. (2015) Automatic Speech Analysis for the Assessment of Patients with Predementia and Alzheimer's Disease. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring*, **1**, 112-124.
- [7] Gautam, P. and Singh, M. (2025) Alzheimer's Disease Classification Using the Fusion of Improved 3D-VGG-16 and Machine Learning Classifiers. *International Journal of Biomedical Engineering and Technology*, **47**, 1-27. <https://doi.org/10.1504/ijbet.2025.143776>
- [8] Clay, F., Howett, D., FitzGerald, J., Fletcher, P., Chan, D. and Price, A. (2020) Use of Immersive Virtual Reality in the Assessment and Treatment of Alzheimer's Disease: A Systematic Review. *Journal of Alzheimer's Disease*, **75**, 23-43. <https://doi.org/10.3233/jad-191218>
- [9] Kritikos, I., Tzannetos, G., Zoitaki, C., Pouloupoulou, S. and Koutsouris, D. (2019) Anxiety Detection from Electrodermal Activity Sensor with Movement & Interaction during Virtual Reality Simulation. 2019 *9th International IEEE/EMBS Conference on Neural Engineering (NER)*, San Francisco, 20-23 March 2019, 571-576. <https://doi.org/10.1109/ner.2019.8717170>
- [10] Yeh, S.-C., Li, Y.-Y., Zhou, C., Chiu, P.-H. and Chen, J.-W. (2017) Effects of Virtual Reality and Augmented Reality on Induced Anxiety. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **26**, 1345-1352.

- [11] Otte, C. (2011) Cognitive Behavioral Therapy in Anxiety Disorders: Current State of the Evidence. *Dialogues in Clinical Neuroscience*, **13**, 413-421. <https://doi.org/10.31887/dcns.2011.13.4/cotte>
- [12] Kritikos, J., Pouloupoulou, S., Zoitaki, C., Douloudi, M. and Koutsouris, D. (2019) Full Body Immersive Virtual Reality System with Motion Recognition Camera Targeting the Treatment of Spider Phobia. In: Cipresso, P., Serino, S. and Villani, D., Eds., *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Springer International Publishing, 216-230. https://doi.org/10.1007/978-3-030-25872-6_18
- [13] Foa, E.B., Chrestman, K.R. and Gilboa-Schechtman, E. (2009) Prolonged Exposure Therapy for Adolescents with PTSD: Emotional Processing of Traumatic Experiences: Therapist Guide. Oxford University Press. <https://doi.org/10.1093/med:psych/9780195331745.001.0001>
- [14] Beidel, D.C., Frueh, B.C., Neer, S.M., Bowers, C.A., Trachik, B., Uhde, T.W., *et al.* (2019) Trauma Management Therapy with Virtual-Reality Augmented Exposure Therapy for Combat-Related PTSD: A Randomized Controlled Trial. *Journal of Anxiety Disorders*, **61**, 64-74. <https://doi.org/10.1016/j.janxdis.2017.08.005>
- [15] Alevizopoulos, A., Kritikos, J. and Alevizopoulos, G. (2021) Intelligent Machines and Mental Health in the Era of Covid-19. *Psychiatriki*, **32**, 99-102. <https://doi.org/10.22365/jpsych.2021.015>
- [16] Kritikos, I., Zoitaki, C., Tzannetos, G., Mehmeti, A., Douloudi, M., Nikolaou, G., *et al.* (2020) Comparison between Full Body Motion Recognition Camera Interaction and Hand Controllers Interaction Used in Virtual Reality Exposure Therapy for Acrophobia. *Sensors*, **20**, Article 1244. <https://doi.org/10.3390/s20051244>
- [17] Thakare, A.E., Mehrotra, R. and Singh, A. (2017) Effect of Music Tempo on Exercise Performance and Heart Rate among Young Adults. *International Journal of Physiology, Pathophysiology and Pharmacology*, **9**, 35-39.
- [18] Hussain, I. and Park, S.J. (2020) HealthSOS: Real-Time Health Monitoring System for Stroke Prognostics. *IEEE Access*, **8**, 213574-213586. <https://doi.org/10.1109/access.2020.3040437>
- [19] Kritikos, I., Mehmeti, A., Nikolaou, G. and Koutsouris, D. (2019) Fully Portable Low-Cost Motion Capture System with Real-Time Feedback for Rehabilitation Treatment. 2019 International Conference on Virtual Rehabilitation (ICVR), Tel Aviv, 21-24 July 2019, 1-8. <https://doi.org/10.1109/icvr46560.2019.8994561>
- [20] Konstantina, A. and Iakovos, K. (2024) Virtual Reality, Electrophysiology & Motion Tracking Technologies in Mental Illnesses. *Annals of Psychiatry and Treatment*, **8**, 23-26. <https://doi.org/10.17352/apt.000060>
- [21] Scheggi, S., Meli, L., Pacchierotti, C. and Prattichizzo, D. (2015) Touch the Virtual Reality: Using the Leap Motion Controller for Hand Tracking and Wearable Tactile Devices for Immersive Haptic Rendering. *ACM SIGGRAPH 2015 Posters*, Los Angeles, 9-13 August 2015, Article No. 31. <https://doi.org/10.1145/2787626.2792651>
- [22] Kritikos, J., Alevizopoulos, G. and Koutsouris, D. (2021) Personalized Virtual Reality Human-Computer Interaction for Psychiatric and Neurological Illnesses: A Dynamically Adaptive Virtual Reality Environment That Changes According to Real-Time Feedback from Electrophysiological Signal Responses. *Frontiers in Human Neuroscience*, **15**, Article ID: 596980. <https://doi.org/10.3389/fnhum.2021.596980>
- [23] Kritikos, I., Sarantopoulos, A., Roumeliotis, A., Vasiliades, J. and Matsinas, I. (2025) Unlocking the Potential of Artificial Intelligence in Pharma Research and Development: Insights from Investor and Researcher Perspectives. *Health Economics and*

Management Review, **6**, 1-16. <https://doi.org/10.61093/hem.2025.2-01>

- [24] Kritikos, I., Caravas, P., Tzannetos, G., Douloudi, M. and Koutsouris, D. (2019) Emotional Stimulation during Motor Exercise: An Integration to the Holistic Rehabilitation Framework. 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, 23-27 July 2019, 4604-4610. <https://doi.org/10.1109/embc.2019.8857548>
- [25] König, A., Köhler, S., Tröger, J., Düzel, E., Glanz, W., Butryn, M., et al. (2024) Automated Remote Speech-Based Testing of Individuals with Cognitive Decline: Bayesian Agreement of Transcription Accuracy. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, **16**, e70011. <https://doi.org/10.1002/dad2.70011>
- [26] Zhao, X., Hu, R., Wen, H., Xu, G., Pang, T., He, X., et al. (2022) A Voice Recognition-Based Digital Cognitive Screener for Dementia Detection in the Community: Development and Validation Study. *Frontiers in Psychiatry*, **13**, Article ID: 899729. <https://doi.org/10.3389/fpsy.2022.899729>
- [27] Jang, S., Choi, S., Son, S.J., Oh, J., Ha, J., Kim, W.J., et al. (2023) Virtual Reality-Based Monitoring Test for MCI: A Multicenter Feasibility Study. *Frontiers in Psychiatry*, **13**, Article ID: 1057513. <https://doi.org/10.3389/fpsy.2022.1057513>
- [28] Min, J., Kim, D., Jang, H., Kim, H., Kim, S., Lee, S., et al. (2025) The Validity of a Smartphone-Based Application for Assessing Cognitive Function in the Elderly. *Diagnostics*, **15**, Article 92. <https://doi.org/10.3390/diagnostics15010092>
- [29] Qiao, Y., et al. (2020) Computer-Assisted Speech Analysis in Mild Cognitive Impairment and Alzheimer's Disease: A Pilot Study from Shanghai, China. *Journal of Alzheimer's Disease*, **75**, 211-221. <https://doi.org/10.3233/JAD-191056>
- [30] Koo, T.K. and Li, M.Y. (2016) A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, **15**, 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- [31] Yan, M., Yin, H., Meng, Q., Wang, S., Ding, Y., Li, G., et al. (2021) A Virtual Supermarket Program for the Screening of Mild Cognitive Impairment in Older Adults: Diagnostic Accuracy Study. *JMIR Serious Games*, **9**, e30919. <https://doi.org/10.2196/30919>
- [32] Rizzo, A.S. (2019) Clinical Virtual Reality in Mental Health and Rehabilitation: A Brief Review of the Future! *Infrared Technology and Applications XLV*, **11002**, 150-158.
- [33] Fraser, K.C., Meltzer, J.A. and Rudzicz, F. (2015) Linguistic Features Identify Alzheimer's Disease in Narrative Speech. *Journal of Alzheimer's Disease*, **49**, 407-422. <https://doi.org/10.3233/JAD-150520>
- [34] Kritikos, J., Makrypidis, A., Alevizopoulos, A., Alevizopoulos, G. and Koutsouris, D. (2023) Can Brain-Computer Interfaces Replace Virtual Reality Controllers? A Machine Learning Movement Prediction Model during Virtual Reality Simulation Using EEG Recordings. *Virtual Worlds*, **2**, 182-202. <https://doi.org/10.3390/virtualworlds2020011>

Abbreviations

AI	Artificial Intelligence
MMSE	Mini-Mental State Examination
VR	Virtual Reality
ICC	Intraclass Correlation Coefficient
LLM	Large Language Model
API	Application Programming Interface