Going to a Virtual Supermarket: Comparison of Different Techniques for Interacting in a Serious Game for the Assessment of the Cognitive Status

Alice E. Martis, Chiara Bassano, Fabio Solari, and Manuela Chessa(✉)

Department of Informatics, Bioengineering, Robotics and System Engineering - DIBRIS, University of Genoa, Genoa, Italy
{fabio.solari,manuela.chessa}@unige.it

Abstract. An increasing number of people suffers from cognitive impairments, also related to aging. Several approaches are used to evaluate the mental status of people affected by cognitive diseases, and there is a growing interest toward approaches that allow a quantitative and personalized evaluation of such impairments. Such approaches comprise serious games and VR-based cognitive assessment systems. Nevertheless, few works attempt to understand how people interact in such systems, and which human-computer interaction modalities are to be preferred when targeting impaired people. The aim of this work is to quantitative and qualitative compare two solutions to play in a virtual supermarket (a PC and a tablet-based solution). The obtained results can be used as a starting point to design VR-based serious games to be used instead of questionnaire-based approaches, thus improving both clinical evaluation performances and patients' motivation.

Keywords: Human-computer interaction · Serious games
Neuropsychological assessment · Virtual reality
VR-based cognitive assessment

1 Introduction

The constant increase of the average age of people has determined a higher incidence of neurodegenerative pathologies related with aging. One of this is dementia, a disease which compromises the abilities of attention, concentration, memory, reasoning, calculation, logic, and orientation, with repercussions not only on the individual but also on his/her family. Until some decades ago, dementia was considered a condition without return but, with the progress of research on central nervous system and in the neuropsychological field, there has been a greater understanding of this pathology, a better differentiation of its forms and, consequently, a therapeutic approach. For this reason, it is essential to use effective screening tools to provide early diagnosis [1,2].

At present, the most widely used and validated test is the Mini-Mental State Examination (MMSE) that allows to evaluate cognitive functions and their modifications over time [3]. It consists of eleven questions referring to seven different

S. Battiato et al. (Eds.): ICIAP 2017 International Workshops, LNCS 10590, pp. 281–289, 2017.
https://doi.org/10.1007/978-3-319-70742-6_26

cognitive areas (orientation, registration, attention and calculation, recall, language and praxis), it is easily administered by the physician and it takes about 10 min to be executed. However, it also has some limitations: it shows poor sensitivity to detect mild/early dementia and individual performance can be influenced by the presence of hearing and visual impairment or by poor education [4,5].

In the last few years, efforts have been made to develop additional tools for the assessment of cognitive functions through the use of Virtual Reality (VR) in order to find a valid alternative to pencil-and-paper tests. In fact, VR allows to reproduce complex situations of daily living where psychopathological reactions and cognitive functions of patients can be more reliably evaluated than laboratory situations. In addition, through VR, it is possible to reproduce environmental and social situations that can stimulate the subject in a similar way to the corresponding context [6], allowing to modulate the intensity and duration of the experience according to the program's needs and the needs of the subject. For these reasons VR-based cognitive assessment represents a tool with a great ecological validity [7,8], meant as the ability to actually measure what it is supposed to [9].

To date, there are many serious games that use both non-immersive and immersive VR to simulate daily activities and that are able to evaluate the cognitive status of elderly people. Regarding non-immersive VR, we must mention the works of Zucchella et al. [10] and Vourvopulos et al. [11]. In the first one, authors have created a first-person game for touch screen platforms. It is set in a virtual apartment in which the subject must perform 5 tasks, each of them evaluating a different cognitive function. The second research presents a game set in a virtual city in which the user can move through the use of a joystick to reach places of interest (pharmacy, supermarket, bank and post office) and perform the required tasks. Regarding immersive VR, we can mention the study of Parsons et al. [12], which consists of a virtual supermarket where the user wears a head-mounted display and interacts with the objects of the scene directly with his/her own hand thanks to markers placed in a wearable glove. Another important work is the one conducted by Tarnanas et al. [8,13], in which a fire evacuation is simulated. The subject can move using a treadmill and can interact with objects using the Leap Motion[1], a hand tracker device.

Following the state of the art tendencies, the work presented in this paper aims to propose an alternative to pencil-and-paper traditional tests with a greater ecological validity. We have created a simple virtual supermarket and we have developed it for two different platforms, tablet and pc. The main difference between the two devices is the interaction modality: in one case users have to tap and scroll on the screen; whereas in the other case they use a standard mouse. Our primary aim is to assess whether the different type of human-computer interaction can affect the performance of subjects, using both subjective and objective evaluation parameters. Therefore, we want to determine which device can better perform an analysis of the cognitive status of the elderly.

[1] https://www.leapmotion.com.

2 Materials and Methods

2.1 Platforms

For our study, we have developed two different versions of the same game: the first one runs on a 13,3-inch MacBook Pro with a 2,5 GHz Intel Core i5 processor and Intel HD Graphics 4000 1536 MB graphics processor; while the second one runs on a 10,1-inch Samsung Tab 4 with a 1,2 GHz Quad-Core processor and Android 5.0.2 operating system.

The game was created with *Unity 3D*[2], a cross-platform game engine for the development of 2D and 3D videogames for desktop, web and mobile devices. Most of the items in the game were made with *SketchUp*[3], a 3D modeling software commonly used to design very realistic 3D objects or to download them from a free library (3D Warehouse) and export them in a Unity-compatible format.

2.2 Participants

The game was tested on 32 volunteer healthy subjects with an age range of 21–78 years (mean 39.8 ± 18.8). We divided the participants into two groups: *Under60* group (22 people in an actual age range of 21–49 and average age 28.1 ± 6.4) and *Over60* group (10 people in an age range of 60–78 and average age 65.7 ± 6.6). Neither people in the *Under60* group nor in the *Over60* group were affected by any degenerative or cognitive disease.

All the participants were asked to try the game with both devices and to report which one they preferred to use and why. The first device they had to try has been randomly chosen. Therefore, half of the participants started with the tablet and the other half with the computer.

2.3 Experimental Procedure

The serious game consists of a virtual supermarket with two shelves and a fruit counter as shown in Fig. 1. The subject has to perform two tasks. The first one consists in buying all the items (in our experiments 10) on a shopping list, the items are randomly chosen when the game starts, but a control ensures that they are all different. The second task consists in paying the exact total amount.

Each experiment consists of:

- a Training with the demo scene;
- a First Trial with a device randomly chosen between pc and tablet;
- a Second Trial with the other device.

The Second Trial usually began immediately after the end of the first one. After each trial the subject was asked to write down all the bought items. The question was first asked at the end of the First Trial, so the subject, initially, did not know that he/she had to memorize the list.

[2] https://unity3d.com.
[3] http://www.sketchup.com.

Fig. 1. A view of the Shopping Task scene (left) and of the Payment Task scene (right)

2.4 Game Design

We designed the game in order to keep the interface as simple as possible. The user cannot freely navigate the scene but he/she has to turn the camera toward the selected shelf using the arrows at the bottom of the screen. Only when he/she is in front of the desired shelf, he/she can interact with the items on it, for example by clicking or tapping on one of them a pop-up window with the product name is displayed and the player is asked if he/she wants to add that article to the cart or not.

The game is composed by three scenes, a menu, a demo scene and the main scene. The menu allows the subject to choose whether to try the demo scene or to go directly to the game. This way the demo scene can be skipped if the user is performing the Second Trial. The demo scene is a simplified version of the main scene and its aim is to explain the game to the player and to allow him/her to learn how to interact with objects and buttons. We have used it in a trial session before each experiment.

Shopping Task. Once the demo is completed, the main scene is loaded. The subject has to type the ID assigned to him/her in an input field and to press the *Start!* button. At this point the game begins and a shopping list appears in the upper right corner of the screen (Fig. 1, left). The player has to buy only the things displayed on the shopping list, if he/she does not recognize an object he/she can click on it and its name appears, then he/she can choose whether or not to put it in the cart.

The subject can interact with the scene through several buttons:

- *Cart* (top left corner): by pressing it a dropdown list with all the selected items appears and the user can choose to delete a product from the cart simply clicking on its name;
- *Shopping List* (top right corner): which shows or hides the shopping list;
- *Payment* (top center): which enables the payment screen and disables the shopping scene.

When the player presses on *Payment* button, a pop-up appears asking him/her if he/she wants to proceed with the payment, as once accepted, he/she will not be allowed to go back to the supermarket view.

Payment Task. In this task the subject is asked to pay the precise sum displayed, clicking on Euro bills and coins (Fig. 1, right). He/She can reset the selected amount, in case of error, or confirm the payment by pressing respectively the *Cancel* and the *Confirm* buttons.

2.5 Evaluated Parameters

The performance of the subjects has been evaluated through several parameters. For both tasks we measured the execution time and computed a score. The ShoppingScore (*SS*) takes into account the number of the bought items that are (*CI*) and are not (*WI*) in the list and also the number of items deleted from the cart (*DI*). *SS* can vary between 0 and 10 and it is obtained thanks to the following equation:

$$SS = CI - \alpha * WI - \beta * DI \tag{1}$$

Different weights (α and β) are associated with different errors: if player selects an incorrect item it is a mistake, but if he/she realizes it and corrects it he/she will be less penalized. The PaymentScore (*PS*) is set to 0 if the payed amount is incorrect and 10 if it is correct, otherwise it is computed, taking into account the number of times the player has reset the total, which is considered an error (E), as:

$$PS = 10 - 0.5 * E \tag{2}$$

In both cases, a low score can be related to an impairment in solving the task, and a high score means that the user completed it easily. Another evaluated parameter is the number of remembered items, at the end of the game the subject was asked to write down all the items he/she remember he/she had bought.

Finally, at the end of the test, participants were asked to express a preference, between interaction with pc and with tablet, and to explain the reason of their choice.

3 Results

In order to evaluate which device could be more suitable for our purpose, we made three different types of analysis: the first one was the comparison of parameters measured with the tablet version and the pc version of the game; the second one consisted in comparing the performances obtained in the First and the Second Trial, regardless the device type; in the last analysis we compared the performances of the *Under60* group versus the *Over60* group, taking into account both the type of device and the difference between First and Second Trial. We also evaluated the statistical significance of the compared data groups by performing a *two-sample t-test* for all of the three cases, in order to assess if differences in results were actually linked to different conditions or not.

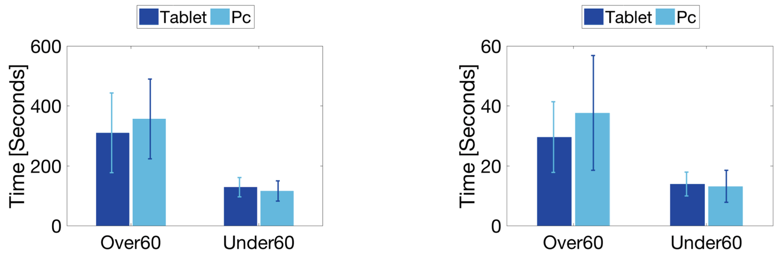
Table 1 summarizes the average of the parameters considered in the first and the second analysis. Comparing tablet and pc performances it can be noticed that there are not significant differences neither for the execution times nor for the scores of both tasks. Regarding the comparison between First and Second Trial, times and scores do not significantly differ one from another. Execution times do not decrease in the Second Trial as we would be expected, but as the standard deviations are very high this may be due to an internal variability of the data. Furthermore, the average number of remembered items does not considerably vary during the Second Trial despite subjects were aware of having to memorize the shopping list. Many subjects complained that first time they had to write down the list, their bad results were due to the fact that they did not know they had to remember items in the list; while the second time they got confused and risked to write items from the first list. So results are not directly related to concentration but to subjective short term memory capability.

Table 1. Mean and standard deviation of evaluated parameters obtained dividing participants in two different conditions: results obtained with the tablet and the pc; results obtained in the First and Second Trial.

	Tablet	Pc	First Trial	Second Trial
Shop.Score/10	9.6 ± 0.9	9.6 ± 0.6	9.5 ± 1.0	9.8 ± 0.5
Paym.Score/10	9.9 ± 0.2	9.9 ± 0.1	9.9 ± 0.2	9.9 ± 0.1
Shop.Time [sec]	185.5 ± 114.7	191.4 ± 136.9	196.8 ± 132.1	180.0 ± 119.8
Paym.Time [sec]	18.8 ± 10.3	20.8 ± 16.1	20.8 ± 13.2	18.8 ± 13.8
Rememb.Items/10	-	-	8.6 ± 1.4	8.4 ± 1.5

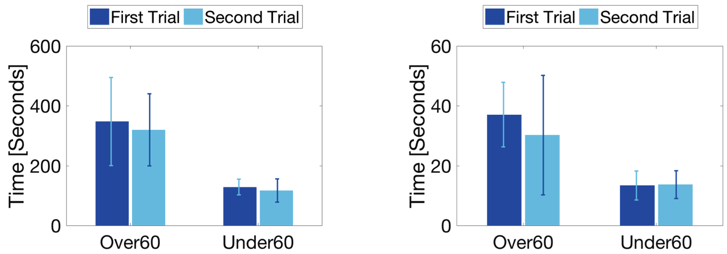
Significant differences were found between the performances of *Over60* and *Under60* groups regarding the mean execution time of both tasks (see Fig. 2): as expected, *Under60* group times are much lower ($p < 0.05$ in both cases). Considering each group separately, we can notice that in the younger group results with the two devices in both tasks are comparable, while in the elderly group performances may look slightly better (but not significantly) in the case of the tablet. Also variability is lower in the case of *Under60* group, maybe because more subjects were considered, or because they are more confident with technologies, on average. However, considering all the data, average scores of both tasks (see equations (1) and (2)) are particularly high for both groups: 9.2 ± 1.2 and 9.8 ± 0.4 are the shopping scores respectively of *Over60* and *Under60* group; 9.9 ± 0.3 and 9.9 ± 0.1 are the payment scores respectively of *Over60* and *Under60* group.

Figure 3 shows execution times (3(a) and (b)) and number of remembered items (3(c)) regarding the First and the Second Trial for both age groups. Average shopping time slightly decreases for *Over60* group (347.6 ± 147.6 First vs 319.3 ± 120.5 Second) as well as for *Under60* group (128.3 ± 26.2 First vs 116.7 ± 38.6 Second), but the difference is not significant probably because of

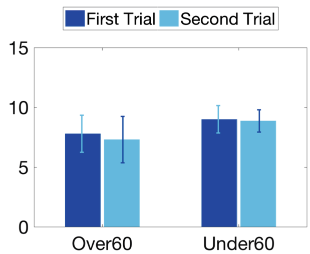


(a) `ShoppingTime` (b) `PaymentTime`

Fig. 2. Comparison of the times required to complete the shopping (a) and the payment (b) task. Comparison between *Over60* and *Under60* groups and between tablet and pc.



(a) `ShoppingTime` (b) `PaymentTime`



(c) `RememberedItems`

Fig. 3. Comparison of the times required to complete the shopping (a) and the payment (b) task and of the number of remembered items (c). Comparison between *Over60* and *Under60* groups and between First and Second Trial.

the high variability of data. The average number of remembered items in the First and Second Trial remains quite constant inside the two groups, but it is significantly higher ($p < 0.05$) in the *Under60* group with respect to the *Over60* group (7.8 ± 1.5 vs 7.3 ± 1.9 for *Over60*, 9.0 ± 1.1 vs 8.9 ± 0.9 for *Under60*).

The last evaluated aspect has been the preferences of the subjects for one device over the other. In general, the 50% of participants said they preferred to use the pc because the screen was larger and the objects were more visible, the 30% of them preferred the tablet for the more intuitive interface and the last 20% enjoyed both devices. In the case of elderly, preferences did not match the objective data, indeed they obtained slightly better performances with the tablet.

4 Conclusion

Our work fits with the current research on designing new innovative methodologies for the assessment of the cognitive status of the elderly using serious games approach. Unlike other works, our focus is not on the design of the simulation environment itself but on the evaluation of the best human-computer interface to be used. Indeed, we developed a simple virtual supermarket and exported it into two devices, a tablet and a pc, in order to compare the performances with the two types of interaction (touch screen and mouse). The game was tested on 32 volunteer healthy subjects without any cognitive disorder, each of them was asked to try the game in both modalities. The performances of the subjects were evaluated by measuring the execution times of the proposed tasks and computing a score for each one of them. We analyzed and compared the acquired data in three different ways: first we compared the performance obtained with the tablet and the pc, then we compared the results obtained in the First and Second Trial and finally we compared the performance of the elderly (*Over60* group) over the young people (*Under60* group), based on the device used and the trial.

From the analysis of the results, no significant differences have been arisen between tablet and pc or between First and Second Trial. This could be due to the high variability of data, heterogeneity in the dataset and limited number of participants. Regarding the latest analysis, it has emerged that the performances of young people were significantly better than those of the elderly, both in terms of execution times and scores. Considering only the results of the elderly group, however, there are small but not significant differences between the performances obtained with the two different devices.

For these reasons we can state that the type of interaction does not significantly affect the performance of the subjects. Both devices, therefore, are potentially valid for developing a serious game for assessing cognitive functions of the elderly. However, the game could be implemented on both platforms and the choice of the device could be left to the patients according to their preferences.

The presented study is a preliminary work that could be used as a starting point for many future works. First of all, it is our strong interest the improvement of the layout and of the usability of the game in the two platforms and the implementation of the game even in an Immersive Virtual Reality in order to obtain a more natural experience and interaction. We have also planned to collaborate with doctors in order to acquire data from cognitive impaired people and old-age people. The cooperation with the medical staff will provide us a better understanding of the problem of assessment of people cognitive status, so we will be able to develop new features for our application in order to create an *ad hoc* tool for doctors. Finally, a comparison and correlation with the score given by the MMSE will be required to understand whether such a game can be used to assess the cognitive status of people.

References

1. Prince, M., Bryce, R., Ferri, C.: World Alzheimer Report 2011: The benefits of early diagnosis and intervention. http://www.alz.co.uk/research/WorldAlzheimerReport2011.pdf. Accessed 20 June 2017
2. Fox, C., Lafortune, L., Boustani, M., Brayne, C.: The pros and cons of early diagnosis in dementia. Br. J. Gen. Pract. **63**(612), e510–e512 (2013)
3. Folstein, M.F., Folstein, F.E., McHugh, P.R.: Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. **12**, 189–198 (1975)
4. Wind, A., Schellevis, F.G., Van Staveren, G.E., van Eijk, J.T.M.: Limitations of the Mini-Mental State Examination in diagnosing dementia in general practice. Int. J. Geriatr. Psychiatry **12**(1), 101–108 (1997)
5. Kurlowicz, L., Wallace, M.: The Mini-Mental State Examination (MMSE). J. Gerontological Nurs. **25**(5), 8–9 (1999)
6. Rizzo, A.A., Schultheis, M., Kerns, K.A., Mateer, C.: Analysis of assets for virtual reality applications in neuropsychology. Technol. Cogn. Rehabil. **14**, 207–239 (2004)
7. Spooner, D.M., Pachan, N.A.: Ecological validity in neuropsychological assessment: a case for greater consideration in research with neurologically intact populations. Arch. Clin. Neuropsychol. **21**(4), 327–337 (2006)
8. Tarnanas, I., Schlee, W., Tsolaki, M., Müri, R., Mosimann, U., Nef, T.: Ecological validity of virtual reality daily living activities screening for early dementia: longitudinal study. JMIR Serious Games **1**, 1–14 (2014)
9. Bronfenbrenner, U.: The Ecology of Human Development. Harvard University Press, Cambridge (2009)
10. Zucchella, C., Sinforiani, E., Tassorelli, C., Cavallini, E., Tost-Pardell, D., Grau, S., Nappi, G.: Serious games for screening pre-dementia conditions: From virtuality to reality? A pilot project. Funct. Neurol. **29**, 153–158 (2014)
11. Vourvopoulos, A., Faria, A.L., Ponnam, K., Bermúdez i Badia, S.: RehabCity: design and validation of a cognitive assessment and rehabilitation tool through gamified simulations of activities of daily living. In: Advances in Computer Entertainment Technology Conference (2014)
12. Parsons, T.D, McPherson, S., Interrante, V.: Enhancing neurocognitive assessment using immersive virtual reality. In: 1st Workshop on Virtual and Augmented Assistive Technology (VAAT), 2013, pp. 27–34. IEEE (2013)
13. Tarnanas, I., Tsolaki, M., Nef, T., Müri, R., Mosimann, U.: Can a novel computerized cognitive screening test provide additional information for early detection of Alzheimer's disease? Alzheimer's Dement. **10**(6), 790–798 (2014)