Effects of Social Conditions in a Virtual Exergame for Cognitive Rehabilitation in Multiple Sclerosis

(Article begins on next page)

Effects of Social Conditions in a Virtual Exergame for Cognitive Rehabilitation in Multiple Sclerosis

Keivan Ameri, Edoardo Battegazzorre
Department of Control and Computer Engineering
Politecnico di Torino
Turin, Italy
edoardo.battegazzorre@polito.it

Chiara Galletti
Rehab Technologies Lab
Istituto Italiano di Tecnologia
Genoa, Italy

Andrea Bottino
Department of Control and Computer Engineering
Politecnico di Torino
Turin, Italy

Jessica Podda, Andrea Tacchino, Giampaolo Brichetto
Scientific Research Area
Italian Multiple Sclerosis Foundation
Genoa, Italy

Lorenzo De Michieli, Giacinto Barresi
Rehab Technologies Lab
Istituto Italiano di Tecnologia
Genoa, Italy

***Abstract*—Rehabilitation is essential for individuals with multiple sclerosis (MS) to improve their quality of life and mitigate the progression of the disease. Cognitive deficits, which commonly affect MS patients, significantly impact daily functioning and well-being. This paper focuses on the development of a digital version of the Tangram puzzle using virtual reality (VR) to promote logical reasoning, attention, and social interactions to enhance engagement in cognitive rehabilitation for MS patients. The developed simulation integrates social presence in the form of a non-playable character (NPC) intending to improve player performance and motivation to ultimately promote engagement and adherence to treatment. A user study involving different game modes with the NPC in either a supporting or interfering role, was conducted to evaluate the effectiveness and usability of the Tangram VR exergame. After collecting and analyzing questionnaire scores and performance data, our results suggest that the application was well-received by users, and the introduction of an NPC had an impact in terms of execution times and sense of frustration, depending on its behaviour. Nonetheless, only a limited set of significant differences between modes was found, suggesting that further study is necessary to validate the results fully.**

***Index Terms*—Multiple Sclerosis, Cognitive Rehabilitation, Virtual Reality, Social Interaction, Engagement, Upper Limb Rehabilitation**

I. INTRODUCTION

Rehabilitation plays a critical role in enabling people with disabilities to actively engage in a transformative process through which they acquire the skills and knowledge necessary for optimal physical, psychological, and social functioning. In multiple sclerosis (MS), a progressive disease characterized by the deterioration of physical and cognitive disabilities, rehabilitation becomes a fundamental component for improving the quality of life of those affected and their families. MS increasingly impacts both the physical and psychological aspects of patients' lives, affecting virtually all facets of their existence. As symptoms greatly vary from person to person and diversely affect various cognitive and motor domains [1], people with MS need to undergo a comprehensive and targeted rehabilitation process aimed at controlling and mitigating the progression of the disease. Mainly, cognitive deficits occur in around half of MS patients [2] and can have a significant impact on an individual's daily functioning and quality of life. As a result, the present work aims to develop a digital solution to promote logical reasoning and attention, along with being exposed to different types of social interactions to foster engagement in the task.

II. BACKGROUND AND SCOPE

A. VR-based Cognitive Rehabilitation

A key aspect of rehabilitation for MS patients is cognitive rehabilitation. In people with MS (PwMS), areas such as attention, memory, learning, planning, mental flexibility, problem-solving, and language recall are commonly affected [3], with direct consequences on self-esteem, participation in social activities, and withdrawal from work and relationships [4]. In recent years, technological advances have made it possible to integrate novel computerized systems into the standard rehabilitation practice. Specifically, virtual reality (VR) has emerged as a promising tool for rehabilitation in PwMS, offering a valid alternative to traditional methods. By shaping the virtual environments and their settings, it is possible to increase engagement and motivation by turning exercises into interactive, immersive experiences [5]. A pivotal point to further promote engagement in the task is to incorporate gamification elements, in order to introduce the fun of an entertaining game while challenging the user's capabilities.

These systems also allow leveraging algorithms and artificial intelligence to collect objective data on the user's performances. Moreover, they can be coupled with biometric sensors to capture and evaluate the patient's psychophysiological state during the session. In the present work, the main objective is to develop a VR-based exergame, namely a game aimed at promoting exercise, which focuses on training cognitive abilities in PwMS with the Tangram game, along with targeting additional skills for upper limb (UL) rehabilitation.

B. Social Interactions for Engagement

In this framework, stimulating engagement in the game is critical to optimize the exercise outcomes, as higher engagement has been directly related to increased in-game

Fig. 1: Overview of the virtual environment during the three game modes. During *Neutral Mode*, players complete the Tangram puzzles by themselves in the room, withouth NPC (*left*). During *Support Mode*, the NPC sits on the couch in front of the player, giving advice, motivation and support (*center*). In *Interference Mode* the NPC impatiently moves in the room, discouraging and bothering the player (*right*).

performances [6]. In the rehabilitation context, enhanced engagement could translate into higher efficacy of the training session, thus improving the overall quality of the rehabilitation protocol and fostering long-term treatment adherence. An effective strategy to promote engagement is to introduce social dynamics in the virtual setting. Indeed, it has been recently shown that social interaction is the key factor that determines the level of engagement of gamers [7]. In the present work, a non-playable character (NPC) is introduced in the game world. Its presence is made interactive through dialogues and comments directed to the player about the actions undertaken. The NPC can either help or disturb the player during the game, making suggestions or criticizing their work. The main novelty of this work is the introduction of social interactions during an exergame for cognitive rehabilitation. The underlying hypothesis is that the NPC enhances the player's performance, which in turn affects their level frustration. The goal of the experiments is to analyze how the behavioral parameters of the NPC influence patient adherence to treatment to support and encourage continued participation effectively. The effectivness of a digital coach has been proved effective in motivating users to adhere to physical activity and wellness programs like in the work presented in [8]. However, here the digital coach behaves in the same way for all patients. To overcome this limitation, we test two different personalities for the NPC (supporting and interfering), which potentially allows, in future works, to match the patient's personality to the virtual avatar's, which has been proved as an effective solution in the context stroke rehabilitation [9], although with physical robots rather than virtual characters. Through these research and development efforts, the aim is to create a system that promotes sustained engagement in treatment, leading to better patient outcomes and overall success in the rehabilitation process. Considering the characteristics of the target users, additional focus is placed on the development of the application to create a user-friendly and usable game that enhances player immersion and presence, which ultimately helps improve their level of engagement in the experience.

To evaluate the effectiveness and usability of Tangram VR, a user study has been conducted involving volunteers who played the exergame in three different conditions, namely alone (*Neutral Mode*), with a supporting NPC (*Support Mode*), and playing with a disruptive NPC (*Interference Mode*).

III. Materials and Methods

A. *Tangram VR*

Tangram is a puzzle game wherein the player is provided with a set number of pieces of different shapes. The objective is to place all the provided pieces properly to form a specific pattern. The Tangram game provides a logic challenge that requires logical thinking, reasoning, and problem-solving skills. It promotes cognitive effort and focus; thus, it is particularly appropriate as an exercise for cognitive rehabilitation. Moreover, manipulating tiles on the board requires precise movements of the upper limbs to properly attach the pieces, training both the accuracy and strength of the utilized arm. These characteristics make the Tangram game well-suited as a combined cognitive and motor rehabilitation exercise in PwMS, effectively addressing the diverse symptoms associated with the condition. In fact, Tangram has been previously emplyed in the context of MS rehabilitation [10]. For our experiment, a VR implementation of the Tangram game was developed using the Unity game engine (version 2021.3.15). . In Tangram VR, there are seven tiles of different geometric shapes: a square, a parallelogram, two small triangles, a medium triangle, and two large triangles, which can form a total of 30 solvable patterns, all inspired by animals (Fig. 2).

Fig. 2: The seven pieces used in Tangram VR: arranged to form the initial square (*left*) and arranged to form the shape of a fox (*right*).

The simulation takes place in a virtual environment replicating an apartment (Fig. 1, *left*), with the player sitting in front of a table wherein there is the Tangram board. In *Support* and *Interference* modes (described later in this section), a Non-Playable Character (NPC) is also present in the room (Fig. 1, *center* and *right*), who interacts with the player by talking and moving around. The simulation implements a text-to-speech module to voice the NPC, leveraging the

API offered by Microsoft Azure [1]. Moreover, speech audio is synchronized with the NPC's lip movements to enhance realism.

B. Participants

Ten healthy volunteers (5 women and 5 men, with an average age of 27.0 ± 1.95 years old) were recruited according to ENACT01 protocol (229/2022) approved by the Ethical Committee of Liguria Region (Italy) on November 14th, 2022.

C. Experimental Setup

The experimental sessions took place at the premises of Istituto Italiano di Tecnologia (IIT). The experimental setup included a Meta Quest 2 headset and Meta Touch controllers, connected via cable to a Predator Triton 300 laptop with NVIDIA GeForce RTX 2070 graphics card and Windows 11. (Fig. 3).

Fig. 3: The experimental setup: Meta Quest 2 headset and Meta Touch controllers.

D. Game Rules and modes

The goal of Tangram VR is to solve 6 different patterns of the Tangram game across 3 different modes (18 patterns total). At the beginning of the game, the seven pieces are arranged to compose a square, and the player needs to manipulate them to compose the target figure properly .

A preview of the shape that needs to be completed is located in the center of the table. The player can grab and move the tiles using the controllers provided by Meta Quest 2 (although interaction with the hand-tracking system is also supported). When the user is about to release the Tangram piece close to its intended final position, a snapping system triggers, which automatically attaches the tile to the board correctly. The snap is signaled by changing the color of the tile (see Fig. 4).

Tangram VR features 3 different game modes, which modulate the social context wherein the game is played thanks to the inclusion or exclusion of a Non-Playable Character. The game modes are:

1) *Neutral Mode*: the player is alone in the room, there is no NPC.
2) *Support Mode*: a male NPC is sitting on the couch in front of the player. He actively assists the user in solving the various figures, offering advice on shape resolution, and providing motivational support.

[1]https://learn.microsoft.com/en-us/azure/developer/intro/azure-developer-overview

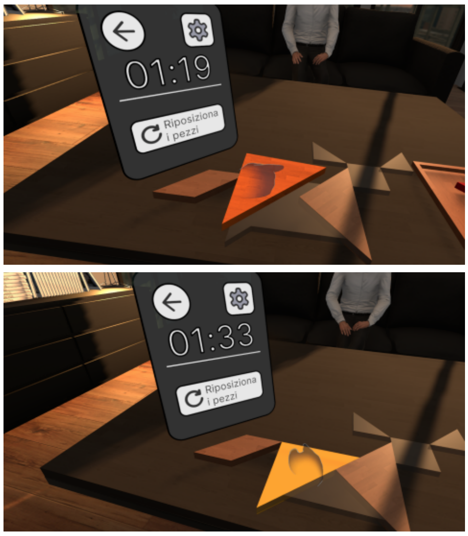

Fig. 4: Snapping system. When a piece is correctly placed in the vicinity of its intended endpoint (top), it automatically snaps to the correct position (bottom) to streamline the task.

3) *Interference Mode*: a male NPC is in the room with the player. He either wanders around the room or sits on the couch. He continuously criticizes the player's work, attempting to demoralize them with belittling voice and piercing gaze.

The goal of the three different modes is to test how the social context influences the in-game performances and how the behaviour exhibited by the NPC impacts the player's focus and motivation in solving the puzzle.

Once the game starts, the system selects a Tangram pattern for the first level. The player has a time limit of 5 minutes to complete each level, after which the attempt is considered as failed. When a pattern is successfully completed, the player proceeds to the next level, where a new pattern is selected from the pool of 30, without repetitions. A three-star score is given if the in-game time is less than 45 seconds, two-star if it is less than 2 minutes, and one-star if the shape is completed in less than 5 minutes. The best time for each shape is recorded. If the same shape is proposed to the user in the following levels, the game duration is compared to the best recorded time to motivate the player to improve their performance and promote engagement throughout the entire game (however, in the experiment described in this paper, the system was tweaked never to select the same pattern twice).

E. Experimental Procedure

The study involves a within-subjects design with a 3-level variable that defines the type of social interaction. Each level defines an experimental condition where the NPC displays a different behaviour. At the beginning of the experimental session, the player undergoes a tutorial to become familiar with the system control and learn the game rules. Here, the NPC explains the basics of the Tangram game and shows the player how to move the tiles to complete the various patterns. The NPC guides the player step-by-step to complete their first pattern. Afterwards, the game session properly begins, and the 6 patterns are presented consecutively. The order of presentation of the different patterns is randomized across all individuals to level any possible learning effect. Moreover, the order of game modes is also randomized across individuals. At the end of each game mode, users fill out three different questionnaires:

1) *Post-test questionnaire:* ad-hoc questionnaire for evaluation of the application. Questions are evaluated on a 5-item Likert-type scale and relate to immersion in the game world, physical and mental effort required by the task, and the influence of the virtual NPC on the experience (see Table 3). The post-test questionnaire features 37 items, 15 of which are ad-hoc questions for this application, 21 are taken (or slighlty altered) from the Exergame Enjoyment Questionnaire [11] and 1 is a variation of a question from SUS.
2) *NASA Task Load Index:* 6-item standardized questionnaire to evaluate Mental Workload [12], [13].
3) *System Usability Scale (SUS)*: 10-item questionnaire designed to evaluate system usability [14].

IV. Results and Discussion

A. Analysis of Questionnaires

First off, statistical analysis was conducted on questionnaires to analyze the differences in scores among the three conditions. The Friedman test was used to compare the scores of the different experimental sessions (i.e.: game modes) since the data were ordinal and not normally distributed. Afterwards, a post-hoc comparison has been performed to investigate the data further and gain additional insights. The Nemenyi test was used to compare two terms of each pair of conditions used (*Neutral-Support*, *Neutral-Interference*, and *Support-Interference*).

For the SUS questionnaires, (Table 1), we did not find any statistical difference between conditions. Nonetheless, the global average score for all modes is 88.17 which, according to [14], [15], is above "Excellent", meaning that participants were engaged with the application and rated it highly usable. Regarding individual SUS items, statistically significant differences ($p < 0.05$) between game modes are found just for the item 3 on ease of use ($\chi^2 = 7.6$, $p = 0.022$) with average values of 4.70 ± 0.46 in *Neutral*, 4.50 ± 0.50 in *Support* and 5.00 ± 0.00 in *Interference*. No post-hoc comparison is significant.

Regarding the NASA-TLX questionnaire (Table 2), significant differencess across game modes are found for two items:

- item 4 (Performance) ($\chi^2 = 7.3$, $p = 0.026$) with 3.70 ± 2.57 in *Neutral*, 6.90 ± 4.20 in *Support*, 7.50 ± 5.26 in *Interference*
- item 6 (Frustration) ($\chi^2 = 14.1$, $p = 8.54e-4$) with 3.30 ± 2.93 in *Neutral*, 4.00 ± 4.19 in *Support*, 6.40 ± 4.67 in *Interference*.

Furthermore, the Nemenyi post-hoc test on the NASA-TLX highlights significant differences:

- item 4 (Performance) between *Neutral* and *Interference* ($p = 0.027$)
- item 6 (Frustration) between *Neutral* and *Interference* ($p = 0.007$) and between *Support* and *Interference* ($p = 0.019$).

TABLE 1: SUS questionnaire results after each mode. Mean and standard deviation over all subjects are reported for all the three conditions. Items marked with asterisk (*) are those where significant differences were found between conditions.

SUS Questionnaire	Condition		
	Neutral	Support	Interference
1. I think that I would like to use this system frequently.	4.5 ± 0.5	4.2 ± 0.7	4.1 ± 1.2
2. I found the system unnecessarily complex.	1.4 ± 0.7	1.6 ± 0.8	1.3 ± 0.5
3.* I thought the system was easy to use.	4.7 ± 0.5	4.5 ± 0.5	5.0 ± 0.0
4. I think that I would need the support of a technical person to be able to use this system.	1.3 ± 0.6	1.3 ± 0.6	1.3 ± 0.5
5. I found the various functions in this system were well integrated.	4.5 ± 0.8	4.2 ± 0.7	4.4 ± 0.7
6. I thought there was too much inconsistency in this system.	1.2 ± 0.4	1.1 ± 0.3	1.1 ± 0.3
7. I would imagine that most people would learn to use this system very quickly.	4.4 ± 0.5	4.6 ± 0.5	4.6 ± 0.5
8. I found the system very cumbersome to use.	1.9 ± 0.9	1.9 ± 0.8	1.8 ± 0.9
9. I felt very confident using the system.	4.5 ± 0.5	4.5 ± 0.5	4.1 ± 0.8
10. I needed to learn a lot of things before I could get going with this system.	1.2 ± 0.6	1.4 ± 0.9	1.2 ± 0.4
GLOBAL SCORE	89.0 ± 8.7	86.8 ± 7.8	88.8 ± 8.4

TABLE 2: NASA-TLX questionnaire results after each mode. Mean and standard deviation over all subjects are reported for all the three conditions. Items marked with asterisk (*) are those where significant differences were found between conditions.

NASA-TLX Questionnaire	Condition		
	Neutral	Support	Interference
1. Mental Demand	12.5 ± 2.0	10.9 ± 3.3	13.4 ± 5.1
2. Physical Demand	4.7 ± 3.0	4.7 ± 2.9	5.7 ± 3.8
3. Temporal Demand	6.6 ± 3.6	14.3 ± 19.2	9.2 ± 4.2
4.* Performance	3.7 ± 2.6	6.9 ± 4.2	7.5 ± 5.3
5. Effort	9.2 ± 4.0	8.7 ± 4.3	12.0 ± 4.7
6.* Frustration	3.3 ± 2.9	4.0 ± 4.2	6.4 ± 4.7

No significant differences are found by comparing the items from the customized Post-test questionnaire (Table 3). Interestingly, the ease of use seems higher in *Interference*, which is also the most difficult in terms of mental demand with a negative impact on self-evaluation of the performance and frustration. The SUS results seem counterintuitive if compared with the ones of the NASA-TLX, demonstrating how certain dimensions of usability can be dissociated from specific aspects of mental workload in a socially engaging context. This remarks the need for further investigations to deepen this aspect. In any case, NASA-TLX item 4 (Performance) highlights a significant difference in perceived performance between the *Neutral* and *Interference* modes, confirming that the NPC acts as a distractor in the latter mode. Moreover, NASA-TLX item 6 (Performance) confirms that the NPC helps relieve frustration in *Support* mode while it heightens it in *Interference* mode.

B. Objective Data

We collected completion times and concentration scores in each mode and level from the application logs for all

TABLE 3: Post-test questionnaire administered after each experimental condition.
Mean and standard deviation over all subjects are reported for all the three conditions.

Post-Questionnaire	Condition		
	Neutral	Support	Interference
1. How much time did the game last? [minutes]	11.3 ± 3.9	12.4 ± 5.4	13 ± 5.9
2. How would you rate your performance?	3.9 ± 0.8	3.7 ± 0.8	3.4 ± 1.1
3. I felt like I lost track of time while playing.	3.6 ± 1.1	3.5 ± 0.9	3.8 ± 0.6
4. I was focused on the game.	4.7 ± 0.5	4.7 ± 0.5	4.9 ± 0.3
5. I felt a strong sense of being in the world of the game to the point that I was unaware of my surroundings.	3.7 ± 1.1	3.9 ± 0.7	3.7 ± 1.1
6. I did not feel a desire to make progress in the game.	1.3 ± 0.5	1.6 ± 0.9	1.2 ± 0.6
7. I did not feel like I wanted to keep playing.	1.5 ± 0.5	1.9 ± 0.8	1.6 ± 1.0
8. I felt that this game provided an enjoyable challenge.	4.5 ± 0.7	4.2 ± 0.6	4.5 ± 0.7
9. I felt that the game was excessively difficult.	2.3 ± 0.9	2.0 ± 0.6	1.9 ± 0.8
10. I felt a sense of accomplishment from playing the game.	4.2 ± 0.7	4.0 ± 0.6	4.2 ± 0.6
11. I felt that the game reacted quickly to my actions.	4.3 ± 0.6	4.4 ± 0.5	4.4 ± 0.5
12. I felt in control of the game.	4.1 ± 0.5	4.1 ± 0.7	4.3 ± 0.6
13. I felt very confident playing the game.	4.2 ± 0.7	3.8 ± 0.9	3.8 ± 0.7
14. I consider playing the game as physical exercise.	2.8 ± 1.2	3.0 ± 1.3	3.3 ± 1.2
15. I felt excited about the physical activities in the game.	2.7 ± 0.8	3.0 ± 0.9	2.9 ± 0.9
16. I felt excited about the mental challenge in the game.	4.1 ± 0.8	3.8 ± 0.9	4.0 ± 1.1
17. The exercise in this game made me feel good.	3.8 ± 1.1	3.5 ± 1.1	3.4 ± 0.9
18. I would rather not be exercising, even though the exercise was accompanied by game elements.	2.1 ± 0.8	2.2 ± 0.9	2.0 ± 0.9
19. I felt that the physical activity was too intense for me.	1.4 ± 0.5	1.5 ± 0.9	1.3 ± 0.5
20. I felt that the mental effort required was exhausting.	1.9 ± 0.8	2.0 ± 0.6	1.8 ± 0.9
21. I felt that playing the game was beneficial for my physical well-being.	3.0 ± 0.9	3.0 ± 1.1	2.9 ± 0.9
22. I felt that focusing on the game was beneficial for my mental state.	3.9 ± 0.9	3.6 ± 0.9	3.9 ± 0.9
23. I would prefer that this physical activity was not accompanied by game elements.	1.4 ± 0.7	1.3 ± 0.5	1.2 ± 0.4
24. The mental effort was more intense than the physical effort during the game.	3.3 ± 1.3	3.4 ± 1.2	3.1 ± 1.4
25. I feel that the muscles of my arm need rest.	1.9 ± 1.1	2.0 ± 1.0	1.7 ± 0.8
26. The situation was making me feel the desire to improve my performance.	4.3 ± 1.2	4.3 ± 0.8	4.5 ± 0.7
27. I felt the need to rest or change arm between sessions.	1.6 ± 0.8	1.7 ± 1.0	1.2 ± 0.4
28. I felt that it was more difficult to concentrate as the game progressed.	1.5 ± 0.7	1.7 ± 1.0	1.4 ± 0.5
29. I felt that my performance was decreasing as the game progressed.	1.9 ± 0.8	1.7 ± 0.9	1.6 ± 0.5
30. I felt that it was difficult to understand how the game works.	1.4 ± 0.9	1.0 ± 0.0	1.2 ± 0.4
31. I felt that it was easy to familiarize myself with different game elements across the sessions.	4.1 ± 0.8	4.3 ± 0.6	4.5 ± 0.7
32. I felt that my pace and performance greatly depended on the game session.	2.9 ± 0.9	3.2 ± 0.7	2.6 ± 0.9
33. I prefer a virtual environment without any characters during this exercise.	2.7 ± 1.2	1.9 ± 1.2	3.2 ± 0.9
34. I would prefer this game to have several but shorter sessions.	2.0 ± 1.3	2.4 ± 1.2	2.2 ± 1.2
35. I felt that it was enjoyable to test different configurations of the tangram.	4.3 ± 0.8	4.3 ± 0.8	4.5 ± 0.7
36. I think that the presence of a virtual character can be useful for motivating me during the exercise.	2.8 ± 0.9	3.7 ± 1.1	2.8 ± 1.5
37. I think that the presence of a virtual character is annoying during the exercise.	2.2 ± 1.0	2.2 ± 0.9	3.3 ± 1.2

experimental subjects. Concentration is defined as the amount of time the player is focused on the puzzle (rather than looking away at the NPC or other parts of the room) divided by the total time to complete a level. Each mode's data set for each individual was compared with the Kruskal-Wallis H-test to find statistically significant differences across modes (Fig. 5). Although we did not find significant differences in completion times or concentration metrics, the trend of completion times suggests that social presence has an impact on performance. Specifically, the supporting NPC acts as a facilitator, while it acts as a distractor in *Interference* mode. However, further study with more experimental samples is needed to validate these results.

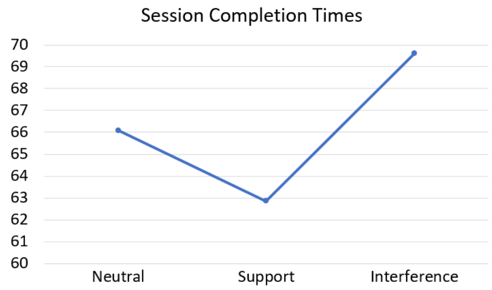

Fig. 5: Average completion time (in seconds) for each mode.

C. Correlations

As a mean to further analyze the data, we computed Spearman rank correlation coefficients between all questionnaire items and task completion times, and we found several significant correlations. In the following we report the most meaningful ones:

- Completion time and NASA-TLX item 4 (Performance) ($\rho = 0.506$, $p = 0.002$)
- Total time and Post-test item 1 ($\rho = -0.397$, $p = 0.017$)
- Completion time and Post-test item 2 ($\rho = -0.517$, $p = 0.001$)
- Completion time and Post-test item 8 ($\rho = -0.425$, $p = 0.01$)
- Completion time and Post-test item 15 ($\rho = 0.443$, $p = 0.007$)
- Completion time and Post-test item 24 ($\rho = -0.455$, $p = 0.005$)
- Completion time and Post-test item 30 ($\rho = -0.432$, $p = 0.008$)
- Concentration and NASA-TLX item 5 (Effort) ($\rho = -0.633$, $p = 3.43e-5$)

- Game mode and Post-test item 37 ($\rho = 0.373$, $p = 0.025$)

The correlation between NASA-TLX item 4 (Performance) and Post-test item 2 with completion times as well as between Total time and Post-test item 1 confirm that subjects were able to reliably assess their own performance, meaning that the offered explanations are clear and that feedbacks offered by the application were unambiguous. The correlation between Post-test item 8 and completion times also suggests that subjects who had more fun took less time to complete the task. Correlation between Post-test items 15 and 24 and completion times also confirm that Tangram VR is primarily perceived as a tool for cognitive rather than physical rehabilitation, as subjects who were more enthusiastic about the physical aspects of the game took more time, while those who preferred the cognitive aspects completed the task faster. Moreover, the correlation between NASA-TLX item 5 (Effort) and concentration, as well as that between time and Post-test item 30, demonstrate that subjects who put cognitive effort in the task were faster and more concentrated. Finally, we found a correlation between the game mode and Post-test item 37. This result was expected since the NPC was rated as more annoying in the *Interference* condition when compared to the other two.

V. Conclusion

In conclusion, this study aimed to develop a virtual reality-based exergame for cognitive rehabilitation in individuals with multiple sclerosis. The exergame targeted cognitive abilities through a digital version of the Tangram puzzle game while incorporating gamification elements and social interactions through an NPC. The user study results demonstrated the VR-based exergame's potential to promote engagement and adherence to treatment. Results from questionnaires suggest that the application was well received by users, and the introduction of social dynamics had an impact on player performance. Moreover, the NPC affects the players' sense of frustration, by acting either as a facilitator or a distractor, depending if they assume, respectively, a supporting or interfering behaviour. Execution times and their correlation with questionnaire scores suggest that adopting a friendly NPC supporting the user has the potential to enhance patient outcomes and overall success in the rehabilitation process. However, it is essential to acknowledge certain limitations. Most prominently, the sample size of the user study was small, warranting further investigations with a larger and more diverse participant pool, including the targeted end-users (i.e.: patients affected by MS). A larger pool is likely to highlight more significant differences among experimental conditions. Moreover, refining the NPC's behavior and interactions can provide a more personalized and adaptive gaming experience. Integration of real-time physiological monitoring, such as heart rate and electrodermal activity, can further enhance the assessment of user engagement, stress, and emotional states during gameplay. Additionally, long-term follow-up studies are essential to evaluate the sustained effects of VR-based cognitive rehabilitation interventions and assess their impact on the overall well-being and quality of life of individuals with MS. Features such as range and speed of movement of the upper limbs will have to be taken into account to tailor the treatment for specific patients' condition and integrate it in their overall rehabilitation program.

VI. Acknowledgments

This study is supported by ENACT Project, a FISM-IIT project [2], in collaboration with the VR@POLITO [3] and CGVG research groups [4] of the Politecnico di Torino, and in synergy with RAISE - Robotics and AI for Socio-economic Empowerment [5]. RAISE is funded by the European Union - NextGenerationEU. Views and opinions expressed are those of the authors alone and do not necessarily reflect those of the EU or the European Commission, and they cannot be held responsible for them.

References

[1] I. Kister, T. E. Bacon, E. Chamot, A. R. Salter, G. R. Cutter, J. T. Kalina, and J. Herbert, "Natural history of multiple sclerosis symptoms," *International journal of MS care*, vol. 15, no. 3, pp. 146–156, 2013.

[2] S. McIntosh-Michaelis, M. Roberts, S. Wilkinson, I. Diamond, D. McLellan, J. Martin, and A. Spackman, "The prevalence of cognitive impairment in a community survey of multiple sclerosis," *British Journal of Clinical Psychology*, vol. 30, no. 4, pp. 333–348, 1991.

[3] M. L. B. Ferreira, "Cognitive deficits in multiple sclerosis: a systematic review," *Arquivos de neuro-psiquiatria*, vol. 68, pp. 632–641, 2010.

[4] G. C. DeLuca, R. L. Yates, H. Beale, and S. A. Morrow, "Cognitive impairment in multiple sclerosis: clinical, radiologic and pathologic insights," *Brain Pathology*, vol. 25, no. 1, pp. 79–98, 2015.

[5] M. G. Maggio, M. Russo, M. F. Cuzzola, M. Destro, G. La Rosa, F. Molonia, P. Bramanti, G. Lombardo, R. De Luca, and R. S. Calabrò, "Virtual reality in multiple sclerosis rehabilitation: A review on cognitive and motor outcomes," *Journal of Clinical Neuroscience*, vol. 65, pp. 106–111, 2019.

[6] C.-H. Chen, V. Law, and K. Huang, "The roles of engagement and competition on learner's performance and motivation in game-based science learning," *Educational Technology Research and Development*, vol. 67, pp. 1003–1024, 2019.

[7] V. H.-H. Chen, H. B.-L. Duh, P. S. K. Phuah, and D. Z. Y. Lam, "Enjoyment or engagement? role of social interaction in playing massively mulitplayer online role-playing games (mmorpgs)," in *Entertainment Computing-ICEC 2006: 5th International Conference, Cambridge, UK, September 20-22, 2006. Proceedings 5*. Springer, 2006, pp. 262–267.

[8] T. Kari and P. Rinne, "Influence of digital coaching on physical activity: motivation and behaviour of physically inactive individuals," in *Bled eConference*. University of Maribor Press, 2018.

[9] A. Tapus, C. Ţăpuş, and M. J. Matarić, "User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy," *Intelligent Service Robotics*, vol. 1, pp. 169–183, 2008.

[10] J. Gich, J. Freixenet, R. Garcia, J. C. Vilanova, D. Genís, Y. Silva, X. Montalban, and L. Ramió-Torrentà, "A new cognitive rehabilitation programme for patients with multiple sclerosis: the 'ms-line! project'," *Multiple Sclerosis Journal*, vol. 21, no. 10, pp. 1344–1348, 2015.

[11] A. Fitzgerald, S. Huang, K. Sposato, D. Wang, M. Claypool, and E. Agu, "The exergame enjoyment questionnaire (eeq): An instrument for measuring exergame enjoyment," 2020.

[12] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Advances in psychology*. Elsevier, 1988, vol. 52, pp. 139–183.

[13] S. G. Hart, "Nasa-task load index (nasa-tlx); 20 years later," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, no. 9. Sage publications Sage CA: Los Angeles, CA, 2006, pp. 904–908.

[14] J. Brooke, *"SUS-A quick and dirty usability scale." Usability evaluation in industry*. CRC Press, June 1996.

[15] J. Brooke, "Sus: a retrospective," *Journal of usability studies*, vol. 8, no. 2, pp. 29–40, 2013.

[2] https://www.enactproject.eu

[3] https://vr.polito.it

[4] https://areeweb.polito.it/ricerca/cgvg/

[5] https://www.raiseliguria.it/)