

Short Communication

Usability of the Virtual Supermarket Test for Older Adults with and without Cognitive Impairment

Stelios Zygouris^{a,*}, Sofia Segkouli^a, Andreas Triantafyllidis^a, Dimitrios Giakoumis^a, Magdalini Tsolaki^{b,c}, Konstantinos Votis^a and Dimitrios Tzovaras^a

^a*Centre for Research and Technology Hellas/Information Technologies Institute, Thessaloniki, Greece*

^b*School of Medicine, Aristotle University of Thessaloniki, Greece*

^c*Greek Association of Alzheimer's Disease and Related Disorders, Thessaloniki, Greece*

Received 29 November 2021

Accepted 19 April 2022

Pre-press 2 May 2022

Published 13 May 2022

Abstract. This study conducted a preliminary usability assessment of the Virtual Supermarket Test (VST), a serious game-based self-administered cognitive screening test for mild cognitive impairment (MCI). Twenty-four healthy older adults with subjective cognitive decline and 33 patients with MCI self-administered the VST and then completed the System Usability Scale (SUS). The average SUS score was 83.11 (SD = 14.6). The SUS score was unaffected by age, education, touch device familiarity, and diagnosis of MCI. SUS score correlated with VST performance ($r = -0.496$, $p = 0.000$). Results of this study indicate good usability of the VST.

Keywords: Age-related memory disorders, cognitive assessment screening instrument, computer games, dementia, mild cognitive impairment, neurocognitive tests, screening, self-evaluation

INTRODUCTION

Virtual reality (VR) and serious game applications have exhibited strong potential for differentiating between healthy and cognitively impaired older adults [1, 2] and their use as digital biomarkers has been proposed [3]. The Virtual Supermarket Test (VST) is a VR application with a low degree of immersion that has been used effectively to detect mild cognitive impairment across different populations and regions [4–11]. Empirical data from seven previous VST studies [4–10], such as lack of drop-outs and overall positive comments about the usability of the VST by study participants, indicate that the VST is usable and acceptable by its target

audience. The successful validation of the latest fully self-administered VST version [4, 5] hinted at high usability as all participants were able to successfully operate the application despite not receiving any kind of help or guidance from an examiner. At the same time, a formal quantitative assessment of the VST's usability had not been conducted.

This study is part of a series of studies to assess a new computerized testing paradigm for older adults with subjective cognitive decline (SCD), based on a self-administered serious game-based test, the VST. Previous studies have evaluated the diagnostic effectiveness [5] and the neurophysiological correlates [4] of the VST. The present study conducts a preliminary assessment of the VST's usability, considering the importance of usability in technology adoption [12]. It focuses on assessing the usability in a sample of older adults with mild cognitive impairment

*Correspondence to: Stelios Zygouris, Centre for Research and Technology Hellas, 6th km Charilaou – Thessaloniki, 57001 Thessaloniki, Greece. E-mail: szygouris@iti.gr.

(MCI) and SCD as they are more likely to use an app to self-screen their cognition [5] compared to healthy older adults with no cognitive concerns. The main aim of this study is to examine the usability of the VST serious game application and how it is affected by age, education, diagnosis, in-game performance and familiarity with touch devices such as tablets and smartphones. It is hypothesized that the usability of the VST application will be high overall and it will be affected by performance in the VST application, age, education, diagnosis and familiarity with touch devices. This study was conducted in the auspices of the European Union Horizon 2020 project ALAMEDA where the VST will be used for cognitive assessment of participants. The quantitative assessment of VST's usability through the use of an established scale was deemed necessary prior to its inclusion as a self-administered cognitive assessment instrument in the multi-center study of ALAMEDA.

MATERIALS AND METHODS

Virtual reality serious game-based cognitive screening test

The VST application has been described in detail in previous studies [4–10]. It features an interactive training session followed by a test trial that is repeated three times for a total of three test trials per VST administration. It features complex metrics including assessment of mistakes and speed, trajectory analysis and assessment of practice effects. It has been found to outperform gold standard pencil and paper cognitive screening tests, such as the Montreal Cognitive Assessment, in differentiating between MCI patients and healthy controls [5, 7–10]. It is aimed at activating a multitude of cognitive functions with the emphasis placed on executive function and navigation as demonstrated through correlations of VST performance with performance on established pencil and paper tests assessing these domains [7, 9, 10]. The need for simultaneous activation of different cognitive functions makes the VST challenging enough to correspond to the ability level of the target population (healthy older adults with SCD and MCI patients) while reducing ceiling effects. Figure 1 provides a screenshot of the VST digital environment.

System Usability Scale

The System Usability Scale (SUS) comprises ten questions scored in a 5-point Likert scale. Its score

ranges from 0–100 and the average SUS score (50th percentile) according to SUS scores normal distribution is 68 [13, 14]. SUS scores between 71 and 84 are considered good and SUS scores of 85 and above are considered excellent [14, 15]. The SUS has been translated and validated for use in Greek adults.

Participants

This study was conducted in parallel to the study for the validation of the latest, fully self-administered version of the VST [5]. The study sample comprised healthy older adults with SCD and older adults with amnesic MCI. Exclusion criteria were: diagnosis of dementia or another major neurological or psychiatric disorder, illiteracy, health issues such as motor and vision difficulties that could interfere with the use of the VST, treatment with cholinesterase inhibitors or other drugs that can affect cognitive performance, alcoholism or drug abuse. Recruitment was conducted between April 2018 and April 2019 at the day centers of the Greek Association of Alzheimer's Disease and Related Disorders (GAARDR) in Thessaloniki, Greece. MCI patients were paired randomly with older adults of similar age and education level, without MCI symptoms but with subjective memory complaints, to ensure a balanced sample. All participants provided informed consent. Diagnosis was conferred by a neurologist after a full neurological, neuropsychological and laboratory assessment. MCI diagnosis was based on Petersen criteria [16–18]. Participants with no objective impairment in neuropsychological testing were allocated to the SCD group.

Study protocol

Familiarity with touch devices was assessed with a binary question where participants were instructed to answer “Yes” if they could operate the touch device and do simple tasks (e.g., find a specific contact in the contacts app) unassisted. Participants were asked to complete the VST exercise on their own. The examiner prepared the tablet and launched the VST app but did not assist participants. After the administration of the VST, participants were asked to complete the Greek version of the SUS, which is a widely-used valid tool for usability assessment [19]. The overall procedure lasted approximately 45 minutes. This study protocol for assessing the usability of the



Fig. 1. The Virtual Supermarket Test (VST) digital environment.

VST in parallel with its diagnostic utility, was specifically selected so the participants were not exposed to the VST app more than once before completing the SUS questionnaire. The study was conducted in accordance with the Declaration of Helsinki ethical guidelines and was approved by the Scientific and Ethical Committee of the GAARDR.

Statistical analysis

Statistical analysis was conducted with the SPSS (version 24.0) statistical software package [20]. Similar to a previous study [4], user performance on the VST was calculated using its core metrics: a) total mistakes conducted during the 3 VST test trials and b) average time needed to complete the 3 VST test trials. Comparisons between user groups were performed using Mann-Whitney U test. Results of these tests with a p value of less than 0.05 were considered as statistically significant. Pearson's correlation analysis was used to determine the relationship between SUS score, VST variables, age, and education.

RESULTS

Demographic characteristics of participants

The sample included 24 healthy older adults with SCD and 33 MCI patients. Fifteen participants were male and 42 were female. Mean age was 67.21 years ($SD=5.095$) ranging from 54 to 75 years. Subjects had a mean of 13.44 ($SD=2.739$) years of formal education ranging from 9 to 21 years of formal education. No significant differences were observed between the SCD and MCI groups in age and education while significant differences were observed in Montreal Cognitive Assessment scores as expected. Similarly, no significant differences were discerned between participants with touch device familiarity and participants with no touch device familiarity in age and education. Demographic characteristics of participants are presented in Table 1.

SUS results

The average SUS score was 83.11 ($SD=14.6$).

Table 1
Demographics characteristics of MCI patients and SCD

	MCI (<i>n</i> = 33) <i>M</i> (<i>SD</i>)	SCD (<i>n</i> = 24) <i>M</i> (<i>SD</i>)	<i>p</i>
Age, y	68.39 (4.53)	65.58 (5.47)	0.403
Education, y	13.06 (2.88)	13.96 (2.49)	0.123
Gender (M/F)	10/23	5/19	0.423
MoCA	25.09 (2.97)	27.87 (1.74)	0.002*

M, mean; SD, standard deviation. Age, education, and Montreal Cognitive Assessment (MoCA) scores were tested by Mann-Whitney U test, gender by Chi-square test (Fisher's exact). * $p < 0.05$.

Table 2
Correlations between SUS score, age, education, and VST variables

		SUS
Age, y	Pearson Correlation	-0.070
	Sig. (2-tailed)	0.607
	N	57
Education, y	Pearson Correlation	-0.010
	Sig. (2-tailed)	0.940
	N	57
Average Duration for 3 trials	Pearson Correlation	-0.496**
	Sig. (2-tailed)	0.000
	N	57
Sum of Errors for 3 trials	Pearson Correlation	-0.250
	Sig. (2-tailed)	0.060
	N	57

**Correlation is significant at the 0.01 level (2-tailed).

Group differences in SUS score

No significant differences were found between the SCI and MCI groups in SUS scores. Similarly, no significant differences in SUS scores were found between participants with touch device familiarity and participants without touch device familiarity.

Correlations between SUS score, VST variables, age, and education

Significant correlation ($r = -0.496$, $p = 0.000$) was found between SUS score and average time needed to complete the 3 VST test trials. No significant correlations were found between SUS score, age, education and total mistakes conducted during the 3 VST test trials. Correlations are presented in Table 2.

DISCUSSION

Computerized cognitive screening test designers have shifted their focus towards brevity and ease of administration during recent years [21]. Many tests feature gamification elements and tasks in virtual environments are being integrated in computerized

tests [22]. Newer tests are often designed to be self-administered at home or in the clinic [23]. The lack of an examiner can render these tests vulnerable to the effect of poor usability on test performance. Thus, test designers have been conducting usability assessments on computerized tests that are either self- or examiner-administered [24, 25]. This follows the paradigm of conducting usability assessment of serious games and other applications that are designed to be used by older adults, often without any assistance [26, 27]. As the VST is a cognitive screening test that is modeled on a serious game, has received considerable research interest and has been found to perform very well across different populations, we conducted a preliminary assessment of its usability using the SUS. This assessment was conducted in order to provide quantitative support to our indirect qualitative assessment of its usability through our interaction with participants in all VST studies [4–10].

Virtual tasks have been used extensively to assess cognitive deficits [28–31]. The usability of virtual shopping tasks has received significant research interest however most of these studies were conducted on more immersive VR apps, utilizing either VR goggles or specialized equipment such as projection screens and sensors, and a large part of the usability assessment was specific to the immersive nature of these apps [28, 30, 31]. Relevant literature suggests that older adults may perform better on non-immersive platforms [31]; however, the usability of on non-immersive virtual tasks such as the VST has not been studied extensively. Assessing the usability of such tasks is crucial as, contrary to immersive VR, they do not necessitate expensive specialized equipment and thus they can easily be used by a large number of older adults.

The present study provides preliminary data supporting the usability of the latest, fully self-administered version of the VST for cognitive assessment. Our primary finding is that VST displays good usability [15] among older adults with or without cognitive impairment, as assessed by SUS, with its usability score (83.11) being at the 90th percentile of SUS scores normal distribution [13, 14]. VST performance (average time to compete the three VST test trials) affected usability. Usability was unaffected by age, education, touch device familiarity, and diagnosis of MCI. The correlation between usability, as expressed through SUS score, and VST performance, as expressed through average time to complete the exercise, raises the question of the causal relationship between these variables. It is unclear whether

participants who took longer to complete the VST test trials provided a lower usability score as they were frustrated by the long time it took them to complete the exercise or if there are underlying usability issues that affected their performance thus resulting in a longer time to complete the VST test trials.

Strengths of this study include a robust diagnostic procedure for assigning participants into groups [5] and the use of a well validated serious game application (i.e., the VST). Furthermore, the study sample is representative of the people who reach out to brain health services and present to memory clinics for assessment [5].

Limitations of this study include the inclusion of significantly more female participants. Furthermore, the study sample was recruited from a day centre for people with cognitive disorders thus all participants, even those who did not present with cognitive impairment, had concerns about their cognitive functioning. Therefore, they may differ from a sample of healthy older adults recruited from the wider community in the sense that they are expected to have a particular interest in brain health and thus be more engaged during the administration of the VST. Future directions comprise studies with more participants including healthy older adults with and without SCD for the quantitative and qualitative assessment of both usability and acceptability through standardized questionnaires and focus groups.

CONCLUSION

This study demonstrates that VST has good usability, and its usability remains unaffected by age, education, touch device familiarity, and diagnosis of MCI. Study results provide a strong incentive for further, more detailed usability assessment of the VST.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to all the study participants who gave generously of their time.

FUNDING

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No GA101017558.

CONFLICT OF INTEREST

The authors have no conflict of interest to report.

REFERENCES

- [1] Mohammadi A, Kargar M, Hesami E (2018) Using virtual reality to distinguish subjects with multiple- but not single-domain amnesic mild cognitive impairment from normal elderly subjects. *Psychogeriatrics* **18**, 132-142.
- [2] Atkins AS, Khan A, Ulshen D, Vaughan A, Balentin D, Dickerson H, Liharska LE, Plassman B, Welsh-Bohmer K, Keefe RSE (2018) Assessment of instrumental activities of daily living in older adults with subjective cognitive decline using the Virtual Reality Functional Capacity Assessment Tool (VRFCAT). *J Prev Alzheimers Dis* **5**, 216-234.
- [3] Gold M, Amatniek J, Carrillo MC, Cedarbaum JM, Hendrix JA, Miller BB, Robillard JM, Rice JJ, Soares H, Tome MB, Tarnanas I, Vargas G, Bain LJ, Czaja SJ (2018) Digital technologies as biomarkers, clinical outcomes assessment, and recruitment tools in Alzheimer's disease clinical trials. *Alzheimers Dement (N Y)* **4**, 234-242.
- [4] Iliadou P, Paliokas I, Zygouris S, Lazarou E, Votis K, Tzouvaras D, Tsolaki M (2021) A comparison of traditional and serious game-based digital markers of cognition in older adults with mild cognitive impairment and healthy controls. *J Alzheimers Dis* **79**, 1747-1759.
- [5] Zygouris S, Iliadou P, Lazarou E, Giakoumis D, Votis K, Alexiadis A, Triantafyllidis A, Segkouli S, Tzouvaras D, Tsiatsos T, Papagianopoulos S, Tsolaki M (2020) Detection of mild cognitive impairment in an at-risk group of older adults: can a novel self-administered serious game-based screening test improve diagnostic accuracy? *J Alzheimers Dis* **78**, 405-412.
- [6] Paliokas I, Kalamaras E, Votis K, Doumpoulakis S, Lakka E, Kotsani M, Freminet A, Benetos A, Ellul I, Polycarpou M, Zygouris S, Megalookonomou V, Tzouvaras D (2020) Using a virtual reality serious game to assess the performance of older adults with frailty. *Adv Exp Med Biol* **1196**, 127-139.
- [7] Boz HE, Limoncu H, Zygouris S, Tsolaki M, Giakoumis D, Votis K, Tzouvaras D, Öztürk V, Yener G (2020) A short assessment tool for small vessel disease with cognitive impairment: The Virtual Supermarket (VSM). *J Alzheimers Dis Rep* **16**, e047040.
- [8] Eraslan Boz H, Limoncu H, Zygouris S, Tsolaki M, Giakoumis D, Votis K, Tzouvaras D, Öztürk V, Yener GG (2020) A new tool to assess amnesic mild cognitive impairment in Turkish older adults: virtual supermarket (VSM). *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn* **27**, 639-653.
- [9] Zygouris S, Ntovas K, Giakoumis D, Votis K, Doumpoulakis S, Segkouli S, Karagiannidis C, Tzouvaras D, Tsolaki M (2017) A preliminary study on the feasibility of using a virtual reality cognitive training application for remote detection of mild cognitive impairment. *J Alzheimers Dis* **56**, 619-627.
- [10] Zygouris S, Giakoumis D, Votis K, Doumpoulakis S, Konstantinos N, Segkouli S, Karagiannidis C, Tzouvaras D, Tsolaki M (2015) Can a virtual reality cognitive training application fulfill a dual role? Using the Virtual Super Market cognitive training application as a screening tool for mild cognitive impairment. *J Alzheimers Dis* **44**, 1333-1347.
- [11] Sexton C, Solis M, Aharon-Peretz J, Alexopoulos P, Apostolova LG, Bayen E, Birkenhager B, Cappa S,

- Constantinidou F, Fortea J, Gerritsen DL, Hassanin HI, Ibanez A, Ioannidis P, Karageorgiou E, Korczyn A, Leroi I, Lichtwarck B, Logroscino G, Lynch C, Mecocci P, Molinuevo JL, Papatriantafyllou J, Papegeorgiou S, Politis A, Raman R, Ritchie K, Sanchez-Juan P, Sano M, Scarmeas N, Spuru L, Stathi A, Tsolaki M, Yener G, Zaganas I, Zygouris S, Carrillo M (2022) Alzheimer's disease research progress in the Mediterranean region: The Alzheimer's Association International Conference Satellite Symposium. *Alzheimers Dement*. doi: 10.1002/alz.12588.
- [12] Yen PY, Bakken S (2012) Review of health information technology usability study methodologies. *J Am Med Inform Assoc* **19**, 413-422.
- [13] Brooke J (1996) SUS: A 'quick and dirty' usability scale. In *Usability Evaluation in Industry*, 1st ed. CRC Press.
- [14] Brooke J (2013) SUS: a retrospective. *J Usability Stud* **8**, 29-40.
- [15] Bangor A, Kortum P, Miller J (2009) Determining what individual SUS scores mean: adding an adjective rating scale. *J Usability Stud* **4**, 114-123.
- [16] Petersen RC (2004) Mild cognitive impairment as a diagnostic entity. *J Intern Med* **256**, 183-194.
- [17] Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, Ritchie K, Rossor M, Thal L, Winblad B (2001) Current concepts in mild cognitive impairment. *Arch Neurol* **58**, 1985-1992.
- [18] Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, Smith GE, Jack CR, Jr. (2009) Mild cognitive impairment: ten years later. *Arch Neurol* **66**, 1447-1455.
- [19] Katsanos C, Tselios N, Xenos M (2012) Perceived usability evaluation of learning management systems: a first step towards standardization of the system usability scale in Greek. *Proceedings of the 16th Panhellenic Conference on Informatics*, pp. 302-307.
- [20] IBM Corp (2016) IBM SPSS Statistics for Windows (V 24.0)
- [21] Zygouris S, Tsolaki M (2015) Computerized cognitive testing for older adults: a review. *Am J Alzheimers Dis Other Dement* **30**, 13-28.
- [22] Zygouris S, Tsolaki M (2015) New technologies and neuropsychological evaluation of older adults: issues and challenges. In *Handbook of Research on Innovations in the Diagnosis and Treatment of Dementia*, Bamidis PD, Tarnanas I, Hadjileontiadis L, Tsolaki M, eds. IGI Global, USA, pp. 1-17.
- [23] Tsoy E, Zygouris S, Possin KL (2021) Current state of self-administered brief computerized cognitive assessments for detection of cognitive disorders in older adults: a systematic review. *J Prev Alzheimers Dis* **8**, 267-276.
- [24] Fredrickson J, Maruff P, Woodward M, Moore L, Fredrickson A, Sach J, Darby D (2010) Evaluation of the usability of a brief computerized cognitive screening test in older people for epidemiological studies. *Neuroepidemiology* **34**, 65-75.
- [25] Robillard JM, Lai J-A, Wu JM, Feng TL, Hayden S (2018) Patient perspectives of the experience of a computerized cognitive assessment in a clinical setting. *Alzheimers Dement (N Y)* **4**, 297-303.
- [26] Palumbo V, Paterno F (2021) Micogito: a serious gamebook based on daily life scenarios to cognitively stimulate older adults. in *Proceedings of the Conference on Information Technology for Social Good*, Association for Computing Machinery, Roma, Italy, pp. 163-168.
- [27] Hassandra M, Galanis E, Hatzigeorgiadis A, Goudas M, Mouzakidis C, Karathanasi EM, Petridou N, Tsolaki M, Zikas P, Evangelou G, Papagiannakis G, Bellis G, Kokkotis C, Panagiotopoulos SR, Giakas G, Theodorakis Y (2021) A virtual reality app for physical and cognitive training of older people with mild cognitive impairment: mixed methods feasibility study. *JMIR Serious Games* **9**, e24170.
- [28] Arlati S, Di Santo SG, Franchini F, Mondellini M, Filiputti B, Luchi M, Ratto F, Ferrigno G, Sacco M, Greci L (2021) Acceptance and usability of immersive virtual reality in older adults with objective and subjective cognitive decline. *J Alzheimers Dis* **80**, 1025-1038.
- [29] Tsai CF, Chen CC, Wu EH, Chung CR, Huang CY, Tsai PY, Yeh SC (2021) A machine-learning-based assessment method for early-stage neurocognitive impairment by an immersive virtual supermarket. *IEEE Trans Neural Syst Rehabil Eng* **29**, 2124-2132.
- [30] Mrakic-Sposta S, Di Santo SG, Franchini F, Arlati S, Zangiacomi A, Greci L, Moretti S, Jesuthasan N, Marzorati M, Rizzo G, Sacco M, Vezzoli A (2018) Effects of combined physical and cognitive virtual reality-based training on cognitive impairment and oxidative stress in MCI patients: a pilot study. *Front Aging Neurosci* **10**, 282.
- [31] Plechatá A, Sahula V, Fayette D, Fajnerová I (2019) Age-related differences with immersive and non-immersive virtual reality in memory assessment. *Front Psychol* **10**, 1330.