POLITECNICO DI MILANO
Doctoral Programme in Bioengineering
Department of Electronics, Information and Bioengineering

Virtual reality-based multidomain interventions for older adults with Mild Cognitive Impairment

Supervisor: Prof. Giancarlo Ferrigno
Co-supervisor: Dr. Marco Sacco
Tutor: Prof. Andrea Aliverti

PhD Student
Sara Arlati, 861414

XXXI PhD Cycle

Summary

In recent years, there has been a notable growth of the aging population [1]. Aging alone cannot be defined as a pathology, however it is often related to the decline of one or more systems, thus leading to a reduced functioning in the motor, sensory, and cognitive domains [2]. When the decline in the cognitive domain is greater than expected for an individual's age and education level, the diagnosis is often of Mild Cognitive Impairment [3]. This neurological condition is characterized by an impairment of one or more cognitive domains according to standardized tests; nonetheless, it is not so severe as to compromise the autonomy and the quality of life of the individual [4]. MCI has a prevalence in older adults over 65 varying from 3% to 19% [5]. It often represents a prodromal stage of dementia, but sometimes it reverts back to a normal cognitive status [6]. Because of this reason, the MCI population represents the optimal target population for which to administer interventions aimed at improving cognitive outcomes or halting the progression of symptoms of dementia.

Currently, pharmacological treatment of MCI is not common [7]. It is generally limited to people who have a great risk of developing dementia, as it implies uncertain prognosis, frequent side-effects, and thus ethical concerns. On the other hand, non-pharmacological treatments appear more promising. In particular, researchers are focusing on multi-domain interventions that target possible risk factors for dementia; these are: the provision of healthy lifestyle guidance (e.g., about Mediterranean diet or the administration of physical exercise), the treatment of possible co-morbidities (i.e., obesity, cardiovascular disorders), and the stimulation of cognitive functions [8, 9, 10]. These interventions have been proven effective in controlling psychological and physical comorbities associated to MCI, by having a positive impact not only on cognitive and functional outcomes, but also on isolation, depression, sleep quality, weight gain, hyperlipidemia, and vascular risk functions.

In this context, Virtual Reality (VR) has recently emerged as a promising tool [11]. Due to its nature, it could provide many benefits to rehabilitation [12], and to cognitive interventions too. First, it allows for the creation of ecological training scenarios that elicit natural behaviours, and thus facilitate the transfer of the acquired capabilities to real-life [13]; additionally, it is a flexible medium, thus it allows for the easy generation of controllable and adaptive training programs, which could be customized in terms of level of difficulty and feedback according to each participant's needs [14]. Lastly, VR has been proven to be engaging and to increase the motivation to train. In the context of therapy and rehabilitation, in which obtaining results is often linked to repeating the same task over time,

this may represent an essential feature. Indeed, more motivated users have better performances, and gain better results [12, 15].

It appears that one of the factors causing the user to feel engaged in a VR-based scenario is Sense of Presence (SoP). SoP is defined as the feeling of “being there” in a computer generated environment [16]; it has been proven to be dependent on many objective and subjective factors, among which there is the degree of immersion provided by the device. More immersion generally results in higher SoP, and thus in the increased motivation to train of the users [17], [18]. However, some drawbacks should also be taken into account: higher immersion – e.g. when using head mounted displays (HMDs) – often causes cyber-sickness, i.e., a series of negative symptoms resulting from the mismatch among sensory cues [19].

Due to these reasons, this thesis addresses the topics of feasibility and (whenever possible), effectiveness of VR-based interventions aimed at improving cognitive outcomes in older adults suffering from MCI. The work has been carried out following a pathway going from the less immersive application, to the most immersive ones. Additionally, in the case of immersive VR technologies, being aware of their possible side-effects, required the conduction of additional trials involving non-vulnerable populations (i.e., healthy young adults).

2D Virtual Reality-based combined Physical and Cognitive Training for MCI Patients

This Chapter addresses the effectiveness of a program composed by Physical Exercise (PE) and Cognitive Training (CT) from a multi-domain perspective. Two studies are presented.

The first pilot trial foresaw the design and the implementation of a system based on a projected touch-screen and a cycle-ergometer [20, 21, 22]. Three scenarios were developed to support the execution of PE and CT; (1) a park in which the users travelled while performing a HR-based aerobic PE; (2) an urban scenario with congested cross-roads aimed at training visuo-spatial abilities; (3) a virtual supermarket scenario in which the users had to search for a specific aisle and specific grocery items among distractors (Figure 1).

The protocol of this pilot study foresaw the administration of 15-20 minutes of cycling and about 25 minutes of CT (cross-roads + supermarket) for six weeks, 3 times a week. Ten volunteers aged > 65, and with mild to moderate dementia

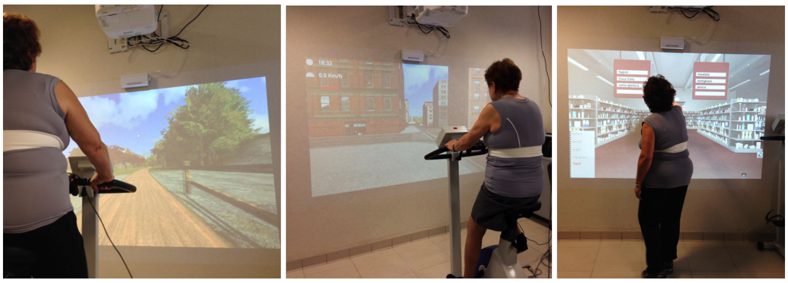


Figure 1: Participants of the pilot study while performing the exercises in the park, in the road-crossing, and in the supermarket scenario.

were randomized into an intervention and a control group. The latter received no treatment.

Analyses conducted post-training revealed a tendency toward improvement in cognitive outcomes for the intervention group, assessed through standardized tests. Significant improvements were found in physiological outcomes (i.e., reactive oxygen species concentration, +8% in the experimental group, −4% in the control group; creatinine, −49% in the experimental group, no differences in the control group), indicating that the intervention, though short and administered to a small sample, had a significant effect on biomarkers indicating an initial impairment of the nervous system due to Alzheimer's Disease.

Finally, the intervention was largely accepted by all participants, who had good adherence, and reported to feel engaged and interested by the use of an innovative means. They also reported a subjective improvement of their Quality of Life (QoL), resulting from the reduced anxiety they had during the accomplishment of activities of daily living (ADLs).

A second study was then carried out with a system resembling the same features of the previous one, with the exception of minor improvements to the interactions and the graphical representation of items. The study was designed as a randomized controlled trial with a 2x2 factorial design, in which the first randomization decided whether the participant received CT, the second whether he/she received PE. Twenty participants aged over 60 with subjective or objective cognitive complaints were recruited in each group and evaluated at the baseline (*t0*), post-training (*t1*), and at 4-months follow-up (*t2*).

The protocol foresaw the administration of training twice a week for 12 weeks. Three groups ($n = 20$ each) received CT, PE and CT+PE; controls ($n = 20$) received no treatment. In this case, significant improvements in cognitive outcomes were found for the groups who had undergone CT, and in particular in long-term memory. Indeed we found differences for *Short Story Delayed Recall* between controls and CT+PE at t1 ($F(3, 76) = 2.82, p = 0.045$; $p = 0.032$ for post-hoc analysis); and between controls and CT+PE ($p = 0.024$), and controls and CT ($p = 0.047$) at *t2*. Also *Phonological Fluency* ($F(6, 152) = 2.254, p = 0.041$) and depression (Geriatric Depression Scale, $H(3) = 1.826, p = 0.05$; $p = 0.046$) improved over time for CT group ($p = 0.008$).

In terms of treatment acceptance, my research team and I recorded positive results, even though the fact of having included less impaired participants for a longer period of time may have caused them to be a little more critical toward the system. In any case, 95% of the participants we interviewed at *t2* said that they would continue with the training, and many reported a perceived improvement of their QoL.

Both the results of these trials were encouraging, though the question of what is the optimal intervention – in terms of frequency, duration, intensity – to improve cognitive outcomes still remains open. The fact that PE alone did not elicit any improvements in our second study was in contrast with previous outcomes [23]. In future studies, designing a more customized treatment, depending on the person's physical and cognitive status may be of help. Research should continue searching for the most effective combination of multi-domain interventions for each specific individual. On the other hand, also improving the VR-based training scenario could be useful, e.g., increasing SoP to elicit higher motivation, thus trying to reach better performances and outcomes [12, 15].

Training with higher Sense of Presence: Cycling in immersive virtual reality

This Chapter explores the possibility of increasing SoP during the accomplishment of PE. Since introducing a navigational component in immersive VR could cause the arousal of sickness, a first preliminary trial was conducted enrolling healthy young adults [24]. In this study, the same scenarios using cycling (i.e., the park and the road-crossing) were modified to run on a HMD. We then performed a comparative study with a large projected screen with the aim of assessing whether SoP would increase with the use of the HMD, and if cyber-sickness would arise.

Thirty-three healthy young adults were enrolled and performed 10 minutes of cycling while wearing the HMD, and while looking at a projected screen in front of them. We found that almost all of them ($n = 26$ out of 30 included) preferred the experience with the HMD. This was probably the result of higher SoP (mean IPQ score: 56.8 ± 10.39 vs 36.5 ± 8.32, $t = -12.06, p < 0.001$) and higher motivation elicited by this condition. However, also cyber-sickness levels were higher in the HMD condition (SSQ median/q1-q3: $30.00/25.75 - 34.00$ vs. $22.00/21.00 - 25.25$, $z = -4.71, p < 0.001$): this occurred because of the expected lateral and forward accelerations that were missing during bends and brakes.

These results showed that the occurrence of cyber-sickness could be somehow masked by the higher engagement of participants, who despite the occurrence of some side-effects preferred the experience with the HMD. Nonetheless, side effects were considered too high to administer this intervention to vulnerable individuals.

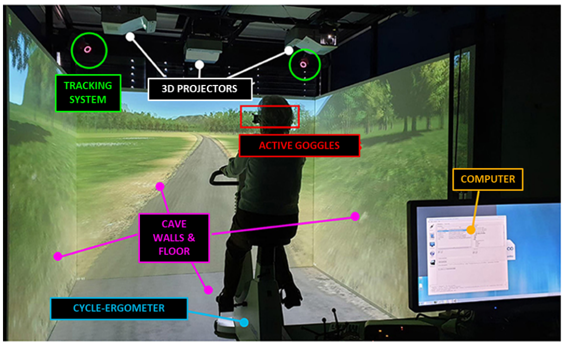


Figure 2: One participant cycling in the CAVE [25].

Thus, another immersive device was considered for the development of a cognitive intervention based on cycling for older adults: the Cave Automatic Virtual Environment (CAVE) [25]. In particular, in the CAVE, we implemented an exercise allowing for the training of cognitive functions using the Dual Task (DT) paradigm. DT is based on the concurrent administration of a physical and a cognitive task [26]; it is motivated by the fact that the simultaneous execution of cognitive and motor tasks can cause a decline either in the physical or in the mental execution, or even in both [27].

The setup comprised a cycle-ergometer placed inside a room-sized CAVE with 3 walls, plus the floor projected with stereoscopic 3D projectors (Figure 2). Participants (5 older adults either with a normal cognitive status or MCI) were asked to cycle in the park for 15 minutes and to recognize some target animals appearing along the route throughout the trial.

All participants, interviewed after the experience, reported a good system usability (SUS score: 76.88 ± 17.0 out of 100), good levels of SoP, and no side-effects. Levels of flow, i.e., the feeling arising when an individual feels that his/her capabilities are challenged to the right extent, were also very positive (SFS score: 4.33 ± 0.75 out of 5). The small sample surely reduced the generalizability of the obtained results, but the homogeneity of the collected data might allow hypothesizing that criticisms would not emerge when extending the experiment to a larger sample of participants.

In conclusion, the CAVE-based training may be promising to administer VR-based PE to older adults, either in concurrency with CT or not.

A HMD-based environment for cognitive training: feasibility and evaluation of naturalness of an immersive supermarket

After having sought the best way to administer PE to older adults, the feasibility of an intervention taking place in an immersive supermarket to train visuo-spatial abilities was also explored.

To do this, a virtual environment containing an aisle and a cash-register scenes was deployed for the HTC Vive HMD (Figure 3). As for the park, a preliminary study enrolling healthy participants was performed [24]; also in this case, we obtained very good results in terms of SoP (ITC-SOPI spatial presence: 3.78 ± 0.42, engagement 3.80 ± 0.33, naturalness 4.15 ± 0.42). Few complaints were recorded in terms of oculomotor disturbances (SSQ-TS: $18.70/3.74 - 38.34$; SSQ-O: $22.77/0 - 39.80$), but the correct placement of the HMD solved almost all of them.

Figure 3: The two scenes of the immersive virtual supermarket.

Prior to evaluating the feasibility of the immersive supermarket on a sample of older adults, we performed a study aimed at assessing whether the naturalness of the reaching movement was preserved in immersive VR [28, 29]. The final goal was investigating whether interacting with VR caused the user to employ higher cognitive resources (due to the elicitation of unnatural behaviours [30]), and thus to distract them from the main (cognitive) task they should perform.

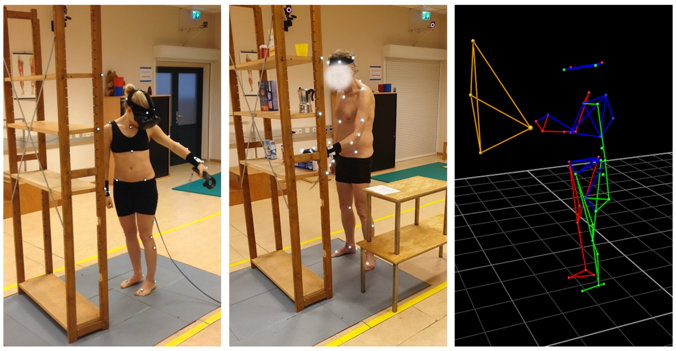


Figure 4: Participants reaching for grocery items in the VR and RWC conditions. On the right side, a screenshot of the motion capture system software [29].

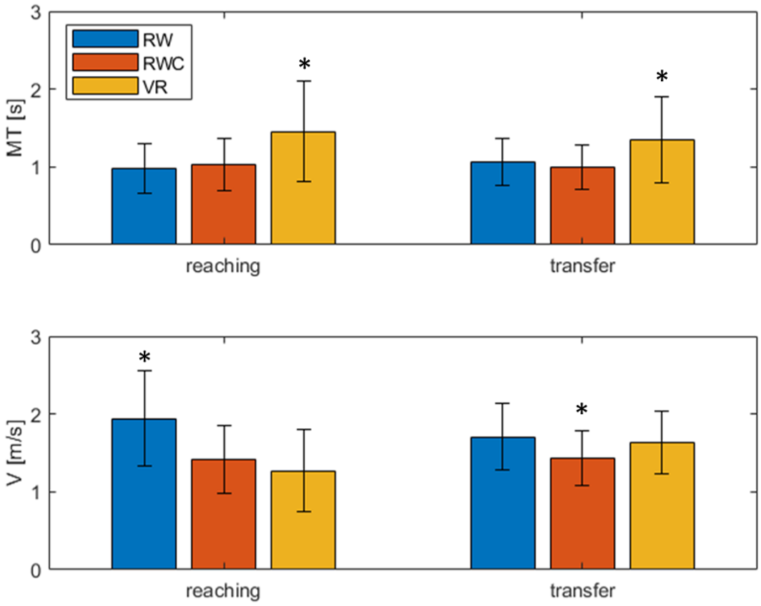


Figure 5: Movement times during the reaching phase in the three conditions of testing [28].

We thus compared 3 different conditions, in which participants had to reach and transport grocery items presented on a simplified version of the supermarket shelves. The 3 conditions were real world (RW), real world while holding the HTC Vive controller (RWC), and virtual reality (VR). Ten healthy adults aged < 40, and 3 healthy older adults (> 65) were enrolled (Figure 4). Their movements were captured with a stereophotogrammetric motion capture system.

Analyzing data collected for young adults, we found, in agreement with previous studies [31, 32, 33, 34, 30, 35], that movement times were significantly longer in VR, and we confirmed that this was not dependent on holding the controller in participants with no motor impairment. We found that the reaching (Figure 5) and the transport phase were influenced by the fact of being in VR, with the exception of transfer peak velocity that was different for RWC only.

On the other hand, we found that curvature, interjoint coordination, and almost all joint ranges of motion were not affected by the fact of wearing an HMD. This could be considered as a promising result in accordance with other findings [36],

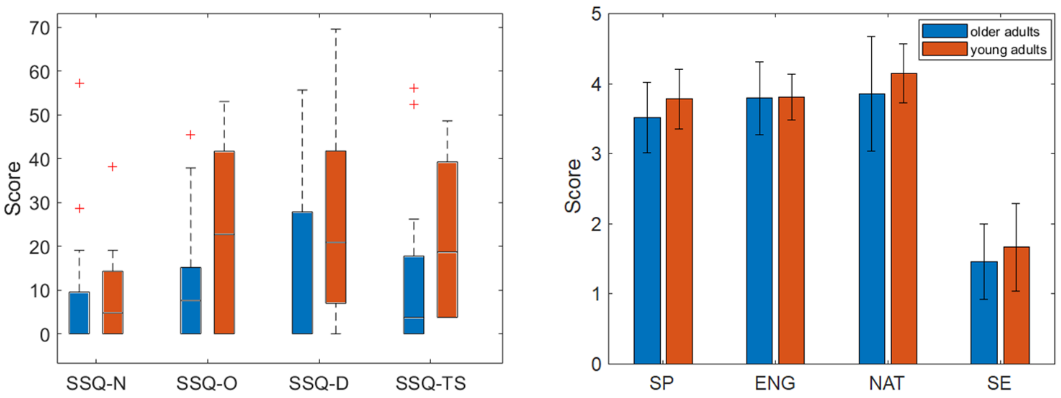


Figure 6: The comparison of SSQ and ITC-SOPI scores obtained for the older and the young adults trying the Virtual Supermarket.

as movement pattern and joint angle synergies were preserved.

Older adults reached even slower than young participants when immersed in VR, and we hypothesized this could be due to the resources required by the target searching, the reduced visual fidelity (vision declines with age [37]), and to the reduced familiarity they had with technology [38]. Due to the very small sample, however, these results have to be treated cautiously.

Clearly, further studies should better address these issues, and try to introduce the investigation of objective performances (e.g., accuracy or precision of the output), the measurement of objective variables (e.g., cognitive workload through EEG), and exploit all the available technological means to try to render the most realistic sensations as possible (e.g., haptic devices).

Lastly, given the good results of the preliminary study on healthy young adults, both in terms of good usability and low cyber-sickness, and the acceptable similarity in between the movements performed in virtual and physical reality (indicating possibly a limited cognitive workload), a feasibility study was conducted enrolling target users [39].

The same scenario used for young adults (with minor modifications aimed at simplifying the interactions with the controllers) was administered to a sample of 57 older adults with either MCI or subjective cognitive complaints. Participants navigated in the VE for 15 minutes and then answered to questionnaires aimed at investigating their subjective experience.

The intervention demonstrated feasibility and acceptablility (TAM3 score: $>$ 5.33 out of 7), and was capable of eliciting high SoP (ITC-SOPI spatial presence: 3.51 ± 0.50, engagement 3.85 ± 0.68, naturalness 3.85 ± 0.82). Levels of realism were high, and only slight side-effects were recorded (SSQ score: $3.74/0 - 16.83$). No differences emerged when comparing the results obtained by the young and the older adults (Figure 6). Thus, also the employment of HMD for the administration of CT programs not implying navigational components (e.g., the cycling) should be considered for further investigations for the implementation of cognitive interventions.

Conclusions and future work

This work has assessed the clinical impact of a VR-based intervention providing PE and CT to older adults with MCI, and the feasibility of immersive VR interventions aimed at improve cognitive outcomes in the same population. The followed pathway allowed testing of the feasibility of such interventions implying an increasing SoP, while always balancing for their potential side effects.

Therefore, the final outcomes of this thesis are: (1) VR could be a promising means to administer CT and PE, and its potential in terms of customization and adaptation of the treatment should be exploited to provide each individual with the best program as possible; (2) immersive VR was largely accepted by older adults, though a good design of the applications is essential.

The limitations of these studies partially reduced the generalizability of the obtained results to the entire MCI population, and especially to older adults with more severe impairments. Nonetheless, the promising outcomes we obtained suggest that the application of (immersive) VR to cognitive interventions is worthy of further investigations.

Future works should focus on the development of VR-based applications able to incorporate novel technologies and exploit the new knowledge emerging from previous studies [11]. Additionally, controlled trials are needed to provide evidence that these interventions are really effective in improving the cognitive status of older adults with MCI.

Acknowledgements

I would like to thank Prof. Giancarlo Ferrigno for the important opportunity he gave me, and for his precious suggestions throughout these years of PhD. I am also thankful to Dr. Marco Sacco, for his constant support to my work, and for his trust in me. Many thanks also to my tutor, Prof. Andrea Aliverti.

This work is the result of valuable collaborations with several partners. Therefore, I would like to thank people who participate in the experimental campaigns presented this thesis: Simona and Flamina from the IRCSS Fondazione Santa Lucia; Elisa and Andrea from the Istituto Auxologico Italiano; Noël, Cheriel, Lise and all the other researchers from the Sint Marteeskliniek. Many thanks also to Gabriele from GPEM, who helped me in working with VICON.

I owe much to all my colleagues at CNR-STIIMA, where most of this work has been carried out; to Luca, Andrea, and Stefano who had supported me from the first moments at CNR, and to all the others who shared with me good and bad moments. Working at the CNR gave the chance to meet a lot of people, and I am lucky to say that many have become good friends; Loris, Daniele, Piri, Giugo, Francesca, Franceschino, Pietro, Jack, Enrica, Luca, Elena, and Vera deserve a special mention.

Thanks to the NearLab people, who always invited me when there was something to eat. Thanks to all my historical friends: Lolla, Carol, Cla, Marghe, to Yvan, to my "fagiane" team, Anna, Cinzia, Fabia, Panna, and Lara, and to my namesake (and *life*sake) Sara.

Finally, I owe I big thank to all my family, and to mum and dad in particular. I know you have always been by my side, even when we were far.

Contents

Chapter 1

Cognitive Decline in Aging

1.1 Aging

In recent years, there has been a notable growth of the aging population. According to the World Health Organization, the proportion of people older than 65 years is increasing faster than that of all the other age groups, and is expected to triple to 1.5 billion by the mid-21st century [1]. Aging is a process leading to the structural and functional deterioration of many physiological systems, even in absence of specific pathologies [2].

Although aging alone is not a pathology, it may cause the decline in one or more systems, thus leading to reduced functioning in the motor, sensory, and cognitive domains. Moreover, aging is characterized by a status of increased vulnerability, in which not only the risk of an adverse event (e.g., minor infection and falls) is increased, but also the body response to a small insult results in a disproportionate change in the health condition (from independent to dependent lifestyle or from lucid to cognitive impaired status) [40].

Given the increased life-expectancy of recent years, many studies have tried to examine the process of aging and how the decline in cognitive, motor and sensory processes occur. This has led to a number of models of aging, though they all can be categorized according to two main leading theories [41]. The first theory includes all those models focusing on the internal causes of aging, thus to the concept that lifespan is genetically determined (Figure 1.1, above). According to them, specific genetic programs lead to a breakdown in the cell replication, immune or endocrine mechanisms, which subsequently result in the death of the individual. This type of aging has been identified as *primary aging*, since it is dependent on the individual. Primary aging is the process leading to slowed movements, fading vision, impaired hearing, reduced ability to adapt to stress, decreased resistance to infections [42].

The theory suggests that aging is due to external causes, rather than internal ones (Figure 1.1, below). Lifespan would be thus indefinite without the damages caused by the environment, viruses and bacteria, foods, and pollutants. This series of models, presenting the so-called *secondary aging*, could explain why some systems can be severely affected by pathologies, while others keep a high level of functioning throughout the whole lifespan [43].

In spite of those conflicting theories, the search for a single cause of aging has

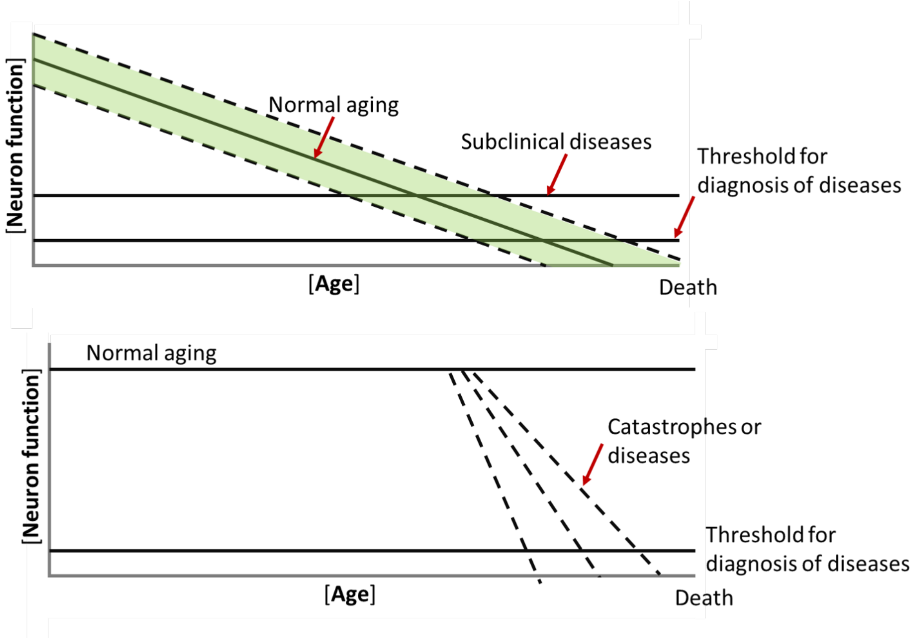


Figure 1.1: The two models of aging; the first (above) model suggests aging occurs as a consequence of an inevitable decline in the nervous system. The second model (below) states that the normal functioning of the neuronal system is kept until a catastrophic event or a specific pathology arises. Adapted from [44].

been replaced in the '90s by the view of aging as an extremely complex, multifactorial process [45, 46], in which both primary and secondary aging factors contribute to different extents. This interaction between primary and secondary factors is particularly relevant for medicine and for research, since – in spite of the inevitable decline due to the genetic makeup – secondary aging can be kept under control. In particular, nutrition, exercise, insults and pollutants are factors that may be addressed to limit the physiological age-related decline. For instance, exercise programs can improve the cardiovascular fitness, thus helping in controlling obesity, motor and cognitive functions. The same is true for nutrition, and preventive health care measures [44].

1.2 Mild Cognitive Impairment and dementia

For what concerns cognitive functions, the beginning of the decline can be associated with the onset of the so-called Mild Cognitive Impairment (MCI). MCI is a syndrome defined as "cognitive decline greater than expected for an individual's age and education level, but that does not interfere with the accomplishment of the activities of daily living (ADLs)" [3]. It is thus different from dementia, in which cognitive deficits are more severe and widespread and have a significant effect on daily function [4].

The term Mild Cognitive Impairment has been firstly introduced in association with stage 3 of the Global Deterioration Scale for aging and dementia (GDS). The GDS, published in 1982, identified 7 clinical stages ranging from normality (1) to very severe cognitive decline (late dementia, 7). People at GDS stage 3 have subtle deficits in cognition and may present some deficits in executive functions affecting complex occupational and social activities. Instead, a GDS score of 4 corresponds to the definition of *mild dementia*. Nowadays, a general definition of MCI is still considered controversial, as the criteria for diagnosis neither specify methods to assess cognitive or functional capacity, nor provide clear cut-offs for cognitive or functional scales to differentiate MCI from normal aging, and MCI from mild dementia [4].

Currently, there are also some attempts to broaden the definition of MCI to include deficits in several cognitive domains other than memory; the clinical phenotypes of amnestic (i.e., the patient has memory complaints) and non-amnestic MCI, and the further classification in single- or multiple-domain MCI have been introduced [47].

Because of these uncertainties, different longitudinal studies have reported a prevalence of MCI in older adults over 65 varying from 3% to 19%, depending on the diverse diagnostic criteria [5]. The same is worthy also for incidence, which was found to be 8‰ to 58‰ per year, and for the risk of developing dementia – especially Alzheimer's Disease (AD) – once diagnosed with MCI, which ranges from 11 to 33% over two years [6]. On the other hand, MCI could also regress: in one year after the first visit, up to 44% of the individuals diagnosed with MCI register the return to a normal cognitive status [6, 48].

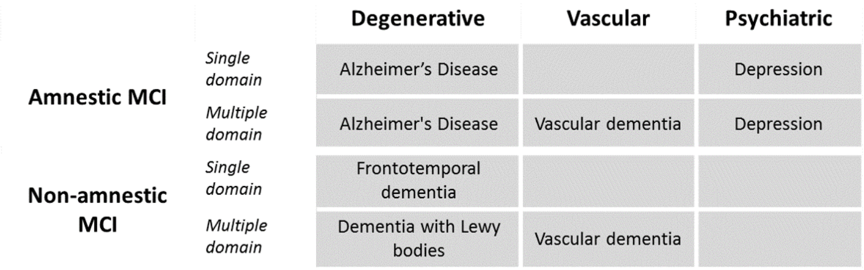
1.2.1 Pathophysiology of cognitive decline

Dementia arises when specific pathological processes occur in addition to the physiological depletion of cognitive resources. The progression of these processes leads to a condition characterized by the progressive decline of different functions, and also of personality tracts and behaviours, thus resulting in a great impairment in the accomplishment of the ADLs, and in the social participation of the affected individual [49]. The most common cause of dementia is AD; MCI is often considered its prodromal stage, even if no clear relationship between these two statuses has been established yet [50].

Pathogenesis of AD is related to the deposition of a protein, the β-amyloid, both at intra and extracellular level [51]. β-amyloid damages neurons through a direct mechanism leading to cell apoptosis, and by stimulating immunity cells normally deputed to central nervous system protection (i.e., astrocytes and microglia) to produce inflammatory and toxic neuron mediators [52]. The most common morphological characteristic of a brain belonging to an AD patient is the marked atrophy, which is mainly due to the neurons' degeneration and to the subsequent decrease of dendritic spines and of synaptic junctions [53]. Compared with people affected by AD, those with MCI have an intermediate amounts of AD pathological findings, with amyloid deposition often limited to the mesial and temporal lobes [54].

In addition to the neurodegenerative damage, cardiovascular diseases have been shown to be related to the occurrence of MCI. The importance of white-matter lesions and small lacunar infarcts is becoming more and more apparent in individuals with cognitive impairments [55]. Findings from the Religious Order Study have demonstrated that the cerebrovascular involvement in MCI is indeed intermediate between what seen in aging and in early AD phases [56].

This evidence has led to the hypothesis that each type of progressive dementia has its own sub-type of MCI (i.e. amnestic and non-amnestic, single or multi-domain) [47]. If confirmed, this would mean that each form of pre-dementia predicts a specific prognosis, with specific effects on survival times. With regard to this theory, Petersen et al. [47] suggested the outline presented in Figure 1.2.



		Degenerative	**Vascular**	**Psychiatric**
Amnestic MCI	*Single domain*	Alzheimer's Disease		Depression
	Multiple domain	Alzheimer's Disease	Vascular dementia	Depression
Non-amnestic MCI	*Single domain*	Frontotemporal dementia		
	Multiple domain	Dementia with Lewy bodies	Vascular dementia	

Figure 1.2: Outline of the syndrome of MCI. The Figure shows that amnestic and non-amnestic types of MCI may be interpreted as prodromal stages of different types of dementia, also depending on the cause. Other medical disorders, such as metabolic and nutritive deficits, upper airways' obstruction and head trauma, can contribute to the development of dementia. Adapted from [47].

As it can be seen from the Figure 1.2, also the presence of behavioural and psychological signs, including depression, predicts a high likelihood of progression toward the status of declared dementia [57]. It is thus plausible that, in the future, the definition of MCI will include also non-cognitive symptoms that may represent the prodrome of disorders that belong to dementia types, but are different from AD, such as fronto-temporal and Lewy-body dementia [4].

1.2.2 Criteria for diagnosis

In spite of the increasing interest in dementia research, currently there are no clear cut-off points that allow distinguishing normal cognition from MCI [58], and MCI from mild dementia. However, an International Working Group has tried to propose some core clinical criteria for the identification of MCI [49]. These are:

1. *presence of concerns regarding a change in cognition*; there should be evidence of concern about a change in cognition, in comparison to the person's previous status. Such concerns can be communicated by the patient him/herself, from a care-giver, or can be observed by a clinician with experience in the field.

2. *impairment in one or more cognitive domains*, i.e. memory, language, executive functions, attention and visuo-spatial skills. The impairment must be greater than what expected for a person of that age, and with that educational level.

3. *preservation of independence in functional abilities*, in a MCI subject, the accomplishment of complex functional task is less efficient, requires more

time, and the person is more prone to make errors; however, autonomy of functioning in daily life, with minimal aids or assistance, is preserved.

4. *the person is not demented*: all the measured cognitive impairments must not be significantly affecting the social and occupational functioning of the individual.

MCI can also be accompanied by other changes, such as impairments in balance and coordination [59].

Since it remains fundamental to assess whether there is an objective cognitive decline, and if so, the degree of such a decline, psychometric tests are currently the way to make a diagnosis. The most widespread scales used for the diagnosis of MCI and dementia are the Mini-Mental Evaluation Scale (MMSE) [60, 61] and the Montreal Cognitive Assessment (MoCA) [62]. Both have high levels of sensitivity for moderate-to-severe cognitive impairment, but showed lower sensitivity for mild dementia and MCI; however, their easy and quick administration still makes them the preferred tests in clinical settings [4]. In this context, much research has underlined the increasing need of sensitive and user-friendly cognitive tests for clinicians. However, in absence of these tests, a variety of validated clinical neuropsychological measures is available to assess several cognitive abilities. Since MCI may affect different cognitive functions, it is indeed recommended to administer a battery of tests, rather than only one. Analysing more scores together allows, in fact, obtaining a more complete clinical picture of the patient [63]. In general, the scores obtained by individuals with MCI are 1 to 1.5 standard deviations below the mean for their age- and education-matched peers on culturally-appropriate normative data (i.e., regarding each impaired domain(s), when available) [64].

Biomarker-based diagnosis

Beside the diagnosis that could be obtained from the administration of standardized psychometric tests, there are some recently-introduced clinical research criteria which incorporate the use of biomarkers [65, 49]. Those criteria are dedicated only to research settings, due to some limitations they have, mainly because of their recent advent: (1) more research is still needed to confirm whether the use of a specific biomarker is effective in identifying cognitive decline, and its extent; (2) there is a limited standardization of procedures, and cut-off points must still be defined; finally, (3) access to biomarkers is limited in non-dedicated research settings.

Biomarkers, especially in the case of MCI due to AD, can be categorized into several different classes. Some biomarkers directly reflect the pathological status by highlighting the presence of key proteins deposited in the brain during the course of AD, such as the β-amyloid protein and tau [65]. In particular, β-amyloid can be detected and quantified in the cerebrospinal fluid and in the plasma of the patient, and reflects the presence of amyloid plaques in the brain, i.e., one of the hallmarks of AD [66, 67].

Other biomarkers provide less direct or nonspecific evidence of AD by tracking a variety of indices of neuronal injury; they can include a number of structural and functional measures, including brain atrophy, and hypometabolism or hypoperfusion obtained with magnetic resonance (MRI), positron emission (PET) and single-photon emission computed tomography (SPECT) imaging [68, 69] .

A third category of markers is constituted by the ones related to biochemical changes. Indeed, the progression of dementia – and of AD in particular – is characterized by numerous biochemical events, including oxidative stress, and inflammation [70].

1.2.3 Risk factors

Several known risk factors for the development of dementia have been identified throughout years, and, in parallel to the aging factors, they can be grouped under two categories: the *non-modifiable* and the *modifiable* risk factors [71].

Among the *non-modifiable* risk factors, there are aging and the person's genetic characteristics. The most consistent risk factor for the development of Alzheimer-type dementia has been identified in the allele APOE $\epsilon 4$ [72]; carriers of such an allele are more than 2 times as likely to manifest dementia. Those who are homozygotes for APOE $\epsilon 4$ have a 4-fold increase in the probability of developing AD [73].

For the secondary, i.e. *modifiable* risk factors, several comorbidities have been identified; they include diabetes mellitus, hypertension, hypercholesterolemia, heart disease (history of coronary artery disease, atrial fibrillations, heart failure, or valvular heart disease) [74]. Being less educated seems to contribute to the development of dementia [75], so as having had a traumatic brain injury or having sleeping disturbances [74]. In addition to these personal characteristics and clinical conditions, there are some other factors correlated with the person's lifestyle that seem to negatively influence the progression of dementia. These are: lack of physical exercise, of cognitively stimulating activities, and of social engagement; smoking, alcohol consumption, and unbalanced diet [74].

The relationships between the above-mentioned factors and the development of different forms of dementia are still uncertain, and further studies and randomized clinical trials are needed to shed new light on these topics. Nevertheless, it is also true that there is enough evidence from a population-based perspective to affirm that it is possible to reduce the risk of cognitive decline acting on the modifiable risk factors. Physical exercise, management of cardiovascular pathologies, a healthy diet, lifelong learning and cognitive training seem to be able to halt or, at least, to slow down the decline both from a normal cognitive status to MCI and from MCI to severe dementia [71].

With the increasing population age worldwide, and the increasing incidence of dementia, it becomes fundamental that all the interventions aimed at preventing or managing cognitive decline act for modifying one, or – preferably – more of the risk factors, thus promoting a healthy, active and social engaging lifestyle [76].

1.3 Management of cognitive impairment in older adults

Currently, there is no therapy proven to reverse the neurological damage due to dementia [7]. However, the increased risk of developing dementia while being affected by MCI requires an appropriate management, also considering the psychological influence of knowing to be cognitively impaired. Many older adults show the so-called *excess disability* after MCI diagnosis, meaning that they withdraw from any activities, even if they would be able to accomplish them [77]. This behaviour, not only contributes to make dementia symptoms appearing worse, but also reduces social participation and engagement in cognitively stimulating activity, making the progression toward the pathological condition faster.

Both pharmacological and non-pharmacological interventions have been developed and tested to try and halt the occurrence of new symptoms or make MCI patients to return to a normal cognitive status [78]. Pharmacotherapy is currently not common, and generally limited to people who have a great risk of developing dementia. Medical side-effects and uncertain prognosis, together with ethical concerns [7], limit the pharmacological treatment of MCI. For instance, acetylcholinesterase inhibitors (i.e., galantamine) are among the most employed drugs to improve cognition; however, they increase the death rate, and have no sure effect on the conversion from MCI to dementia: an adequate evaluation of risks and benefits must thus be done prior to prescription. Moreover, even if their functioning may appear promising, their effectiveness seems to be limited only in APOE $\epsilon 4$ patients with amnestic MCI [79].

On the other hand, considering non-pharmacological treatments, results are more encouraging. Positive outcomes have been obtained in several trials, and meta-analyses have started showing the efficacy of: periodic monitoring of patients, providing lifestyle guidance [8], treating lifestyle-related diseases [9], and training cognitive functions [10]. Considering *secondary aging* factors, these treatments have been proven effective in controlling psychological and physical co-morbities associated to MCI, by having a positive impact on isolation, depression, sleep quality, weight gain, hyperlipidemia, and vascular risk functions [80]. A few of these interventions, their administration and effectiveness are better described in the following paragraphs.

1.3.1 Cognitive interventions

Cognitive interventions may bring an important benefit to people suffering from MCI, because these individuals have a high need for treatment, and they could retain the cognitive capabilities to learn and apply sets of new strategies. If designed properly, cognitive interventions could optimize the functioning of persons with MCI, thus reducing their limitations and preventing the anxiety resulting from their condition awareness, and from their cognitive failures [81]. Different forms of cognitive interventions have been described in literature. Clare et al. [82] have identified three approaches to cognitive intervention. These are:

- *cognitive stimulation*, which refers to the involvement in group activities specifically designed to increase cognitive and social engagement in a non-specific way; cognitive stimulation may include supervised leisure activities, group discussion, reality orientation, or reminiscence therapy.

- *cognitive rehabilitation*, which foresees individually-tailored programs defined together with the patient and his/her care-givers; cognitive rehabilitation focuses on specific ADLs that are defined as "aims" by the older adult him/herself.
- *cognitive training*, which involves a guided practice on a set of standardized tasks designed to reflect particular cognitive functions such as memory, attention or problem-solving. It includes repeated problems and exercises aimed at training specific cognitive abilities in different conditions, and strategies to exploit spared cognitive capacities to improve the impaired ones (e.g., memory-training techniques that rely on visual imagery to support episodic memory).

Cognitive training (CT) is the intervention that has been tested most often. Programs based on CT should address each individual's impaired cognitive domain, and consider both the ecological validity and the impact of the intervention on the person well-being. The mechanisms with which CT can contribute to improve cognitive functioning are still not well understood, but are thought to involve the enhancement of neurobiological processes that support cognitive function, including synaptic reinforcement, long-term potentiation, and the activation of diverse neural networks [83]. In recent years, evidence has shown that CT changes both brain structure [84] and function [85]. Nonetheless, the final aim of CT should be an improvement in cognition which generalize to functional outcomes, i.e. a perceivable improvement in the activities of daily living [86]. With respect to this outcome, evidence in literature is mixed, and researchers call for more studies aimed at identifying barriers and facilitators to transferring abilities acquired during the training period to real life [81].

In CT field, the current tendency is going toward the implementation of tailored training, rather than applying the one-size-fits-all paradigm. This is favoured by the progressive substitution of paper and pencil-based training with Computerized Computer Training (CCT) and Virtual Reality (VR)-based training programs, which can both adapt the level of difficulty and the typology of the proposed tasks depending on the cognitive profile of each user [86].

A Cochrane review [81] has reported a positive effect of CCT and VR-based programs in the cognitive domains of attention, executive functions and memory, assessed through standard psychometric tests. Improvements on psychological aspects, such as mood, depressive symptoms and anxiety were also found. On the other hand, the assessment of ADLs demonstrated no significant improvements, highlighting the need of better investigating the transfer of gain also in the case of computer-based programs, (often) characterized by high ecological validity.

Nowadays, CCT interventions are very diverse and span from the use of commercial software targeting specific cognitive domains (e.g., NeuroPsychological Training, Brain Fitness) to ad-hoc developed applications. Batteries of exercises generally include identification of target sounds, memorization, visual matching, target detection, discrimination and localization, mathematical operations, tasks to train reaction times. VR-based study make use of multisensory stimulation in environments simulating living environments; exercises in this case dealt with route finding, memorization, performance of simple ADLs [87].

1.3.2 Physical exercise

Physical inactivity is one of the modifiable risk factors for the development of AD, and dementia in general [23]. Vice versa, doing physical exercise (PE) is known to improve general health reducing the occurrence of infarction, stroke, and diabetes and facilitating neurogenesis and improving cerebral perfusion. Because of these reasons, it is plausible to hypothesize that PE may be helpful also in preventing the occurrence of AD, and dementia in general [88]. Different studies have started investigated this field, and, as for CT, results appear to be promising. In 2016, Stephen et al. [88] reviewed the results of 24 longitudinal observational studies to determine whether PE can be considered as a preventive intervention with respect to AD development. Eighteen studies out of 24 confirmed this hypothesis and the results of previous reviews of the literature, showing that being physically active throughout life can contribute to the maintenance of a good cognitive status in old age, and also that increasing the level of PE during mid- or even late-life may be to be beneficial.

Interventions based on the provision of PE – also for a limited period of time – have also proven effective in improving cognitive outcomes. In 2013, a RCT was performed enrolling 86 older adults with subjective memory complains [89]; participants were randomized in three groups: two intervention groups performing either resistance or aerobic training twice a week for 26 weeks and a control group doing stretching and balance exercises. Results showed improved verbal memory for the aerobic training group, and improved reaction times in spatial memory tests for both intervention groups. These outcomes were in agreement with previous studies [90, 91]. In particular, the first reported that the group who performed 6 months of aerobic training 4 days/week improved in both executive functions and attention domains, whereas the control (stretching) group remained stable. The second study, instead, showed improvements in global cognition and logical memory in MCI patients undertaking aerobic exercise, strength and postural training for 6 months.

However, as for other types of cognitive interventions, evidence is mixed, with studies obtaining no significant improvements [23] or showing that PE is not superior to other preventive interventions, e.g., health education program [92]. This is probably due to high heterogeneity of the conducted studies. Also for PE, in fact, it remains difficult to define which type(s) of PE, and which intensity, frequency, and duration are needed to obtain an effective reduction of the risk of developing dementia.

1.3.3 Multi-domain interventions

The shortage of high-quality evidence in the field of dementia prevention has not allowed identifying all the factors (i.e., training frequency, duration, intensity) contributing, either positively or negatively, to the effectiveness of the non-pharmacological interventions [93]. Thus, the definition of standard protocols aimed at halting or delaying the cognitive decline in older adults is still lacking. Given this, and given the multifactorial aetiology of dementia, researchers have recently shifted their focus toward longer (e.g., 2 to 6 years) interventions targeting multiple risk factors simultaneously.

The conduction of three large RCTs addressing different risk-factors, and concluded in the last years, constitutes the demonstration that the multi-domain approach appears today as the most effective strategy to try and halt the occurrence of cognitive impairments in the older population [94].

The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) [95], ended in 2014, was the only study demonstrating the effects of a multidomain approach made of physical exercise, diet, cognitive training, social stimulation, and vascular risk factors management, in improving or maintaining cognitive functioning in older adults at risk of dementia. FINGER enrolled 1260 participants showing the presence of modifiable risk factors (old age, low education, sex [M], high body mass index and cholesterol, no physical activity, assessed with CAIDE Dementia Risk Score $\geq$ 6 pts) and cognitive performance slightly lower than what expected for age-matched normative samples. At the end of the intervention period, the between-group difference in neuropsychological test battery was in favour to the intervention: improvement in the total score was 25% higher in the intervention group than in the control group.

Moreover, the intervention had a significant beneficial effect on processing speed, executive function and complex memory tasks, and reduced the risk of cognitive decline acting on non-cognitive outcomes such as body-mass index (BMI), diet, PE and quality of life (QoL). Finally, sub-group analysis demonstrated that also participants carrying APOE $\epsilon 4$ allele benefitted from the intervention [96]. This meant that APOE $\epsilon 4$ appeared not to hinder the intervention-related improvement, thus opening the way to a series of practical positive implications for further non-pharmacological treatments, also for APOE-related genetic susceptibility to dementia.

On the other hand, the two other RCTs – the French Multidomain Alzheimer Preventive Trial (MAPT) [97] and the Dutch Prevention of Dementia by Intensive Vascular care (preDIVA) [98] – obtained no significant results in their primary outcomes, i.e., cognition and dementia incidence, respectively.

In the former study, researchers enrolled 1680 frail elder individuals, and tested the efficacy of omega-3 polyunsaturated fatty acid supplementation either alone, or in combination with a multidomain intervention composed of cognitive training, advice on diet, and physical activity. In spite of non-significance of the primary outcome, i.e., general cognition assessed via a psychometric test battery, positive results were found through post-hoc analyses by pooling together both groups receiving the multidomain intervention. Both groups showed a beneficial effect on MMSE and orientation tests. Furthermore, the multidomain intervention was found to be effective among individuals with an increased risk of dementia defined either as a CAIDE (Cardiovascular Risk Factors, Aging, and Incidence of Dementia) score $\leq$ 6, or presence of amyloid-β in PET images.

In preDIVA, the effectiveness of a 6-year multidomain cardiovascular intervention was assessed in comparison with standard care, in a sample of unselected population of older people with all-cause dementia. The intervention consisted in 3 visits per year to a general practice, in which customized lifestyle advices were given to the study participants, after the assessment of cardiovascular risk factors: smoking habits, diet, physical activity, weight, and blood pressure. Control group received standard care. Also in this case, the intervention did not result in an overall decrease of dementia incidence, which was the primary outcome, but additional analyses showed that the intervention had a protective effect for non-Alzheimer dementia, and that treating hypertension was beneficial to dementia occurrence.

In conclusion, what emerges from these trials, other smaller RCTs, and observational cohort studies, is that multidomain interventions that target several risk factors at the same time might be the key for an effective prevention. Nonetheless some essential aspects, such as age, severity of symptoms of the target population, or the intervention intensity, still have to be determined in order to identify the (potentially different) optimal treatment(s) addressing at-risk and early-symptomatic individuals with diverse types of dementia [93].

1.4 Virtual reality and rehabilitation

VR is a computing technology that generates simulated or artificially 3-dimensional environment, which imitates reality; thanks to input and output devices, VR users are brought to believe that they actually perceive sensory information that is similar to that of the real world, and are invited to respond with natural behaviours.

A growing number of devices is currently available for interaction with VR (e.g., joysticks, gloves, surfaces, etc.), and for stimuli presentation (e.g., screens, audio headsets, speakers, haptic interfaces); however, VR system can be roughly classified into the following 3 categories [99].

- *Non-immersive systems*; in this case, the user experiences 3D environment using a conventional monitor, and usually interacts using keyboard, joysticks, trackball and mouse.

- *Semi-immersive systems* comprise relatively high-performance graphics computing systems, coupled with a large screen monitor, or a large projected screen, or a multiple television projection system. Stereographic imaging, in this case, can be achieved using shutter glasses.

- *Immersive systems* ideally would consist of a set of displays (visual, auditory, haptic) and a tracking system. Immersive VR is currently lacking a clear definition in literature, but the minimum requirements for a system to be considered immersive are: (1) a display allowing for stereoscopic vision, and (2) a tracking system enabling the adjustment of the point-of-view in the virtual scene, according to the user's head orientation in space. Up to now, two different categories of devices are believed to be fully-immersive: head mounted displays (HMDs) and the Cave Automatic Virtual Environment (CAVE).

Going from desktop-based system to immersive VR technologies, users can theoretically experience increasing levels of Sense of Presence (SoP). SoP, i.e., the feeling of "being there" in an environment that is different from the actual physical locale, is influenced by *immersion* (the feeling of being part of the virtual world, interacting directly with the environment) and *involvement* (the focusing of one's attention toward a specific set of stimuli). The elicitation of SoP depends on many factors, such as the control the user has, the quality and the synchronization of the sensory stimulation, the realism of the virtual scene. Subjective characteristics may also have an influence: the willingness of the user to be distracted from reality [16] or the susceptibility to motion sickness [100] could play a role in the experience of SoP.

Because of the presence of subjective factors, and of the fact that none of the above-mentioned factors is effective in isolation, obtaining SoP is not a straight process. Rather, high SoP is achieved only considering the complex interaction of the many elements involved.

1.4.1 The potentialities of virtual reality in rehabilitation

VR has recently emerged as a promising tool in many domains of therapy and rehabilitation. It has been defined as *more than a simple linear extension of existing information and communication technology for human use* [14], because of its advantages and potentialities with respect to other media. In the field of rehabilitation and training, the following points of strength have been identified [14].

- *High ecological validity*, i.e., the degree of similarity that a test or training system has with respect to the real world, and its value for predicting or improving functioning in ADLs. Indeed, VR allows for the creation of realistic environments in which challenges requiring real-world functional behaviours can be easily included, tested and trained in a systematic fashion; this is generally not possible in standard laboratories dedicated to rehabilitation.
  An important point to highlight is that ecological validity is often related to the graphical realism of the virtual scene. However, this element alone could not be enough; or, vice versa, very unrealistic scenarios could be considered ecologic as humans are capable of suspending disbelief and of embodying in a non-existing character with no particular effort [12]. In essence, as long as the VR scenario resembles the real world and its challenges, and the system responds well to user interaction, ecological validity could be considered preserved, and probably enhanced with respect to standard real-life laboratory-based scenarios (which are usually simplified due to space and cost-related constraints).

- *Controlling stimuli delivery*: VR gives researchers the possibility of controlling the amount of stimuli to be conveyed to the end user. Virtual Environments (VEs) can be developed to present simulations that allow for the rehabilitation of physical, cognitive or psychological processes under a range of stimulus-conditions that are not easily controllable in real-world. This flexibility can be used to increase or decrease the level of difficulty of the proposed tasks, according to user's capabilities, and his/her progress throughout time. It can also support a learning-by-doing approach, in which cues are given to the patient *prior* to a response, in order to elicit an error-free performance. Finally, creating adaptive environment that are adjustable according to various patients' needs may also constitute an added-value [101].

- *Real-time feedback and performance record*; in rehabilitation, as in all learning scenarios, performance feedback is necessary to foster learning and skill acquisition processes. VR embeds the possibility of computing the performance in real-time – considering objective measures collected during the simulation by appropriate sensors – and to return feedback that would not be occurring (e.g., sounds, red signs, etc.), or that would be impossible to achieve in reality (e.g., seeing an avatar mimicking the patient's movements from different point-of-views [102]). The information collected during the training session can also be stored, analysed and reviewed in the form of a complete digital record of performance, which remains available for the therapists.

- *Independent practice and spontaneous behaviours*; VR can be assumed to be able to provide theoretically high levels of SoP. When SoP is high, the suspension of disbelief is enhanced and a particular psychological state, defined as *flow*, occurs [103]. In such a state, the user's attention is directed toward the VE, he/she feels that challenges and skills are progressively balanced, and the user experience can become somehow less conscious, with, for instance, the perceived removal of the therapist from the room, or the feeling to be elsewhere with respect to a clinical lab. Flow has been hypothesized to influence task performance, and to create a unique window in which individuals perform in an independent and autonomous fashion, giving the therapists the possibility to assess how personal compensatory or problem-solving strategies are spontaneously applied when challenging situations occurs.

- *Increased motivation and treatment adherence*; one of the main limitations affecting rehabilitation programs deals with the repetitiveness of the proposed tasks; bored patients have less motivation to train, and thereby reduced outcomes that often result in decreased treatment adherence and increased possibility of withdrawals. In such a context, VR has emerged as a means to overcome this limitation. Whether exploring a novel world or performing familiar actions, people naturally consider VEs to be fun and more stimulating. Likewise, studies have shown that the employment of innovative technologies causes users to have more positive reactions to their experiences [104]. Additionally, the excitement produced by VR and relevant technologies may be intensified if the program includes entertaining and gaming elements, and if a scoring system – based on actual performance – exists. Whether increased motivation leads to increased treatment outcomes is still an open issue, whereas the positive correlation between VR-based programs and patients' motivation has been repeatedly demonstrated.

- *Safety* represents an important concern whenever patients need to be trained in ADLs that are intrinsically dangerous, such as driving, crossing streets, cooking, and operating mechanical or industrial equipment. Clearly, practising in a simulated environment preserves the users (and the therapist) from all potential risks; also, it offers the patients the possibility of performing a self-assessment and of increasing their awareness of limitations, thus allowing a more informed decision about whether they feel confident enough to go on performing that activity in real life [105].

- *Telerehabilitation* constitutes one of the possibilities enabled by VR-based treatments [106]. Having the chance to exercise at home, in a controlled way, would mean to increase client access, to enhance rehabilitation outcomes and to reduce costs related to transfers and to the treatment itself. A remote connection with the therapist could allow periodic monitoring and adaptation of the exercise, both to increase the difficulty, and to insert new scenarios in order to avoid boredom. Possibilities are many, but the actual implementation of telerehabilitaion is still limited by the potential risks connected to lack of supervision, security of the collected data, and the concerns of the patients in terms of being disconnected or abandoned by the therapist, or, again, in terms of privacy.

As mentioned, the advantages offered by VR-based treatments are many, but some existing issues must also be considered [14]. First, limitations in terms of software and hardware development still constitute a barrier to the application and the diffusion of VR technologies in the field of rehabilitation. Lack of platform compatibility, of appropriate human-machine interfaces allowing for a natural interaction, or for the restitution of a realistic feedback, or, again, that are very expensive, still halt the large diffusion of VR technologies in many clinical settings.

Second, as the development of technologies proceeds, researchers may be tempted to increase the complexity of their applications and of their interfaces; however, this may cause the occurrence of usability problems, as a result of the lack of understanding of the meaningfulness and of the functioning of the VR application. If not taken into account, low usability, especially when dealing with vulnerable individuals (e.g., older adults, people with disabilities, and children), can negatively affect treatment acceptance [107].

Finally, another important aspect to consider in order for VR to become a safe and widespread tool for rehabilitation is the potential occurrence of side-effects. *Cybersickness*, i.e., a form of motion sickness including nausea, eyestrain, dizziness and disorientation occurring during the VR experience [108], and *aftereffects*, i.e., disturbed balance and locomotion, flashbacks, fatigue, drowsiness occurring as a consequence to the exposure to a non-real world, are the two categories of symptoms to be aware of when developing VR-based rehabilitation programs [109].

These limits, and some others (e.g., the difficult exploitation of collected data), do exist, and further research in the field cannot neglect them. However, given the expected evolution of the current systems, and the results obtained until now, it appears worthy to continue investigating the potential of VR. Indeed, from the one hand, VR and gaming industries are evolving fast and thus better technologies would be probably available soon; on the other hand, developers should continue their work not disregarding an appropriate design of the applications, which should be user-centered, targeted to respond to specific patients' needs, and to be used by them; taking care of possible side effects and reducing them as much as possible is also mandatory [110].

1.4.2 Virtual reality for cognitive interventions

The potential opportunities of VR for rehabilitation can be applied to the field of cognitive stimulation, and in general, to all those interventions dedicated to MCI and early-dementia patients [104]. Several studies have ascertained the feasibility of this approach, even though the involved population can be considered somehow critical. Older adults are indeed less prone to accept new technologies: lack of familiarity, mistrust, privacy concerns, difficulties in operating the equipment, and comfort-related complaints are some of the motivations underlying this fact. Evidence remains mixed, with studies indicating that patients had high sense of control and enjoyment while interacting with the VE, and few others reporting low acceptance [111, 11].

An analysis of the state-of-the-art highlighted that different cognitive domains, and diverse risk factors, have been targeted with VR-based application. Among the others, the first category of applications targets memory impairments. In [112], a VR system proposing path recall has been shown to improve memory functions, after 6 months of training. Man et al. [113] investigated the effectiveness of a VR-based memory training for older adults with questionable dementia; the VE foresaw the visualization of a living environment, and users were asked to move around and remember objects present in such locales. The group that trained in VR had better outcomes, with respect to the control group, who trained with a therapist using colour-print images. Other studies did not focus on the effectiveness of the intervention, but showed the potentialities of VR-based training on episodic memory [114], and assessed treatment acceptance [115], also when using HMDs [116].

Dealing with visuo-spatial and orientation abilities, a VR-based application in which people at the early stages of AD had to navigate within a virtual building is presented in [117]. Results showed that older adults were able (with training) to find the correct route in the VE, which was displayed via a HMD; benefits in real-life and enjoyment were also reported. In [118], a virtual museum including activities aimed at training basic and executive functions, and encouraging a light physical exercise is presented. The treatment resulted in positive outcomes for amnestic MCI patients undergoing the training for 5 months, though their degree of improvement was variable from subject to subject.

Different types of VR system allow for the simulation and training of ADLs, with the aim of improving a person's ability to live independently and trying to facilitate the transfer of the acquired capabilities to real world-scenarios. A single-case study investigating whether it is possible to increase autonomy in cooking activities has been conducted with positive results, both in terms of transfer to real life, and of retention of the acquired capabilities over time.

Many other VR applications were focused on the assessment of cognitive abilities, rather than on providing cognitive interventions [119], or targeted different populations with different cognitive complaints (e.g., stroke patients [120] or subjects with traumatic brain injury [121]).

In addition to cognitive training, few studies have been exploited VR to support physical activity in older adults with MCI or mild dementia. Some of them tackled balance and gait issues, and fear of falling. In [122], a prospective randomized pilot study was conducted in an assisted living facility enrolling 22 mild-AD volunteers. At the end of 8 weeks, the group who trained with VR and the Wii Balance Board, and the walking group showed the same improvements in all balance and gait-related assessed variables. The same positive results were obtained also on mild AD individuals living at home, with the supervision of a caregiver [123].

In [124], instead, authors evaluated the feasibility of a PE-based intervention and its effects on cognitive outcomes; 10 people with dementia were invited by a trainer to exercise their upper limbs by blowing up blue bubbles with "EyeToy" application. Results were positive both in terms of technology acceptance, with people with more technical attitude appreciating the game and perceiving it as a tool to improve their health-status, and of cognitive outcomes, with a significant increase of MMSE score after the trial. In [125], researchers showed that cybercycling (i.e., cycling while looking at a virtual path) resulted in better cognitive function than traditional exercisers requiring the same effort, suggesting that simultaneous cognitive and (light) physical exercise had greater potential for preventing cognitive decline.

1.5 Aims

Within the above-presented context, the general aim of this PhD work was investigating the potentialities of VR-based technologies as instruments to counteract cognitive decline in older adults. More in details, the two main objectives pursued while working on this thesis have been:

Aim 1: The assessment of the impact, in terms of clinical outcomes and acceptance, of a 2D VR-based system providing PE and CT to older adults with MCI or subjective cognitive decline (SCD).

Aim 2: The evaluation of the feasibility and of the acceptance of immersive VR interventions for the administration of both cycling-based PE (**a**), and CT (**b**).

The distinction in two *Aims* has been made because 2D technologies were considered already known in their potentialities and limitations, and "safe" enough as previous experiences (§1.4.2) confirmed. The novelty of such investigation lies in the combination of VR-based PE, administered using a cycle-ergometer, and CT.

For what concerns immersive VR technologies, instead, the approach was different, because their employment for older adults' rehabilitation is still an open field. Surely, the recent advent of improved devices has paved the way to their application, but scientific contributions on the topic are sparse [126]. Consequently, it was fundamental to focus first on these technologies' acceptance, by performing studies aimed at evaluating their usability, and the SoP and the side-effects that they may convey (§1.4.1).

1.6 Thesis outline

This thesis presents the work I did from 2015 to 2019 at the Institute of Intelligent Industrial Systems and Technologies for Advanced Manufacturing (STIIMA) of the Italian National Research Council (CNR) and involved the design, the development and the validation of VR-based applications dedicated to the physical and/or to cognitive training of older adults with cognitive impairments.

To achieve the objectives mentioned in §1.5, the whole work has been developed following a pathway that goes from the design and development of a "classical" 2D VE, to the assessment of the feasibility of interventions exploiting totally immersive VR devices. Such a path resulted in the development of the 3 main systems that are presented in Figure 1.3. The Figure also shows their classification in terms of provided intervention (either only cognitive or both cognitive and physical), and level of immersion.

To better clarify how the evolution toward more immersive and more complex scenarios occurred, Figure 1.4 represents the timeline of all the sub-activities that have been performed to complete the work reported in this thesis, in the form of a Gantt diagram; such a timeline also shows the importance of preliminary tests that must be conducted prior of administering new interventions to vulnerable populations (i.e., §3.2 and §4.2), as older adults with cognitive impairments are.

The design, development, feasibility and clinical tests (whenever it had been possible to make them) of each system are addressed in each of the following Chapters.

More in details, Chapter 2 describes two studies addressing *Aim 1*, and evaluating the effects of PE and CT, administered via a 2D-projected screen, in older

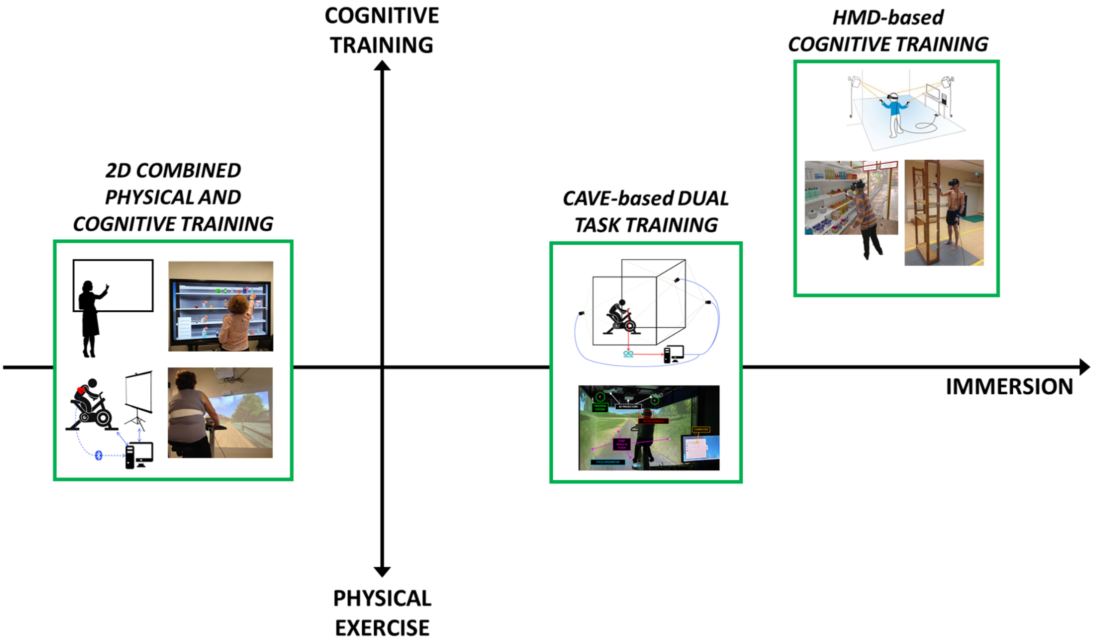


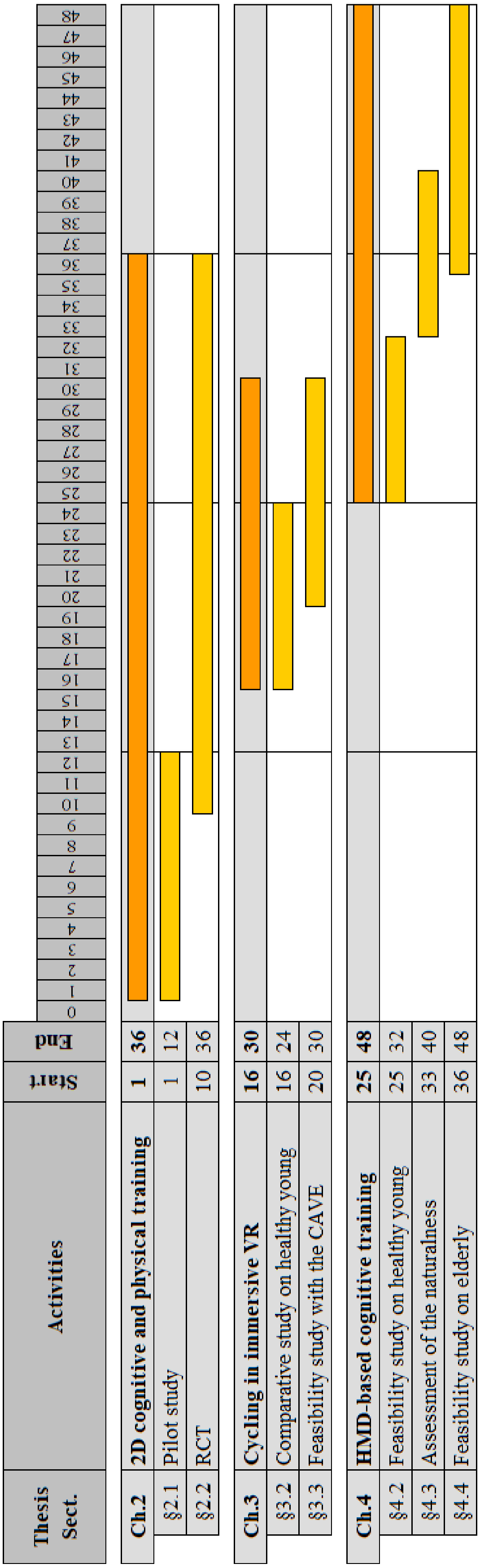
Figure 1.3: A schema depicting the three main systems discussed in this thesis and their classification in terms of intervention provided and immersion of the VR system.

adults with MCI. The Chapter describes the scenarios we developed (a park for the PE, road-crossings and a supermarket for CT) and the results obtained in two randomized clinical trials (RCTs) foreseeing of their use.

Chapter 3 investigates the possibility of making the PE (i.e. the park) scenario more immersive, using a phone-based HMD first, and a CAVE system after (*Aim 2a*). As more immersion is sometimes related to the occurrence of side-effects, particular attention has been paid on the evaluation of cyber-sickness before performing a trial enrolling a vulnerable population. This Chapter also describes what *dual-task* is, and how it was integrated in the CAVE-based system to enhance the effectiveness of cyber-cycling training.

Chapter 4 reports the results of two feasibility studies performed on healthy young subjects, and on older adults with MCI, using the Virtual Supermarket, i.e., a HMD-based version of the supermarket scenario described in Chapter 2 (*Aim 2b*). The Chapter also proposes a study assessing whether the movements performed in immersive VR occur in a natural way. The aim of this latter investigation was to determine if the interaction with immersive HMD technologies requires a high-cognitive workload, and thus if it distracts individuals with cognitive impairments from the main (i.e., cognitive) task they should perform during the training.

Finally, Chapter 5 draws the conclusions of all the work, and outlines future works and perspectives.



Thesis Sect.	Activities	Start	End
Ch.2	**2D cognitive and physical training**	**1**	**36**
§2.1	Pilot study	1	12
§2.2	RCT	10	36
Ch.3	**Cycling in immersive VR**	**16**	**30**
§3.2	Comparative study on healthy young	16	24
§3.3	Feasibility study with the CAVE	20	30
Ch.4	**HMD-based cognitive training**	**25**	**48**
§4.2	Feasibility study on healthy young	25	32
§4.3	Assessment of the naturalness	33	40
§4.4	Feasibility study on elderly	36	48

Figure 1.4: A Gantt diagram showing the activities timeline.

Chapter 2

2D Virtual Reality-based combined Physical and Cognitive Training for MCI Patients

2.1 Introduction

This Chapter presents two different studies addressing *Aim 1* (§1.5), i.e., the assessment of the impact of a multi-domain intervention, administered using 2D VR technologies, in samples of older adults with MCI. Both studies aimed at informing about the feasibility of the proposed interventions, and provided preliminary estimates of treatment clinical effects.

In details, the first study (§2.2) was a pilot trial that evaluated the impact of the combination of PE and CT (vs. no intervention) according to different criteria: (1) cognitive outcomes, (2) physiological outcomes, and (3) treatment feasibility and acceptance.

The second study (§2.3) foresaw the enrolment of a larger sample, and addressed the potential effectiveness of CT and PE in counteracting the progression of cognitive decline, either in conjunction, or when administered separately. Treatment acceptance, and participants' satisfaction were evaluated also in this case. User's acceptance of technology is indeed a critical success factor for the adoption of the proposed innovative means [107, 127]. It is thus fundamental to give it the proper attention, in particular when dealing with a vulnerable population, as older adults with MCI are.

2.2 Goji: a pilot study on older adults with MCI

This section presents the results of a research project found by the Italian Ministry of Health [1] that CNR-STIIMA carried out together with Fondazione Santa Lucia (Rome, Italy). Results of this work have been previously described in [20, 21, 22].

---

[1] Progetto di Interesse Invecchiamento CNR: Goji – an Advanced Virtual Environment Supporting Training of Physical and Cognitive Activities to Prevent Dementia Occurrence in Elderly with Minor Cognitive Disorders

2.2.1 Methods

The training system

Three different scenarios have been designed to support older adults while performing PE and CT. The design of the intervention, and of the required VR applications, were carried out through a focus group. Such a focus group was composed of two neuropsychologists working with MCI and AD patients, two biomedical engineers, a psychologist, a designer, a humanist with yearly expertise in the study of human factors.

The final decision was of simulating 3 activities taken from daily living. The 3 designed scenarios thus foresaw: (1) riding a bike in a park while performing the PE, (2) crossing streets avoiding moving cars, and (3) going shopping in a supermarket. The last two scenarios implied the use of cognitive functions, and, in particular, of visuo-spatial abilities, which we believed that could be easily targeted with the use of VR.

The hardware devices composing the training system were a cycle-ergometer (Cosmed Euro-Bike 320), a smart garment, aimed at measuring the hearth rate in real time (Wearable Wellness System, Smartex), a finger touch projector (EB-1430WI, Epson) and a PlayStation controller anchored on the cycle-ergometer handlebars. All the VEs (further descripted in the following paragraphs) were developed using Unity.

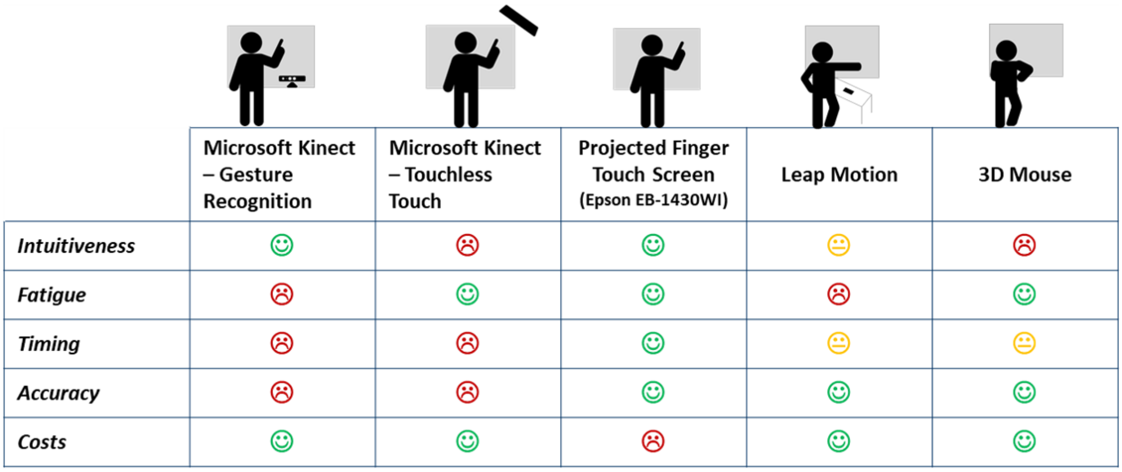
The choice of the hardware equipment The rationale behind the choice of a cycle-ergometer lied first in its higher safety; with respect to the treadmill, i.e., the other device used in PE studies allowing for the modification of the workload, the use of a cycle-ergometer is associated with a lower risk of injury, especially in case of older and frail users [128]. Moreover, previous studies have shown that performing stationary cycle exercises could improve balance, weight shifts and gait, and general lower limb functioning, thus significantly reducing the risk of falling [129, 130, 131].

Dealing with the projected screen, its choice was the result of a preliminary qualitative test conducted on healthy volunteers [22]. Given the premise that having a large display depicting the supermarket shelf would have been more realistic, 5 healthy young adults were asked to try different interaction technologies, and express their judgment in terms of intuitiveness, timing, fatigue, and accuracy. Costs were also considered in the analysis.

For all participants, the task was to pick 5 products displayed on the projected shelf, as indicated in the shopping list (also shown in the corner of the screen).

The technologies included in the analysis were the following: Microsoft Kinect v1, used both with the *gesture recognition* [132], and the *touchless touch* [133] protocols, a projected finger-touch screen, a Leap Motion sensor [134], and a 3D mouse. The results of such a test are summarized in Figure 2.1.

In spite of higher costs, the finger-touch projector emerged as the only feasible solution. Devices causing fatigue, and not able to guarantee an appropriate response both in terms of time and accuracy had to be excluded because our target users were expected to be less familiar with technologies, and more prone to



	Microsoft Kinect – Gesture Recognition	Microsoft Kinect – Touchless Touch	Projected Finger Touch Screen (Epson EB-1430WI)	Leap Motion	3D Mouse
Intuitiveness	☺	☹	☺	😐	☹
Fatigue	☹	☺	☺	☹	☺
Timing	☹	☹	☺	😐	😐
Accuracy	☹	☹	☺	☺	☺
Costs	☺	☺	☹	☺	☺

Figure 2.1: Results of the usability test [22].

experience physical and cognitive fatigue [135].

The park scenario. While performing the PE, the user had to ride the cycle-ergometer, facing the projected screen and wearing the smart garment. Though some studies have reported the training of some cognitive abilities while just paying attention to a virtual scene [125], the dual aim of this VE was increasing the user's engagement, while providing him/her with information on the exercise, i.e., speed, covered distance, revolutions per minute, time elapsed and heart rate.

The VE represented a trail in the park that the user travelled in first person and that flew according to cycling velocity (Figure 2.2). To pass this piece of information from the bike to the computer running the Goji application, we created an ad-hoc protocol exploiting the SDK provided by the cycle-ergometer manufacturer in order to allow the exchange of data between the bike and the VR environment.

The path to follow within the virtual scene was predefined, and foresaw turns and slopes to increase realism. However, only slight bends were present along the whole path to try to avoid the occurrence of cyber-sickness due to the expectation of lateral accelerations [136]. The users could not deviate from the predefined path. Such a path to follow was created by a series of nodes that were then dynamically interpolated at run-time using quaternion spherical linear interpolation (slerp).

To further increase the realism of the projected scene, and thus the SoP, few animations were added: trees' leaves and grass moved as the wind blew, and wild animals appeared in the sky or on the trail sides. We also added auditory elements, i.e., a rustle simulating the cycling on an untarmacked road, and, sometimes, birds' sounds.

To control the participant effort during the PE, the cycle-ergometer workload was adjusted according to the heart rate (HR) of the user, measured through the garment. A digital controller integrated in the VE allowed keeping the HR of the participant in a specific target range, so that he/she could train with constant effort, and avoid over-exhaustion. The server allowing for the data exchange between the garment and the computer running Goji application was developed by CNR-IMM.

Figure 2.2: A screenshot of the park scenario.

The range for acceptable heart rates (HR_t, target heart rate) for each participant has been computed using the Equation 2.1 and Equation 2.2. Such a range was set to obtain a training effect throughout time, in agreement with ACSM/AHA guidelines [137].

$$HR_{t\_max} = 0.75(220 - age) \tag{2.1}$$

$$HR_{t\_min} = 0.55(220 - age) \tag{2.2}$$

For the algorithm to work correctly, participants had to keep their cycling velocity between $RPM = 50$ and $RPM = 70$. To indicate the correct ranges, we used a color code showing RPM in green when within the range, and red when outside. While being in the range, if a participant's HR was not between the two values computed with the formulas, the workload was either increased (+5 W, if $HR_{current} \leq HR_{t}\_min$) until a maximum of 75 W, or decreased (-5 W, if $HR_{current} \geq HR_{t}\_max$) until a minimum of 30 W.

Before the training session started, the operator supervising the entire session had the chance to set the duration of the physical training, choosing between 15 or 20 minutes. After that the selected duration has elapsed, the application loaded automatically the road crossing scenario.

The road crossing scenario. After the completion of the physical training, the park turned into an urban scenario. In this second scene, the user still had to ride the cycle-ergometer, but the task required only cognitive abilities, i.e., the workload was set to 0; the participant, in fact, had to face the crossing of five traffic-congested and non-regulated crosswalks (Figure 2.3), thus training his/her visuo-spatial and attentional abilities.

In details, the sub-tasks to accomplish in this scenario were: (1) cycling to reach the border of the sidewalk, (2) braking when being near it, (3) checking on both sides if there are cars moving closer and, if the way was clear, (4) restarting cycling to reach the following cross. To reach the crosswalk and proceed forward

Figure 2.3: The road crossing scenario.

the user has to pedal on the cycle-ergometer, while to brake he/she must stop cycling and press the X button on the PlayStation controller. The brake was only digital because it was impossible to access the wheel compartment and create a physical brake able to stop the inertial rotation of the wheel; moreover, this would have compromised the ergometer medical CE certification.

The controller (its right joystick, in particular) was also used to turn the user's point-of-view. In this way, when the sidewalk border was reached, the participant could check on both sides if there were cars moving closer.

Cars were generated by the application at fixed interval time, but their velocities were randomly set. To signal the occurrence of an accident, the sound of a car braking abruptly and of glass breaking were reproduced. After the accident, the participant was taken to a safe position, and could proceed forward as the cars disappeared from that cross. The number of total accidents and the total time needed to reach the supermarket were saved on a XML file, containing also the data regarding the shopping scenario.

The shopping scenario. The last cognitive task took place in a virtual supermarket, in which the participant was required to buy some grocery items as indicated on a shopping list. Such a list was either displayed on the left side of the projected screen, or presented to the participant and then hidden with the aim of targeting also memory functions.

The tasks proposed by this scenario had to be performed while standing in front of the projected screen. In particular the users had first to tap on the correct aisle (*aisle task*), and then on the product to buy, choosing it among other items on the shelves (*shelf task*). Both the aisle and the products were 2D buttons. In between the aisle and the shelf task, an animation in a 3D scene brought the user from viewing the aisles to standing in front of the shelves.

Both tasks had 5 different levels of increasing difficulty. They were independent so the operator could set, at the beginning of each virtual shopping session, the level of difficulty for both the aisle and the shelf tasks. For the aisle scenario, the parameters used to implement the increasing difficulty are shown in Table 2.1.

As shown in the Table, starting from the third level, a word that looks similar either semantically or orthographically to the name of the target item was placed in one of the non-target aisles, with the aim of misleading the participant if he/she did not pay enough attention. An example of orthographically similar word is the misleading word *menta* [mint] used instead of the target item's name *mela* [apple];

Table 2.1: Parameters to set the difficulty of the aisle task for each level [21]

# level	# aisles	#elements/aisle	#misleading word
1	2	1	no
2	2	3	no
3	2	3	yes
4	3	4	yes
5	4	4	yes

.

an example of a semantically similar word is *Pepsi* used instead of *Coca Cola*.

For the implementation of the 5 levels of difficulty in the shelf task, the algorithm foresaw setting different parameters:

- *distractors number (#)*; this value defined how many distractors appeared on the shelf in addition to the target item.
- *discount (yes/no)* determined if discounted items were present on the shelves or not. Each item had a correspondent discounted version.
- *small products (yes/no)* determined if a small version of the target product (and also other items) were present on the shelves or not. Only some items had a classification by dimension (i.e. bottles and jars); others had not (fruits, clothes, etc.).
- *variability (%)*: this value identified the percentage of the items placed on the shelves by picking completely random object from the whole database of products (i.e., products that were not *similar*).
- *similarity (#)*: the percentage of items that were not random; the highest complexity is thus obtained having the majority of items (i.e., low *variability*) very similar among them (i.e., high *similarity*). Similarity was computed by determining if each item the belonged (value 1) or not (value 0) to a certain category, and then comparing the resulting binary numbers using an algorithm based on the Hamming distance. A similarity value of 1 meant that the two items have one category in common (i.e., they were both fruits or bottled items); a similarity value equal to 2 meant having two categories in common (i.e. products were fruit and round-shaped).
- *Shelf area (center/border/all)* determined the area of the shelf in which the target item was placed; central positions were considered easier to identify than lateral ones.

The detailed parameters used to implement the 5 levels of the shelf task are shown in Table 2.2. An example of two possible shelves is reported in Figure 2.4.

For both the aisle and the shelf tasks, if the participant committed an error or did not interact with the VE for more than 20 seconds, the application gave an hint to help him/her to make the right choice. In the aisle scenario, the first hint was the audio replay of the name of the item to search for, coupled with the blurring of a wrong aisle sign. This type of hint was repeated for each error (or hesitation),

Table 2.2: Parameters to set the difficulty of the shelf task for each level [21].

# level	# distractors	# discount	small items	variability	similarity	shelf area
1	9	no	no	0.6	1	center
2	9	no	no	0.6	1	border
3	9	no	no	0.5	1	all
4	15	no	yes	0.3	1	all
5	25	yes	yes	0.2	2	all

until just one aisle remained active; the last aisle was also highlighted in yellow. For the shelf task, hints were given as follow: (1) the name of item to collect was repeated to the user; (2) the target item on the shelf blinked 3 times; (3) the target item was highlighted in yellow. For each session of the virtual shopping, the current level of difficulty, the total completion time, the target items and possible errors or hesitations were stored on a XML file.

Participants

Older adults with subjective complaints in one or more cognition domains were evaluated for the enrolment in the study. Inclusion criteria were: one or more scores in neuropsychological tests determining compromised visuo-spatial abilities; mild to moderate cognitive impairment, according to MMSE criteria [61]. Exclusion criteria were: severe cognitive and/or functional impairment; cardiovascular pathologies preventing the performance of the physical training in safety conditions; acute pain of lower back or extremities; peripheral neuropathy; rheumatic and orthopaedic diseases; inability to provide informed consent.

Around 200 individuals were examined for eligibility; each person fitting the inclusion criteria underwent an assessment of risk factors (cardiovascular pathologies and others comorbidities) and an effort electrocardiogram to exclude risk of adverse events during the PE, due to preexisting pathologies. All decisions about eligibility were made before the block-randomization performed on the basis of cognitive decline.

Ten individuals (4M/6F) were randomized to an intervention (EXP) group ($n = 5$, 2M/3F) and a control (CTR) group ($n = 5$, 2M/3F) that did not receive any treatment. Three older adults with MCI and 2 with mild dementia (i.e., an impairment of cognition that significantly affects instrumental skills of daily living), §1.2) were present in each group.

Individuals participating in the study provided informed written consent. The study was performed in agreement to the Declaration of Helsinki, and was approved by the IRCCS Fondazione Santa Lucia medical ethics committee (Rome, Italy).

Experimental protocol

EXP group performed the VR-based PE and CT for 6 weeks, 3 sessions/week. The training session followed the same procedure for each participant; it lasted approximately 40-45 min: 15 (for the first three weeks) or 20 minutes of cycling, about

Figure 2.4: An example of the shelves proposed to the user in Level 1 (above) and in Level 5 (below). In the latter case, the target is a bottle of water, thus due to the similarity implementation, the shelf is filled with many bottled items.

5 minutes to cross 5 crossroads, and 20 minutes of shopping at the supermarket. Before to the beginning of the program, participants performed a training session to get to know the technologies and the interaction methods in the three VEs.

As mentioned, the cycling exercise intensity was set to keep the participant heart rate around 65-70% of the maximal frequency, which was calculated on the basis of the age-predicted value (ref. Equation 2.1, and Equation 2.2). During the training session, the participants were instructed to keep the cycling velocity between 50 and 70 RPM. All the participants started from Level #1 for both the aisle and the shelf tasks; levels of difficulty were then increased accordingly between the two tasks; in particular, the increase in difficulty occurred when the participant completed 3 subsequent shopping sessions receiving at most one hint, i.e., making 1 error or waiting once for more than 20 seconds. When the participant completed Level #5 three times without receiving any hint, the CT was re-started from the first level in which he/she failed for the first time. In this second round, the shopping list was shown only at the beginning of the supermarket exercise, and then hidden.

Outcomes

The study comprised the assessment of different domains by evaluating different cognitive, physiological and subjective outcomes. The evaluation of psychometric variables was performed by psychologists of IRCCS Fondazione Santa Lucia, whereas physiological response (i.e., oxidative stress) was evaluated by researchers of CNR-IBFM.

Cognitive assessment. Cognitive abilities of each participant were assessed pre- and post-intervention by the same person (a neuropsychologist), in order to avoid inter-rater variability. The entire neuropsychological battery was composed by an initial screening test for cognitive abilities (MMSE), and a battery of specific tests for each subdomain:

- *Episodic verbal memory*: Immediate Recall and Delayed of Rey Auditory (RAVLT_I and RAVLT_D) [138]; in the RAVLT_I, a list of 15 words was presented to the participant, and he/she had to memorize and repeat it. The task was repeated 5 times and the number of correct words was recorded. The RAVLT_D required the recall of the same words after 15 minutes.

- *Visuo-spatial functions*: Rey-Osterrieth Complex Figure Test (ROCFT) [139], Attentional Matrices Test (AM) [140], and Trail Making Test A (TMT-A, [141]); ROCFT was used to assess constructive apraxia, and required to copy a geometric drawing composed of 18 sub-elements: the correct execution and collocation of the sub-elements were both criteria for scoring. AM investigated selective visuo-spatial attention by requiring the deletion, quickly as possible, of a few target numbers among distractors. In the TMT-A, the participant had to perform a visuo-spatial search in order to connect a series of 25 numbers presented in a random distribution on a sheet (with a single pen stroke).

- *Executive functions*: Frontal Assessment Battery (FAB) [142]; it evaluated categorization, programming, inhibition, sensitivity to interference and environmental autonomy with a series of sub-tests; the Trail Making Test B (TMT-B, [141]) required matching alternatively numerical and alphabetical stimuli, making use of visual-motor coordination, set-shifting and mental flexibility.

- *Lexical finding skills*: Verbal Fluency test (VF) [143]; in this test, the participant had to name as many words belonging to a given semantic category in a minute as he/she could.

- *Autonomy in daily living activities*: Functional Activity Questionnaire (FAQ) [144]; the participant had to make a choice among levels of performance of 10 common daily activities.

The participants objective performances, i.e., the committed errors, the time required to perform the tasks in the 3 environments, the number of accidents, have been analysed and discussed in [21]. Due to the small sample, it was impossible to correlate these performances to the cognitive status of the participants, thus this information has been omitted from this thesis.

Oxidative stress. Reactive Oxygen Species (ROS) production rate was determined adopting a recently developed mini-invasive method exploiting electron paramagnetic resonance [145]. Moreover, a capillary blood sample (10 μL) was used in order to assess blood reducing capacity (Total Antioxidant Capacity, TAC) by mean of a commercial potentiostat electrochemical analyser equipped with a redox sensor. Finally, urine samples were collected and 8-hydroxy-2-deoxy guanosine (8-OH-dG) and 8-isoprostane (8-iso) were assessed by commercially enzyme immunoassay kit as markers of oxidative DNA damage and of lipid peroxidation, respectively. Creatinine concentration was assessed by commercial enzymatic assay kit; in this way, the standardization of urinary parameters on its excreted quantity was made possible, even in absence of 24h of urine-collection. Further details can be found in [20].

Treatment acceptance. Acceptance is defined as "the question of whether a system is good enough to satisfy all the needs and requirements of potential stakeholders" [146]. In the clinical scenario, acceptance is often intertwined with the concept of treatment acceptability, i.e., the degree to which non-professionals stakeholders found an intervention to be fair, reasonable, intrusive, and consistent with treatment expectations [127, 147]. It refers to the general subjective evaluation of the applied procedures, and it represents an important dimension of any treatment, beside its efficacy and effectiveness [127].

According to the most common framework used for the evaluation of acceptance, i.e., the Technology Acceptance Model (TAM) [148], when a user is presented with a new technology, two main factors influence his/her behavioural intention to use such a technology:

- *Perceived usefulness* (PU), defined as the extent to which a person believes that using a system will enhance his/her [job] performance;
- *Perceived ease of use* (PEOU), i.e., the degree to which the person believes that using the system will be free from any effort.

In this study, the assessment of PU and PEOU was made interviewing the participants in EXP group at the end of the intervention. Both closed and open questions were used to explore participants' subjective perceptions and feelings toward the whole intervention. More specifically, closed questions aimed at investigating the subjective perception of the obtained results, the physical and cognitive effort required to complete the tasks, and the potential sources of difficulties in the training program. Open questions were administered in the form of a semi-structured interview, in order to indulge each participant's attitude. Before the questionnaire administration, the purpose of the interview and the assurance that all opinions were valuable and confidential were disclosed to the participants to encourage their honest feedback.

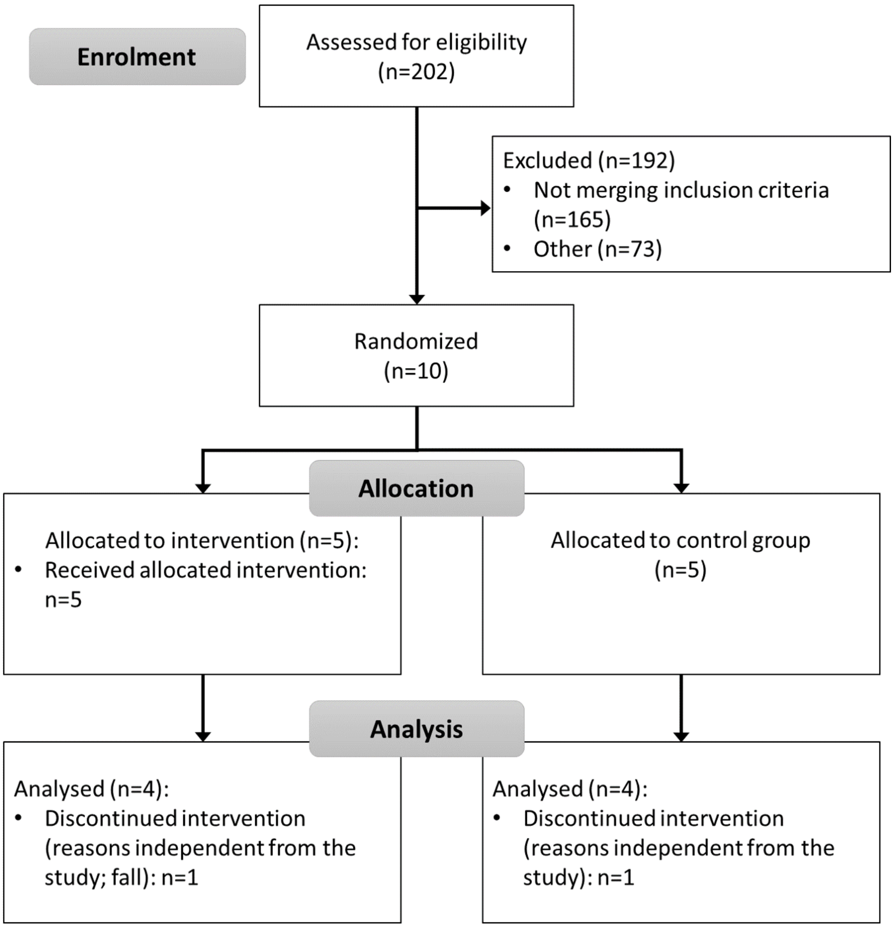


Figure 2.5: CONSORT flow chart for Goji pilot study.

Statistical analysis

GraphPad Prism package (GraphPad Prism 7, GraphPad Software Inc., San Diego, CA, USA) was used for all the statistical analyses. We used non-parametric tests, as the sample size was small. The effect of the combined physical and cognitive VR-based training on cognitive and physiological outcomes between the CTR and the EXP groups were compared using Mann Whitney U-test for independent samples.

Non-parametric Wilcoxon matched-pair signed-rank test was used to compare the change in the scale scores from pre- to post-treatment within the EXP and CTR groups.

For cognitive and functional scores Intention-To-Treat and Last-Observation-Carried-Forward (ITT-LOCF) analyses were used; $p<0.05$ was considered as statistically significant.

Regarding evaluation of acceptance, we transcribed the answers given by EXP group participants during the interview. Descriptive statistics (median and quartiles) were used to summarize the information from closed questions. A thematic analysis was then performed to synthesize participants' subjective perceptions, starting from the information gathered through open questions and free comments [149].

2.2.2 Results

The socio-demographic and clinical characteristics of the sample were comparable to those of participants taking part in other studies targeting cognitive decline (age= 73.3 ± 5.6 years; schooling= 7.6 ± 4.4 years; MMSE= 23.0 ± 3.4, range: 19.3-27.9).

A drop out of one participant both in EXP and CTR group was recorded for reasons independent from the study. Adherence in the EXP group was considered *good* (more than 75%, reaching 89% in two cases) for 3 participants, and *partial* (more than 50% and less than 75%) for 1 (67%). The study CONSORT flow diagram is presented in Figure 2.5.

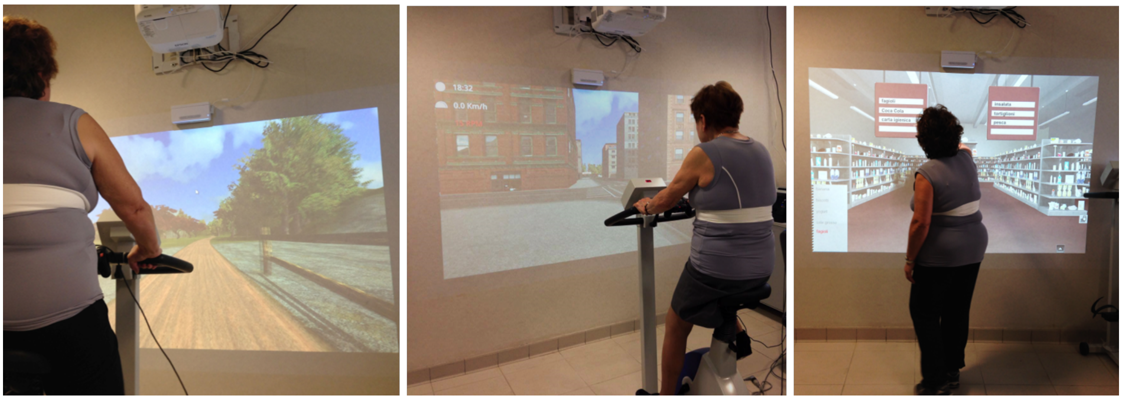


Figure 2.6: participants in EXP group while performing the PE, the road-crossing and the shopping [21].

Cognitive response

The sample was composed of individuals with a deterioration of global cognition that was classifiable as *mild* or *mild-to-moderate*.

All participants had a slight impairment in the majority of the assessed cognitive sub-domains at the baseline: visuo-spatial and executive functions were moderately impaired in both groups. Before the study, no significant differences were found between EXP and CTR group in terms of demographics (age 72.00 ± 5.15 for EXP, 76.60 ± 6.43 for CTR) and of neuropsychological assessment (Table 2.3).

After the study, the majority of the variables either improved in EXP, or declined less, as shown in Figure 2.7. However, none of these differences reached significance in statistical tests comparing time-points pre- and post-interventions, and EXP and CTR.

Both groups showed a decline in ADL post-intervention. For the EXP, such decline was more pronounced (FAQ score: 1.80±4.60) with respect to CTR (1.00± 1.73). However, this result was heavily influenced by the evaluation of a single participant in the EXP group who markedly worsened post intervention. Such a

Domain	Test	EXP	CTR
Cognition	MMSE	23.1 ± 3.6 (19.3 - 27.4)	23.0 ± 3.5 (20 - 27.9)
Memory	RAVLTD$_I$	31.4 ± 10.8 (22.5 - 50.1)	36.4 ± 13.8 (25.9 - 60.1)
	RAVLTD$_D$	5.4 ± 3.8 (1.9 - 10.8)	5.1 ± 4.3 (2.6 - 12.8)
Visuo-spatial abilities	ROCFT	23.5 ± 12 (2.8 - 32)	21.2 ± 10.3 (7.0 - 34.5)
	CDT	1.5 ± 0.7 (1.0 - 2.0)	1.5 ± 1.3 (0 - 3.0)
Attention	AM	45.1 ± 6.2 (37.8 - 51.3)	37.9 ± 16.1 (14.5 - 49)
	TMT-A	43.8 ± 16.8 (26 - 70)	57,5 ± 29,1 (34 - 96)
Executive functions	TMT-B	175.3 ± 23.1 (150 - 206)	249 ± 58 (208 - 290)
	FAB	13.1 ± 2.7 (10.5 - 17.2)	12 ± 4.7 (6.3 - 15.9)
Language	VF	33.8 ± 8.2 (24.5 - 43.5)	20.7 ± 3.1 (16.4 - 23.9)
ADL	FAQ	8.8 ± 7,2 (2 - 20)	10.6 ± 8.1 (0 - 21)

MMSE, Mini-Mental State Examination;
$RAVLTD_I$ and $RAVLTD_D$, immediate and delayed Rey's Auditory Verbal Learning Test;
ROCFT, Rey-Osterreith Complex Figure Test;
CDT, Clock Drawing Test;
AM, Attention Matrices;
TMT-A and TMT-B, Trial Making Test;
FAB, Frontal Assessment Battery;
VF, Verbal Fluency;
FAQ, Functional Assessment Questionnaire.

Table 2.3: Psychometric variables at the baseline, for both the experimental (EXP) and the control (CTR) groups.

participant lost 10 FAQ-points at the end of the program, due to a health issue independent from our program. The same person was re-evaluated 3 months after, and showed an improvement in the FAQ scores of 6 points, compared to the baseline assessment. Excluding that participant as an outlier, the EXP group obtained a slight improvement of 0.25 ± 0.5 in FAQ.

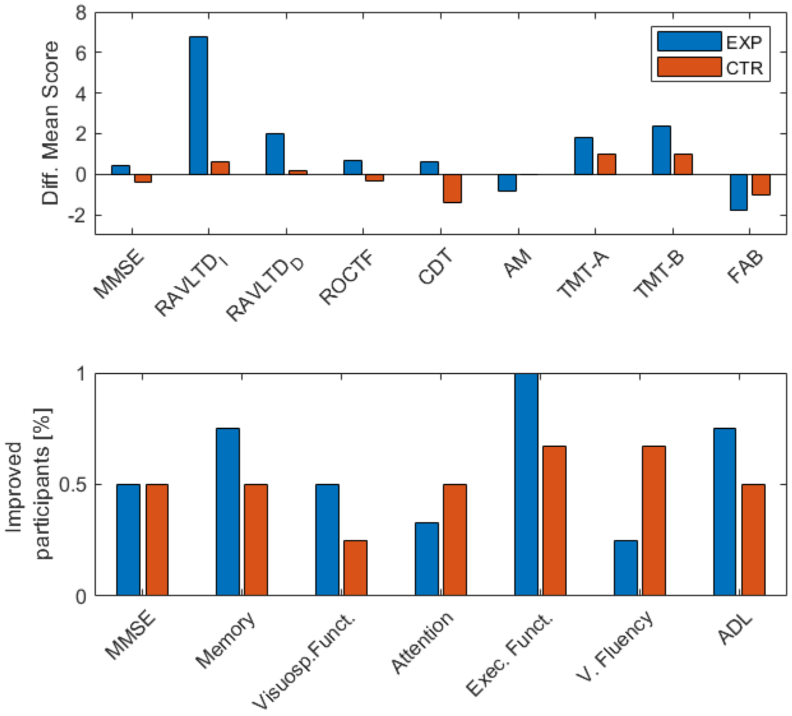


Figure 2.7: **Top**: *Mean change scores between the baseline and the end of experimental phase (calculated as POST - PRE scores. For TMT tests and FAQ, the changes were calculated as PRE - POST as lower scores correspond to better performances).* **Bottom**: *percentage of participants improved or stable at POST.*

Oxidative stress

At the baseline, no difference between EXP and CTR group were present in terms of physiological outcomes.

An increase (+8%) in ROS production was recorded post intervention in the CTR group; instead, in the EXP group ROS production decreased (−4%). Comparing the two groups at the end of the training, ROS production rate resulted statistically ($p < 0.05$) lower in EXP with respect to CTR (2.03 ± 0.23 vs. $2.58 \pm 0.33 \mu mol \cdot min^{-1}$, respectively) (Figure 2.8, A).

A similar trend was observed in 8-iso concentration: after training, 8-iso was significantly lower in the EXP group (379.30 ± 54.82 vs 495.20 ± 44.56 $pg \cdot mg^{-1}$ creatinine, and $p < 0.05$) (Figure 2.8, C). 8-OH-dG concentration increased significantly ($p < 0.01$) in the CTR group with respect to its baseline values (5.80 ± 1.40 vs 10.32 ± 1.28 $ng \cdot mg^{-1}$ creatinine). On the contrary, it significantly decreased

($p < 0.001$) in the EXP group post intervention, also compared to CTR (5.14 ± 1.05 vs. 10.32 ± 1.28 $ng \cdot mg^{-1}$ creatinine, respectively) (Figure 2.8, D).

No differences were found comparing the relative values post-treatment in the CTR and EXP groups in the TAC level (CTR: 149.75 ± 30.44 vs. post: 147.25 ± 22.20 nW; EXP: 152.25 ± 25.55 vs. post: 163.50 ± 42.35 nW) (Figure 2.8, B).

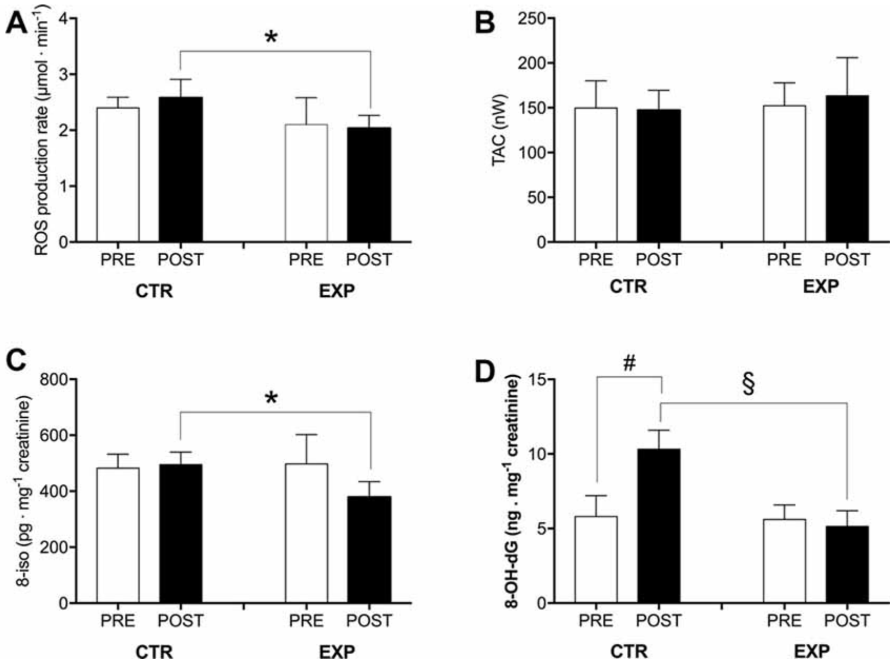


Figure 2.8: The influence of physical and cognitive training on oxidative stress is shown by the histogram plots of: A) ROS production rate (µmol· min^{-1}), B) total antioxidant capacity (TAC; nW), C) lipids peroxidation (8-iso; pg · mg^{-1} creatinine) and D) DNA damage (8-OH-dG; ng · mg^{-1} creatinine) obtained from capillary blood and urine samples in the control (CTR) and experimental (EXP) groups before and after 6 weeks. Results are expressed as mean ± SD. Statistically significant differences symbols: $*$: p<0.05, #: p<0.01, and §: p<0.001 [20].

Treatment acceptance

Answers to closed questions are presented in 2.4 according to the proposed Likert scale ranging from 0 (*not at all*) to 4 (*very much*). In general, participants reported a neutral position towards the training and a positive confirmation of their expectations consisting in an improvement or, at least, in the maintenance their current status.

Data gained from the interviews about PU and PEOU are presented according to the themes identified within the thematic analysis, i.e., expectations, acceptance, and continuation of training.

Expectations and general satisfaction. Participants had different expectations about the intervention; one reported to have accepted to enroll in the study just because he was encouraged by his daughter to do so. Others were expecting an improvement or, at least to maintain their current status.

Table 2.4: Results (median and q1-q3) of the questionnaire administered after the sixth week of training. The evaluation was based on a Likert scale whose elements indicated: not at all (0), not really (1), neutral (2), somewhat (3) and very much (4). Adapted from [21] and [20]

		Score
General satisfaction	Are you satisfied with this training?	2.0 (2.0-2.5)
	Did this training meet your expectations?	2.5 (2.0-3.0)
Park/ road crossing scenarios	Did you enjoy pedaling?	2.5 (2.0-3.0)
	Did you get tired?	1.0 (0.0-1.0)
	Was it difficult to keep the required velocity?	0.0 (0.0-0.0)
	Did you feel comfortable wearing the smart garment?	4.0 (4.0-4.0)
	Was it comfortable to use the joystick?	3.0 (1.8-4.0)
	Did you enjoy looking at the park while pedaling?	3.0 (2.8-3.0)
	Did you get sick?	0.5 (0.0-0.5)
	Did you get bored?	0.0 (0.0-0.0)
	Did you get anxious?	0.0 (0.0-0.0)
Supermarket scenario	Did you enjoy doing the shopping?	2.3 (1.8-3.0)
	Was it comfortable to interact with the touch screen?	2.0 (2.0-3.0)
	Was the shopping list clear?	3.0 (2.8-3.5)
	Was the guiding voice clear?	3.0 (2.9-3.5)
	Were the hints useful?	3.0 (3.0-3.0)
	Did you encounter specific complications?	1.5 (1.0-1.5)
	Did you get sick?	0.0 (0.0-0.0)
	Did you get bored?	0.0 (0.0-0.0)
	Did you get anxious?	0.0 (0.0-0.0)

All the participants reported having benefitted from the intervention (*"I was expecting worse results... I noticed that I do not forget things anymore. [...] I had an [car] accident last week, but I am not anxious; I am not anxious anymore, I will continue driving". / "The program encouraged me to do more, to remember more things and to exercise more"*). Thus, the usefulness (PU) of the system was positively rated. There were several aspects that all participants liked and that were not expected, probably indicating that closed questions were not detailed enough to capture participants' real feelings. The use of innovative technologies and of VR in particular was considered an added value to the treatment for all of them. None, except one, reported to get (a little) bored during the sessions; however, some criticisms were made with respect to the proposed tasks and their difficulty (*"I would have changed the task: I do the shopping in real life, too". / "I would have liked more difficult tasks as a roundabout with many entrances. [...] Give me something more difficult to do, create more unpredictable scenarios: they are more engaging!"*).

Acceptance. The intervention was judged acceptable and enjoyable by all the participants. The interaction with the technology was easy and intuitive, indicating a good system's PEOU. Only few side effects were reported (*"The bends were too sharp. Sometimes I felt dizzy." / "Once, I had headache at night. I usually do not suffer from it, so I remember it".*) None of the participants complained about the difficulty of the physical exercise, and 3 out of the 4 participants stated to be satisfied with their physical status at the end of sixth week of training (*"I like the training, I even lost one kilogram!"*). One participant complained that sometimes he was tired before the training and that he would have liked to exert less effort on those days. The setup was defined comfortable and changing clothes to wear the smart garment was not considered an issue.

Difficulties related to the cognitive tasks were mainly related to graphical issues or products' design, which did not match the packages the participants used to buy (*"I would like clearer products; I could not recognize the package of mozzarella." / "I suggest clearer labels on the items; I could not distinguish some of them. Green bottles represent sparkling and not still water in Rome!"*). Natural and sparkling water bottles were often confused because the colour code was not interpreted correctly. Also, participants explained that they felt uncomfortable in picking a non-discounted product (as written in the shopping list) when, on the shelf, there was the discounted version of the same item, because this was judged *"against common sense"*. With the exception of modifying some products to make them more easily recognizable and of introducing new scenarios, all the participants were satisfied and were happy to have attended the sessions, even in the case their adherence was demanding for the participants themselves or their care-givers ("*I had to drive 26 kilometers to reach the clinic!" / "I do not drive, my daughter drives me here . . . Sometimes she could not, this is why I missed some sessions".*).

Continuation of training. All the participants stated that they would continue with the training, if possible (*"If I were sure to get better and better, I would do it to prevent new symptoms!"*). One said that she had many other activities to do, so training was a duty that she must find the time for. All the participants were asked if they would like to continue the training at home. Three embraced this possibility as a great improvement, which could make them save a lot of time. One participant refused this possibility arguing that she needed a strict schedule to be sure to adhere to the treatment ("*I prefer to come to the hospital because at home I have other duties distracting me. I am not sure if I would do it.*")

2.2.3 Discussion

This work constituted a preliminary evaluation of the impact and of the acceptance of an innovative non-pharmacological intervention aimed at slowing down the occurrence of symptoms of dementia in older adults with MCI. The VR-based protocol that we proposed revealed able to promote a tendency towards improvement in individuals with mild and mild-to-moderate cognitive impairment.

The positive tendencies we found for cognitive outcomes were in agreement with recent reviews [87, 150] that showed an effect of computerized and VR-based training of cognitive abilities on MCI population. Probably, the lack of significance in statistical tests was due to the small sample size that amplified the heterogeneity within and between groups. Nonetheless, a weak effect of the program could be noticed on global cognition, on visuo-constructive and visuo-spatial abilities, and on executive functions.

In addition to the tendency highlighted by neuropsychological tests, we found significant differences in physiological variables. This was a relevant outcome, as much evidence is now emerging and confirming the role of OxS in the progress of AD-related dementia [151].

Concerning participants' acceptance of technology, qualitative data retrieved from questionnaires and participants' comments showed good appreciation. The use of the VR technologies appeared to have played a role in this judgement: in fact, study participants defined the system a nice "novelty", and an engaging and enjoyable way to approach cognitive training.

Feedback on the VEs and the general adherence were positive. The only participant showing adherence lower than 75% was the only one without a driving license; her absences were related to the unavailability of her driver.

Given this, we may conclude that the training program was well-accepted by all participants in the EXP; this hypothesis was confirmed by the fact that: (1) the majority of participants would continue with the program, also if administered at home, and (2) the only drop-out was due to the consequences of a fall not related with this study. This was a satisfying results, as acceptance is a key element for the successful employment of innovative technologies in therapy, especially when dealing with older populations, who may not be familiar with innovative means [135, 148].

Another important result we obtained within Goji was related to the fact that participants enrolled in the EXP group reported to *feel better*, and to be less anxious in their daily life, also when performing potentially risky activities (e.g., driving). This feeling, even if not measured by objective or psychometric tests, resulted in a perceivable improvement of the participants' QoL. In fact, it contributed to reduce the so-called excess disability [77], and may have enabled also higher self-efficacy [152].

Given our experience in this study, we may suggest, for future experiments, not to limit both the intervention and the assessment to a single domain. Rather, it would be interesting to include the evaluation of physiological parameters and of subjective perceptions [153, 154, 155], because they both contribute in estimating the effects of a newly designed intervention. With respect to the subjective outcomes, it is clear that the clinical effectiveness remains the objective to pursue, but also perceived improvements may play a key role: reducing the excess disability could, in fact, generate a positive loop, in which the more confident the individual feel, the more he/she could act to increase his/her social participation, his/her cognitive reserve and improve his/her general lifestyle, thus promoting the minimization of modifiable risk factors (§1.2.3).

Limitations

We acknowledged that this clinical trial had limitations, most of which could be attributed to the fact that it was a pilot study; the sample was small, and the treatment was limited in time. Moreover, we adopted a multi-domain approach that was not compared with other intervention(s) requiring an active involvement of participants, e.g., a matched cycle-ergometer training program, but without the VR component, or a social activity in group. Therefore, neither the impact of each single component (CT or PE), nor their complementary effect could be evaluated. Moreover, we did not perform a long-term follow-up, thus we could not inform about the retention of benefits or the progress of the pathology throughout time. Finally, we did not take into account the physical exercise that participants performed beside our intervention, during the training period.

On the other hand, the high adherence to the intervention, the supervised exercise program, and the use of validated outcomes from different domains could be interpreted as points of strength. In conclusion, the proposed technologies and the whole training program were considered worthy of further investigations (§2.3).

2.3 A multidomain intervention for dementia: assessing the role of physical and cognitive training in a randomized controlled trial

Starting from the observations made in the pilot study, and given its promising results, the system we developed in Goji was improved with the aim of conducting a second, and more structured feasibility study exploiting the same technologies of the pilot (§2.2.1). This second study was conducted within a *Young Researchers* project[2] that foresaw the involvement of my research group and of IRCCS Fondazione Santa Lucia.

The administration of VR-based physical and cognitive training was kept, but the sample characteristics, the typology of interventions (i.e. only physical, only cognitive, and both) and the training time were varied as described in the following paragraphs. The reasons for these changes are due to the need of lengthening of the training period, in order to make it more similar to other studies that revealed effective, and to the possibility of evaluating the synergic effects of PE and CT.

2.3.1 Methods

Training system improvements

The training system resembled the characteristics of the system presented in §2.2.1, though some modifications have been made both to hardware and software components (Figure 2.9).

First, the cycle-ergometer has been replaced with another one (Cosmed K100) allowing for the direct measurement of the heart rate, using either a proprietary (Ergoline) or a commercial (Polar) chest-band. This allowed for more participants

---

[2] Progetto di ricerca finalizzata – Bando Giovani Ricercatori: A comprehensive preventive program for dementia tailored on the neuropsychological profile of persons with Mild Cognitive Impairment: cognitive stimulation, physical intervention and healthy nutrition, a randomized controlled trial. Funded by the Italian Ministry of Health

Figure 2.9: Some of the improvements made to the setup used in the previous study.

to be included, because they could use the same sensor while only exchanging the band for hygienic reasons.

The PlayStation controller was replaced to improve the interaction within the cross-road scenario. More in details, to implement the braking function, we mounted a 3D-printed brake lever on the cycle-ergometer handlebars. It was equipped with a potentiometer measuring the pushing force. The braking action was thus made more natural. On the other hand, increasing the naturalness of the interaction was not possible for what concerned the change of the point of view (the flat screen had to remain fixed in front of the user). Therefore, a single joystick (as shown in Figure 2.9) has been mounted on the handlebars by means of a 3D printed case. Both the joystick and the potentiometer signals were collected by an Arduino Due board.

The VR application was changed accordingly, adding a specific co-routine for the retrieval and the management of the stream of data from the Arduino board. Additionally, an ad-hoc section dedicated to the operator, and accessible prior to application login, was inserted in the application to allow for the calibration of the potentiometer.

Finally, the projected touch screen was replaced with a finger-touch television: this increased sharpness of colors and the sensitiveness to touch.

For what concerns the software application, the park scenario was improved in terms of graphics, and the path to follow was made smoother to avoid the expectations of lateral acceleration during the exercise.

This was expected to reduce the mismatch between the visual and the (expected) vestibular feedback, thus reducing also the risk of the occurrence of sickness [108]. The new path following algorithm was implemented using LeanTween, a Unity package also including functionalities for path interpolation [156].

Dealing with the supermarket scenario, the modifications we made majorly entailed the generation of shelves and the increase of the number of levels. In particular, shelves were computed (randomly) in advance using an separate ap-

Figure 2.10: The path to follow in the park, in the older (left) and the new (right) versions of Goji.

plication, stored within the training application, and retrieved when needed; in this way, all the users could practice with the same exact scenarios. The levels of difficulty were significantly increased (from 5 to 56) with the aim of creating different scenarios more challenging and “adaptable” to participants’ capabilities, with respect to the few levels developed for the previous trial. Easier levels were made available so that the intervention could be extended to people with more severe symptoms. More complicated levels (probably reachable only from healthy participants) foresaw the entire shelf filled with products (40 items) and a list of 10 items to remember. The characteristics of all the levels of difficulty can be found in Appendix B.

Having increased the levels of difficulty, we also introduced an algorithm for the automatic handling of the transition to the next (or the previous level). The algorithm, created following psychologists’ instructions, foresaw passing to the next level when no errors (or hesitations) were made for three repetitions in a row of the same level; on the other hand, if the user committed more than two errors while completing a single shopping list, he/she had to go back to the previous level. The possibility of going back to previous levels constituted a novelty of this second version of the application too.

To ease the process of configuring the PC on which the VE runs, some parameters were made available for modifications in an external file; these parameters were: the repetitions without errors required to move on to the next level; the maximum number of errors allowed to move on to the next level; the minimum number of errors in a single trial to go back to the previous level; the maximum duration of the shopping and the park session; the COM ports on which the bike and the Arduino board had to be set; the minimum and maximum percentage of the heart rate causing the workload to decrease, or increase, respectively.

Finally, discounted products were made more recognizable (Figure 2.9); also, to avoid the discomfort signalled by the participants of the first trial who had to bought non-discounted items when the discounted version was present on the shelf, both product versions (i.e., with and without discount) were displayed on the shelf only when the product to pick was discounted. Otherwise, only the non-discounted version was placed on the shelves.

Participants

A cohort of 320 individuals composed of older community-dwelling individuals attending seniors' centres or Alzheimer's Evaluation Units (AEVs), and with objective (MCI) or subjective (Subjective Cognitive Decline, SCD) complaints in one or more cognitive domains was assessed for enrolment. The assessment included a clinical evaluation of risk factors (i.e., physiological parameters, comorbidities, lifestyle, depression, vascular risk factors); the evaluation of the cognitive status through MMSE, and other cognitive-related tests; a test to assess the autonomy in the Instrumental Activities of Daily Living (IADL); the Functional Activities Questionnaire (FAQ); and the categorization of each individual with MCI according to the presence (or absence) of impairments in one or more cognitive domains.

Inclusion criteria were: age≥ 60; having an impairment in one or more cognitive domain (MMSE≥ 26, or MMSE≥ 28 if years-of-schooling were ≥ 16; or a score$\geq$1.5 standard deviations lower than age-, sex-, and years-of-schooling-matched normative sample), but without a significant functional impairment (FAQ< 10 and IADL$> 80\%$); absence of co-morbidity. Exclusion criteria included MMSE$\leq$20; presence of neurologic or psychiatric diagnosis; history of cardiovascular diseases; suffering from a brain damage, or seizure; history of alcohol or drug abuse; sensorimotor dysfunctions and/or inability to perform PE.

With respect to the first study, enrolment criteria were slightly relaxed to ensure the access to a larger sample of individuals. Eighty individuals fitted the defined criteria and were enrolled in the study. They all provided informed written consent. This study fulfilled the requirements of the Declaration of Helsinki, and was approved by the IRCCS Fondazione Santa Lucia medical ethics review committee (Rome, Italy).

Experimental protocol

The study was designed as a randomized controlled trial with a 2x2 factorial design. Eighty participants have been randomized twice: the first randomization decided whether the participant had to receive CT, the second whether the participant had to receive PE. This resulted in the following four groups:

CTR: controls, who received no treatment,
CT: participants who received only Cognitive Training,
PE: participants who received only Physical Exercise,
CT+PE: participants who received both.

The CT consisted in 20 minutes of training in the supermarket, using the scenario described in 2.3.1, plus in 20 minutes of training with a commercial clinical-tested software (*Brainer* Professional Brain Trainer). Brainer provided a set of exercises aimed at training different cognitive domains, i.e. complex attention, executive functions, memory, language, visuo-spatial abilities. These exercises and their level of difficulty were chosen from time to time by the neuropsychologist supervising the training. Each new exercise always started from the easiest level.

The PE comprised 20 minutes of cycling on the cycle-ergometer while navigating in the park scenario. The target heart rate was set to vary from 55% to 75%

of the maximal frequency calculated on the basis of the age of the participant (see Equation 2.1 and Equation 2.2). In addition to the cycling activity, 20 minutes of light physical exercises dedicated to increase the strength of both the upper and lower body were performed. For the accomplishment of this part of the training, participants were guided by a video showing a therapist performing the exercises.

Participants belonging to the CT+PE group performed 40 minutes of CT and 40 of PE with the modalities described above. They also performed the road-crossing scenario.

For all the intervention groups, the training lasted 12 weeks and the treatment was administered twice a week. All the participants, including controls, received indications about a healthy nutrition plan and the Mediterranean diet.

Outcomes

As for the pilot study (§2.2.1), the effects of the proposed multi-domain intervention were evaluated from the cognitive and the subjective point-of-view. The assessment of cognitive capabilities foresaw the administration of MMSE, ROCFT, FAB and FAQ test already presented in §2.2.1. In addition to these, the following tests were performed in order to better investigate the cognitive and functional functioning of the enrolled participants.

- *Forward* and *Backward Digit Span tests* (DF, DB) [157], to assess the verbal memory span;
- *Short-Story Immediate and Delayed Recall tests* (SSI, SSD) [143] introduced for the evaluation of the verbal memory functions, in substitution of RAVLT test;
- *Line Cancellation Test* (LCT) and *Multiple Feature Target Cancellation* (MFTC) [158] test, i.e., tools that explore attention disorders, and in particular conjunction search disturbances; these tests replaced TMT-A and AM tests;
- *Phonemic and Semantic Fluency Test* (PF, SF) [159] that investigated word spontaneous production and recall, respectively;
- *Naming from Description test* (ND) [143], that assessed verbal fluency;

Moreover, some emotional and mood-related variables were evaluated; this assessment was introduced to estimate whether subjective cognitive symptoms arose from mood disorders; this battery foresaw the administration of:

- *Geriatric Depression Scale* (GDS) [160];
- *Apathy Evaluation Scale* (AES) [161];

All these tests were administered by the neuropsychologists of IRCCS Fondazione Santa Lucia, where the study has been carried out, at 3 time-points: *t0* (pre-intervention), *t1* (post-intervention), and *t2* (4-month follow-up).

The questionnaire for the evaluation of the subjective perception toward the intervention, of PU and PEOU was slightly modified with respect to its first version (§2.2.1) according to the suggestions of a psychologist working in the field of user-experience evaluation; this questionnaire was administered at *t1* to all the participants belonging to the CT, PE and CT+PE groups; specific sections were omitted according to the activities performed by each participant.

The questionnaire comprised a closed-question part dedicated to the investigation of (1) specific aspects related to the usability and the acceptance of the training, and of the technologies (i.e., pleasantness of the VEs, intuitiveness of input devices, etc.), and (2) the feelings that participants had during the experience (refreshed/tired, relaxed/stressed, bored/engaged, etc.). In addition, the questionnaire comprised a semi-structured interview whose aim was to unveil the reasons underlying the participation to the training, the met/unmet expectations at the end of the program, and the suggestions for improvement.

Statistical analysis

The statistical analysis of the psychometric outcomes has been performed using SPSS. The comparison of baseline characteristics was made using ANOVA or Kruskal-Wallis tests, depending whether the assumptions for the former were met (e.g., the data were continuous, and not ordinal).

For the analysis of study outcomes at t1 and t2, we used mixed-factor ANOVAs, considering *time* as a within factor, and *group* as between factor. When normality assumptions were not met, either the data was transformed using a logarithmic transformation, or independent-sample Kruskal-Wallis test was used. In all cases, we applied the Intention-To-Treat–Last Observation Carried Forward (ITT-LOCT) approach for handling missing data.

Post-hoc analyses were performed only when a significant *group* effect or *group* **time* interaction emerged, as only in this case the change could be attributed to the participation to the training program.

2.3.2 Results

Eighty older adults have been enrolled in the study and randomized in the four groups mentioned in §2.3.1 (each with n=20). The characteristics of the sample, and the outcomes of the cognitive, functional and behavioural assessments are reported in Table 2.5;. AES score resulted higher in the PE group with respect to CTR group.

Drop-outs

A CONSORT flow chart for this study is presented in Figure 2.11. Twenty of the enrolled individuals either refused to participate to the study, or interrupted the training program for personal reasons (i.e., physical worsening, dislike of training activities, others). Of these, 8 were in the PE group, 5 in the CT group, 4 in the CT+PE group, and 6 were controls. One adverse event (i.e., a fall) precluded the continuation of training in the CT+PE group. The evaluation at *t1* was concluded by 57 participants (12 PE, 15 CT, 16 CT+PE, 14 CTR). Forty-nine participants participated to the evaluation phase at *t2*.

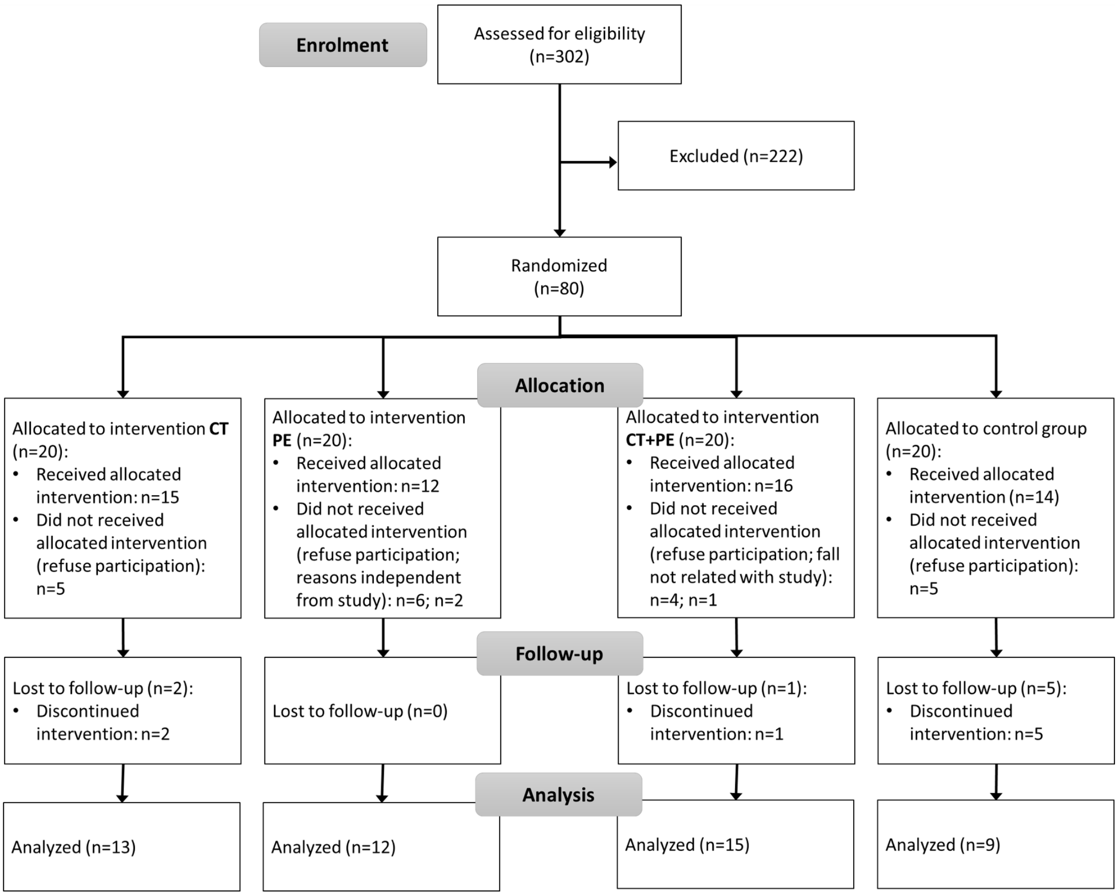


Figure 2.11: CONSORT flow diagram for this study.

Table 2.5: Characteristics of the sample. Values are presented as mean (SD), with the exception of female percentage. PE = physical exercise, CT = cognitive training, CTR = control group.

	PE	**CT**	**PE+CT**	**CTR**
Females (%)	16 (80)	13 (65)	14 (70)	15 (75)
age [yrs]	71.80 (6.54)	70.60 (5.04)	70.65 (6.48)	68.50 (5.96)
schooling [yrs]	13.70 (4.79)	11.40 (4.37)	12.50 (4.66)	12.25 (4.41)
MMSE	26.90 (1.32)	27.03 (1.38)	26.31 (2.19)	25.94 (2.03)
DF	5.80 (0.86)	5.63 (0.98)	5.32 (1.02)	5.60 (0.94)
DB	4.07 (1.31)	3.88 (1.24)	3.77 (1.27)	4.12 (1.28)
SSI*	-0.34 (1.55)	-0.14 (1.42)	-0.59 (1.36)	-0.05 (1.31)
SSD*	-0.20 (1.28)	0.11 (1.54)	0.72 (1.20)	-0.12 (1.48)
LCT [s]	31.62 (5.64)	32.32 (6.08)	30.88 (8.86)	28.38 (7.46)
MCTC [s]	27.75 (13.64)	24.28 (14.68)	29.98 (13.62)	28.37 (17.57)
MCTC [a]	0.92 (0.08)	0.94 (0.07)	0.93 (0.06)	0.92 (0.07)
FAB	16.06 (2.07)	16.11 (1.75)	15.77 (1.99)	15.01 (2.36)
PF	12.51 (5.04)	11.85 (4.50)	11.42 (4.15)	11.39 (3.48)
SF	17.55 (3.77)	17.68 (5.24)	15.46 (3.92)	17.21 (4.85)
ND	36.28 (1.75)	36.65 (1.51)	36.34 (1.59)	36.28 (1.89)
ROCFT	31.55 (5.72)	31.70 (5.34)	30.15 (5.61)	30.50 (5.68)
FAQ	1.60 (1.50)	2.30 (2.08)	1.80 (2.40)	2.30 (2.45)
GDS	1.26 (1.19)	0.60 (0.94)	1.30 (1.45)	1.40 (1.35)
AES	26.78 (5.12)§	29.15 (5.30)	30.60 (6.27)	33.35 (8.10)§

MMSE, Mini-Mental State Examination;
DF, Digit-span Forward; DB, Digit-span Backward;
SSI, Short Story Immediate recall;
SSD, Short Story Delayed recall;
LCT [s], Line Cancellation Test (seconds);
MFTC [s, a], Multiple Features Target Cancellation (seconds, accuracy);
FAB, Frontal Assessment Battery;
PF, Phonological Fluency;
SF, Semantic Fluency;
ND, Naming from Description;
ROCFT, Rey-Osterreith Complex Figure Test;
FAQ, Functional Assessment Questionnaire;
GDS, Geriatric Depression Scale,
AES, Apathy Evaluation Scale.
* Standardized score according to Italian population means.
§ Significant difference, $p<0.05$.

Cognitive and emotional outcomes

Cognitive, functional and behavioural outcomes at the baseline, post-intervention and at follow-up are reported in Table 2.6.

For brevity, only significant statistical analyses are reported. A more detailed description of all the neuropsychological outcomes can be found in [162]. For SSD, there existed a significant main effect of *group* $F(3, 76) = 2.82, p = 0.045$. The subsequent analysis disclosed a difference between controls and CT+PE at *t1* ($p = 0.032$); and between controls and CT+PE, and controls and CT at *t2*, with $p = 0.024$ and $p = 0.047$, respectively. However, since there was no significant improvement with respect to the baseline, results have to be treated cautiously.

For PF, we found a significant *time*group* interaction ($F(6, 152) = 2.254, p = 0.041$) and a main effect of *time* ($F(2, 152) = 11.635; p < 0.001$). Post-hoc analysis revealed that existed a significant difference between *t0* and *t2* for CT group, who improved significantly over time ($p = 0.008$). A significant improvement occurring between *t0* and *t2* was however also recorded for CTR ($p = 0.001$). Also SF registered a significant effect of *time* ($F(2, 152) = 4.51, p = 0.0013$), but neither *group* ($F(3, 76) = 1.127, p = 0.343$), nor *group*time* interaction ($F(6, 152) = 2.12, p = 0.054$) resulted significant.

GDS resulted in a almost significant difference occurring at *t1* ($H(3) = 1.826, p = 0.050$). Post-hoc analysis was however significant, and revealed that the difference occurred between CT and control groups ($p = 0.046$). In any case, this result was not maintained at *t2*.

Table 2.6: Cognitive, functional and behavioral outcomes at t0 (baseline), t1 (end of training), t2 (follow-up). Results of post-hoc analysis are also shown. PE = physical exercise, CT = cognitive training, CTR = control group.

		PE	**CT**	**PE+CT**	**CTR**
MMSE	t0	26.90 (1.32)	27.03 (1.38)	26.31 (2.19)	25.94 (2.03)
	t1	26.92 (1.10)	27.15 (1.44)	26.33 (2.15)	25.67 (2.78)
	t2	26.72 (1.21)	26.83 (1.40)	26.34 (2.92)	25.34 (2.00)
DF	t0	5.80 (0.86)	5.63 (0.98)	5.32 (1.02)	5.60 (0.94)
	t1	5.75 (1.00)	5.48 (1.03)	5.88 (0.90)	5.92 (1.26)
	t2	5.65 (1.00)	5.64 (1.23)	5.49 (0.87)	5.88 (1.07)
DB	t0	4.07 (1.31)	3.88 (1.24)	3.77 (1.27)	4.12 (1.28)
	t1	3.99 (1.29)	3.88 (1.23)	3.87 (1.50)	3.84 (1.01)
	t2	4.05 (1.39)	4.21 (1.65)	3.83 (1.72)	3.95 (0.89)
SSI*	t0	-0.34 (1.55)	-0.14 (1.42)	-0.59 (1.36)	-0.05 (1.31)
	t1	0.06 (1.97)	0.41 (1.79)	0.69 (1.55)	-0.18 (1.43)
	t2	0.33 (1.79)	0.72 (1.66)	0.84 (1.74)	0.00 (1.23)
SSD*	t0	-0.20 (1.28)	0.11 (1.54)	0.72 (1.20)	-0.12 (1.48)
	t1	0.14 (1.63)	0.48 (1.82)	0.77 (1.41)$^\phi$	-0.15 (1.43)$^\phi$
	t2	0.31 (1.47)	0.55 (1.78)$^\#$	0.87 (1.50)§	-0.08 (1.41)$^{\S,\#}$
LCT [s]	t0	31.62 (5.64)	32.32 (6.08)	30.88 (8.86)	28.38 (7.46)
	t1	32.52 (8.72)	31.66 (7.15)	28.93 (7.96)	30.29 (4.81)
	t2	33.34 (12.46)	34.20 (8.31)	32.20 (10.32)	30.28 (3.96)
MFTC [s]	t0	27.75 (13.64)	24.28 (14.68)	29.98 (13.62)	28.37 (17.57)
	t1	26.07 (14.82)	29.16 (16.69)	25.99 (14.52)	21.55 (9.97)

	t2	24.43 (12.46)	25.01 (8.87)	28.06 (15.83)	19.21 (8.44)
MFTC [a]	t0	0.92 (0.08)	0.94 (0.07)	0.93 (0.06)	0.92 (0.07)
	t1	0.92 (0.08)	0.94 (0.07)	0.93 (0.06)	0.92 (0.07)
	t2	0.92 (0.06)	0.96 (0.04)	0.93 (0.06)	0.92 (0.07)
FAB	t0	16.06 (2.07)	16.11 (1.75)	15.77 (1.99)	15.01 (2.36)
	t1	15.74 (1.61)	16.36 (1.68)	15.81 (2.05)	15.93 (2.15)
	t2	15.88 (2.16)	16.37 (1.70)	15.96 (1.86)	16.11 (1.41)
PF	t0	12.51 (5.04)	11.85 (4.50)§	11.42 (4.15)	11.39 (3.48)§
	t1	13.05 (5.40)	13.20 (3.61)	12.61 (3.54)	12.66 (3.83)
	t2	14.10 (6.08)	14.10 (4.28)§	11.31 (4.15)	14.06 (4.23)§
SF	t0	17.55 (3.77)	17.68 (5.24)	15.46 (3.92)	17.21 (4.85)
	t1	18.44 (4.59)	18.60 (4.32)	17.90 (5.06)	16.98 (5.50)
	t2	17.94 (5.23)	19.38 (5.41)	15.55 (5.42)	16.83 (5.35)
ND	t0	36.28 (1.75)	36.65 (1.51)	36.34 (1.59)	36.28 (1.89)
	t1	36.28 (1.51)	36.88 (1.37)	36.88 (1.29)	35.98 (1.88)
	t2	36.46 (1.69)	36.88 (1.35)	36.42 (3.63)	36.07 (2.23)
ROCFT	t0	31.55 (5.72)	31.70 (5.34)	30.15 (5.61)	30.50 (5.68)
	t1	31.39 (5.48)	31.72 (5.59)	30.73 (5.65)	30.61 (5.65)
	t2	31.51 (5.40)	31.42 (4.89)	30.97 (8.14)	31.37 (4.29)
FAQ	t0	1.60 (1.50)	2.30 (2.08)	1.80 (2.40)	2.30 (2.45)
	t1	1.65 (1.90)	2.65 (2.20)	1.60 (2.26)	2.25 (2.65)
	t2	1.85 (2.00)	2.50 (2.28)	1.70 (1.81)	2.35 (2.89)
GDS	t0	1.26 (1.19)	0.60 (0.94)	1.30 (1.45)	1.40 (1.35)
	t1	0.84 (0.95)	0.30 (0.57)§	0.85 (1.49)	1.45 (1.57)§
	t2	0.84 (1.12)	0.50 (0.68)	1.00 (1.41)	1.20 (1.40)
AES	t0	26.78 (5.12)	29.15 (5.30)	30.60 (6.27)	33.35 (8.10)
	t1	26.05 (5.15)	27.85 (5.18)	30.35 (6.13)	32.80 (7.57)
	t2	25.72 (4.48)	28.2 (5.08)	31.20 (6.04)	31.35 (6.59)

MMSE, Mini-Mental State Examination;
DF, Digit-span Forward; DB, Digit-span Backward;
SSI, Short Storty Immediate recall;
SSD, Short Story Delayed recall;
LCT [s], Line Cancellation Test (seconds);
MFTC [s, a], Multiple Features Target Cancellation (seconds, accuracy);
FAB, Frontal Assessment Battery;
PF, Phonological Fluency;
SF, Semantic Fluency;
ND, Naming from Description;
ROCFT, Rey-Osterreith Complex Figure Test;
FAQ, Functional Assessment Questionnaire;
GDS, Geriatric Depression Scale,

AES, Apathy Evaluation Scale.
* Standardized score according to Italian population means.
$^{\phi,\S,\#}$ Significant difference, $p<0.05$.

Subjective outcomes

We collected 38 questionnaires investigating treatment PU and PEOU. Of these, 9 were from older adults who were part of the PE group, 16 were from CT group and 12 from PE+CT.

As the park and the supermarket VEs were investigated with two different questionnaire's sections, we reported the results of closed questions grouped according to scenario, rather than to group (Table 2.7). For the themes identified by the thematic analysis, i.e., motivation and expectations, acceptance, and perceived improvements, a specific reference to group is made whenever significant.

Motivation and expectations. Mainly, older adults decided to participate to the training for two different reasons, which were probably dependent on their perceived cognitive status. Twelve out 38 reported that they wanted to take part in the training because they had started perceiving a slight decline of their functions (*"I was anxious about forgetting important things" / "I had some inconveniences in the activities of daily living" / "I wanted to contrast memory decline"*), especially of memory. Of these, 3 were also worried by other AD cases in their families, one was afraid of isolation, one of 'feeling old' after the death of his wife, another of becoming lazy.

Study participants also decided to take part in the training because it appeared to them a good prevention method ($n = 8$; *"It seemed to me a beautiful idea, and it could be good for prevention" / "I take care of my cognitive abilities" / "I heard a lot about aging-related decline"*); or a good way to discover something more about their current cognitive status ($n = 9$; *"I am curious about my mind" / "I expected a complete check-up of my brain functions"*). Many participants specifically mentioned that they wanted to support research ($n = 12$; *"Taking part in scientific research is important" / "I am trying to improve the quality of life for future generations" / "I would like to know if there is a relationship between physical activity and cognitive status"*).

In terms of expectations, 8 participants reported that they wanted to improve their cognitive functions, one aimed at obtaining a physical improvement, one thought to improve on both sides. Two participants explicitly mentioned that they expected many exercises (more than the ones proposed in the program), and more VR. Six participants were interested in discovering something more about themselves (*"I was curious, I wanted to challenge myself" / "I wanted to know if I had to do something for my cognitive capabilities"*). All the other participants reported not to have clear expectations, in few cases the participation to the training was suggested by their general practitioner, or by other family members, especially their sons and daughters.

Table 2.7: Results (median and q1-q3) of the questionnaire administered after the sixth week of training. The evaluation was based on a 5-points Likert scale (-2 to 2), if not specified differently; positive values were used for positive opinions.

		Score
General satisfaction	Are you satisfied with this training?	2 (1, 2)
Park scenario	Did you enjoy cycling?	0 (0, 1)
	Was it difficult to keep the required velocity?	1 (0, 1)
	Did you feel comfortable wearing the chest-band?	0 (0, 0)
	Was the park realistic?	-1 (-1, -1)
	The duration of the cycling was too long (2) / too short (-2)	0 (0, 0)
	During the cycling, you felt...(#)	
	rested / neutral / tired	10 / 7 / 5
	relaxed / neutral / stressed	18 / 1 / 3
	calm / neutral / shaken	22 / 0 / 0
	engaged / neutral / bored	10 / 7 / 5
	concentrated / neutral / distracted	16 / 2 / 4
Supermarket scenario	Did you enjoy doing the shopping?	0 (0, 1)
	Was it difficult to complete the shopping?	1 (0.75, 1)
	Was the shopping scenario realistic?	0 (0, 1)
	Was it easy to interact with the touch screen?	1 (1, 1)
	Was the shopping list clear?	1 (1, 2)
	Was it easy to recognize shopping items?	1 (0, 1)
	Was the guiding voice clear?	1 (1, 2)
	Were the hints useful?	1 (1, 2)
	The duration of the shopping was too long (2) / too short (-2)	1 (0, 1)
	During the shopping, you felt...(#)	
	rested / neutral / tired	15 / 7 / 6
	relaxed / neutral / stressed	20 / 4 / 4
	calm / neutral / shaken	24 / 0 / 4
	engaged / neutral / bored	12 / 10 / 6
	concentrated / neutral / distracted	24 / 2 / 2

Acceptance. Results in terms of acceptance were positive, but few complaints in term of repetitiveness and duration of tasks arose. Of the 22 participants who performed PE in the park scenario, 13 reported very positive feedback saying that it was *"a new experience"*, that was *"nice to be distracted from the physical effort"*, and *"useful to have indicators showing the performances"*; one person said it was a *"relaxing environment"*. On the other hand, the remaining 40% of the participants reported to get bored ($n = 5$), to be used to perform PE, or to prefer walking or cycling on a real bike. Seven participants suggested to improve the variability of the scenario introducing moving objects, or to add other scenarios. None of the participants reported any symptoms during the park experience, with the exception of a little shortness of breath in few cases, and only at the beginning of the training.

For what concerns the supermarket, all 29 participants agreed that it was an activity of daily living, 12 appreciated the fact of training while performing a familiar task (*"I was continuously stimulated"*), but 10 found it boring and monotonous especially after a while. Other participants complained that *"having to remember the exact order of the shopping items was unfair"*, and the duration was too long to keep concentrated throughout the trial. One participant reported to feel anxious.

Usability was judged good, with no difficulties in the use of touch screen. Products, list, hints and the guiding voice were evaluated positively, with the exception of comments related to the complexity of the task itself (*"When products are many, it should give me more time" / "When products are a few, they are too sparse on the shelves"*). All participants agreed that attention and concentration were fundamental to accomplish the task correctly (*"It seemed easy, but there is that something creating difficulty" / "I think many people could find it difficult"*), though the duration was judged too long in 7 cases. No one reported side-effects, but the animation guiding the user from the aisle to the front of the shelf was annoying for someone ($n = 4$); one participant said he would preferred to stay farther away from the screen, and another one that she would have liked to perform the shopping while staying sit.

Perceived improvements and continuation of training. In general, both those who expressed expectations and those who did not appreciated the training program, and evaluated the experience positively (*"I do not want to leave prevention programs anymore" / "I was annoyed by the feeling of being old, the training helped me in reversing this sensation"*). It was interesting to notice that also participants who were not expecting anything from the training, reported to have been positively surprised to have benefitted from it ($n = 9$; *"I had no expectations [...] now I pay more attention!" / "I have high self-esteem, I did not expect anything, but now it seems to me that I can concentrate more*).

To questions investigating the changes occurred after the training, 28 participants answered reporting positive feedback. Of these, 3 participants (2 in the PE group, and 1 in CT+PE) said that they improved their physical shape (*"[The training] strengthen my muscles, I think I walk better now" / "I think I am more active now"*). The other comments were mainly related to improved attention and memory (*"I can go to the supermarket without making a list!"*).

For all of the participants, with the exception of 4, the experience was a way to increase the awareness of their status and a way to challenge their capabilities (*"I realized I was often distracted" / "I realized that I had so many duties that I do not concentrate on my needs anymore" / "The training stimulated my curiosity" / "It highlighted my limits and my potentialities" / "[This experience] pushed me to do more, I'd like to make some exercise with my tablet now"*). Regarding the others, 1 participant in the PE group reported to had an increase in the pain in the upper limbs, 1 in PE+CT said that the training was not frequent enough to get results, and the other 2 (CT and PE) said that they already had an active life, therefore participating in the study made a little difference for them.

Thirty-six participants out of 38 said that they would be willing to continue with the training (*"I trust to improve"* / *"It's a way to take care of myself"* / *"It worked!"* / *"I hope to improve more, so I can help researchers with their work"*). The two saying that they would not, reported *"to be tired"* in one case (CT+PE) and that it was too difficult to reach the hospital in the other. The first person also said that she would not continue at home, whereas the second would be happy to do so.

In total 15 people said that they would refuse to continue the program at home; beside the already mentioned participant, all the others explained that they preferred to *"to meet the nice personnel of the hospital"* and *"to have a fixed appointment to attend during the week,* [at home] *I would be afraid of postponing it"*, or that they were not familiar with the use of personal computers.

2.3.3 Discussion

This study aimed at verifying the impact, in terms of cognitive and emotional outcomes, and the acceptance of a VR-based physical and cognitive treatment in a population of individuals with subjective or objective cognitive complaints. At the baseline, we enrolled homogeneous groups, in terms of gender, age and year of schooling. Nonetheless, it had to be noticed that females were present in a percentage that was significantly higher than males; this was in agreement to what found in previous studies, i.e., that women take more care of their health, and therefore are more prone to participate in scientific trials aimed at improving the physical and psychological well-being [163, 164].

One important note regards the significant decrease of the sample throughout the study: one third of the participants did not complete the evaluation at *t1*, and almost 40% did not complete the evaluation at *t2*. These results did not seem to be related to the nature of proposed treatments: of the 23 participants not evaluated at *t1*, only 6 (4 in the PE, and 2 in the PE+CT) interrupted the training; all the others refused to participate prior to the beginning of the administration. The same is worthy for *t2*, in which the highest number of individuals retired from the study was in the control group (5 vs. 0, 2, and 1 in PE, CT, CT+PE, respectively). The presence of PE could have influenced the decision of leaving the program: it may be hypothesized that the effort required was considered too high for older adults who were not prone and/or physically prepared to start a physical exercise program [165].

As for the highest rate of drop-outs, also questionnaires and interviews regarding the intervention's acceptance collected more negative feedback than what measured in the previous trial (§2.2.1). Complaints were however mainly marginal, and the general experience was appreciated; usability was judged good, and, above all, 95% of the participants who were interviewed said that they would continue with the training. A possible hypothesis for the higher number of complaints could be found in: (1) the highest number of participants, and (2) the typology of impairment they had. In the first study, in fact, we enrolled patients attending the hospital because of an already-present diagnosis of AD, made on the basis of cerebro-spinal fluid markers; in this RCT, instead, participants were frequenters of community centres, and most of them were even unaware of suffering from a (mild) cognitive deterioration. Thus, it is plausible that they were less intrinsically motivated in undertaking a treatment, and the highest number of refusals before the beginning of the program supports this hypothesis. On the other hand, it could be that their cognitive or physical status was superior, and thus they did not feel challenged enough by the tasks required by the program: few participants confirmed this in the interviews. A more personalized treatment, both in terms of PE and CT could constitute a solution to try and solve this issue.

Regarding the objectives of the study in terms of cognitive and emotional outcomes, results showed that the proposed treatment could have good potentialities, though they should be treated cautiously, because samples were not large, and did not allow for the generalization of outcomes.

However, in contrast with our previous study (§2.2), few improvements in memory were recorded. We found significant differences between controls and treated participants in the CT and in CT+PE groups in long-term memory recall tests; this data emerged after the training and were maintained also at follow-up. This was also reported by Coyle et al. [87] who identified memory as one of the most improved domain after CCT or VR-based training. However, other studies also highlighted improvements in executive functions and attention, which we did not find. Perhaps enrolling individuals with more compromised abilities could have highlighted further differences.

With respect to our previous study, the differences obtained in the outcomes could be linked to the different duration of the intervention, both with respect to the single session (40 vs. 20 minutes), and to the whole treatment (3 times/week for 6 weeks vs. 2 times/week for 12 weeks).

PE alone did not result in any effect in our study, and this was in contrast with other studies' findings (§1.3.2). This may be attributed (excluding study limitations) to the type of aerobic exercises that we administered. On the one side, the customization of the intervention only on the basis of heart rate may be not enough to personalize the treatment according to each individual's needs, as already trained participants found it easy, whereas non-trained participants ended (the first sessions) with being very tired. On the other hand, the whole quantity of PE administered could have been too less, as it was less than what recommended by the World Health Organization (WHO) [166]. WHO indeed recommends 150 minutes of PE at moderate intensity, or 75 at vigorous intensity, during the week to improve cardio-respiratory fitness, depression, and cognitive decline.

In the future, beside the change and the customization of the treatment, additional elements could be added to similar interventions based on PE, in order to to improve their effectiveness; recent studies suggest that a having the chance to have social interactions in a pleasant atmosphere could improve the effectiveness of the PE on cognitive outcomes [167]. In mice, the presence of pleasant stimuli in the training environment had a suppressing effect on the accumulation of β-amyloid [168].

For what concerns the changes in emotional outcomes, a difference was recorded between CT and control groups in the depression scale at *t1*; this agreed to what reported by Coyle et al. in their review on the effects of computerized training in MCI populations [87]. However, it was not maintained at follow-up; a possible hypothesis may be sought in the interruption of the contacts between the therapists and the participants, who received "less attention" between *t1* and *t2*.

Unfortunately, the impossibility of performing analysis on oxidative stress in this study did not give us the chance of gathering data related to the physiological outcomes of the proposed program.

The study had indeed some limitations, i.e., the small number of participants in each group, and the absence of a placebo for controls. Additionally, it could be that multi-domain interventions had different effects in older adults with high or low risk of developing dementia, and we did not verify the effectiveness of this intervention in high-risk groups.

In spite of these limitations, the results of this study – coupled with previous one's – allowed concluding that our intervention was feasible, and surely acceptable in older adults with MCI. The good acceptance rate, the subjective perceptions of well-being, and also the improvement in a few cognitive variables, represented a promising result thus making this intervention – with appropriate modifications – worthy of further investigations.

2.4 Conclusions

The two studies described in this Chapter presented the application of VR technologies for the implementation of a multi-domain intervention dedicated to older adults with MCI.

With respect to the *Aim 1* of this thesis, we obtained promising results. The intervention was largely accepted, and participants enjoyed to be part of the experimental studies. Additionally, most of them reported an improved perception of well-being, which was an essential element to reduce the excess disability, and to maintain a good QoL.

Dealing with objective outcomes, in spite of the lack of statistical significance (with the exception of few cases), cognitive, emotional and physiological outcomes have mostly highlighted a tendency toward improvement throughout the training in the experimental groups of both studies. Consequently, though the data we obtained must be treated carefully from a clinical point-of-view, they were useful to inform judgement on the feasibility of this type of intervention (i.e., including PE performed on a cycle-ergometer), and the future directions of technological and scientific developments.

Future studies should thus continue addressing the use of VR for cognitive interventions. Possible improvements could be: the design of a more customized training program (e.g., using the International Physical Activity Questionnaire to determine the initial bike workload or the slope with which it has to be increased/decreased); the introduction of more variability in the VE accompanying the cycling; the introduction of social elements [169]. Another interesting aspect to investigate would be introducing scenarios capable of engaging more the participants by eliciting higher SoP (as described in Chapter 3 and 4), with the final aim of increasing the participants' motivation to train, their performances and thus their treatment outcomes [15].

Finally, more from the clinical point-of-view, the implementation of RCTs with adequate sample sizes remains fundamental to provide evidence on the effectiveness of non-pharmacological intervention to manage MCI. The possible inclusion of more impaired individuals should also be considered to generalize the results to the entire population of older adults with cognitive impairments.

Chapter 3

Training with higher Sense of Presence: Cycling in immersive virtual reality

3.1 Introduction

As explained in §1.4.1, VR allows users to feel present in a simulated environment by immersing the individual in a computer-generated world, and allowing him/her to interact and navigate within the virtual scenarios. However, as there exist systems providing different levels of immersion, it has been argued that the closer the system is to human perception, the higher is its potential of eliciting SoP [170].

Achieving high SoP represents one of the objectives that VR developers should pursue, as having the perception of being present in the VE has been demonstrated to increase users' motivation, thus possibly improving also performances [15, 16]. In the field of cognitive training, and rehabilitation in general, this aspect acquires particular relevance, as higher motivation results in higher treatment adherence, and thus in better outcomes [171, 14].

Though one may think that higher immersion would necessarily result in higher SoP, this relationship is not always straight-forward. Even if the majority of data shows that more advanced and sophisticated VR devices providing a higher immersion result in an increase in SoP [17], [18], few exceptions exist [172, 173]. Additionally, subjective factors have been proven to strongly contribute to SoP, as *involvement* is in turn strongly dependent on the user's willingness to be distracted from the real world [16].

Finally, it has been demonstrated that there exist a negative correlation between SoP and cyber-sickness [19]. Thus, to elicit SoP, the design of the VR scenario must envision appropriate expedients to limit potential physical drawbacks, especially in the case of immersive VR. The main of these drawbacks is undoubtedly *cyber-sickness*, i.e., the specific type of sickness arising due the navigation in VEs (§1.4.1). It comprises a set of symptoms that anyone can experience alternatively and/or to different extents; the most common are: nausea, disorientation, dizziness, visual disturbance and sweating [174].

The main cause underlying cyber-sickness arousal is the conflict between sensory perceptions, and, in particular, the mismatch between visual and vestibular signals [108, 19]. This mismatch can occur frequently in navigational environments as, due to the restricted walkable area, different metaphors are used to move around, thus causing the visual system to observe more movement than what implied by the vestibular system's response.

Other aspects to consider when designing immersive applications are the ergonomics of the solutions (e.g., the weight of the headset), or the postural instability related to the increased body sway normally occurring during the experience of immersive VR [135]. Lack of usability and user-friendliness, which tends to make users not to accept the proposed technological solution, may also impact [107, 135].

Because of all these reasons, standard guidelines for development of acceptable applications for healthy individuals may not be enough for older adults with MCI, and preliminary tests on non-vulnerable populations are mandatory to get rid of (or reduce to the minimum) potential design errors that may reduce the VR system's acceptance.

3.1.1 Aims

Given the above-mentioned premises, this Chapter presents the activities that have been carried out to pursue the *Aim 2a* of this thesis: to evaluate the acceptance of immersive VR technologies for implementing cycling-based PE (§1.5).

The achievement of such an objective foresaw the implementation of an immersive version of the park and the road-crossing scenarios presented in Chapter 2, with the aim of making them more engaging, also for older adults.

Since no previous studies investigating the effects of wearing a HMD while cycling were found prior of conducting this research, the execution of a feasibility study on healthy young adults was believed to be an appropriate safety measure. Also, the implementation of a preliminary test was considered useful to detect possible issues of the system (i.e., latency, bad interaction, etc.).

3.2 Cyber-cycling: HMD or large projected screen? A comparative study on healthy young adults

The first attempt made to move the scenarios presented in Chapter 2 to immersive VR has been made considering the park and the urban scenarios. This situation was indeed the most critical since both scenarios included navigation, and thus a condition eliciting the mismatch between visual and vestibular feedback. On the other hand, the fact of having a complete control over the forward-velocity may partly counterbalance these negative feelings [16]. Moreover, the fact of being sit on a cycle-ergometer limited the risk of falls thus allowing for the implementation of a safe scenario.

Given these premises, a comparative study enrolling healthy young adults was performed [175] with the aims of (1) assessing the acceptance of a HMD-based cycling, and (2) determining the pros and cons of wearing a headset versus looking at a large projected screen. As secondary aims, we investigated the effects of time on both negative and positive effects of the experience, namely, occurrence of symptoms and their severity, and positive recall of the experience.

3.2.1 Methods

Study design and participants

A within-subject repeated-measurements study was designed. Participants were required to cyber-cycle in two different experimental conditions, i.e., in front of a Large Projected Screen (LPS) and while wearing a HMD. Participants were all healthy young adults who gave their informed written consent. Exclusion criteria were: severe vision deficits, mobility issues, history of navigation sickness, cognitive deficits. The sample characteristics are shown in Table 3.1.

Table 3.1: Characteristics of the sample. $*$ =missing values [175].

Participants	N=33
Age [yrs]	31.00 ± 4.94
Sex [M/F]	23/10
Impaired vision [Y/N]	21/9
Familiarity with VR [Y/N]	6/26
Bike users [Y/N]	28/4

To evaluate the effect of time and reminiscence, and to better inform toward possible combination of treatments in the case of MCI training programs, we applied a balanced randomization scheme. Its goal was studying whether there existed differences in the two conditions when performed (1) with 5-days-washout or (2) on the same day. The resulting scheme is presented in Table 3.2.

Table 3.2: The study randomization scheme. Adapted from [175].

	5-days washout	same day
HMD then PS	group 1	group 3
PS then HMD	group 2	group 4

Equipment

The two VEs used in this study were the park and the road-crossing scenarios presented in §2.2.1. To create the HMD condition, we used a Samsung GearVR HMD equipped with a Samsung S6 smartphone, whereas for the LPS condition, we used a projector (EB-1430WI, Epson). The dimensions of the LPS were kept constant for all the trials at (1.30 x 2.35 m). It was placed in front of the cycle-ergometer at 1.22 m from its base.

The cycle-ergometer was connected to a PC using a serial port, to synchronize the visual flow according to the cycling velocity. For the HMD condition only, an additional TCP client-server connection was set up to allow the PC sending the velocity data to the application running on the smartphone, via a wireless network.

The potential existence of delays was verified prior to begin any test; no latency was perceivable.

The interactions with the virtual scene occurred differently for the 2 conditions. In LPS condition, the PlayStation controller's joystick was used to look around, while the X button was used to brake. In HMD condition, the rotation of the point of view occurred naturally, by turning the head; the Samsung Gear VR implemented a gyroscope to improve the measurements performed by the smartphone itself. The braking function was implemented using the only interaction means of this HMD, i.e., the touch-pad on the right side. This type of (unnatural) interaction was expected to distract the participant from the immersive experience, because it was inconsistent with reality. However, also the use of PlayStation controller was not natural, thus allowing the comparison of the 2 conditions of testing.

In both cases, pressing the braking button caused the visual flow to stop smoothly in 0.5 m. Restarting to cycle after braking caused the visual flow to restart according to the cycling velocity.

Study protocol

We orally informed the participants about the two environments' interaction modalities before the beginning of the test. All were given a schedule for the performance of the two conditions, according to their group belonging (Table 3.2). All participants navigated first in the park for 5 minutes, and then in the urban scenario for the time required to cross 5 cross-roads (about 5 minutes). The time of the experience was intentionally kept quite short in order to limit the extent of symptoms, if any would occur [176]. We required them to keep a cadence in between 50 and 70 RPM, in agreement with what we had done in previous projects with older adults (§2.2.1). Finally, the workload was set to 20 W to induce no fatigue.

Measures

The subjective experience was assessed through the following standard questionnaires. We had two main outcomes: SoP and cyber-sickness, plus a third one aimed at detecting users' motivation or anxiety.

- *Igroup Presence Questionnaire* (IPQ) [177]; this questionnaire aimed at investigating user's SoP. It is composed of 14 items investigating the sub-domains of *Spatial Presence*, *Involvement* and *Experienced Realism.* Answers are given through 7-item Likert scales, and range from 13 (minimum SoP) to 91 (maximum SoP).

- *Simulator Sickness questionnaire* (SSQ) [178] assessed potential adverse effects arising during the experience in the VR environment. SSQ is composed of 3 different sub-scales: nausea (N), oculomotor disturbances (O), and disorientation (D); each sub-scale comprises 7 symptoms that have be rated as *none*, *slight*, *moderate*, or *severe.*

- *Intrinsic Motivation Index* (IMI) [179]; we used the sub-scales of *Interest and Enjoyment* (INT-ENJ, 7 items) and *Tension and Pressure* (TEN-PRES, 5 items) in order to evaluate whether these factors were positively or negatively linked to SoP or cyber-sickness; each item was measured through a 7-point Likert scale.

We also collected participants' comments during the test. Comments that were judged relevant to better frame the user-experience in terms of SoP or side-effects were examined through thematic analysis [149]. Lastly, at the end of the two conditions of testing, participants were asked which device they preferred.

Statistical analysis

Statistical analyses were performed using SPSS, and considering a significance level equal to α=0.05. Chi-squared tests were performed to establish whether the difference in the device preference was generalizable. Such a test was performed on the entire sample, and on a sub-portion of it that did not include frequent VR users; in this case, the goal was to exclude the effect of familiarity with VR technologies.

Normality of the collected data was checked using Kolmogorov-Smirnov test. Paired t-tests were run to compare SoP in the two conditions. Unpaired t-tests were also used to compare SoP in sub-groups defined according to: (1) device's preference, (2) group, (3) personal characteristics. A 2x4 mixed ANOVA was performed to evaluate the effects of the interaction *condition*group* and the two main effects of *condition* and *group*. Cohen's d was used to assess effect size.

As cyber-sickness is known to have a skewed profile [178], the comparison of the two conditions in each group was performed using Wilcoxon signed-rank tests.

Correlations of SoP with cyber-sickness, engagement and anxiety were evaluated using Pearson's correlation coefficient. A square root transformation of SSQ was performed before running such a statistic test.

3.2.2 Results

Data from 3 participants were excluded from the analysis; distribution and reasons are presented in the CONSORT flow diagram (Figure 3.1).

Device preference. We found a significant difference in device's preference between participants who preferred the HMD ($n = 24$) with respect to the LPS ($n = 6$): $\chi^2 = 10.8, p < 0.001$. Significance was preserved also when excluding frequent VR users ($n = 25$): $\chi^2 = 6.76, p < 0.01$).

Sense of presence. In LPS condition, IPQ mean score was 36.5 ($SD = 8.32$, range: 20-50), whereas with HMD, we obtained 56.8 ($SD = 10.39$, range: 35-91). The difference between the two values reached statistical significance ($t = -12.06, p < 0.001$). All participants reported higher SoP in the HMD condition, irrespective to the group which they belonged to (Table 3.3).

Running ANOVA, we found a main effect of *condition* ($F(1, 26) = 134.19, p < 0.001$). Neither the *condition*group* interaction ($F(3, 26) = 0.297, p > 0.05$), nor the main effect of *group* ($F(3, 26) = 1.97, p > 0.05$) were significant.

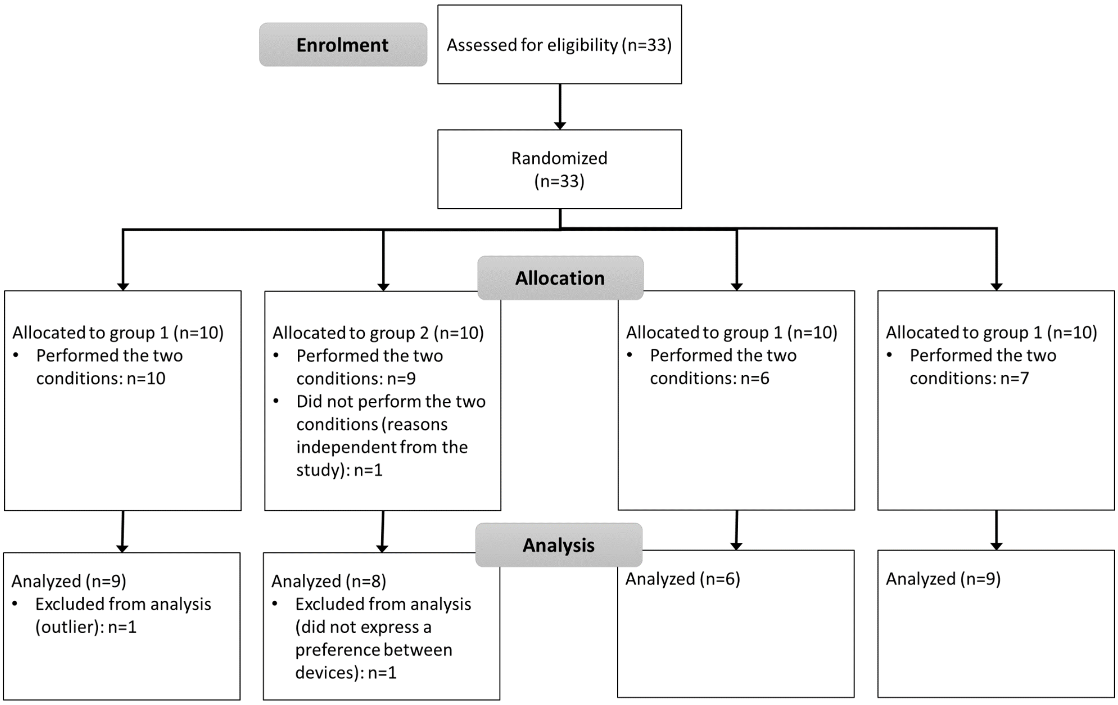


Figure 3.1: CONSORT flow diagram for the HMD vs. LPS comparative study.

Table 3.3: Sense of presence; results are presented as mean value (SD). p-values corresponds to ANOVA post-hoc analysis results. For effect size, Cohen's d values are reported. Adapted from [175].

Group	Cond.	SoP	p	Effect size
1	LPS	37.00 (9.17)	$p < 0.05$	1.90
	HMD	55.89 (10.68)		
2	LPS	33.00 (6.16)	$p < 0.05$	1.98
	HMD	52.00 (12.10)		
3	LPS	42.29 (7.63)	$p < 0.001$	2.16
	HMD	60.86 (9.48)		
4	LPS	24.43 (0.10)	$p < 0.05$	3.97
	HMD	57.86 (8.71)		

Cyber-sickness. SSQ subscales and total scores are reported in Table 3.4. All participants experienced greater discomfort in tge HMD condition. Considering all groups, SSQ-TS in HMD condition was 30.00 ($q1 - q3 : 25.75 - 34.00$), while in PS, it was 22.00 ($21.00 - 25.25$); such a difference was statistically significant ($z = -4.71, p < 0.001$). No differences in SSQ-TS were found when comparing the LPS and HMD between groups; almost all sub-scales resulted in a significant difference within the same group in the 2 conditions: in the HMD condition, oculomotor disturbance was higher for all groups, nausea and disorientation were higher in 3 out of 4 groups.

SoP and SSQ only correlated in the HMD condition ($r = -0.39, p < 0.05$). No other significant correlations were found, neither with motivation, nor with anxiety (Figure 3.2). Also, no effects of gender, familiarity with VR and order of exposure were recorded.

Participants' comments

The participants' comments were categorized in 4 different themes.

- *realism*: the naturalness of the virtual environment and of the exercise, i.e., how similar to reality the virtual experience was perceived to be (Table 3.5);
- *interaction*: the perceived self-efficacy within VR. Both opinions concerning interactions and their quality, and the (lack of) perception of control were identified as part of this theme (Table 3.6);
- *involvement*: the sense of being immersed and engaged during the experience; these comments were especially related to the interruption of the so-called "flow" (Table 3.7).
- *physical drawbacks*: physical discomfort and negative feelings (Table 3.8).

3.2.3 Discussion

The aims of this study were assessing whether an immersive VR-based intervention requiring cycling would be acceptable for older adults, and what advantages the use of HMD would have with respect to a LPS. A preliminary study on healthy young adults was considered mandatory as older adults – especially when showing symptoms of cognitive impairments – have to be considered a vulnerable population.

*Table 3.4: Scores of the Simulator Sickness Questionnaire; results are presented as median (min-max). * = significance at Z-test, p<0.05 [175].*

Group		SSQ-N	SSQ-D	SSQ-O	SSQ-TS
1	LPS	9.54* (0.00-28.72)	0.00* (0.00-30.32)	0.00* (0.00-41.76)	3.74 (0.00-29.92)
	HMD	19.08* (0-57.24)	22.74* (0.00-45.48)	27.84* (0.00-33.66)	29.92 (3.74-74.80)
2	LPS	4.77 (0.00-28.62)	0.00* (0.00-7.58)	0.00* (0.00-41.76)	1.87 (0.00-22.44)
	HMD	19.08 (19.08-114.48)	18.95* (0.00-29.92)	55.68* (13.92-125.28)	33.66 (3.74-97.24)
3	LPS	19.08* (0.00-28.62)	3.79 (0.00-30.32)	20.88* (0.00-55.58)	16.83 (0.00-33.66)
	HMD	38.16* (19.08-57.24)	15.16 (0.00-30.32)	48.72* (13.92-83.52)	35.53 (14.96-59.84)
4	LPS	0.00* (0.00-28.62)	0.00* (0.00-15.16)	0.00* (0.00-41.76)	0.00 (0.00-26.18)
	HMD	28.62* (0.00-171.72)	30.32* (7.58-121.28)	41.76* (27.84-222.72)	37.40 (14.96-187.00)

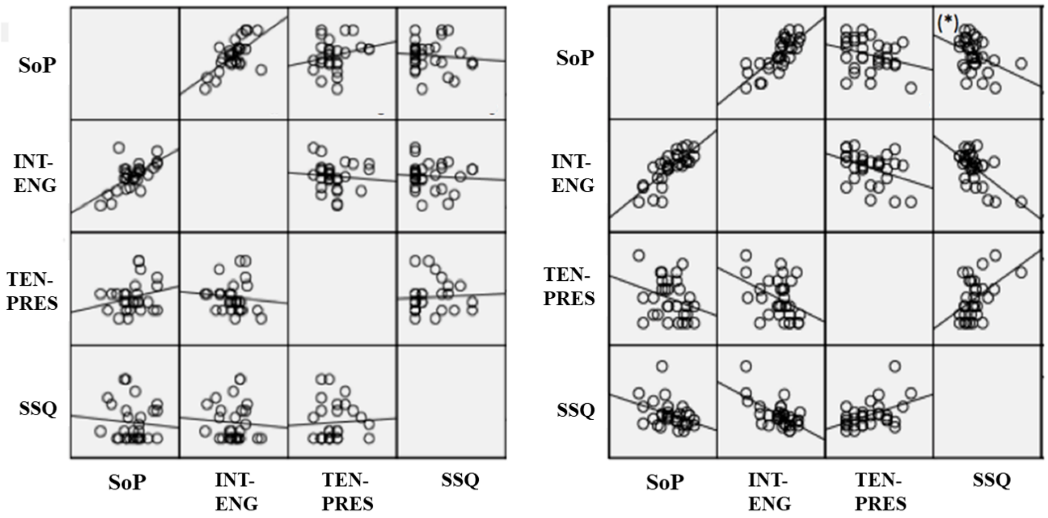


Figure 3.2: The correlations between the variables investigated in the study. Only SoP and cyber-sickness were significantly (negatively) correlated in the HMD condition (*). Adapted from [175].

Table 3.5: Comments labelled as realism.

Realism	
Both conditions	*"I believe there are missing elements which can make the environment more realistic: walking people, animals."* *"During the test, I felt detached from the ground."* *"After braking, the restart is not immediate."* *"The change in bike's direction is not so natural."* *"I would add the noise of the wind and some walking persons to make the experience more realistic."* *"There are too few living beings in the environment."* *"Between the two scenarios there is no continuity, and this is not realistic."* *"I did not like some aspects of the graphics, which made the environment less realistic."* *"The environment is not very detailed."* *"The movement of cars is unrealistic."* *"The city is aseptic and fake."*
LPS condition	*"In the first scenario, I could not look right and left and this made me feel less involved."*
HMD condition	*"It would be better to see the bicycle's handlebars, my hands and the other things usually included in the field of view when you look down."* *"Compared to the projector, vision with the HMD is sharper."*

Table 3.6: Comments labelled as interaction.

Interaction	
Both conditions	*"It would be nice if something happens at your passage (for example something that rolls and should be avoided)."* *"I wanted to change the bicycle direction."* *"The bike handlebars could not be turned; you could set the bicycle on a rotating platform to allow it to rotate."* *"I would have liked to receive physical feedback when the bicycle direction changes."* *"I would have liked to be able to choose the path to follow."* *"The control I had on the virtual environment was too limited."* *"In the second scenario I would prefer to have greater reality, greater freedom of movement, in other directions too."*
LPS condition	*"The change of the visual perspective with the joystick is too quick and unnatural."* *"The commands to change the point of view are unrealistic and increase the difficulty of the task."* *"It is not natural to control the point-of-view with the joystick."* *"It is impossible to handle the gaze direction with the joystick."*
HMD condition	*"Braking is not so natural and distracts from the experience."* *"Exercising with HMD is more realistic and immersive: the environment surrounds the person and thus its visual exploration is more immediate."*

Table 3.7: Comments labelled as involvement.

Involvement	
Both conditions	*"The exercise is repetitive."* *"I found the exercise boring and therefore not very engaging."* *"The bike handlebars could not be turned; you could set the bicycle on a rotating platform to allow it to rotate."* *"The urban environment is repetitive and does not help to feel part of it."* *"Environments would be more engaging with well-defined goals."* *"It is to be improved, but I found the experience very fun."*
LPS condition	–
HMD condition	*"The experience with HMD was engaging."*

The results of this comparative study were mixed. Outcomes showed that healthy young adults largely preferred the HMD, and that this device was able to convey higher SoP with respect to the LPS [180]. On the other hand, we found

Table 3.8: Comments labelled as physical drawbacks.

Interaction	
Both conditions	*"I did not feel comfortable when the bike curved without my control."*
PS condition	*"I had headache."*
	"I had a bit of dizziness."
HMD condition	*"I had nausea and dizziness."*
	"I had a slight sense of nausea."
	"During the curves the trajectory was fragmented and this caused malaise."
	"The background was somewhat blurry and it caused a slight annoyance."
	"The device is heavy."
	"I felt good on the straight path, but not in the curves."
	"I did not feel comfortable, maybe because I did not see well."
	"When I brake, I feel myself pushed out of my body."

that physical drawbacks were more pronounced in the HMD condition.

This allowed hypothesizing that, though cyber-sickness probably compromised (partially) the experience with the HMD, its related symptoms were not sufficient to overcome the positive feelings elicited by higher SoP, which eventually led the users to prefer the most immersive visualization device.

Different factors appeared to contribute to the relationship occurring between SoP and cyber-sickness. Among the others, Weech et al., in their recent review [19], identified some factors majorly contributing to the arousal of side-effects: *sensory mismatch* and *vection*, *display factors*, and *navigation control*. The first two factors are strongly related, because sensory mismatch is often the result of vection, which, by rendering the illusion of self-motion, often causes visual and vestibular cues to be in conflict. Such a conflict may have arisen during our study: the vestibular perception of lateral and frontal acceleration could not be reproduced in any way. This was noted by study participants, who underlined to feel worse during bends and brakes, i.e., when the expectations of vestibular feedback was higher. It is possible that this has also resulted in the interruption of the flow of the experience, and thus of the feeling of presence. Mismatched multimodal cues may, in fact, cause a sudden transition from the virtual to the real world [181].

Interruptions in presence (not only related to cyber-sickness), which were identifiable in our participants' comments (e.g., *"I could not turn the handlebars and this distanced me from the virtual environment"*; textit"I felt good on the straight path, but not in the curves"; textit"raking is not so natural and distracts from the experience"), have been already noticed to occur in previous studies. The idea of users realizing presence only when it is interrupted is currently discussed in literature [19, 182, 183]. These "breaks" in presence have also been suggested as an objective way to measure SoP [181].

Dealing with *display factors*, the literature has shown that visual display characteristics could influence both SoP and cyber-sickness [19]. In particular, stereoscopy has been demonstrated to cause cyber-sickness introducing a conflict between vergence (the distance of a virtual 3D object) and accommodation (its focusing distance). As a result, experiencing a VE using a HMD could result in an increased visual discomfort, e.g., eyestrain, blurred vision. The results of the present study confirmed this fact, since all group reported high visual discomfort in the HMD condition, and a few participants also explicitly mentioned that their vision was blurred. In addition to the vergence-accommodation conflict, it has to be considered that Samsung GearVR is not the headset currently providing the best quality images; higher frame rate and wider field-of-view would have probably improved the experience in terms of both SoP and drawbacks [19].

Regarding *navigation control*, and interaction in general, the design of the VE proposed in this study limited the possibility of moving freely around the environment. This is known to decrease SoP, as control represent one of the key factors contributing to presence [16]. Users highlighted this aspect in their comments, and reported the experience to be somehow "restricted" in both conditions. The fact that the participants could decide to move forward at their own velocity was not sufficient to elicit positive feelings, as complaints were almost the same occurring in other studies in which the participants were passive [110, 184]. On the other hand, the possibility of looking around with the HMD was appreciated ("[the environment] *visual exploration is more immediate*"), but this did not contribute to reduce sickness: SSQ scores were significantly higher for all groups.

Concerning other factors potentially influencing SoP and side-effects, we found no correlation. This was in contrast with previous studies showing that engagement, anxiety, sex, and gaming experience influence the quality of the experience [19, 185, 186]. Regarding engagement, it has been found that the more the participant feels involved and motivated in the experience, the higher is the SoP experienced [185, 186]; the same is worthy for anxiety, which appeared to be positively linked to SoP [187, 188]. Nonetheless, it may be possible that the type and the nature of the proposed scenarios influenced our outcomes, as in this study the whole experience was short, representing common daily scenarios, and not particularly touching, especially for young adults.

Regarding gaming experience and group allocation, we found no correlation of these two variables with SoP and cyber-sickness; this was also in contrast to what reported in the review of Weech et al. [19] and in the study of Ling et al. [189]. This may be due to the different way in which questions related to past experience with VR have been posed – as in this study just a simple and generic question was used –, and to the fact that washout period was not long enough to strengthen or weaken the feelings perceived during the first condition experience. Further studies are needed to assess exactly how this factor could influence SoP and perceived symptoms; the effect of gender has to be further investigated too. Evidence in literature are currently mixed, with some studies reporting females suffering more from cyber-sickness [190], and experiencing higher SoP [79], and others showing no differences (as in our case) [190, 191].

Summarizing our results, we may say that, in spite of the clear preference that the enrolled participants expressed for the HMD and the higher levels of SoP it elicited, the high scores obtained in the SSQ [178] must raise some concerns with respect to the feasibility of and HMD-based cycling intervention on older adults.

Some modifications to improve the design of the whole system, according to participants' suggestions, were considered essential for future studies. Bends and brakes were too sharp, and were identified as the main cause of malaise. Thus, eliminating immediate stops (after braking) and reducing bends curvature may work in reducing the onset of physical side-effects [192]. Moreover, it seemed plausible that adding some interactions with the environment could be useful to increase users' engagement, and consequently to reduce motion sickness [193, 192]. These pieces of information were considered for the development of a new training system, also based on cyber-cycling, for older individuals. Its features are described in Section 3.3.

3.3 CAVE-based cyber-cycling: a usability study on older adults

The results of the previous studies suggested that (1) older adults may enjoy participating in VR-based physical and cognitive training (§2.2.1, 2.3.2), but also that (2) wearing a HMD while cycling is not free from the risk of experiencing cybersickness (§3.2.2). Due to this, we did not believe that a HMD-based setup was the optimal way to provide older adults with a more immersive cycling experience. Consequently, we directed our attention toward another device able to convey immersive scenarios: the Cave Automatic Virtual Environment (CAVE).

Moreover, following the suggestions given by the young adults enrolled in the study described in §3.2, and the advice of the personnel of the Istituto Auxologico Italiano, we worked to design and introduce interactions within the VE.

In particular, as we wanted to frame this application in the field of cognitive interventions, we decided to transform the park scenario in a scenario in which performing Dual-Task (DT) training [25, 194].

3.3.1 Dual task training

DT is based on the *concurrent* administration of physical and cognitive tasks; it is motivated by the fact that the simultaneous execution of cognitive and motor tasks can cause a decline in the execution of one of them, or even in both, depending on the cognitive demand [27]. In patients with declared dementia, DT performance declines prior and faster compared to single-task performances, thus making DT deficits highly-specific and sensitive indicators of cognitive decline [59]. In fact, people with cognitive impairments show a significant decrease in motor performances during DT; maximal muscle power, postural control and gait are all affected by the execution of DT, and the extent of their impairment becomes more apparent as cognitive demands become higher [26]. DT training could thus represent a promising methodology not only to improve attention-related outcomes, but also to improve motor outcomes (e.g., gait parameters or fall risk), in patients with MCI and dementia [195]. These benefits could be of key importance for maintaining autonomy in ADLs: it is enough to think that successful locomotion requires the ability of performing simultaneously cognitive tasks, which may cause interference in the performance of gait (e.g., paying attention to moving people or vehicles, or reacting to a stimulus) [196].

Evidence of the effectiveness of DT training can be traced in many studies in literature, though populations undertaking those trials were diverse; positive outcomes have been obtained in older frequent fallers [197], post-stroke [198] and Parkinsonian patients [199], for both what concerned cognitive and motor parameters.

However, previous DT program were based either on the use of a treadmill [200], or implemented stepping-in-place strategy, e.g., on a balance board [201]. However, these two methods could not ensure that the training is free from any risk (e.g., slipping or falling). Thus, a cycle-ergometer, as in our case, may represent a better solution. Indeed, there are studies reporting lower risk of injury for cycle-ergometers than for treadmills, especially in case of elder and/or frail users [128]. In addition, the effects of training with walking or cycling-based exercises are comparable; both paradigms have been reported to improve balance, weight shifts, gait,and lower body extremity functioning [130, 129, 131]. Finally, the pattern of cycling can be easily be re-conducted to the gait pattern. They are both cyclical, involve the reciprocal flexing and extension of the leg joints, and cause the alternative activation of agonist and antagonist muscles [202].

The following paragraphs describe the study we performed in collaboration with Istituto Auxologico Italiano (Milan, Italy) enrolling a small sample of older adults with cognitive complaints, with the aim of assessing the feasibility of a DT intervention taking place in a CAVE, and using a cycle-ergometer; the developed system was named *Positive Bike* [25, 194].

3.3.2 Methods

Participants

For the usability and acceptance assessment of Positive Bike, 5 [1] older adults both with normal cognitive status and MCI were enrolled.

Though there was not the complete certainty that the proposed CAVE-based solution would be acceptable for this population, the previous experiences conducted in the Istituto Auxologico Italiano, and the fact of not having neither bends and brakes (as better described in §3.3.2), nor a device completely occluding the real world sight, supported the idea that conducting the study involving target users could be reasonably safe.

The only inclusion criterion was age $\geq$ 60. Individuals with motor issues, severe vision deficits, severe dementia (MMSE<19), or that were unable to provide informed written consent were excluded.

The sample was composed of 3 females and 2 males. The mean age was 70.0 ± 11.7, and the mean years of schooling were 11.0 ± 5.6. Mean MMSE score was 25.66 ± 3.31; 3 participants were in the normal cognition range (MMSE score$\geq$24), and 2 had MCI (MMSE between 19 and 23, according to MMSE criteria [61]).

Before the trial, all participants had to give their informed written consent. The study was approved by the Ethical Committee of the Istituto Auxologico Italiano

---

[1] Five is the minimum number of users required to perform a usability study [203].

and carried out in its premises.

Equipment

As mentioned, Positive Bike exploited a CAVE. Its application was developed using Unity rendering engine. MiddleVR Unity plug-in [204] provided the functionalities to make the VR application to communicate with all the CAVE system modules; it allowed projection on the CAVE walls, and the exploitation of motion data retrieved from the CAVE tracking system as inputs.

Within the CAVE environment, 4 stereoscopic projectors (Full HD 3D XGA DLP, Optoma) projected the VE onto 3 walls, plus the floor. The right-eye and left-eye images were combined together by active goggles, thus allowing for the perception of depth. A stereo-photgrammetric system [205] enabled tracking the position of passive reflective markers, and the correction of the spatial distortion of the simulated environment, which is eventually displayed with a 1:1 scale ratio (3.3).

For this study, we equipped both CAVE goggles and an X-Box joystick with an asymmetrical set of markers, thus allowing the system to recognize their position and orientation within the game area. This information was used to adjust the user's point of view, and to "transform" the X-Box joystick in a laser pointer allowing for the interaction with virtual 2D elements (i.e., buttons) displayed on the CAVE walls.

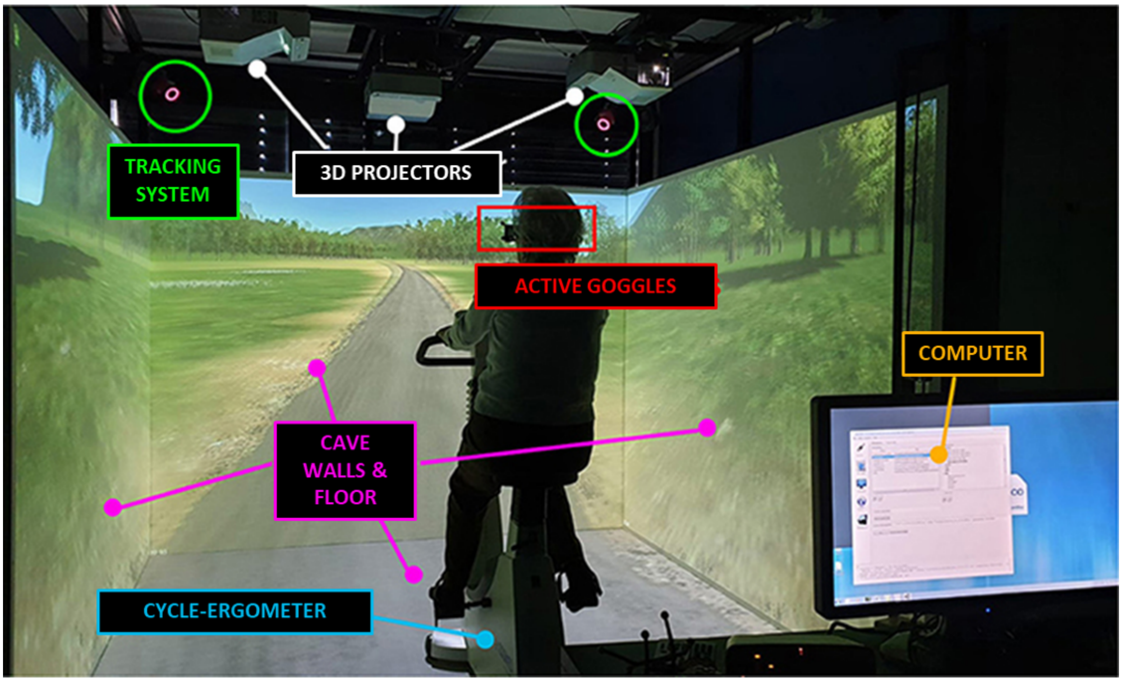


Figure 3.3: One of the study participants cycling in the CAVE [175].

The cycle-ergometer (Cosmed EuroBike 320) was placed in the middle of the CAVE. The control of bike velocity and workload was implemented as did previously §2.2.1. However, with respect to previous studies (§2.3.1 and §3.2.1), some modifications have been made.

First, the application was modified to implement a cognitive task. Such a cognitive task consisted in the identification of target elements appearing, among distractors, at the two sides of the cycling route. Targets and distractors could be animals or street furniture; in the case of animals, users had to identify all the animals whose name starts with a specific letter (e.g. for the letter "c", "cat" is a target, whereas "giraffe" is not). Instead, in the case of street furniture, participants had to look for a a specific a colour (Figure 3.4).

Second, a pushing button was anchored on the cycle-ergometer handlebars for the detection of the user's interaction, and an Arduino DUE board that was used to stream the button digital output to the computer. Participants had to press this button to signal the recognition of a target. The answer was considered valid if given before the target disappeared on the CAVE lateral wall. A green tick appeared on the target to signal a correct answer, whereas a red cross meant that the selected item was not a target; no feedback was given in case of no interaction.

Third, the path was made straight, and the naturalistic elements were simplified in order to limit the occurrence of cyber-sickness, and not to interfere with the identification of the targets, respectively.

Finally, also the written text showing the RPM to the user was eliminated. Feedback about the correct cycling velocity (between 50 and 70 RPM) were given via auditory feedback: acute and grave earcons communicated that the cycling velocity was too high and too low, respectively.

Figure 3.4: The two types of game in Positive Bike: recognition of the target animal, or of the colour of the furniture [25].

The application also allowed the therapists to perform the treatment customization. These were: the game type (animal/ street furniture); the target category (names starting with C/G/T/S; or orange/ blue/ yellow/ violet); level (2/3 targets each 30 seconds); workload (20/30/40/50 W); duration (15/20 minutes).

Study protocol

For this feasibility study, all participants followed the same protocol cycling 15 minutes with a fixed workload of 20 W. All participants were required to keep a constant velocity during the task and were told about the auditory signals' meaning. Additionally, all trained with *animal* task, letter *C*, and the lowest level of difficulty. To render the 3D scene appropriately, participants wore active goggles and an asymmetrical structure on which a few markers were placed, around their neck. The latter was used to estimate the user's point of view in the virtual scene.

Wearing it on the neck avoided the abrupt change of the point-of-view (for instance, in the case of head turning) and allowed to look at the target items also when they were displayed on the CAVE lateral walls.

Measures

The outcomes to investigate in this study were defined according to the requests of clinicians and psychologists from Istituto Auxologico Italiano. They also administrated the questionnaires, and conducted the analysis of the quantitative scales. The primary outcome was the system usability, i.e., the degree to which a specific person is able to use a given system to achieve specific goals effectively, efficiently and satisfactorily within a well-defined context of use.

- Usability was evaluated with the *System Usability Scale* (SUS): a questionnaire composed of 10 items with 5 response options ranging from strongly agree to strongly disagree [203];
- the *Short Flow State scale* [206] was used to assess the degree of engagement that patient had during the task and, indirectly, the perceived sense of control.
- finally, a *formative evaluation* was made via a semi-structured interview focused on 4 areas: usability, sense of presence, cyber-sickness [2] and expectations. The first two areas were further divided in sub-sections: respectively, utilization (effectiveness), learning (efficiency), and pleasantness (satisfaction) for usability, and spatial presence, engagement and realism for SoP.

3.3.3 Results

SUS median score was 76.25 (iqr: 25.65; q1-q3: 63.75 – 89.38) that according to Bangor et al. [207] was in the range in between *good* and *excellent*, and corresponds to a B score (range 76.2 – 78.8) [208]. Short Flow State scale median score was 5 (1; 4-5); all the domains investigated in the scale are shown in Figure 3.5.

The outcomes of the formative evaluation were divided into positive (Table 3.9) and negative feedback (Table 3.10). Two independent experts performed the classification; their agreement rate was satisfactory (Cohen's $k = 0.85$; standard error = 0.1).

3.3.4 Discussion

This study presents an attempt to implement a DT paradigm exploiting immersive VR for the administration of a cognitive task, and a cycle-ergometer for the performance of the motor exercise. We obtained promising outcomes, as the system resulted usable and user-friendly. In particular, according to what stated by Bangor et al., who tried to categorize SUS results, we could implement no modifications to our setup [207].

---

[2] It was not possible to administer the SSQ as the Ethical Committee refused to include a questionnaire that has not been validated in Italian yet.

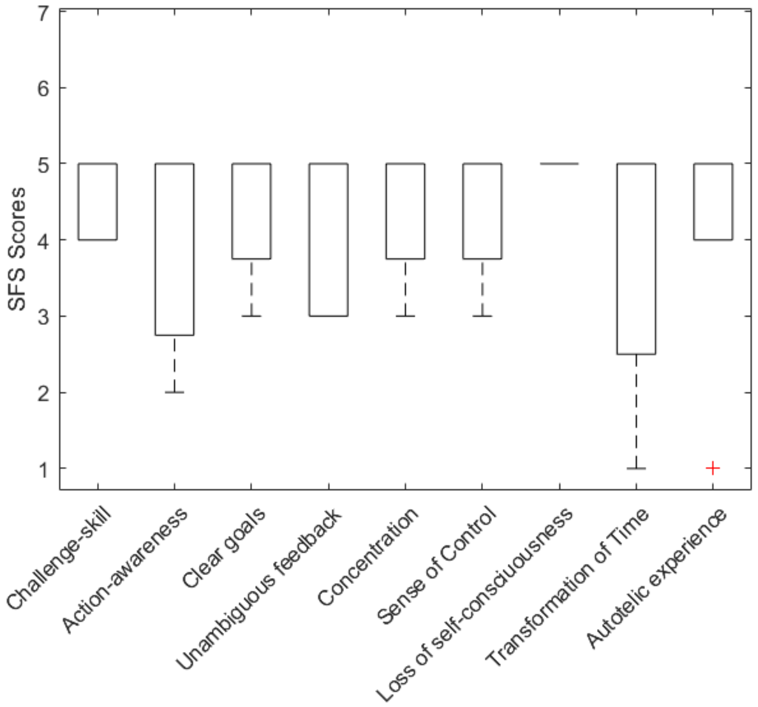


Figure 3.5: Short Flow Scale variables distribution.

Also in terms of flow, the system could be considered satisfactory; the results of SFS we recorded were comparable or higher than the ones reported in previous studies presenting physical training protocols for older adults [209, 210]. Csikszentmihalyi [103] described flow as the sensation that a person perceives when he/she acts with total *involvement*. Flow is also strongly related to well-being, as it emerges when there exists a good balance between challenge and personal skills, i.e., when a situation is characterized by high sense of *control*. Both these elements are contributors of SoP (§3.2.3), thus Positive Bike revealed able to generate this feeling in our study participants. Additionally, – looking at the single dimension of flow "loss of self-consciousness" that obtained a rating of 5 out of 5 ($SD = 0$) – one could argue that participants indeed forgot the context in which they were performing the exercise, i.e., the hospital and the rehabilitative program. Therefore, thanks to high SoP and the positive alienation from the context (and the fact that no adverse effects were reported), the use of Positive Bike as a DT training tool could be considered worthy of further investigations.

Though we did not measure cyber-sickness using a numerical survey instrument (SSQ), and thus we could not compare the results of this study with previous ones, we did not record any negative feedback in terms of nausea, visual symptoms, or dizziness. CAVE may thus represent a valuable way to administer cognitive interventions with high levels of SoP, even if including a navigation task.

Since the design of Positive Bike was novel, no comparisons could be made with other systems. We traced only one study involving a cycle-ergometer-based system, and reporting the results of an usability test in older adults. In such a work, Holland et al. [211] described a tablet-based application dedicated to the

Table 3.9: Positive feedback after the CAVE experience [25].

Topic	**Sub-topic**	**Feedback**
Usability	Utilization	*“Both the motor and cognitive tasks were easy.”*
	Learning	*“There was no problem in learning the use of the system.”*
	Pleasantness	*“The 3D glasses were comfortable.”*
		“The cycle-ergometer was manageable.”
		“The environment was beautiful.”
SoP	Spatial presence	*“The feeling was to be in the real park.”*
		“I had the feeling of being suspended.”
		“The environment was relaxing.”
	Engagement	*“I was focused on the task.”*
		“I think I have been pedaling for 5 minutes.”
		“I have forgotten you (the examiners) were here too.”
	Realism	*“The environment was realistic.”*
Cyber-sickness	Side-effects	No one reported side effects.
Expectations		*“This system could be useful for several types of patients.”*
		“I think it would be easier to train with this tool.”

physical training of chronic obstructive pulmonary disorders patients. Thanks this application, the users were able to take part in a video-conference with other peers (and the therapist) while cycling. This system was rated $SUS = 81$, i.e., obtained a score that was closer to ours, although slightly higher.

We may hypothesize that this difference was due to the higher familiarity that people generally have with tablets, whereas CAVE was surely a pretty new technology for everybody. Additionally, the active involvement of the users was probably higher in the case of a video-conference, also considering that patients were directly supervised by a clinician; this fact could have contributed in enhancing the perceived system usability too [212].

In spite of positive outcomes, we also believed that some aspects of Positive Bike could be improved, also according to participants’ criticisms 3.10). We collected few complaints regarding animals’ characteristics, which were told to hamper their clear recognition. First groups of comments were related to their size. Animals with a small size (e.g., frog, owl) were placed on a rock; however, this was judged not sufficient to allow our participants to recognize them as a potential target. Secondly, issues emerged when the animals were rotated; third, some targets were judged difficult to be recognized per se (e.g., two participants named a zebra “horse”, and a turkey “swan”); lastly, earcons were not always clear, and users found difficult to remember what they meant, especially at the beginning of the experience.

Table 3.10: Negative feedback after the CAVE experience [25].

Topic	**Sub-topic**	**Feedback**
Usability	Utilization	*“It is difficult to recognize small animals.”* *“Some similar animals can be confused.”* *“It is not easy to identify animals rotated backward.”*
	Learning	*“The sound of the bike might be confused with the feedback about speed.”* *“When frequency increased the exercise becomes more difficult.”*
	Pleasantness	*“Animals are repetitive.”*
SoP	Spatial presence	*“I had the feeling that animals bumped me.”*
	Engagement	*“I felt passive in the environment.”*
	Realism	*“The environment was nice, but did not look very real.”* *“Some animals were out of context.”*
Cyber-sickness	Side-effects	One patient was tired before the end of the task.
Expectations		*“There is no difference between this type of treatment and another.”*

To try and solve these issues, the following modifications may be of help. First, the animal target should always be placed along the route to face the participant; second, small animals should be replaced with others, bigger in size, and with distinguishable features (e.g., deer, goat, etc.). All the potential sources of confusion should be eliminated.

Finally, beside these changes to the software application, the whole intervention has to be enriched with a training phase, whose aim is to make participants to familiarize with their targets, and with the application's auditory feedback. Introducing such a training could be of help also for including in the study older adults with more severe cognitive deficits.

The main limitation of this study was the small sample size; this clearly affected the generalizability of the results we obtained, but the homogeneity of the data about the experience allowed hypothesizing that no serious criticisms would emerge in future studies.

In the near future, the above-mentioned issues will be corrected, and a training phase during which participants would be better instructed about the tasks and the types of feedback will be arranged. A protocol for further experiments has been already designed, and is available in Pedroli et al. paper [213].

Future works include also the integration of sensors to monitor the person physiological status during the training; a heart rate or a breath rate monitor could be easily integrated in the setup to ensure patients' safety throughout the sessions (even if they will be always supervised), and to measure their potential progress during the training program.

3.4 Conclusions

This Chapter has presented two attempts of building systems exploiting immersive VR to enrich the experience of physical training in the context of multidomain cognitive interventions. The aim of these studies was finding a proper design able to elicit high levels of SoP in the participants, while balancing for possible side-effects of immersive technologies (*Aim 2a*, §1.5).

It emerged that the CAVE-based experience was acceptable for older adults, thus informing about the potential feasibility of a structured training program including this type of activity.

On the other hand, the use of the HMD still had to be considered carefully, because risks of arousal of cyber-sickness may be higher. Nonetheless, it remains worthy to continue investigating the use of HMDs, as it has been shown to promote higher SoP. Also, increasing the quality of the device, reducing bends and slopes, introducing tasks requiring active participation for the users, and limiting the velocity while moving forward could be all ways to further reduce the occurrence of cyber-sickness symptoms [108].

Chapter 4

A HMD-based environment for cognitive training: acceptance and evaluation of naturalness in an immersive supermarket

4.1 Introduction

Having found a good compromise for the accomplishment of immersive VR-based physical or DT training, the second objective (in terms of increasing SoP) pursued in this thesis was related to the implementation of an immersive scenario able to reprise the characteristics and the aims of the shopping task presented in §2.2.1, i.e., an activity able to train MCI patients' visuo-spatial abilities (*Aim 2b*).

As mentioned in Chapter 3, an immersive supermarket – contrary to the park environment – could potentially exclude the navigational component that is often responsible of the occurrence of side-effects. Nonetheless, it would be better exploited in a standing position, thus introducing the risk of trips and falls (also due to increased body sway normally occurring during the "complete" immersion [176]).

Given this, and the release of new and high-quality HMDs that also allowed for the interaction with virtual objects, an attempt to try and make the supermarket scenario more immersive, and thus to exploit the potentialities of higher SoP also for cognitive training, was judged worthy to be made. Contrary to the Samsung GearVR used in §3.2.1, the more recent VR sets had also controllers to be held in the hands permitting the navigation (using teleport), and the interactions in the virtual scene.

As for the cycling task, a first preliminary study aimed at evaluating the usability of the proposed system and potential software criticisms, the onset of potential side-effects, and the quality of user-experience was needed. A preliminary study enrolling healthy young adults was thus performed [24](§4.2).

In addition, prior to conduct an acceptance study on older adults with MCI (§4.4), we performed another study aimed at assessing whether the interactions within the immersive VR environment we developed occurred naturally. In fact, due to system limitations, interactions in VR often require the user to act differently than how he/she would behave in natural world: this is also the case of controllers used for our application. These unnatural responses may affect a user's SoP and, above all, user's performance, both in terms of single objects' manipulation, and in terms of the achievement of the general task aims.

These lower performances are the consequence of the higher cognitive load required by unnatural interactions; each task draws the from a pool of (limited) cognitive resources (as for dual-task training §3.3), those resources cannot be used for a concurrent, and possibly more important task, as the completion of a cognitive exercise. Additionally, when interactions become complex, as in the case of immersive VR (i.e., they require body movements), the enactment of unnatural interactions could also result in awkward body postures and in postural fatigue.

As a consequence, we performed also a study aimed at assessing the naturalness – in terms of kinematics – the reaching gesture in a simplified version of the immersive supermarket (§4.3).

4.2 The Virtual Supermarket: a usability study on healthy young adults

4.2.1 Methods

Participants

Eight volunteers were enrolled among the employees of the Italian National Research Council. Exclusion criteria were: severe visual, cognitive and/or motor impairments, and the inability to provide informed written consent; the only inclusion criterion was to be aged less than 40. All participants provided their informed written consent to participate in the study.

Equipment

The VE had the features of a real supermarket, and it was constituted by two scenes: the *shelf* and the *cash-register* scene. The first scene represented a supermarket aisle of 4x3 meters. With respect to the previous 2D version (§2.2.1), the products were arranged onto two shelf units placed along the two longest walls. The aisle task was removed to avoid the need of navigating in the environment using teleport (the tracked gaming area in which the user could walk was too limited); signs were placed perpendicularly to the shelves to indicate on which side the product to pick was (Figure 4.1).

A shopping list containing 8 grocery items to buy was displayed above the controller in the non-dominant hand. The other controller was thought to handle all the interactions with the virtual objects, i.e., pressing buttons on the User Interface (UI), and grabbing, dragging and dropping 3D grocery items (Figure 4.2). Controllers were always visible in the scene, no other proprioceptive feedback (e.g., avatar, hands) was implemented.

To interact with UI buttons, the user had to press-and-hold the controller's track-pad with the thumb: this generated a pointing laser; when the laser-pointed

Figure 4.1: An overview of the shelf scene [24].

to the button area, the user could click on the button by pressing the controller's back trigger. Instead, for handling operations with 3D virtual objects, the user had to intersect the item with the controller; keeping the back-trigger pressed allowed to drag the object around. Conversely, releasing the back trigger caused the item to drop according to the laws of physics (e.g., apples rolled on the floor).

Figure 4.2: A screenshot showing the two controller while playing. The list was displayed on the controller in the non-dominant hand; the controller in the dominant hand was used to pick products.

Users had to drop all the products on the list into the shopping cart, which was placed on one side of the aisle, below the signs. Participants were left free to put in the cart whatever they want. The correctness of the shopped item was signalled the item's name turning to green on the list, and a tick appearing next to it (Figure 4.2). Follow the items order was not mandatory.

When all the items on the list were placed in the cart, a UI appeared asking the user if he/she wanted to proceed with the payment: by selecting *Continue*, the user was brought into the cash-register scene. In this scene, the aisles were not present anymore, but a cash-register with desk-tape appeared. Participant had to place all the items they bought in the previous scene on the desk, and select the exact amount to pay. Three banknotes and 3 coins were presented to the user (Figure 4.3): he/she had to select them using the same interaction modalities described for UI buttons. To ease the process of paying, each time the user selected a banknote or a coin, its value was automatically subtracted from the displayed amount. The initial amount was generated randomly.

After the conclusion of the payment, another UI requesting the user to continue with another shelf task appeared.

Figure 4.3: The cash-register scene. After having placed all the products on the tape, the user had to select the right amount to pay [24].

Interaction data was collected during the exercise, and saved in a XML file; such data comprised the total time needed to complete each shopping list, as well as the time employed to place the items on the belt. For the shelf scene, we recorded the following errors: *Wrong item Error* (WE) occurring when an item was placed in the cart, but was not on the shopping list; *Dropped item Error* (DE), and *Fallen item Error* (FE), occurring when a product was dropped and recollected, or dropped and left on the floor, respectively. Mistakes committed during the payment were also saved (*Cash Error*, CE).

The application was developed using Unity and deployed for HTC Vive [214] using Virtual Reality ToolKit (VRTK)[215] and SteamVR [216] dedicated plug-ins. More in details, VRTK provided a collection of scripts that simplified the object interaction within VR environments (i.e., it made objects "grabbable" by the attachment of the script *InteractableObject*), while SteamVR allowed for the deployment of the developed application for the HTC Vive VR set. The HMD was equipped with room tracking units (infrared cameras) and two controllers [214]. This specific setup enabled the user to move in the physical space (4x3 meters) while being immersed in a virtual scene.

Protocol

Before starting the experience, one of the investigators provided the participants with information regarding the tasks, and the use of the controllers. Participants with minor visual deficits were left free to wear or not to their glasses under the HMD. All were helped to wear the HMD, and adjust the lens-to-eye distance. In the shelf scene, the shopping lists and items locations on the shelves were the same for all participants. The sequence of shelf and cash-register scene was repeated for 15 minutes.

Measures

The experience of using HTC Vive in the Virtual Supermarket application was evaluated using both quantitative and qualitative data. The main outcome was the usability of the system, evaluated through the *System Usability Scale* (SUS) [203] (also used in §3.3). In addition, participants completed:

- the *Simulator Sickness Questionnaire* (SSQ) [178] (§3.2), and
- the *International Test Commission - Sense of Presence Inventory* (ITC-SOPI) [217], a questionnaire investigating SoP and focused on users' experiences of media, with no reference to objective system parameters. It was composed by 4 sub-scales investigating *spatial presence* (the sense of placement in the mediated environment, and of control over it; 23 items), *engagement* (the sense of being psychologically involved in the virtual experience; 18 items), *naturalness* (the sense of that the mediated environment is lifelike; 8 items) and *negative effects* (the adverse reactions to the presented environment; 6 items).

Participants' comments during and after the exoerience were recorded by the two investigators observing the trials. Total and partial task timings, WE, DE, FE, and CE were collected and stored automatically by the application.

Statistical Analysis

Matlab R2018a Statistical Toolbox was used to perform the statistical analyses. All the variables were checked for normality, with the exception of SSQ that was known to have a skewed distribution [178]; before running correlations, SSQ data were transformed using a square-root transformation [218]. The correlation between usability and cyber-sickness, and between usability and SoP was evaluated using Pearson's correlations.

Repeated measures ANOVA was performed to estimate the difference between trials in terms of time employed to complete the shelf and the cash-register scenes. For the latter, only the positioning of objects on the tape was considered, as the amount to pay was generated randomly, and thus not comparable.

Users' comments and specific behaviours of interest were recorded and analyzed using the thematic analysis method [149].

4.2.2 Results

Demographic data

All participants were graduated and had a good computer expertise. Other demographic data are reported in Table 4.1.

Table 4.1: Characteristics of the sample. (= one participant tried the Oculus Rift, two the Samsung Gear VR) [24].*

Participants	N=8
Age [yrs]	28.75 ± 3.65
Sex [M/F]	5/3
Past experience with HMDs* [Y/N]	5/3
Familiarity with VR	
none	2
sufficient	4
good	2
excellent	0
Grocery shopping	
never	1
once or twice a month	2
less than 50% of days in a month	4
more than 50%	1
everyday	0

Subjective data

Usability. The Virtual Supermarket obtained a score of 81.56 ($SD = 7.19$) that, according to Bangor et al. [207], corresponded to the *acceptable* range, to a grade equal to B. It could be also interpreted as a usability score between *good* (SUS score = 70) and *excellent* (85). The lowest obtained score was 70, the highest 92.5.

Cyber-sickness. SSQ-TS median was 18.70 (q1-q3: $3.74 - 38.34$). Sub-scales frequencies are reported in Figure 4.4; their median values were 4.77 ($0 - 11.93$) for nausea, 22.77 ($0 - 39.80$) for oculomotor disturbances, 20.88 ($10.44 - 34.8$) for disorientation. Cyber-sickness and usability did not correlate.

Cyber-sickness subscales were correlated, as expected; in particular, SSQ-O correlated with SSQ-D, and SSQ-TS ($r = 0.71, p = 0.049$, and $r = 0.96, p < 0.001$, respectively); SSQ-D correlated with SSQ-TS ($r = 0.71, p = 0.014$). The only SoP sub-scale significantly correlated with SSQ score was the side-effects sub-scale, which correlated with SSQ-O ($r = 0.79, p = 0.018$) and SSQ-TS ($r = 0.84, p = 0.01$).

Sense of presence. Results of ITC-SOPI subscales are presented in Table 4.2. Comparison with previous studies' results [219, 217] are also presented in the same Figure. No correlation was found between SoP and usability. In between ITC-SOPI subscales, *spatial presence* and *naturalness* were strongly correlated ($r = 0.88, p = 0.004$).

Objective data

On average, each participant completed the purchase of 40 items, corresponding to the completion of 5 shopping lists (range $4 \div 7$). As 4 was the minimum number of

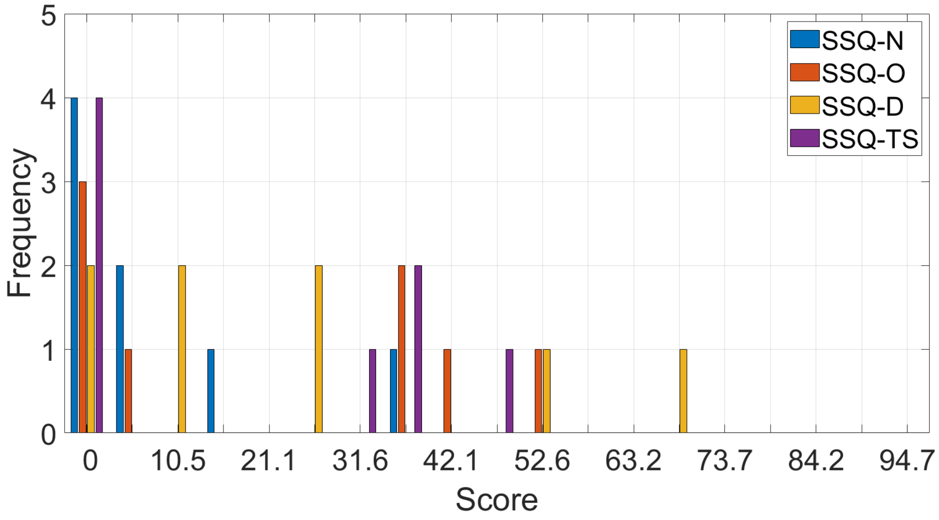


Figure 4.4: Simulator Sickness Questionnaire scores's distribution for total score and sub-scales. N = nausea, O = oculomotor disturbances, D = disorientation, TS = total score.

Table 4.2: Scores of ITC-SOPI sub-scales. Data is presented as median (iqr); also min and max values are reported.

Sub-scale	**Score**	**Range**
Spatial Presence	3.88 (0.41)	3.29-4.65
Engagement	3.85 (0.39)	3.31-4.38
Naturalness	4.10 (0.40)	3.60-5.00
Side effects	1.67 (0.01)	1.00-2.83

trials completed by all the participants, all the statistical analyses were computed considering the first 4 trials for each participant.

Execution times. Regarding the time taken to complete the shelf scene (Figure 4.6, repeated measures ANOVA was statistically significant ($F(3, 21) = 18.19, p < 0.001$). Post-hoc tests confirmed a significant difference between the first and the second trials' execution times ($p < 0.005$), the first and the third trials ($p < 0.05$), and the first and fourth trials ($p < 0.005$). Also for the cash-register scene, we found the same pattern; the first trial ($F(3, 21) = 15.63, p < 0.001$) was longer than the second ($p < 0.001$), the third ($p < 0.005$) and the fourth ($p < 0.05$).

Errors. Shopping items fallen and then recollected from the ground (DE), items remaining on the floor (FE) and wrongly-picked ones (WE) are shown in Table 4.3. For what concerns payment errors (CE), only one person selected the wrong amount to pay, and only once.

Users' comments and observations

Three different themes were identified using the thematic analysis: *control*, *visual drawbacks*, and *realism*.

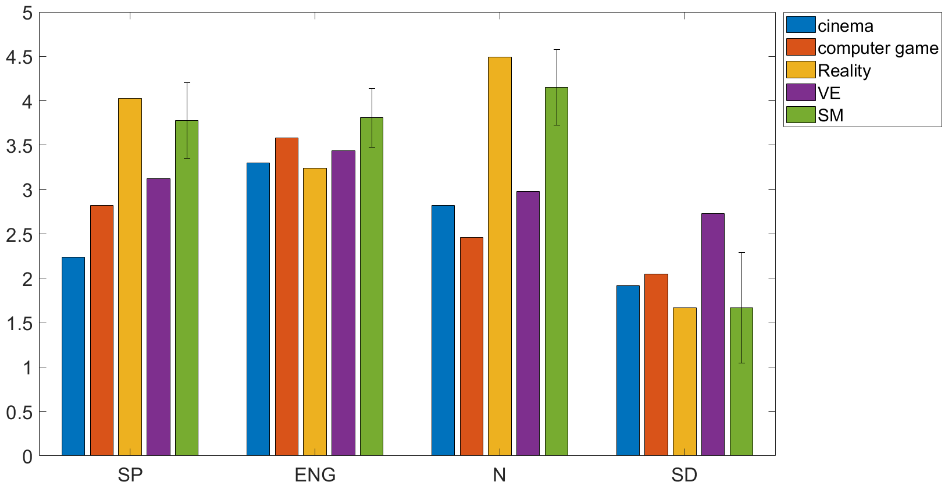


Figure 4.5: Comparison of ITC-SOPI results (mean and SD) with previous studies on different media samples. SP = spatial presence, ENG = engagement, N= naturalness, SD = side-effects; VE = virtual environment as in [219], SM = virtual supermarket.

Table 4.3: Number of errors in the shelf scene per trial. Data are presented as median value (q1-q3). Adapted from [24].

trial	**#1**	**#2**	**#3**	**#4**
WE	0.00 (0.00-0.00)	0.00 (0.00-1.25)	0.00 (0.00-0.25)	0.00 (0.00-0.50)
FE	1.00 (0.75-2.00)	0.00 (0.00-0.25)	1.00 (0.00-3.25)	0.00 (0.00-1.25)
DE	0.50 (0.00-1.25)	0.00 (0.00-0.25)	0.00 (0.00-0.00)	0.00 (0.00-0.05)

Control. Within this category, we listed all the issues related to the interactions with the shopping items and UI elements. The most common criticism was forgetting, just after the beginning of the test, how to interact with UI ("*What should I do to start?*"), and how to grab the objects. Six out of 8 participants forgot the instructions they were given just after the immersion in the virtual scene. It happened also that the track-pad and the back trigger functionalities were swapped: during the first trial, 3 users pressed onto the track-pad to grab the objects, instead of using the trigger.

One person commented that it was difficult to select the elements of the UI, because pressing two buttons simultaneously (the track-pad and the back trigger) was complex.

One participant said that it was awkward not to see one's arms and legs while moving around.

Visual drawbacks. In general, no one complained typical cyber-sickness symptoms; only few comments were addressed to visual quality. One of the participants had some issues with the positioning of the headset ("*When I look down, this* [the HMD] *slipped on my head, and I could not focus items anymore*"), and then complained about the quality of vision in the virtual scene ("*Normally, I see better than this!*"). Another participant repeatedly adjusted the HTC Vive position say-

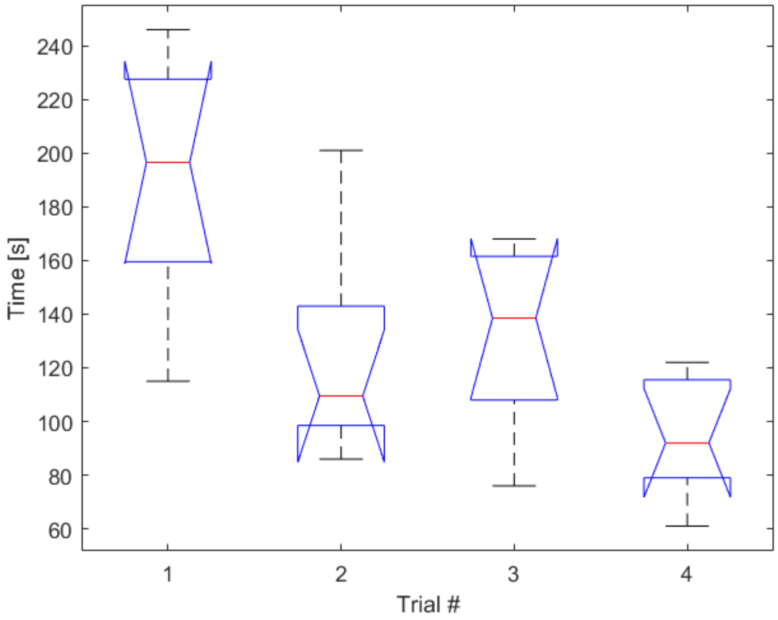


Figure 4.6: A box-plot presenting the time taken to complete the first four trials (shelf scene).

ing that the images were blurred; the same person performed the last trial holding the HMD with one hand.

Some people complained that the labels on some of the product packages were blurred. A person took the spaghetti box and moved it closer while commenting: "*I want to see if the label appearance would sharpen*", while another one complained that he could not find the flour package ("*The name on the box is not clear*").

Realism The majority of our participants' comments dealt with the realism of the two scenes. Regarding items' realism, almost all the remarks were directed toward fruits. Participants highlighted that many fruits were not proportionate ("*Oranges and tangerines have the same size*" / "*I cannot recognize oranges!*") and encountered difficulties in distinguishing them. One of the participants, while standing in front of the apple basket commented: "*I do not understand if these ones are apples, they look like plums!*". A couple of participants reported that it should have been possible to grab the cooking pan by the handle, as it usually happens in reality (in the VE, grabbing occurred in correspondence of the intersection between the controller and the item to shop).

A few remarks regarded bottles. A few participants commented about the lack of realism: "*The bottle fell* [on the floor] *without breaking up into pieces!*". Participants who never dropped a bottle, explicitly asked what would have happened in that case.

One of the participants wanted to recollect an item that rolled under the cart. After having thought about possible options, the participant opted for crouching and recollecting the item avoiding possible collisions with the cart. Clearly, it was also possible to pass through the cart with the controller, but the participant preferred to implement a natural behaviour, as he would have done in the physical world.

Other participants threw few shopping items in the cart. However, this behaviour was recorded only for small and light-weighted objects that – also in reality – would not break into pieces (e.g., apples, bread). No one threw fragile objects (e.g., glass bottles) with the exception of one participant who intentionally wanted to "*make a mess*".

4.2.3 Discussion

The results of this study showed a "more than good" usability of the Virtual Supermarket. It was notable that the lowest score belongs to the highly-acceptable range of usability, and that the standard deviation of the SUS scores we obtained was lower than what recorded in previous studies [208]. This allowed concluding that all participants agreed in judging the system *usable*.

Nonetheless, contrary to previous works, perceived usability was not correlated to any other variables; other authors found correlation between usability and SoP [220, 221], and between usability and cyber-sickness [222]. Also, the correlation between SoP and cyber-sickness was not present: only the side-effects' ITC-SOPI sub-scale correlated significantly with SSQ scores. This was in contradiction with what emerged in the review of Weech et al. [19], who found an inverse relationship between SoP and cyber-sickness. The difference in our outcomes may be due to the reduced sample size, or to the tasks proposed in the VE, which were very different from those mentioned all the above-referenced studies.

Though no correlations with usability were found, both cyber-sickness and SoP scores were satisfactory. Actually, SSQ scores are considered to be acceptable when $< 10 - 15$, which was not the case reported in this study [223, 178]. However, results may have been biased because of the small sample, and because of two participants, who ascribed high values to the items "difficulty focusing" and "blurred vision", both contributing to the oculomotor disturbance (SSQ-O) and to the disorientation (SSQ-D) scales. These 2 participants had issues in keeping the HMD in place during the experience, thus – after the conclusion of the study – we asked them to rate the quality of their vision after having (re)adjusted the HMD properly. In one case, the person reported improvements whereas, in the other, the participant admitted that his quality of vision was always low when using VR devices. Given this, and having obtained a low nausea score, SSQ results (together with participants' comments) were evaluated good enough to allow organizing an experimental campaign to evaluate the acceptance and the usability of the Virtual Supermarket in a sample of older adults with cognitive impairments (after few modifications, see §4.4).

In terms of SoP, our outcomes were positive. The comparison with previous studies (as shown in Figure 4.5) revealed that the Virtual Supermarket scores were comparable (in terms of engagement) or better (for spatial presence, naturalness, and side-effects) with respect to other immersive scenarios (e.g., the one proposed by Nisenfield [219]); our SoP was also comparable to the spatial presence, naturalness, and side-effects experienced in real world [219]. These results may be influenced by the chance of actively participating in a specific task within the virtual world: higher control could, in fact, contribute in creating higher SoP

[219, 16].

Taking into consideration objective measures, some observations could be made. First, usability was the only factor influencing the time needed to complete the shopping. This outcome confirmed what emerged in previous studies regarding the relationship between performance and perceived usability [146].

The first trial lasted significantly more than all the others, in both scenes. This allowed hypothesizing that, when scenarios were presented for the first time, participants needed more time both to explore the environment and to remember how to interact with the environment. Indeed, the majority of the participants had to recall how to use the controllers during the first trial. This behaviour, found also in previous studies, can be interpreted also as the result of a "wow" effect, i.e., the feeling that causes people to feel great excitement or admiration in the first instants of immersion [224, 225]. In any case, the fact that starting from the second trial, the time required to complete to shopping did not decrease anymore could support the assertion that usability was good. Also, no more than one attempt was needed to learn how to interact properly with the virtual scene, meaning that the Virtual Supermarket was indeed easy-to-use.

The issues related to the use of controllers, which we recorded during the first trial, were consistent with what emerged from Coldham and Cook's study [135], who used the HTC-Vive too. In that case, problems were possibly more stressed as the users were older adults, i.e., they are usually attributed less familiarity with immersive VR, and innovative technologies in generals [111]. To solve the issue related to the use of controllers, as suggested by our study participants, the interactions with UI elements, which occurred pressing two buttons simultaneously (the track-pad and the trigger), were simplified prior of conducting further tests.

The factor that was found to be more disturbing by the majority of our participants was related to blurred images. Seeing objects out of focus distracted participants from the experience and created an unpleasant sensation; this may have partially reduced the sensation of immersion, and thus their perceived SoP [16], whose level was however satisfactory throughout the whole experience. To try and solve this issue, more time should be dedicated to the adjustment of the HMD on the head of participants: letting them looking around, and giving them more time to adjust the distance of the lenses from the eyes, and then firmly block the HMD using the straps is fundamental to reduce blurring and slips.

Finally, we noticed that most of the participants' comments were about the *realism* of the VE: this fact underlined how important this aspect is to feel engaged in the proposed tasks. In spite of the requests of improvements (e.g., improve fruit quality, adding breaking glass effects), all participants agreed that the Virtual Supermarket had a good level of realism. This opinion was also confirmed by participants' behaviour, who avoided throwing or dropping fragile items, and avoided the collisions with virtual objects (i.e., the cart). This *avoiding behaviour* indicated that visual cues played a key role in the perception of a realistic environment. Indeed, the good visual quality of the scene helped participants to interpret virtual objects as they were real.

Given the promising results obtained in this usability study, the Virtual Supermarket was considered worthy of further investigations. No adverse events and no symptoms were recorded, with the exception of blurred vision, which was demonstrated to improve if more attention is paid while wearing the HMD.

We have thus worked to improve the interactions in order to ease the process of remembering commands, and of using the controllers (e.g., not having to double-pressing the touch-pad and trigger). Prior to the feasibility study performed enrolling older adults with cognitive complaints (§4.4), we have also conducted a study aimed at evaluating the naturalness of the interactions: its results are described in the following Section (§4.3).

4.3 Assessment of the naturalness of the reaching gesture in physical vs. virtual reality

As already mentioned, one of the main elements contributing to generate SoP while being immersed in VR is *control*, i.e., the control a person has over the task environment or in interacting with the virtual scene [16]. In general, the more the control, the greater the experience of presence. It appears evident that the feeling of *control* is strictly related to the usability of the system, and to the naturalness with which the interactions occur.

Therefore, the design of *natural interfaces* is an essential goal for the implementation of effective 3D interaction techniques. A key theme is whether interaction techniques can be designed to resemble physical world's. In an ideal case, users would be able to apply in VR the same skills they use normally for the accomplishment of their everyday activities.

A few studies have started investigating the issue of natural interactions in VR. Most of them were related to locomotion, as the simulated walking area is usually bigger than the real walking workspace, and many techniques to overcome this issue have been studied through years. Examples encompass treadmills [226], "hamster balls" [227], methods for deceiving users' perception by modifying sensory feedback, and metaphors [228].

In the case of the Virtual Supermarket, locomotion occurs naturally, as the walking area is restricted by the aisle shelves and walls, and fits the gaming area provided by the VR system. On the other hand, what occurs in a less-natural way is product picking; it is performed using the controller, thus hand movements occur differently from reality. Though this problem apparently could be solved using a data-glove, a previous study showed that the lack of haptic feedback and the consequent lack of the perception of the object weight impacted on SoP more than the metaphoric (but easy) interactions occurring when using the controllers [229].

Concerning the movement of the arm, previous studies have tried estimating the degree of similarity between movements performed in the physical and in the virtual world. Their aim was mainly evaluating the kinematic of arm movements for rehabilitative purposes (with the exception of [230]; see further in §4.3.1). In all the cases, however, one could assume that the more similar the movements were, the more natural the interaction was; thus, possibly, the less the cognitive demands would be required.

Within this context, we proposed a study aimed at investigating the kinematic differences of aimed movements between physical reality, and an immersive virtual environment resembling a shelf of the Virtual Supermarket (§4.2.1). The hypothesis underlying the study was that the HTC Vive VR set was able to return to the user appropriate (visual) feedback allowing the interactions – even if performed with the controllers – to occur in a natural way. If this is the case, the similarity of the movements performed in real vs. virtual environment would have been high, meaning that minimum cognitive workload was required to performed the product picking task.

4.3.1 Related works

A few studies investigated the similarities between reaching and grasping [231, 232] performed in real world versus a 2D virtual environment. Differences in terms of hand trajectory and movement times were found both for healthy people and post-stroke patients. The presence of such differences was attributed mainly to the lack of an appropriate perception of depth, due to the use of a 2D environment.

More recently researchers try to overcome this limitation by exploiting stereoscopic environments created thanks to active goggles or HMDs. Nonetheless, this factor alone was probably not enough, as few kinematic differences were recorded in healthy adults, and also post-stroke patients, also in these cases. Reaches in VR were slower [233, 31, 34], more curved [233, 31], less accurate [31], and elicited a different trunk displacement [34]. New hypotheses have thus been formulated to explain these outcomes. In particular, the following were the most discussed: the misperception of depth [231] that stereoscopy was able to reduce, but not to get rid of; the scarce familiarity with the VR technology [231], which caused the users not to rely on previous experiences' cues; the absence of haptics [233, 34].

As misperception of depth emerged also in study employing HMDs, it has to be highlighted that a couple of the above-mentioned studies [34, 31] employed a HMD whose diagonal field of view (fov) was restricted to 50 degrees. Devices that are currently on the market, as the HTC Vive [214] or the Oculus Rift [234] perform significantly better, having 150 degrees of diagonal fov and about 110° of horizontal and vertical fov. However, their performances are still limited with respect to the human eye, which counts on 120° and 200° of vertical and horizontal fov, respectively. Nonetheless, Furmanek et al. [233] obtained results that were not completely satisfying in terms of movement time even with an Oculus Rift v2; in contrast, reach-and-grasp synergies were conserved.

Finally, we noted that all these works made use of simplified environments that were precisely designed with the aim of being highly controllable. This was comprehensible as the authors have focused on the kinematic comparison of gesture features, but it also implied that the full potential of VR had not been exploited.

Given this, we hypothesized that a more ecological environment could be of help in eliciting more natural behaviours. Additionally, being immersed in a recognizable (and engaging) scenario could contribute in enhancing users' motivation, and thus possibly their performances [12, 15].

4.3.2 Methods

The study was designed as a within-subject repeated-measurements study, in which participants performed an aimed reaching movement, followed by the transport of the grabbed item, in 3 experimental conditions:

- the real world (RW),
- the real world while holding the HTC Vive controller (RWC) and
- the virtual reality environment (VR).

The RWC condition was introduced to investigate the effect of holding the controller while performing the task. For each participant, the order of conditions was randomized.

The study took place at the Sint Maartenskliniek in Nijmegen (Netherlands), and was approved the clinic's medical ethics committee. This study has been described in [29, 28].

Participants

In order to be eligible to participate in this study, people had to have a good cognitive status (MMSE>25); and be able to provide informed written consent. Motor and balance issues, severe vision impairments, history of seizure were considered exclusion criteria.

Participants were enrolled and grouped according to an age criterion: young participants were aged >18 and <40. Participants in the older adults' group were aged ≥65.

Equipment

A virtual and the real setup sharing the same features have been used. As the focus of this study was comparing aimed movements performed toward different specific targets, the Virtual Supermarket environment described in §4.2.1 was simplified with the dual aim of: (1) better controlling the variability of the whole scenario, and (2) to have the chance of building a comparable real setting. The whole environment was thus reduced to one shelf unit. Such a shelf unit was constituted by 3 shelves, on which we placed 3 products each; notches on the vertical bars allowed to adjust the heights of the shelves to match the hip (*bottom* shelf), trunk (*centre*) and head (*top*) level of participants. On each shelf, target items were placed in ipsi-, medial and contralateral position.

Figure 4.7 shows the comparison between the real and the virtual scenarios.

The interactions with supermarket items occurred as described in §4.4.1. As in the previous experiment, only the controller was visible in the virtual scene. A list indicating the order in which the participant had to reach for the 9 items was displayed on the side of the cart, toward the user. The cart could be placed on the right or on the left side of the participant, depending on his/her handedness. In RW and RWC, paper lists were printed and placed on a high table simulating the cart. In RW condition, this table was also the place on which "bought" items

Figure 4.7: The real (on the left side) and the virtual shelf units (on the right side) [29].

had to be left. The high table was preferred to a real cart because we did not want people to bend to place the item in the cart (as this did not occurred in VR condition). Thus, using an horizontal surface on which placing the real items at a height that was comparable to the superior edge of the cart appeared as the best solution to compare participants' behaviours during transfer phase.

For this experiment to be effective, the position of each target had to be the same in the 3 conditions. To ensure this, we made use of a stereo-photogrammetric motion capture system (VICON [205]) that allowed retrieving the exact position of reflective markers with a precision of less than 1 millimeter. In particular, we integrated the functionalities provided by VICON into Unity developing environment, thanks to VICON *DataStream* SDK [235]. This SDK allowed to stream position (and orientation, in the case of rigid bodies) from Vicon Nexus to Unity, and thus to exploit the position of real objects to align virtual ones accordingly.

For the data-stream to work correctly, it was necessary to align the two cameras systems' present in our setup (i.e., VICON infrared cameras, and HTC Vive base stations). This was made using *VR Alignment Tool* plugin [236], and then implementing an ad-hoc algorithm to adjust translations and rotations in the VE.

After having aligned the cameras' systems, and enabled the data stream, we used the positions of 4 reflective markers (1 for each shelf + 1 on the opposite side) to adjust the heights of the virtual shelves according to the real ones. Horizontal alignment was obtained by using notches on the shelves, and coding these distances in VR.

VICON system was exploited also to capture user's body movement during the experiment, with a sampling rate of 100 Hz. We used the Full-Body Plug-in Gate model [237].

Protocol

The calibration and the alignment of the cameras' system had to be performed at least once a day.

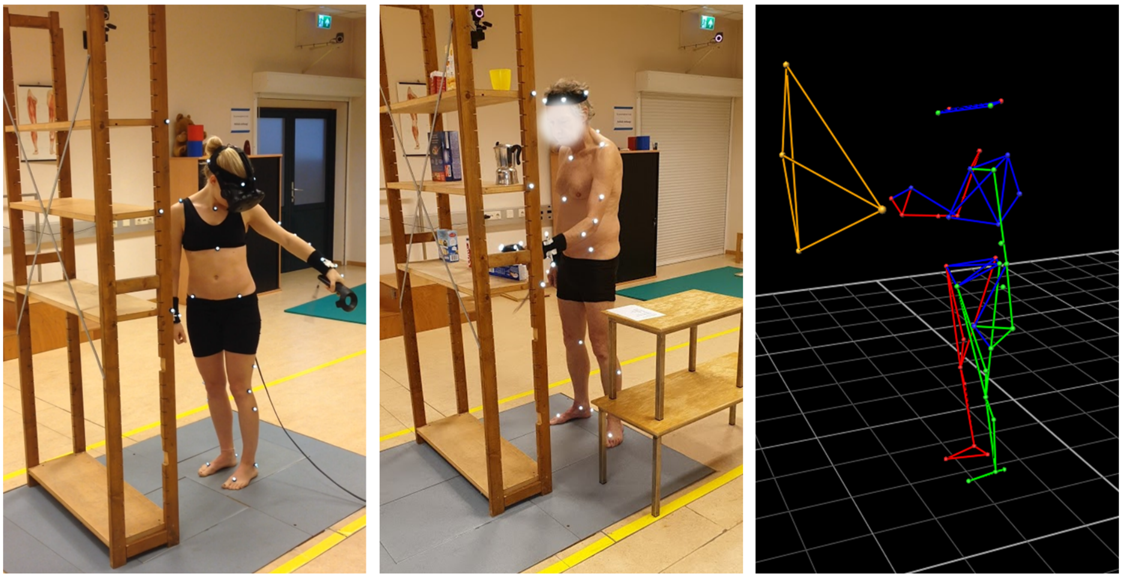


Figure 4.8: Participants performing the reaching and transport task in VR and RWC, and a screenshot showing participant and shelf tracking during the exercise [29].

All participants were required to wear short and tight-fitting clothes to facilitate markers' placement, and avoid measurement errors due to the slip of markers upon clothes. Markers were positioned as described in [237]; then, participants were asked to stand in front of the shelf (either real or virtual) with arms elevated in front of them. The distance at which their middle fingers touched the centre shelf was the one allowing to reach all items without stepping forward. They were also asked to try not to move their feet during the trials.

In each trial, participants had to reach and transport all the 9 items on the list, following the order. Such order was random, but chosen to allow the comparison among the 3 conditions of testing; i.e., if the first product to pick from the real shelf was on the top shelf, contralateral position, the first virtual object to grab was also on the top shelf, contralateral position, and so on for all the trials.

In RW condition, participants had to reach and transport the item adjusting the hand as they would normally do. In RWC, participants had just to use the controller to point toward the item to pick, press the back trigger, and then pretend to transport it on the high table (Figure 4.8, in the middle). Prior to VR condition, participants were told about the interaction modalities, then helped to wear the HMD, and to adjust the lenses. Then, they had to go through a familiarization phase, which foresaw the reaching of 6 items on the shelf. These items were different and placed in different positions with respect to the experimental targets. Giving the participants a little time to familiarize with the VE was considered useful to limit the "wow" effect [225], and to help remembering how to grab items using the controller (§4.2.2).

All participants had to complete 5 trials x 3 conditions, for a total of 15 trials.

If needed, they could sit and rest in between each condition. The whole experience lasted around 1 hour (comprehensive of marker's placement).

Measures

Movement Time during the reaching (MT_r) and the transport phase (MT_t) were the main study outcomes. They were calculated as the time elapsed from reaching movement onset till target reaching, and from target reaching to movement offset. Movement onsets (offsets) were fixed when the hand velocity (RFIN or LFIN, according to [237]) surpassed and remained above (or fell and remained below) 0.2 m/s [33]. Secondary outcomes were:

- *endpoint velocity*, i.e., the maximum hand velocity during reaching (v_r) and transport phase (v_t);
- *endpoint trajectory curvature*, defined as the ratio between the measured endpoint trajectory length and the length of a straight line connecting the hand position at movement onset and the target object.
- relevant *Ranges of Motion* (ROM) for the reaching and transport gesture, namely, shoulder flexion/extension, shoulder abduction/adduction, elbow flexion/extension [238]; backward tilt and rotation of trunk.

Statistical analysis

Collected data were analyzed by means of Matlab2019a scripts exploiting Statistics and Machine Learning Toolbox, and of VICON Nexus to extract joint angles, and to calculate markers' position derivatives.

Data collected for young adults ($n = 10$, §4.3.3) were analysed by means of repeated measures ANOVAs using *condition* (RW/RWC/VR) and target *position* (contralateral, medial, ispilateral, and top, centre and bottom shelves, Figure 4.9) as factors. Before, data were checked for normality using Shapiro-Wilk test; when normality assumption was not met, data were transformed using logarithmic transformation. Mauchly test was performed to assess whether the assumption of sphericity was met. Tukey-Kramer HSD post-hoc pair-wise comparisons were used to assess differences when significant main effects or interactions were observed.

For what concerns the data of older adults, the reduced sample ($n = 3$, §4.3.3) did not allow including these results in the multivariate analysis. However, to obtain a preliminary evaluation of the differences due to age, we compared the data of each condition (excluding position data that resulted statistically different in the previous analysis) using Wilcoxon rank sum test for unpaired samples.

4.3.3 Results

Ten (2 males, 8 females) healthy young adults aged 26.7 ($SD = 5.46$) and 3 healthy older adults (mean age= 69.0, $SD = 2.0$; 3 males) were enrolled for the study. All of them completed the experiment without complaining any issue. Means and standard deviations of all the computed variables are presented in Table C.1 for

Figure 4.9: Shelf positions for right-handed participants. C stands for contralateral, M for medial, I for ipsilateral; T for top shelf, C for centre, B for bottom.

young, and in Table C.2 for older adults (both in Appendix C). Shelf positions were coded as shown in Figure 4.9 for right-handed participants. In the case of left-handed participants ($n = 1$, in the young group), we always considered the target relative to the participant's sides (i.e., for a left-handed individual the target in the ipsi-lateral positions are the ones on his/her left).

Kinematic variables

For what concerns MT, differences between conditions emerged both in the reaching ($F(2, 18) = 17.06, p < 0.001$) and in the transport phase ($F(2, 18) = 6.10, p = 0.009$). In both cases, the only difference that reached statistical significance was VR (Figure 4.10, first row). RW and RWC showed no statistical differences for both MT_r and MT_t.

Also main effects of position were recorded for MT_r ($F(8, 72) = 2.67, p = 0.012$) and MT_t ($F(8, 72) = 2.16, p = 0.040$). Post-hoc analysis then revealed one significant difference for MT_r, i.e., between CC and CB (+20% for CC, $p = 0.02$), and one for MT_t, i.e., between MC and IC (+9%, $p = 0.03$).

Main effects of condition and position were present also for peak velocities: for V_r, we found a condition main effect with $F(2, 18) = 41.11, p < 0.001$ (Figure 4.10, second row, left), and a position main effect with $F(8, 72) = 2.68, p = 0.012$. For V_t, we had $F(2, 18) = 10.07, p = 0.002$ (Figure 4.10, second row, right), and $F(8, 72) = 9.04, p < 0.001$, respectively. Pair-wise comparisons of V_r for target positions highlighted no differences; instead, for V_t, we found significant differences between CB and all the other targets (except MC). In all cases CB had a lower peak velocity during transfer: from -12 to -29%, $p < 0.02$ in all cases.

In terms of curvature, ANOVA highlighted no effects (condition: $F(2, 18) = 2.36, p = 0.12$; position: $F(8, 72) = 0.76, p = 0.63$; condition*position $F(16, 144) = 0.52, p = 0.92$).

Joint angles

Not all the acquisitions were considered for the analysis because, in same cases, the occlusion of the sternum marker (STRN) resulted in the impossibility of computing angle joints.

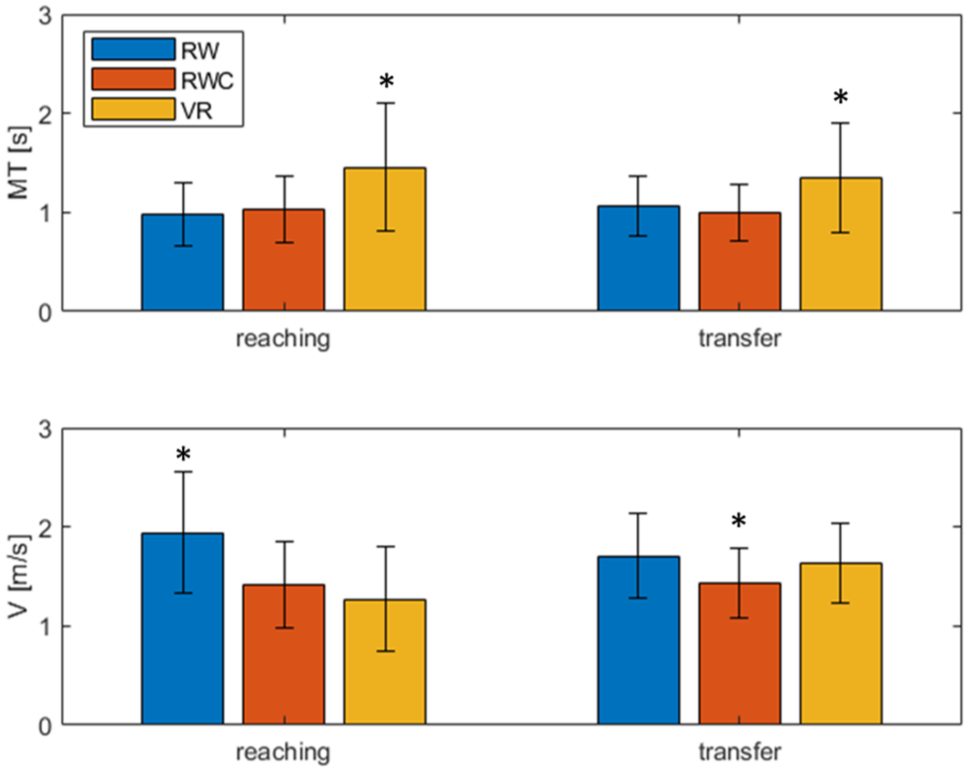


*Figure 4.10: Movement times and peak velocities during the reaching and the transfer phase. * indicates statistically different quantities ($p < 0.05$).*

The analysis of joint RoM disclosed that a few joints were influenced by condition (Figure 4.11). In particular, ANOVA resulted in a significant main effect for shoulder abduction ($F(2, 10) = 10.77, p = 0.003$), thorax tilt ($F(2, 12) = 5.84, p = 0.020$) and thorax rotation ($F(2, 12) = 32.209, p < 0.001$). For trunk only, also an interaction effect was recorded (for tilt: $F(16, 96) = 1.79, p = 0.43$, for rotation: $F(16, 96) = 3.52, p < 0.001$).

For what concerns trunk rotation, post-hoc analysis revealed difference reaching significance for all targets with the exception of MT (between RW and VR), IT (between RW and VR, and RW and RWC), and CB (between RW and VR, and RWC and VR). Instead, for tilt, only one significant difference emerged: CB target, between RWC and RW ($p = 0.043$). Looking deeper at condition effects, it emerged the thorax rotation was larger in VR, whereas thorax tilt and shoulder abduction were reduced in RWC condition.

Finally, for what concern the effects of position alone, we found the following:

- Shoulder abduction: $F(8, 48) = 3.53, p = 0.004$; target in CB was different from CT, MT, IT, CC. In all cases, CB required from 47 to 60% less abduction ($p < 0.040$ in all cases).
- Shoulder flexion: $F(8, 48) = 2.98, p = 0.009$; CB required 40% less shoulder flexion than IC ($p = 0.003$).
- Elbow flexion: $F(8, 48) = 3.17, p = 0.006$; for CB, elbow RoM was reduced of 34% with respect to IC ($p = 0.003$).

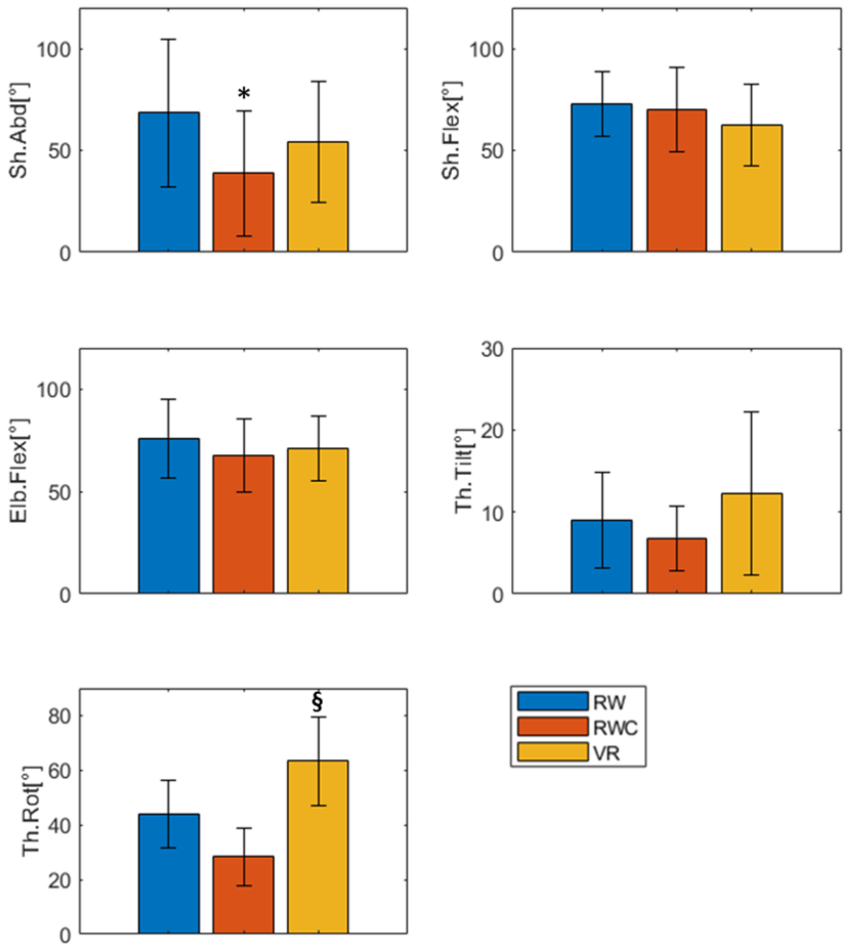


*Figure 4.11: Joints' RoM plotted across condition. Thorax tilt did not result in any difference at post-hoc analysis. *:* $p < 0.05$, *§:* $p < 0.001$.

- Thorax tilt: $F(8, 48) = 1.40, p = 0.21$.
- Thorax rotation: $F(8, 48) = 0.83, p = 0.57$.

Comparison between young and older adults

After the removal of the data corresponding to the target position(s) that resulted statistically different from the others, we compared the young population's data to older adults', in order to identify whether there existed differences depending on age. Results of condition comparisons are presented through box plots in Figure 4.12 for kinematic variables, and Figure 4.13 for joint RoMs.

4.3.4 Discussion

Virtual reality environments dedicated to cognitive interventions should guarantee that all the cognitive resources of the participants are used for the accomplishment of the main task. Therefore, it is mandatory that the actions performed in VE (i.e., grabbing shopping items) do not compete for the same cognitive resources.

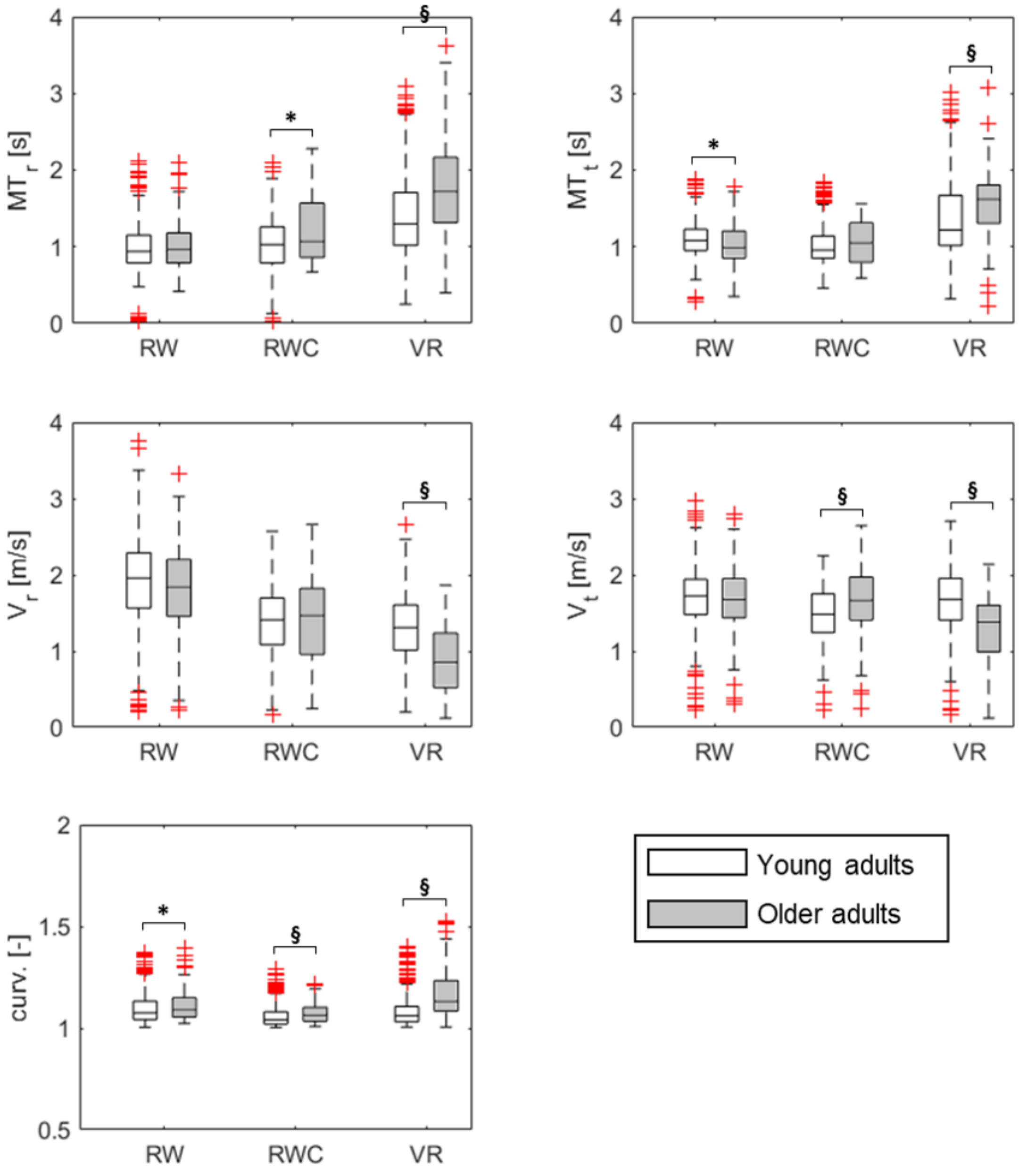


*Figure 4.12: Kinematic variables for young and older adults in the 3 conditions of testing (RW, RWC, VR). *: $p < 0.05$, §: $p < 0.001$ at Wilcoxon rank sum test.*

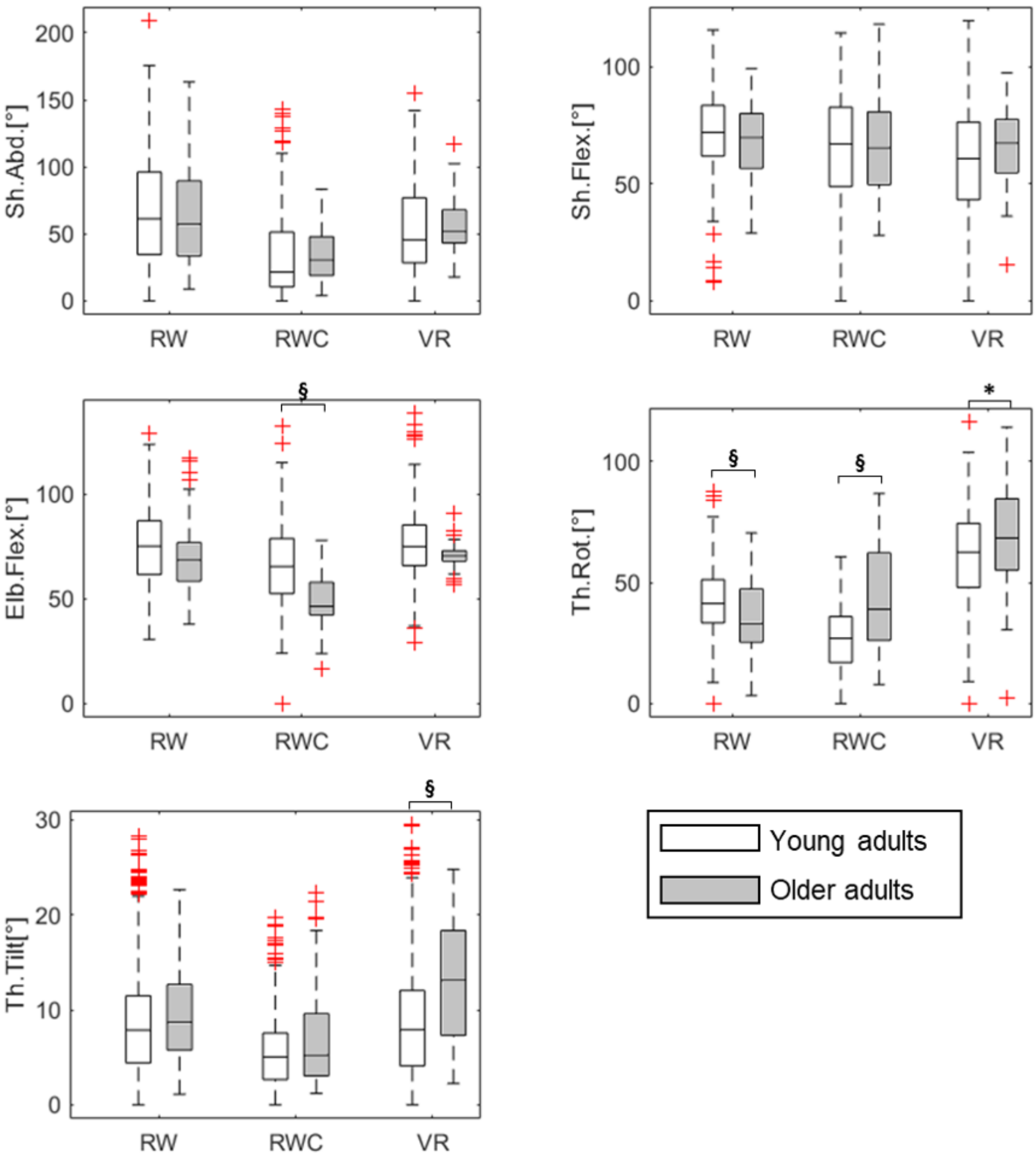


*Figure 4.13: Joint ranges of motion for young and older adults in the 3 conditions of testing (RW, RWC, VR). *:* $p < 0.05$*, §:* $p < 0.001$ *at Wilcoxon rank sum test.*

In such a context, the presented study aimed at comparing the kinematics of the reaching movements made in virtual and in physical reality, assuming that the more natural the interaction occurred – i.e., the kinematics of the movement made in VR was the same of the physical world –, the less cognitive resources were required to reach, grab, and transport the target object [30].

Regarding our main outcomes, we found results that were in agreement with previous works [31, 33, 34, 30, 35]: MT was longer in VR for both the reaching and the transport phase.

Regarding the reaching phase, we obtained results that were comparable with Knaut et al. [31], Stewart et al. [33], and Magdalon et al. [34], and higher with respect to what obtained by Furmanek et al. [233] (who use the more similar HMD). However, all the movements recorded in these studies were shorter than ours (less than 1 second), and also no search was foreseen. In our case, the use of setup requiring longer reaches, and, maybe, of an additional task (i.e., locating the target item) may have enhanced the differences existing among conditions.

Though in the previously mentioned studies it has been argued that one of the elements contributing to slower MTs could be related to the wrong estimation of distances generally occurring in VR [239], Magdalon et al. stated that if joints RoMs were not affected, wrong estimation of distances did not occur.

In contrast to what found by Magdalon et al. [34], in our study, the condition RWC resulted most of the time not statistically significant from RW. This meant that the fact of holding the controller had almost no effect on kinematic variables, whereas the cyber-glove used by Magadalon et al. was demonstrated to slow down the movements even in absence of VR.

The longer MT recorded in VR condition may be attributed to the lower peak velocity during the reaching phase (v_r) [32, 34, 35, 33], though probably more specific measures (e.g., number of peaks, principal component analysis) are required to confirm this statement. The peak velocity during the transport phase, instead, appeared not to influence the duration of the whole movement (with the exception of one target, CB). Thus, we hypothesized that the need of finding the physical space to place the grabbed item on the table had an influence. Indeed, the search of free space on the table occurred concurrently with respect to the movement, thus slowing down could have been of help to find a good spot.

Another hypothesis may be, again, that transporting required less attention and precision with respect to the product (searching and) reaching, thus that it resulted in the same peak velocity, irrespective from the condition. Nonetheless, these hypothesis would not explain both the difference for CB, and the significant difference between RW and RWC recorded for young adults. Also in this case, a more detailed analysis of the velocity profiles conducted on an extensive sample could be helpful to draw more evidence-supported conclusions.

In terms of trajectory curvature, our outcomes were in contrast with the ones obtained by Knaut et al. [31]. They found that participants exhibited more curved trajectory in VR, but only when reaching toward contra-lateral targets. They explained this behaviour arguing that it was more complex to reach targets at the edge of the participants' fov, especially if to do that, individuals had to cross the body mid-line. In our case, we found no difference for any of the targets; however, we may argue that none of the targets could be considered to be in peripheral fov, and also that the distance from the mid-line was limited (less than an arm-reach): thus, this may explain the difference of our results. On the other hand, the differences in MT_r for IB and CT, and for IB and IT may suggest that target items on the contra-lateral side required more time to be reached. Perhaps increasing the distance of ipsi-lateral and contra-lateral targets may contribute in highlighting more this effect.

Other studies reporting a more marked curvature were the ones making use of 2D VR systems [231, 232]; therefore, our results supported the hypothesis that stereoscopy is fundamental to elicit natural behaviours in VR.

Dealing with arm joints, we did not identify any difference in terms of RoM, with the exception of shoulder abduction in RWC condition. Thus, results were encouraging, especially assuming that no differences in RoM meant that no distance over-estimation was occurring [34].

Regarding trunk, we found that rotation was more marked in VR with respect to the physical world. This was in contrast to what reported by Magdalon et al. [34], who recorded less trunk movement, and increased elbow extension and shoulder abduction. In that case, however, participants were sat, and the experiment was designed to make the participants move only onto the horizontal plane.

Knaut et al [31], instead, had targets displaced on a vertical plane, as in our testing scenario. They found no differences for any joint, but, also in this case, the participants were sat, and their upper body movements were limited by the chair back support, as the study was designed to engage the arm. On the contrary, in our case, trunk rotation in VR was probably encouraged by the fact that participants were standing, and that the cart was on their side. Also, no clear dropping point was specified for product release.

Our commitment of recreating an ecological environment has influenced the control we had over the whole setup: future studies may try to better constrain the movement onset and offset points, while keeping high ecology in the environment.

Regarding the influence of the *age*, the small sample and the large SDs made all the obtained results very preliminary, and thus they must be considered cautiously. However, an interesting tendency appeared to emerge from this study, i.e., that age influences the movements performed in VR, both in terms of kinematic variables and joint RoMs. Older adults showed MTs and peak velocities that were mostly comparable to younger ones' when moving in the physical world; for VR, instead, we noted a statistical difference between young and older individuals. Also, trunk rotation and trunk tilt resulted different between the two age-groups only in VR. Few hypothesis may be formulated to try to explain this results, even though they should be confirmed by future studies enrolling more older adults.

The first explanation could be related to the quality of vision. There were studies showing that humans move slower when their fov is reduced [240, 241]. This was also highlighted in a study similar to ours [31]. As normal fov values for the human eyes are higher with respect to fov provided by the HTC Vive (120° vertical, 200° horizontal; for HTC vive, fov is 110° for both the horizontal and the vertical fov [242]), we may assume this fact had an influence. Additionally, in the case of older adults, this fact may have contributed even more, as it had to be summed to the vision-loss that was possibly already more marked in older adults [37]. In the future studies, evaluating sight abilities may be of help in order to clarify whether this hypothesis is true.

Second, we hypothesized that older adults had longer MTs as they were less familiar with immersive technologies. A study has proven that being less used to perform a certain task, caused the people to be less efficient in accomplishing it [38].

Future studies could try addressing these open points, especially considering that cognitive interventions are often dedicated to older adults. In order to improve the results and to reduce the outcomes' heterogeneity in future studies, we suggest introducing strategies aimed at increasing the visibility of both the targets and the reaching hand, in order to return a better proprioceptive feedback; adding haptics; and eliminating the need of searching items. If all the above-mentioned situations were the result of the perceptual uncertainty or additional cognitive resources used during target acquisition in VR, all these strategies should help to improve the comparisons.

Our study constituted just a small investigation in the field, and had many limitations. The main was the small sample, especially for older adults. Other limiting factors included the fact that gender was not balanced in any of the groups, and to the lack of clear constraints for all the movement phases. Also, the presence of shelf lateral bars could have introduced differences in the participants behaviours.

In spite of these limitations, however, we had found that joints' RoMs were mostly not affected, indicating that – even if occurring more slowly – movements performed in VR preserved the same synergies that were present in RW.

Further studies should better address these issues, enlarging the sample, and foreseeing longer familiarization phases. Also the modifications proposed above (e.g., enhance proprioception, haptic feedback, etc.) could contribute in eliciting more natural behaviours [34, 232, 233]. Future studies may also try to consider (1) tasks that are less or more demanding, to better evaluate whether VR influence task performance (2) to include objective variables in their measures (e.g., accuracy or precision, an more deep analysis of the velocity profiles, perhaps considering the number of peaks), and (3) the measurement of physiological signals (e.g., cognitive workload through EEG).

4.4 Usability and acceptance of the Virtual Supermarket in older adults

Given the good results of the preliminary study on healthy young adults, both in terms of good usability and low cyber-sickness, and the (possibly) acceptable cognitive load required by the Virtual Supermarket, we decided to conduct a feasibility study enrolling older adults with MCI or SDC.

Prior to the performance of the test, few modifications were made to the Virtual Supermarket described in §4.2.1 to improve its usability and to eliminate unwanted tasks (i.e., placing the products in the cart on the cash-register tape; see §4.4.1). As for the preliminary study, this study foresaw the experience of the Virtual Supermarket and the subsequent administration of questionnaires aimed at investigating the subjective user-experience [39].

4.4.1 Methods

Equipment

The hardware equipment was the same used in the preliminary study on healthy adults (§4.2.1). As already mentioned, few modifications had been made to improve usability, according to the suggestions given by the first study's participants. First of all, the UI were eliminated, thus eliminating the need of pressing both the track-pad and the trigger simultaneously. The only task in which the pointing laser was still needed, was the payment. Therefore, in the *cash-register* scene, the button functioning was changed as follows: the pressing of back trigger caused the pointing laser to become visible, and its release – if done while pointing at a coin or a banknote – triggered the selection.

Additionally, also according to the psychologists of the IRCCS Fondazione Santa Lucia (Rome, Italy), where the study was performed, the task foreseeing the placement of the products in the cart on the cash-register tape was eliminated. This was made to simplify the cash-register scenario, as such task did not comprehended any cognitive stimulation.

Finally, to further reduce the complexity of the tasks to be performed in the shelf scene, we limited the use of controllers to just one. This was made to avoid the participants to keep awkward body positions, and to limit physical fatigue. As a consequence, the shopping list that before was displayed on the controller in the non-dominant hand, was put on the cart, as shown in Figure 4.14.

Figure 4.14: The shelf scene. The list was displayed on the cart, so users could hold just one controller [39].

Participants

The same cohort of older adults examined for the study presented in Chapter 2 (§2.5) was considered for enrolment in this study. Since this was a usability study, the fact of having been included in another experimental campaign was not believed to influence any of the outcomes. Inclusion criteria were: age ≥ 60; having an impairment in one or more cognitive domain: MMSE ≥ 26, or MMSE

≥ 28 if years of schooling were greater or equal to 16, or a score lower than ≥ 1.5 standard deviations with respect to the age-, sex-, and years-of-schooling-matched normative sample. Exclusion criteria included MMSE$\leq$20, or significant functional impairment (FAQ<9 or loss > 20% of functionality in the Instrumental Abilities of Daily Living); presence of co-morbidities, neurological or psychiatric diagnosis; history of cardiovascular diseases; suffering from a brain damage, or seizure; history of alcohol or drug abuse; history of motion sickness; sensorimotor dysfunctions; previous experiences with immersive VR.

Written informed consent was signed before the experience with the Virtual Supermarket. The study was approved by the Ethical Committee of IRCCS Fondazione Santa Lucia.

4.4.2 Protocol

Before starting the navigation in the supermarket, participants were orally instructed about how to manage the interactions in the two scenes. The immersion lasted 15 minutes and consisted in the subsequent presentation of shelf and cash-register tasks. When the 15 minutes had passed, the game was automatically interrupted at the end of the scene. The difficulty of each list proposed to the participants was comparable.

At the end of the test, participants were administered questionnaires aimed at evaluating their subjective experience; objective outcomes were automatically collected by the application and stored in a XML file (§4.2.1).

Measures

The user experience in the Virtual Supermarket application was evaluated considering both subjective and objective outcomes. Participants were self-administered the following questionnaires:

- the *Simulator Sickness Questionnaire* (SSQ) [178] (§3.2), and
- the *International Test Commission - Sense of Presence Inventory* (ITC-SOPI) [217].

In addition, as the attitude of older adults toward new technologies may be not always positive, we foresaw the administration of the *Technology Acceptance Model* (TAM3) [243] questionnaire. Its goal was to verify whether being less prone to use technology influenced cyber-sickness and SoP. In particular, we used the following sub-scales from TAM3:

- *perceived usefulness* (PU), i.e., how much an individual thinks that exploiting the system of interest will improve his/her performances;
- *perceived ease of use* (PEOU), i.e., the extent to which the use of the system of interest requires no effort;
- *perception of external control* (PEC), i.e., the extent to which an individual believes to have the resources to use the system; question 4 of this sub-scale

was not asked, as it involved the use of two systems concurrently, and had no relevance for this study.

- *computer anxiety* (CANX), i.e., the apprehension or fear arising from having to interact with technology;
- *perceived enjoyment* (ENJ), i.e., the degree to which the use of the system is enjoyable.
- *output quality* (OUT), i.e., the degree to which the system is able to perform the task for which it has been designed for.
- *behavioural intention* (BI), i.e., the willingness of the individual to use the system in his/her future daily life.

The answers were given on a 7-item Likert scale where 1 represented "strongly disagree" and 7 "strongly agree".

Dealing with the objective variables, total and partial task timings, WE, DE, FE, and CE were collected and stored.

Statistical Analysis

Statistical analyses were performed using MATLAB 2019a Statistics and Machine Learning Toolbox. Significance level was set to $\alpha = 0.05$ for all tests. Cronbach's α was used to estimate the internal validity of TAM3 sub-scales, as we removed some items from the original questionnaire [243].

Since SSQ has a skewed distribution [218], correlations between SSQ sub-scales, ITC-SOPI, TAM3 were investigated using Spearman's correlation. Spearman's correlations were used also to assess the relation between SSQ and the sample demographic characteristics (i.e., age, MMSE, years-of-schooling, and gender). Pearson's correlation was used to evaluate the relation in between SoP and technology acceptance, and between these questionnaires' sub-scales and the sample characteristics.

Repeated measures ANOVA was performed to compare the time taken to complete each trial. Only the shelf scene was considered, since in the cash-register scene, the amount to pay was generated randomly, and thus not comparable.

Multiple linear regression was performed to investigate the influence of SoP and of technology acceptance on task execution times, and on errors. To do this, TAM3 scores and ITC-SOPI sub-scales' scores were used as independent variables, while the time needed to complete the task (i.e., to collect the products in shelf scene) or the committed errors in each trial were used as dependent variables.

Finally, assuming that the differences between the first version of the Virtual Supermarket (§4.2.1) and the one used in this study were not so relevant to have strongly influenced SoP and cyber-sickness results, a comparison of ITC-SOPI and SSQ scores obtained for young and older adults was made. T-tests for unpaired samples were used in the case of ITC-SOPI sub-scales, and Mann-Withney U tests were employed to compare SSQ sub-scales and total score.

Figure 4.15: Two participants performing the trial [39].

4.4.3 Results

Fifty-seven older adults aged 71.4 ($SD = 5.85$) were enrolled for participation in this study. Thirty-eight participants were females, 19 were males. The sample's mean MMSE score was 27.38 (1.88): 22 participants were affected by MCI, while 35 had subjective cognitive complaints. Average years of schooling were 13.07 (3.70).

The results of 2 participants have been excluded because, in one case, the person did not answer to the questionnaire, and in the other, the experience in the Virtual Supermarket had to be interrupted because the participant complained a little discomfort and did not want to proceed with the questionnaires. All the other participants were able to reach the end of the test. One adverse effect was recorded as one of the participants fell without consequences during the trial. This person after admitted: *"I was tired, I was trying to lean on one of the shelves"*.

Subjective data

Cyber-sickness. SSQ total score was 3.74 (median, $q1 - q3 = 0 - 16.83$). SSQ sub-scales scored: 0 ($0 - 9.54$) for nausea, 7.58 ($0 - 15.16$) for oculomotor disturbances, 0 ($0 - 27.84$) for disorientation. The distribution of scores for both sub-scales and total score are reported in Figure 4.16. As expected, significant positive correlations were present between all the sub-scales and total score: SSQ-N and SSQ-O ($r_s = 0.40, p = 0.001$), SSQ-N and SSQ-D ($r_s = 0.46, p < 0.001$), SSQ-N and SSQ-TS ($r_s = 0.65, p < 0.001$); SSQ-O and SSQ-D ($r_s = 0.77, p < 0.001$), SSQ-O and SSQ-TS ($r_s = 0.90, p < 0.001$); finally, SSQ-D and SSQ-TS ($r_s = 0.88, p < 0.001$). No correlation existed between SSQ scores and gender, age, schooling, and MMSE of the participants.

Sense of presence Results of ITC-SOPI sub-scales are presented in Table 4.4. In between ITC-SOPI subscales, correlations were found for *spatial presence* and

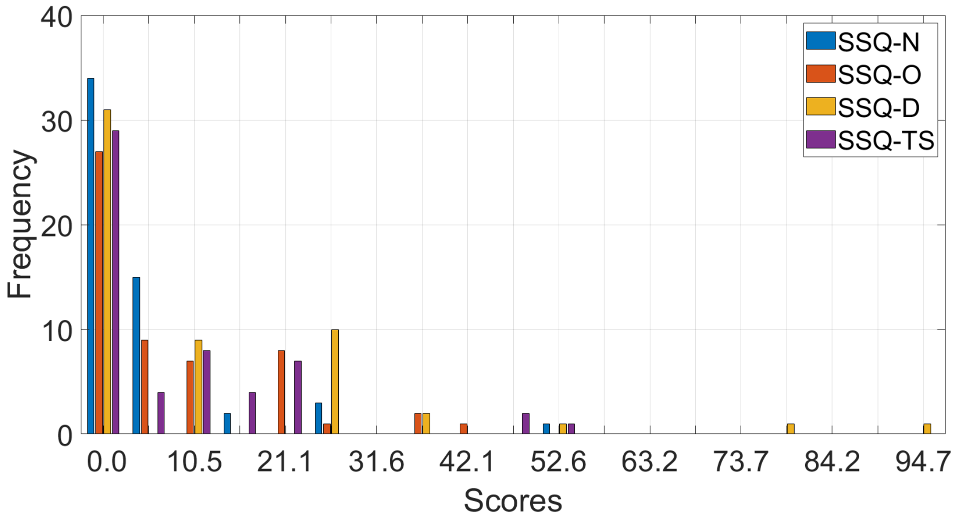


Figure 4.16: Simulator Sickness Questionnaire scores' distribution for total score and sub-scales. N = nausea, O = oculomotor disturbances, D = disorientation, TS = total score.

engagement ($r = 0.58, p < 0.001$); *spatial presence* and *naturalness* ($r = 0.69, p < 0.001$); *engagement* and *naturalness* ($r = 0.39, p = 0.003$). ITC-SOPI *side effects* scale correlated with all SSQ scores and sub-scores;in particular, we found $r_s = 0.49$ ($p < 0.001$) for SSQ-TS, $r_s = 0.50$ ($p =< 0.001$) for SSQ-N, $r_s = 0.34$ ($p = 0.002$) for SSQ-O, and $r_s = 0.34$ ($p = 0.009$) for SSQ-D, respectively. There existed no correlations with participants' demographic characteristics, and with their cognitive status.

Table 4.4: Scores of ITC-SOPI sub-scales; mean (SD).

Sub-scale	**Score**
Spatial Presence	3.51 (0.50)
Engagement	3.85 (0.68)
Naturalness	3.85 (0.82)
Side effects	1.46 (0.53)

Descriptive statistics for TAM3 are reported in Table 4.5. All measures had high internal consistency reliability ($\alpha \geq 0.70$), with the exception of PEOU, whose reliability was slightly inferior, but however acceptable ($\alpha = 0.68$). Weak correlations were present between PU and OUT ($r = 0.30, p = 0.028$), between PEOU and CANX ($r = -0.32, p = 0.019$) and PEOU and ENJ ($r = 0.33, p = 0.015$); a weak correlation also existed between ENJ and BI ($r = 0.28, p = 0.05$). Finally, a moderate negative correlation was present between CANX and ENJ ($r = -0.44, p < 0.001$).

Technology acceptance was correlated with SoP and cyber-sickness in few sub-scales: in particular, PEOU and SD ($r = -0.37, p = 0.006$), and PEOU and SSQ-N ($r = -0.33, p = 0.013$); CANX and SSQ-N ($r = 0.38, p = 0.004$), CANX and SD ($r = 0.40, p = 0.0023$); OUT and ITC-SOPI *naturalness* ($r = 0.32, p = 0.02$).

Years of schooling were correlated to OUT ($r = -0.27, p = 0.047$) and BI

($r = -0.34, p = 0.014$). No other participants' characteristic was correlated to TAM3 sub-scales.

Table 4.5: Mean and standard deviation scores, and reliability coefficient for the administered sub-scales of TAM3.

Sub-scale	**Score**	**Cronbach's α**
PU	5.19 (0.21)	0.93
PEOU	6.28 (0.40)	0.68
PEC	5.61 (0.31)	0.75
CANX	2.12 (0.16)	0.80
ENJ	6.28 (0.12)	0.97
OUT	5.74 (0.18)	0.74
BI	5.33 (0.13)	0.93

Objective data

On average, each participant completed 3.75 ± 1.60 shopping lists, which corresponded to the purchase of 30 items; all the participants completed at least one list in 15 minutes. The maximum number of completed shopping lists (trials) was 8. The number of completed trials resulted significantly and negatively correlated with age: r=-0.56, p<0.001. No significant correlations with years of schooling, MMSE and gender were found.

Execution times. A comparison of each shopping list completion time is presented in Figure 4.17. The statistical comparison of trial's completion time was made including all the participants who completed the first four trials ($n = 31$). Repeated measures ANOVA was significant (Huynh-Feldt $F(2.48, 74.52) = 19.86, p < 0.001$); post-hoc tests revealed that only the first trial's duration was significantly longer than the second ($p = 0.006$), the third, and the fourth ($p < 0.001$ for both).

Using the regression model, it resulted that the time needed to complete the first trial was independent from SSQ, ITC-SOPI and TAM3's sub-scales. However, there existed a positive correlation between age and execution times for the first ($r = 0.32, p = 0.019$), the second ($r = 0.37, p = 0.007$) and the third ($r = 0.49, p = 0.0008$) trial. Also MMSE resulted correlated with execution times, but the relationship was mixed: it was negative for the second trial ($r = -0.35, p = 0.022$) and positive in the fourth ($r = 0.55, p = 0.035$).

Errors. The number of errors is reported in Table 4.6 as number of relative errors with respect to the number of trials completed by all the participants. Twenty-two participants put in the cart at least one wrong object (or a duplicate of an already bought product) during the first trial; 43 participants dropped at least one item in the first trial. No CE were recorded.

Considering just the first trial (that was completed by all the participants), we found that WEs, DEs, and FEs were not dependent neither from SoP, nor from cyber-sickness, nor from technology acceptance. Mean errors (i.e., the total

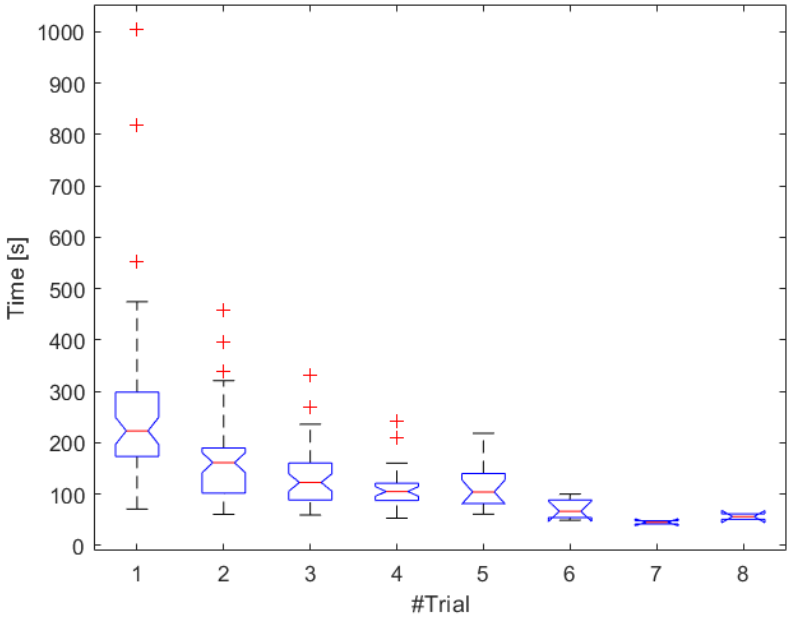


Figure 4.17: The duration of each trial. Red crosses indicate outliers.

Table 4.6: Number of errors relative to the number of trials [39].

Trial	#1	#2	#3	#4	#5	#6	#7	#8
WE	0.89	0.48	0.25	0.48	0.20	0	0	0
FE	0.29	0.26	0.18	0.23	0.27	0.29	0	0.50
DE	1.14	0.66	0.73	0.45	0.40	0.14	1.00	0.50
Rep.	55	50	44	31	15	7	2	2

number of WE, DE, CE divided the number of trials, for each participant) were not significantly related to any participant's characteristics.

Comparison with healthy young adults' results

The comparisons of sub-scales relative to SoP and SSQ are presented in Figure 4.18. None of the t-tests were significant. Also for SSQ sub-scores and total score, no difference were found using Mann-Whitney U test tests.

4.4.4 Discussion

This study aimed at examining the feasibility of a cognitive training intervention based on an immersive VR environment representing a supermarket. To our knowledge this was one of the first studies investigating the effect of immersive VR on the elder population, and the first specifically addressing MCI patients [39].

The outcomes we obtained were very promising. First, we recorded very low levels of cyber-sickness: all median in SSQ values were below acceptable levels, i.e., < 15; SE score was lower for VR with respect to what obtained in a study comparing different rendering means [219]; and the results were comparable with levels

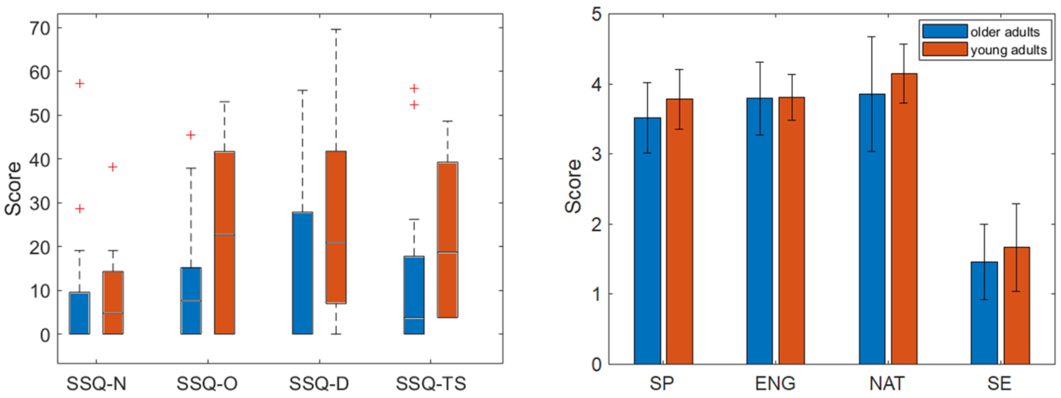


Figure 4.18: The comparison of SSQ and ITC-SOPI sub-scales' values obtained for the older adults and young adults' samples.

previously recorded for young adults. *Disorientation* appeared to be the symptom which affected the most the whole experience, followed by *ocumolotor disturbances* and *nausea*. This was in agreement with previous studies reporting that disorientation symptoms (i.e., dizziness and vertigo) are the major distinguishing features of cyber-sickness [136], i.e. the type of sickness arising in VEs, which is different from motion-sickness, normally arising in driving simulators. However, we could argue that our hypothesis that no conflicting sensory cues would occur in the Virtual Supermarket – as the motion resulted just because of participants' walking in the game area – was confirmed by the outcomes of the study: no severe symptoms arose as no sensory mismatch was induced.

No effect of personal characteristics of the participants was found to influence the perception of cyber-sickness. This is in contrast with the results obtained in previous studies, but there were some possible explanations. First, for what concerns age, it was reported that susceptibility to sickness decreases starting from the age of 21 [244] till becoming nonexistent after 50; however, the study reaching this conclusion was investigating the effect of motion-sickness in real tasks, therefore the results could not be directly transferable to our context. A more recent study reported that older adults did not suffer from any side effects during the exposure to immersive non-interactive VR [111].

The fact that one of the participants left the study without completing the questionnaires, after having complained a not better described malaise, supported the hypothesis that probably subjective factors (e.g., sensitiveness to motion sickness) played a strong role in the arousal of symptoms [175].

Additionally, as the participants enrolled in the study were aged between 60 and 83, it could be that the span was not large enough to highlight difference or, also, that the side-effects arising from the experience were too light to highlight any differences in the sample [245]. Other studies addressing the study of cyber-sickness symptoms made use of environments specifically developed to stimulate the arousal of symptoms [246, 247], and this was surely not our case. The hypothesis that symptoms were too light to highlight correlations with personal factors could be worthy also for gender, – whose effects reported in literature are mixed [190, 191, 248] – and for MMSE score. For the last, it has also to be remembered that, on the one hand, it is a diagnostic criterion for MCI, but, on the other, that it is never administered alone, as it is indeed not very sensitive for the discrimination of MCI [4].

In the future, studies addressing the feasibility of immersive VR-based interventions in older adults should include larger samples, individuals with pre-dementia and/or at higher risk of developing AD, and should consider including additional psychometric tests in the outcomes' analyses. Also, it would be interesting to couple subjective measures (i.e., SSQ) with objective ones: postural instability, EEG, heart rate and breath rate, and stomach upset are all non-invasive techniques that could be used to further study the acceptability and the feasibility of immersive VR programs [244].

In terms of SoP, the obtained results were comparable with those of young adults and superior than those reported in the study of Nisenfeld et al. [219]. The Virtual Supermarket was capable of eliciting very good levels of *spatial presence*, *engagement*, and *naturalness*, and the positive correlation between these three factors confirmed the reliability of the obtained results. The falling of one participant who later reported that his intention was to lean on the shelf was an unfortunate event, but, on the other hand, confirmed that SoP, and realism in particular were very high, as the participant's behavioural response was the same as in the physical world [183]. As for cyber-sickness, we found no effect of age, gender, and cognitive impairment. Regarding gender, as for cyber-sickness, evidence in literature is mixed [248, 249]. In terms of age and cognitive abilities, evidence is even more sparse; there appears to be a consensus among researchers that SoP could be influenced by these personal characteristics [217, 245], but no conclusion could be drawn [11]. Further studies should thus consider these aspects too.

Another relevant variable that this study has addressed is technology acceptance; this factor is fundamental to discriminate whether the user considers the proposed solution meaningful, and if he/she is willing to use it in his/her daily life [243]. According to the results we obtained, the Virtual Supermarket was highly acceptable, with all positive sub-scales scoring higher than 5 out of 7. This result could be linked both the good usability (perceived ease-of-use, PEOU) of the system we proposed, and to the perceived usefulness (PU) of the VR scenario for CT training [250, 251]. Though we did not tested usability, we could report that none of the participants enrolled in the study had difficulties in interacting with the controller; additionally, the decreasing number of errors throughout trials, and the longer time needed to complete only the first trial, supported the hypothesis that learning to use this technology was an easy process. In terms of PU, all the participants were told about the aim of the Virtual Supermarket, therefore they were aware of the potential usefulness of the application, and probably they appreciated this innovative way of performing CT [252].

Regarding technology acceptance, it is interesting to point out also that PEOU was influenced negatively by *computer anxiety* (CANX), nausea (SSQ-N) and side-effects (SD), and positively by *enjoyment* (ENJ). This was in agreement with previous findings: anxiety is often reported to be linked to lower performances (also in reality), which in turn could depend on less comprehension of the systems' interactions modalities or less attention toward it [253]; nausea and general malaise scores were indeed very low for everybody, but they could have contributed in creating a sense of unease that distracted, probably unconsciously, the participants from the system use [254].

On the other hand, being positively involved may result in increased performances and better flow [255], which could increase participant's appreciation toward the environment. In a state of *flow*, in fact, *control* plays a crucial role in the creation of SoP [16]: this hypothesis was also confirmed in this study, as there existed a correlation between the quality of the output (OUT), and the perceived naturalness of the environment.

We found no correlations between technology acceptance and participants' demographic and personal characteristics. This was in contrast to what was reported in the review of Hauk et al. [256], and in a recent study by Huygelier et al. [111], who reported that a negative age-attitude (i.e., the combination of PU and PEOU) association was present in their sample; they also reported this correlation not to be mediated by the cognitive status of the participant. This last outcome was confirmed in our study. The difference in the age-attitude relationship may have been due to the different instrument we used to evaluate it, or, more probably, to the fact that Huygelier et al. evaluated participants' attitude prior and after the experience, which we did not. It could have happened that, also in our case, the positive opinions of participants had been influenced by the just-ended pleasant experience, but we could not draw any conclusions in this sense.

Another interesting fact emerging from the analysis was that OUT and BI were negatively correlated with schooling, as people with higher educational level were more critical toward the general quality of the proposed experience, and less prone to repeat it, with respect to their less-educated peers.

Interestingly, a variable that resulted correlated with age (and not with gender, schooling, or cognitive status) was the time required to complete the first trial; for older participants, being more cautious (even involuntarily) during the first experiences with new technology may have led to needing longer time to get familiar with the environment and, consequently, to a reduced number of completed trials. However, this did not appear to be dependent on anxiety (CANX).

Further studies comprising a larger sample – perhaps with different cognitive diagnosis – and more sessions are thus required to shed light on the influence of age (and of cognitive status) on the feasibility and the acceptance of interactive immersive VR experiences in older adults.

4.5 Conclusions

This Chapter addressed *Aim 2b* of this thesis, i.e., the design and the validation of an immersive VR environment dedicated to the training of cognitive functions in older adults with MCI. The developmental path foresaw two different studies addressing diverse aspects of VR experience prior of the implementation of an acceptance and usability test on the target population.

In particular, the first study allowed determining the acceptance of the Virtual Supermarket environment in healthy young adults. The study revealed that environment was safe in terms of cyber-sickness, and able to elicit high levels of SoP.

Additionally, it helped in identifying potential interaction issues, which, in fact, had been simplified.

The second study assessing the kinematic differences between the movements performed in an ecological VR environment and in the physical world had also promising results. Indeed, joint RoMs were not affected by the fact of wearing an HMD; trunk rotation was an exception, but environmental factors had probably contributed. According to Levin et al. [36], coordination is an important component of dexterous movement, thus having maintained the appropriate segments' RoM was positive.

On the other hand, the possible limited perceptual fidelity that led to increased MTs and lower v_r in VR may disguise a higher cognitive workload, that would not be adequate for anybody, and especially for individuals who need a cognitive intervention [257]. The design of an appropriate VR environment is thus fundamental to elicit natural behaviour, and to let the participants focus only on the main (cognitive) tasks required by their training program.

Finally, a study addressing the acceptance of an immersive VR interventions on older adults with MCI has been presented. With respect to the previous acceptance study performed by Huygelier at al. [111], our experience foresaw also an active and purposeful participation of the older adults, who walked around and interacted with the objects in the scene, instead of being seated and having limited interactions within the virtual scene. Higher *control* over the environment could have contributed in creating a more engaging experience, but further studies are required to better investigate how older adults deal with immersive VR. It remains also to determine whether more cognitively impaired participants could benefit from immersive VR technologies without any safety concerns. Foreseeing the use of a harness should be recommended to ensure the safety of all participants.

In any case, the good outcomes of this study in terms of reduced side-effects, good SoP and high acceptance of the technology go in the direction of supporting the development and the further investigation of immersive VR applications for older adults with cognitive deficits. Future studies should then try addressing the clinical effectiveness of this immersive application.

Chapter 5

Conclusions and future work

This thesis describes immersive and non-immersive VR applications developed to provide PE and/or CT to older adults with cognitive deficits. In particular, we evaluated the impact, also from the clinical point-of-view, of interventions provided with 2D technologies (Chapter 2) and explored the acceptance and the usability of physical and cognitive training supported by immersive interactive VR technologies (Chapter 3 and Chapter 4).

In terms of clinical outcomes, we found some positive results. In the first trial we performed, in spite of the reduced sample, we found a significant reduction of oxidative stress, thus indicating that this type of intervention may be helpful in reducing the concentration of physiological markers predicting Alzheimer's disease (§2.2.1). We also recorded a tendency toward improvement in the Mini-Mental State Examination (MMSE, §1.2.2), in visual-constructive and visuo-spatial tests of attention (though not significant, probably due to the small and heterogeneous sample §2.2.1). In the second trial, we found a significant improvement in long-term memory recall in the groups undertaking CT, whereas PE alone did not seem to provide any effect (but oxidative stress was not tested in this occasion, §2.3.2).

Furthermore, in both studies, we recorded a subjective perception of improvement (§2.2.1 and §2.3.2). This aspect is of primary importance considering that the final aim of all rehabilitative interventions should be the improvement of patients' Quality of Life. Being more confident in ADLs reduces the excess disability, and contributes to relieve the care-givers. Also, increased self-efficacy helps to increase people's engagement in social and cognitive-stimulating activities, thus creating a virtuous circle, and promoting the implementation of all those strategies working against the occurrence of new symptoms.

Our clinical-oriented outcomes are promising, though their sparseness and heterogeneity did not allow drawing any evidence-based conclusion. Instead, they contributed to what had already been highlighted in literature: multi-domain interventions addressing modifiable risk factors could be effective in delaying or halting the progression of dementia [95]. However, the optimal "recipe" indicating which treatments have to administered, their frequency, and intensity has not been determined yet.

Recent findings, also from large RCTs [98, 97], have suggested that probably one of the key elements to improve treatment outcomes is indeed *customization*: people with different diagnosis should be treated differently. For instance, fatty acid supplementation, physical activity and cognitive training [98] worked for individuals at high-risk of developing dementia (CAIDE>6, or with β-amyloid plaques), but not for others. The preDIVA [97] intervention resulted effective mainly on patients with non-AD dementia, and, in particular, on individuals with vascular dementia.

In such a context, VR could represent a powerful means to promote the customization of treatments thanks to its high flexibility, which allows the stimulation of different cognitive domains, supports PE, and, potentially, the implementation of healthy lifestyle plans.

VR has firstly emerged in the rehabilitation sector as a tool able to stimulate the motivation to train, and thus treatment adherence. We believe – in agreement with many other studies – that two key factors increasing the motivation to train are, among others, SoP and technology acceptance. This is the reason that pushed us to investigate whether VR-based treatments providing higher SoP would be acceptable and enjoyable for older adults with cognitive impairments. In fact, if it is true that higher SoP leads to increased adherence, and thus better outcomes [19, 258], it is also important to establish whether there exist factors hindering the employment of VR technologies in the aged population (e.g., cyber-sickness).

Our acceptance and usability studies concerned the evaluation of two different protocols employing two different immersive VR systems: the CAVE (§3.3) and a HMD (§4.4). In both cases, there were several limitations, but our results showed that older adults largely accepted the intervention we proposed. In addition, we found that the use of immersive VR via HMD possibly did not affect movement coordination (§4.3.3), thus indicating that these technologies may be acceptable also from the point of view of cognitive ergonomics.

In contrast to what reported in previous studies [135, 256], we found that older people were not reluctant toward new technologies, but rather that they enjoyed the VR experience, and interpreted these technologies as an innovative way to train and to keep healthy. Clearly, whether the results of our studies are generalizable to a larger population has still to be determined.

Our studies had indeed several limitations. First, our sample were small. Also, we enrolled participants ranging from a healthy to mild cognitively-impaired, therefore we could not draw any conclusions for other populations with more severe symptoms. Finally, we could not draw any conclusion in terms of clinical effectiveness. The interventions we proposed have still to be investigated in well-designed RCTs, comprising an appropriate sample and a controlled administration of each condition, preferably for a long period of time.

Nonetheless, our outcomes were promising, and could be considered a starting point for the creation of new immersive VR-based cognitive training programs for older adults with MCI. In addition, these research experiences have contributed to define a series of possible guidelines, which might be generalized to inform about the design and the assessment of immersive VR applications for vulnerable individuals. A summary of these guidelines combining suggestions for the proper design of applications, and possible outcomes to evaluate in preliminary and clinical studies is presented in Figure 5.1. As it can been noticed, we have considered the design of the navigational and of the non-navigational environments distinctly. Though some recommendations are similar (e.g., avoid visual/vestibular mismatch), our experiences led us to believe that they would better require different approaches. For instance, in non-navigational scenarios (e.g. the Virtual Supermarket), we suggest the use of HMDs, because of their ability to convey the highest SoP, and to isolate the user from the external world; instead, for navigational environments, we would rather focus our attention on a device not completely occluding the view of physical world (e.g., a CAVE), as higher immersion may also be linked to more severe symptoms.

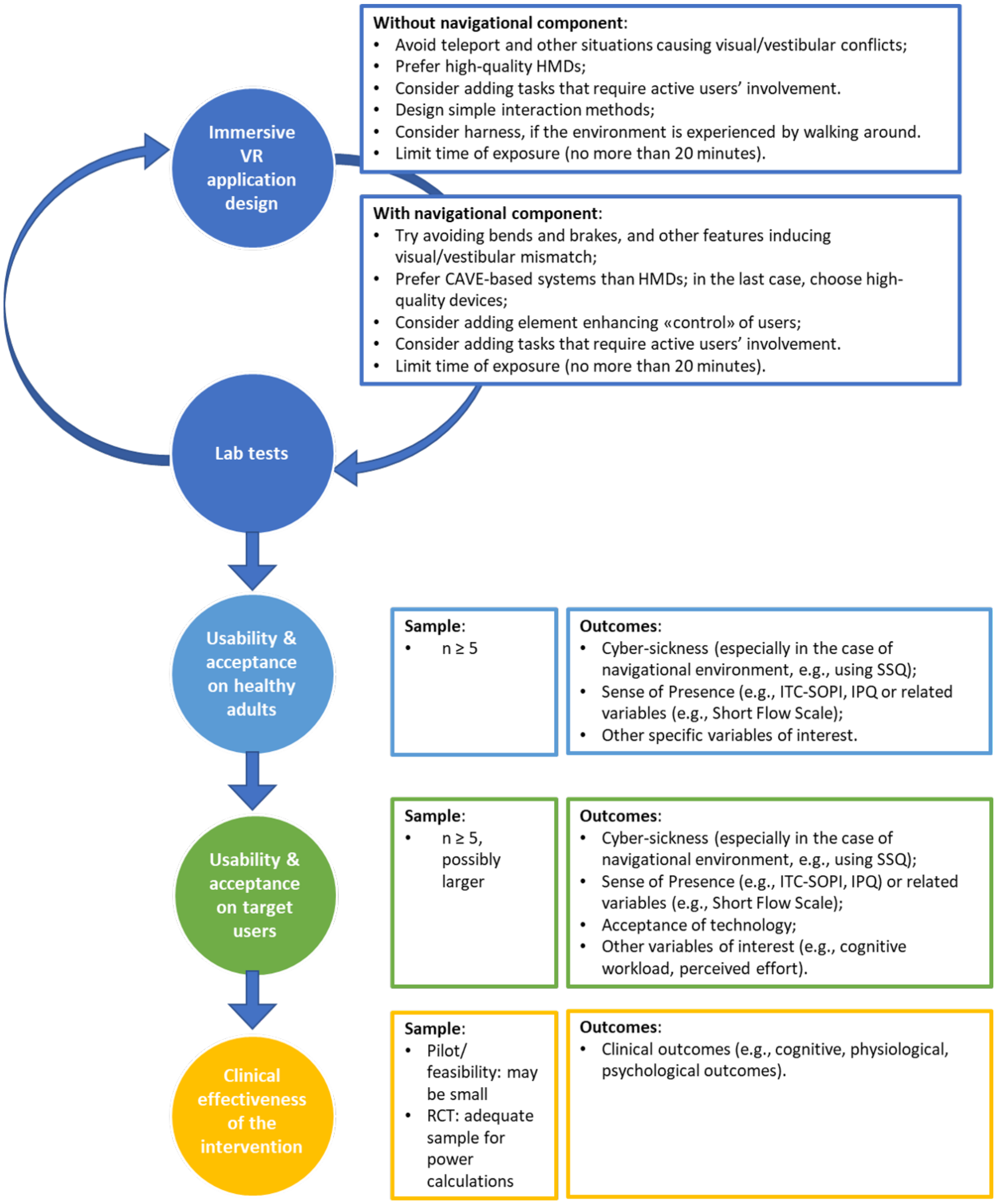
For what concerns preliminary studies, beside lab-based tests that are fundamental to eliminate of possible software malfunctionings, the performance of one or more studies enrolling healthy (young or age-matched) individuals should help addressing all those subjective issues related to the occurrence of cyber-sickness, and, secondly, to the engagement of the users. In our opinion, this represents a key step to provide older adults with a VR application that is safe, acceptable, and enjoyable.

Once all the acceptance/usability tests are concluded, and the application could be considered reasonably free from risks, a clinical trial employing the newly-designed VR-based treatment may start. Even if the primary aim of clinical trials should always remain to establish the effectiveness of a specific intervention, RCTs could be useful also to evaluate the effects of long-term treatment in terms of subjective outcomes, either from the positive (e.g., SoP, motivation, adherence to the treatment) or the negative point-of-view (e.g., arousal of symptoms).

In fact, a prolonged use of the same immersive VR application could either increase motivation, because of the subjective perception of an improvement throughout sessions, or weaken the willingness to proceed forward, because of the repetitiveness of the proposed tasks. Because of this, prior of the performance of a clinical trial, the implementation of new scenarios should be considered in order to reduce the risk of boredom.

It is also possible that, with long-term use, unforeseen side-effects would emerge. For instance, it has been proven that being immersed in VR has an effect on balance, even in young adults with no impairments [259]; and that such malaise increases with longer and repeated exposure [176]. Larger body sway has also been suggested as an objective measure for cyber-sickness [260]. Thus, exploring the possibility of administering VR-based treatments while staying seated or wearing a harness could be valuable options, though this may introduce a few limitations in the control the individual has over the VE, and in the naturalness of the interaction. Another possibility would be to test the effects on static and dynamic balance after a longer (or repeated) use, and thus to determine specific safety limits, or specific populations that should be excluded from these trials (e.g., older adults with cognitive impairments due to Parkinson's disease).

Another fundamental aspect that cannot be neglected for all future developments is the *social* aspect. It has been argued that social participation is an essential element of a healthy lifestyle [261], and including it in non-pharmacological treatment protocols could constitute an added value [262]. Researchers developing VR-based solutions should thus consider including social elements in their applications. Preliminary evidence appeared promising, especially for PE. For instance, it has been shown that performing physical activity in a competitive situation in-



Without navigational component:
- Avoid teleport and other situations causing visual/vestibular conflicts;
- Prefer high-quality HMDs;
- Consider adding tasks that require active users' involvement.
- Design simple interaction methods;
- Consider harness, if the environment is experienced by walking around.
- Limit time of exposure (no more than 20 minutes).

With navigational component:
- Try avoiding bends and brakes, and other features inducing visual/vestibular mismatch;
- Prefer CAVE-based systems than HMDs; in the last case, choose high-quality devices;
- Consider adding element enhancing «control» of users;
- Consider adding tasks that require active users' involvement.
- Limit time of exposure (no more than 20 minutes).

Sample:
- n ≥ 5

Outcomes:
- Cyber-sickness (especially in the case of navigational environment, e.g., using SSQ);
- Sense of Presence (e.g., ITC-SOPI, IPQ or related variables (e.g., Short Flow Scale);
- Other specific variables of interest.

Sample:
- n ≥ 5, possibly larger

Outcomes:
- Cyber-sickness (especially in the case of navigational environment, e.g., using SSQ);
- Sense of Presence (e.g., ITC-SOPI, IPQ) or related variables (e.g., Short Flow Scale);
- Acceptance of technology;
- Other variables of interest (e.g., cognitive workload, perceived effort).

Sample:
- Pilot/ feasibility: may be small
- RCT: adequate sample for power calculations

Outcomes:
- Clinical outcomes (e.g., cognitive, physiological, psychological outcomes).

Figure 5.1: A framework diagram showing a possible approach to the design and the assessment of rehabilitative treatments exploiting (immersive) VR. Different recommendations are given depending on the navigational or non-navigational nature of the virtual environment.

duced young subjects to reach higher levels of energy expenditure [263, 264], and that collaborative games had a positive influence on motivation and self-efficacy, thus promoting adherence to PE programs [265, 263]. The same tendencies were found also in a study enrolling older adults [125].

What appeared important to consider in multi-player applications, once again, was the customization, which should occur according to each person's attitude: results showed that asking to compete to non-competitive individual led to no improvement [125]. Competitive and non-competitive personalities should thus be treated differently.

In our research group, we started working on a social application for DT training. We used the same paradigm described in §3.3, but we gave the older adults the possibility – prior to the beginning of the training – to decide whether they would like to compete or collaborate with other users (up to 4). Collaboration or competition occurs for the cognitive task, whereas the physical exercise is balanced, so that all the users proceed at the same velocity and exerting the same level of effort. Such application, namely *Social Bike*, allows users talking to each other, and thus to exercise together from remote locations too [169]. We hope that in the next future, trials assessing the feasibility and the effectiveness of this intervention, perhaps mediating for the effect of sociability, would be performed.

A further potential benefit that social applications could introduce is the possibility of training at home. Exploiting the potentialities of at-home interventions would mean promoting the continuity of care, and the reduction of time and costs required to go to clinics at a fixed frequency [266]. Introducing the social element in such a framework may result in a few additional advantages: first, people could me more motivated to train, as their peers could exercise with them [262]; second, older adults who perhaps do not feel safe in going out alone, could avoid loneliness and isolation while training [267]; third, a continuous monitoring of the performance could be implemented thus allowing the clinical personnel to track the progression of the cognitive decline through time, and to notice whether a person is or is not adhering to treatment; this may also contribute to reduce anxiety and the feeling of being abandoned by who should take care of him/her. Finally, implementing social platforms allows for the creation of a sort of "virtual neighbors" network, which produce trust in others [268], and allows for the support and the indirect monitoring of older adults [269]: if one of the peers is not present, perhaps because of an adverse event, someone in the community will surely notice. Clearly, also in this case, VR, and digital technologies in general, could constitute means worthy of further research.

In conclusion, we could say that a few attempts to reduce cognitive symptoms in older adults through multi-domain interventions have already been made, and that results have been promising. However, the current evidence is limited, and the current demographic situation calls for further studies and clinical trials.

As the technological development proceeds, higher quality VR-based applications able to increase user's Sense of Presence and strongly limit side-effects, should be developed [11]. Future research should prioritize treatment customization, taking into account specificity of each person suffering from MCI or dementia: his/her symptoms' evolution, and his/her attitude. Future VR-based applications should also be able to incorporate emerging knowledge and new evidence from previous clinical trials in order to improve as much as possible the QoL of subjects with MCI and dementia, and to better support clinicians and care-givers in the treatment of these chronic conditions.

Bibliography

[1] *World report on ageing and health.* World Health Organization, 2015.

[2] T. Finkel and N. J. Holbrook, “Oxidants, oxidative stress and the biology of ageing,” *nature*, vol. 408, no. 6809, p. 239, 2000.

[3] K. Jekel, M. Damian, C. Wattmo, L. Hausner, R. Bullock, P. J. Connelly, B. Dubois, M. Eriksdotter, M. Ewers, E. Graessel, *et al.*, “Mild cognitive impairment and deficits in instrumental activities of daily living: a systematic review,” *Alzheimer’s research & therapy*, vol. 7, no. 1, p. 17, 2015.

[4] S. Gauthier, B. Reisberg, M. Zaudig, R. C. Petersen, K. Ritchie, K. Broich, S. Belleville, H. Brodaty, D. Bennett, H. Chertkow, *et al.*, “Mild cognitive impairment,” *The lancet*, vol. 367, no. 9518, pp. 1262–1270, 2006.

[5] A. Ward, H. M. Arrighi, S. Michels, and J. M. Cedarbaum, “Mild cognitive impairment: disparity of incidence and prevalence estimates,” *Alzheimer’s & Dementia*, vol. 8, no. 1, pp. 14–21, 2012.

[6] K. Ritchie, “Mild cognitive impairment: an epidemiological perspective,” *Dialogues in clinical neuroscience*, vol. 6, no. 4, p. 401, 2004.

[7] S. A. Eshkoor, T. A. Hamid, C. Y. Mun, and C. K. Ng, “Mild cognitive impairment and its management in older people,” *Clinical interventions in aging*, vol. 10, p. 687, 2015.

[8] P. Heyn, B. C. Abreu, and K. J. Ottenbacher, “The effects of exercise training on elderly persons with cognitive impairment and dementia: a meta-analysis,” *Archives of physical medicine and rehabilitation*, vol. 85, no. 10, pp. 1694–1704, 2004.

[9] M. Loef and H. Walach, “Midlife obesity and dementia: meta-analysis and adjusted forecast of dementia prevalence in the united states and china,” *Obesity*, vol. 21, no. 1, pp. E51–E55, 2013.

[10] J. Huntley, R. Gould, K. Liu, M. Smith, and R. Howard, “Do cognitive interventions improve general cognition in dementia? a meta-analysis and meta-regression,” *BMJ open*, vol. 5, no. 4, p. e005247, 2015.

[11] R. I. García-Betances, M. T. Arredondo Waldmeyer, G. Fico, and M. F. Cabrera-Umpiérrez, “A succinct overview of virtual reality technology use in alzheimer’s disease,” *Frontiers in aging neuroscience*, vol. 7, p. 80, 2015.

[12] S. T. Koenig, D. Krch, B. S. Lange, and A. Rizzo, *Virtual reality and rehabilitation.* American Psychological Association, 2019.

[13] M. Bayón-Calatayud, E. Peri, F. F. Nistal, M. Duff, F. Nieto-Escámez, B. Lange, and S. Koenig, "Virtual rehabilitation," in *Emerging Therapies in Neurorehabilitation II*, pp. 303–318, Springer, 2016.

[14] A. â. Rizzo and G. J. Kim, "A swot analysis of the field of virtual reality rehabilitation and therapy," *Presence: Teleoperators & Virtual Environments*, vol. 14, no. 2, pp. 119–146, 2005.

[15] S. Grassini, K. Laumann, and M. R. Skogstad, "The use of virtual reality alone does not promote training performance (but sense of presence does)," *Frontiers in Psychology*, vol. 11, 2020.

[16] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence*, vol. 7, no. 3, pp. 225–240, 1998.

[17] S. Estupiñán, F. Rebelo, P. Noriega, C. Ferreira, and E. Duarte, "Can virtual reality increase emotional responses (arousal and valence)? a pilot study," in *International Conference of Design, User Experience, and Usability*, pp. 541–549, Springer, 2014.

[18] A. MacQuarrie and A. Steed, "Cinematic virtual reality: Evaluating the effect of display type on the viewing experience for panoramic video," in *2017 IEEE Virtual Reality (VR)*, pp. 45–54, IEEE, 2017.

[19] S. Weech, S. Kenny, and M. Barnett-Cowan, "Presence and cybersickness in virtual reality are negatively related: a review," *Frontiers in psychology*, vol. 10, p. 158, 2019.

[20] S. Mrakic-Sposta, S. G. Di Santo, F. Franchini, S. Arlati, A. Zangiacomi, L. Greci, S. Moretti, N. Jesuthasan, M. Marzorati, G. Rizzo, *et al.*, "Effects of combined physical and cognitive virtual reality-based training on cognitive impairment and oxidative stress in mci patients: A pilot study," *Frontiers in aging neuroscience*, vol. 10, 2018.

[21] S. Arlati, A. Zangiacomi, L. Greci, S. G. Di Santo, F. Franchini, and M. Sacco, "Virtual environments for cognitive and physical training in elderly with mild cognitive impairment: a pilot study," in *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pp. 86–106, Springer, 2017.

[22] S. Arlati, L. Greci, M. Mondellini, A. Zangiacomi, S. G. Di Santo, F. Franchini, M. Marzorati, S. Mrakic-Sposta, and A. Vezzoli, "A virtual reality-based physical and cognitive training system aimed at preventing symptoms of dementia," in *International Conference on Wireless Mobile Communication and Healthcare*, pp. 117–125, Springer, 2017.

[23] M. Brasure, P. Desai, H. Davila, V. A. Nelson, C. Calvert, E. Jutkowitz, M. Butler, H. A. Fink, E. Ratner, L. S. Hemmy, *et al.*, "Physical activity interventions in preventing cognitive decline and alzheimer-type dementia: a systematic review," *Annals of internal medicine*, vol. 168, no. 1, pp. 30–38, 2018.

[24] M. Mondellini, S. Arlati, S. Pizzagalli, L. Greci, M. Sacco, and G. Ferrigno, "Assessment of the usability of an immersive virtual supermarket for the cognitive rehabilitation of elderly patients: A pilot study on young adults," in *2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)*, pp. 1–8, IEEE, 2018.

[25] E. Pedroli, L. Greci, D. Colombo, S. Serino, P. Cipresso, S. Arlati, M. Mondellini, L. Boilini, V. Giussani, K. Goulene, *et al.*, "Characteristics, usability, and users experience of a system combining cognitive and physical therapy in a virtual environment: Positive bike," *Sensors*, vol. 18, no. 7, p. 2343, 2018.

[26] M. Schwenk, T. Zieschang, P. Oster, and K. Hauer, "Dual-task performances can be improved in patients with dementia: a randomized controlled trial," *Neurology*, vol. 74, no. 24, pp. 1961–1968, 2010.

[27] H. Pashler, "Dual-task interference in simple tasks: data and theory.," *Psychological bulletin*, vol. 116, no. 2, p. 220, 1994.

[28] S. Arlati, N. Keijsers, G. Paolini, G. Ferrigno, and M. Sacco, "Kinematics of reaching in ecological immersive virtual reality: a comparative study with real world," *Virtual Reality*, (*Submitted*).

[29] S. Arlati, N. Keijsers, G. Ferrigno, and M. Sacco, "A protocol for the comparison of reaching gesture kinematics in physical versus immersive virtual reality," in *2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pp. 1–6, IEEE, 2020.

[30] L. Liu, R. van Liere, C. Nieuwenhuizen, and J.-B. Martens, "Comparing aimed movements in the real world and in virtual reality," in *2009 IEEE Virtual Reality Conference*, pp. 219–222, IEEE, 2009.

[31] L. A. Knaut, S. K. Subramanian, B. J. McFadyen, D. Bourbonnais, and M. F. Levin, "Kinematics of Pointing Movements Made in a Virtual Versus a Physical 3-Dimensional Environment in Healthy and Stroke Subjects," *Archives of Physical Medicine and Rehabilitation*, vol. 90, no. 5, pp. 793–802, 2009.

[32] C. Y. Wang, W. J. Hwang, J. J. Fang, C. F. Sheu, I. F. Leong, and H. I. Ma, "Comparison of virtual reality versus physical reality on movement characteristics of persons with Parkinson's disease: Effects of moving targets," *Archives of Physical Medicine and Rehabilitation*, vol. 92, no. 8, pp. 1238–1245, 2011.

[33] J. C. Stewart, J. Gordon, and C. J. Winstein, "Planning and adjustments for the control of reach extent in a virtual environment," *Journal of NeuroEngineering and Rehabilitation*, vol. 10, no. 1, 2013.

[34] E. C. Magdalon, S. M. Michaelsen, A. A. Quevedo, and M. F. Levin, "Comparison of grasping movements made by healthy subjects in a 3-dimensional immersive virtual versus physical environment," *Acta psychologica*, vol. 138, no. 1, pp. 126–134, 2011.

[35] M. A. Just, P. J. Stapley, M. Ros, F. Naghdy, and D. Stirling, "A comparison of upper limb movement profiles when reaching to virtual and real targets using the Oculus Rift: implications for virtual-reality enhanced stroke rehabilitation," *Virtual Reality & Associated Technologies Gothenburg*, pp. 2–4, 2014.

[36] M. F. Levin, L. A. M. Knaut, E. C. Magdalon, and S. Subramanian, "Virtual reality environments to enhance upper limb functional recovery in patients with hemiparesis," in *Studies in Health Technology and Informatics*, vol. 145, pp. 94–108, 2009.

[37] E. Bouchard Ryan, A. P. Anas, M. Beamer, and S. Bajorek, "Coping with age-related vision loss in everyday reading activities," *Educational Gerontology*, vol. 29, no. 1, pp. 37–54, 2003.

[38] M. Fahle and S. Henke-Fahle, "Interobserver variance in perceptual performance and learning.," *Investigative ophthalmology & visual science*, vol. 37, no. 5, pp. 869–877, 1996.

[39] S. Arlati, S. G. Di Santo, F. Franchini, M. Mondellini, B. Filiputti, M. Luchi, F. Ratto, G. Ferrigno, M. Sacco, and L. Greci, "Immersive virtual reality for cognitive training in elderly with objective and subjective cognitive decline: A feasibility study," *Journal of Alzheimer Disease*, (*Submitted*).

[40] C. López-Otín, M. A. Blasco, L. Partridge, M. Serrano, and G. Kroemer, "The hallmarks of aging," *Cell*, vol. 153, no. 6, pp. 1194–1217, 2013.

[41] K. Anstey, L. Stankov, and S. Lord, "Primary aging, secondary aging, and intelligence.," *Psychology and Aging*, vol. 8, no. 4, p. 562, 1993.

[42] R. G. Cutler, "Antioxidants and aging," *The American journal of clinical nutrition*, vol. 53, no. 1, pp. 373S–379S, 1991.

[43] F. W. Booth, M. J. Laye, and M. D. Roberts, "Lifetime sedentary living accelerates some aspects of secondary aging," *Journal of applied physiology*, vol. 111, no. 5, pp. 1497–1504, 2011.

[44] A. Shumway-Cook and M. H. Woollacott, *Motor control: translating research into clinical practice.* Lippincott Williams & Wilkins, 2007.

[45] M. Tosato, V. Zamboni, A. Ferrini, and M. Cesari, "The aging process and potential interventions to extend life expectancy," *Clinical interventions in aging*, vol. 2, no. 3, p. 401, 2007.

[46] B. T. Weinert and P. S. Timiras, "Invited review: Theories of aging," *Journal of applied physiology*, vol. 95, no. 4, pp. 1706–1716, 2003.

[47] R. C. Petersen and J. C. Morris, "Mild cognitive impairment as a clinical entity and treatment target," *Archives of neurology*, vol. 62, no. 7, pp. 1160–1163, 2005.

[48] M. Ganguli, H. H. Dodge, C. Shen, and S. T. DeKosky, "Mild cognitive impairment, amnestic type: an epidemiologic study," *Neurology*, vol. 63, no. 1, pp. 115–121, 2004.

[49] M. S. Albert, S. T. DeKosky, D. Dickson, B. Dubois, H. H. Feldman, N. C. Fox, A. Gamst, D. M. Holtzman, W. J. Jagust, R. C. Petersen, *et al.*, "The diagnosis of mild cognitive impairment due to alzheimer's disease: Recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease," *Alzheimer's & dementia*, vol. 7, no. 3, pp. 270–279, 2011.

[50] J. C. Morris, M. Storandt, J. P. Miller, D. W. McKeel, J. L. Price, E. H. Rubin, and L. Berg, "Mild cognitive impairment represents early-stage alzheimer disease," *Archives of neurology*, vol. 58, no. 3, pp. 397–405, 2001.

[51] T. W. Mitchell, E. J. Mufson, J. A. Schneider, E. J. Cochran, J. Nissanov, L.-Y. Han, J. L. Bienias, V. M.-Y. Lee, J. Q. Trojanowski, D. A. Bennett, *et al.*, "Parahippocampal tau pathology in healthy aging, mild cognitive impairment, and early alzheimer's disease," *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, vol. 51, no. 2, pp. 182–189, 2002.

[52] J. Stefaniak and J. O'Brien, "Imaging of neuroinflammation in dementia: a review," *J Neurol Neurosurg Psychiatry*, vol. 87, no. 1, pp. 21–28, 2016.

[53] C. R. Jack, R. C. Petersen, Y. C. Xu, S. C. Waring, P. C. O'Brien, E. G. Tangalos, G. E. Smith, R. J. Ivnik, and E. Kokmen, "Medial temporal atrophy on mri in normal aging and very mild alzheimer's disease," *Neurology*, vol. 49, no. 3, pp. 786–794, 1997.

[54] C. DeCarli, "Mild cognitive impairment: prevalence, prognosis, aetiology, and treatment," *The Lancet Neurology*, vol. 2, no. 1, pp. 15–21, 2003.

[55] J. T. O'Brien, T. Erkinjuntti, B. Reisberg, G. Roman, T. Sawada, L. Pantoni, J. V. Bowler, C. Ballard, C. DeCarli, P. B. Gorelick, *et al.*, "Vascular cognitive impairment," *The Lancet Neurology*, vol. 2, no. 2, pp. 89–98, 2003.

[56] D. Bennett, J. Schneider, J. Bienias, D. Evans, and R. Wilson, "Mild cognitive impairment is related to alzheimer disease pathology and cerebral infarctions," *Neurology*, vol. 64, no. 5, pp. 834–841, 2005.

[57] M. P. Copeland, E. Daly, V. Hines, C. Mastromauro, D. Zaitchik, J. Gunther, and M. Albert, “Psychiatric symptomatology and prodromal alzheimer’s disease,” *Alzheimer Disease & Associated Disorders*, vol. 17, no. 1, pp. 1–8, 2003.

[58] C. Brayne and P. Calloway, “Normal ageing, impaired cognitive function, and senile dementia of the alzheimer’s type: a continuum?,” *The Lancet*, vol. 331, no. 8597, pp. 1265–1267, 1988.

[59] L. Bahureksa, B. Najafi, A. Saleh, M. Sabbagh, D. Coon, M. J. Mohler, and M. Schwenk, “The impact of mild cognitive impairment on gait and balance: a systematic review and meta-analysis of studies using instrumented assessment,” *Gerontology*, vol. 63, no. 1, pp. 67–83, 2017.

[60] M. F. Folstein, S. E. Folstein, and P. R. McHugh, “Mini-mental state: a practical method for grading the cognitive state of patients for the clinician,” *Journal of psychiatric research*, vol. 12, no. 3, pp. 189–198, 1975.

[61] A. J. Mitchell, “A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment,” *Journal of psychiatric research*, vol. 43, no. 4, pp. 411–431, 2009.

[62] Z. S. Nasreddine, N. A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings, and H. Chertkow, “The montreal cognitive assessment, moca: a brief screening tool for mild cognitive impairment,” *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695–699, 2005.

[63] A. Nordlund, S. Rolstad, P. Hellström, M. Sjögren, S. Hansen, and A. Wallin, “The goteborg mci study: mild cognitive impairment is a heterogeneous condition,” *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 76, no. 11, pp. 1485–1490, 2005.

[64] J. M. Ringman, L. D. Medina, Y. Rodriguez-Agudelo, M. Chavez, P. Lu, and J. L. Cummings, “Current concepts of mild cognitive impairment and their applicability to persons at-risk for familial alzheimer’s disease,” *Current Alzheimer Research*, vol. 6, no. 4, pp. 341–346, 2009.

[65] H. Hampel, S. Teipel, T. Fuchsberger, N. Andreasen, J. Wiltfang, M. Otto, Y. Shen, R. Dodel, Y. Du, M. Farlow, *et al.*, “Value of csf β-amyloid 1-42 and tau as predictors of alzheimer’s disease in patients with mild cognitive impairment,” *Molecular psychiatry*, vol. 9, no. 7, p. 705, 2004.

[66] K. Blennow and H. Hampel, “Csf markers for incipient alzheimer’s disease,” *The Lancet Neurology*, vol. 2, no. 10, pp. 605–613, 2003.

[67] H. Zetterberg, L.-O. Wahlund, and K. Blennow, “Cerebrospinal fluid markers for prediction of alzheimer’s disease,” *Neuroscience letters*, vol. 352, no. 1, pp. 67–69, 2003.

[68] M. Atiya, B. T. Hyman, M. S. Albert, and R. Killiany, "Structural magnetic resonance imaging in established and prodromal alzheimer disease: a review," *Alzheimer Disease & Associated Disorders*, vol. 17, no. 3, pp. 177–195, 2003.

[69] W. Jagust, "Positron emission tomography and magnetic resonance imaging in the diagnosis and prediction of dementia," *Alzheimer's & dementia*, vol. 2, no. 1, pp. 36–42, 2006.

[70] E. Tönnies and E. Trushina, "Oxidative stress, synaptic dysfunction, and alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 57, no. 4, pp. 1105–1121, 2017.

[71] N. L. Campbell, F. Unverzagt, M. A. LaMantia, B. A. Khan, and M. A. Boustani, "Risk factors for the progression of mild cognitive impairment to dementia," *Clinics in geriatric medicine*, vol. 29, no. 4, pp. 873–893, 2013.

[72] J. Tschanz, K. Welsh-Bohmer, C. G. Lyketsos, C. Corcoran, R. C. Green, K. Hayden, M. Norton, P. P. Zandi, L. Toone, N. West, *et al.*, "Conversion to dementia from mild cognitive disorder: the cache county study," *Neurology*, vol. 67, no. 2, pp. 229–234, 2006.

[73] L. S. Elias-Sonnenschein, W. Viechtbauer, I. H. Ramakers, F. R. Verhey, and P. J. Visser, "Predictive value of apoe-$\varepsilon 4$ allele for progression from mci to ad-type dementia: a meta-analysis," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 82, no. 10, pp. 1149–1156, 2011.

[74] M. Baumgart, H. M. Snyder, M. C. Carrillo, S. Fazio, H. Kim, and H. Johns, "Summary of the evidence on modifiable risk factors for cognitive decline and dementia: a population-based perspective," *Alzheimer's & Dementia*, vol. 11, no. 6, pp. 718–726, 2015.

[75] R. Katzman, "Education and the prevalence of dementia and alzheimer's disease.," *Neurology*, 1993.

[76] N. Schneider and C. Yvon, "A review of multidomain interventions to support healthy cognitive ageing," *The journal of nutrition, health & aging*, vol. 17, no. 3, pp. 252–257, 2013.

[77] L. Teri and J. M. Uomoto, "Reducing excess disability in dementia patients: Training caregivers to manage patient depression," *Clinical Gerontologist*, vol. 10, no. 4, pp. 49–63, 1991.

[78] T. Odawara, "Cautious notification and continual monitoring of patients with mild cognitive impairment," *Psychogeriatrics*, vol. 12, no. 2, pp. 131–132, 2012.

[79] H. Feldman, T. Pirttila, J. Dartigues, B. Everitt, B. Van Baelen, H. Brashear, J. Berlin, W. Battisti, and S. Kavanagh, "Analyses of mortality risk in patients with dementia treated with galantamine," *Acta Neurologica Scandinavica*, vol. 119, no. 1, pp. 22–31, 2009.

[80] C. Pinto and A. A. Subramanyam, "Mild cognitive impairment: The dilemma," *Indian journal of psychiatry*, vol. 51, no. Suppl1, p. S44, 2009.

[81] A. Bahar-Fuchs, L. Clare, and B. Woods, "Cognitive training and cognitive rehabilitation for mild to moderate alzheimer's disease and vascular dementia," *Cochrane Database of Systematic Reviews*, no. 6, 2013.

[82] L. Clare and R. T. Woods, "Cognitive training and cognitive rehabilitation for people with early-stage alzheimer's disease: A review," *Neuropsychological rehabilitation*, vol. 14, no. 4, pp. 385–401, 2004.

[83] T. Li, Y. Yao, Y. Cheng, B. Xu, X. Cao, D. Waxman, W. Feng, Y. Shen, Q. Li, J. Wang, *et al.*, "Cognitive training can reduce the rate of cognitive aging: a neuroimaging cohort study," *BMC geriatrics*, vol. 16, no. 1, p. 12, 2016.

[84] A. Engvig, A. M. Fjell, L. T. Westlye, N. V. Skaane, A. M. Dale, D. Holland, P. Due-Tønnessen, Ø. Sundseth, and K. B. Walhovd, "Effects of cognitive training on gray matter volumes in memory clinic patients with subjective memory impairment," *Journal of Alzheimer's Disease*, vol. 41, no. 3, pp. 779–791, 2014.

[85] M. C. Carlson, J. H. Kuo, Y.-F. Chuang, V. R. Varma, G. Harris, M. S. Albert, K. I. Erickson, A. F. Kramer, J. M. Parisi, Q.-L. Xue, *et al.*, "Impact of the baltimore experience corps trial on cortical and hippocampal volumes," *Alzheimer's & Dementia*, vol. 11, no. 11, pp. 1340–1348, 2015.

[86] A. Bahar-Fuchs, S. Webb, L. Bartsch, L. Clare, G. Rebok, N. Cherbuin, and K. J. Anstey, "Tailored and adaptive computerized cognitive training in older adults at risk for dementia: a randomized controlled trial," *Journal of Alzheimer's Disease*, vol. 60, no. 3, pp. 889–911, 2017.

[87] H. Coyle, V. Traynor, and N. Solowij, "Computerized and virtual reality cognitive training for individuals at high risk of cognitive decline: systematic review of the literature," *The American Journal of Geriatric Psychiatry*, vol. 23, no. 4, pp. 335–359, 2015.

[88] R. Stephen, K. Hongisto, A. Solomon, and E. Lönnroos, "Physical activity and alzheimer's disease: a systematic review," *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, vol. 72, no. 6, pp. 733–739, 2017.

[89] L. S. Nagamatsu, A. Chan, J. C. Davis, B. L. Beattie, P. Graf, M. W. Voss, D. Sharma, and T. Liu-Ambrose, "Physical activity improves verbal and spatial memory in older adults with probable mild cognitive impairment: a 6-month randomized controlled trial," *Journal of aging research*, vol. 2013, 2013.

[90] L. D. Baker, L. L. Frank, K. Foster-Schubert, P. S. Green, C. W. Wilkinson, A. McTiernan, S. R. Plymate, M. A. Fishel, G. S. Watson, B. A. Cholerton,

et al., "Effects of aerobic exercise on mild cognitive impairment: a controlled trial," *Archives of neurology*, vol. 67, no. 1, pp. 71–79, 2010.

[91] T. Suzuki, H. Shimada, H. Makizako, T. Doi, D. Yoshida, K. Tsutsumimoto, Y. Anan, K. Uemura, S. Lee, and H. Park, "Effects of multicomponent exercise on cognitive function in older adults with amnestic mild cognitive impairment: a randomized controlled trial," *BMC neurology*, vol. 12, no. 1, p. 128, 2012.

[92] K. M. Sink, M. A. Espeland, C. M. Castro, T. Church, R. Cohen, J. A. Dodson, J. Guralnik, H. C. Hendrie, J. Jennings, J. Katula, *et al.*, "Effect of a 24-month physical activity intervention vs health education on cognitive outcomes in sedentary older adults: the life randomized trial," *Jama*, vol. 314, no. 8, pp. 781–790, 2015.

[93] M. Kivipelto, F. Mangialasche, and T. Ngandu, "Lifestyle interventions to prevent cognitive impairment, dementia and alzheimer disease," *Nature Reviews Neurology*, p. 1, 2018.

[94] A. Solomon, F. Mangialasche, E. Richard, S. Andrieu, D. A. Bennett, M. Breteler, L. Fratiglioni, B. Hooshmand, A. S. Khachaturian, L. S. Schneider, *et al.*, "Advances in the prevention of alzheimer's disease and dementia," *Journal of internal medicine*, vol. 275, no. 3, pp. 229–250, 2014.

[95] T. Ngandu, J. Lehtisalo, A. Solomon, E. Levälahti, S. Ahtiluoto, R. Antikainen, L. Bäckman, T. Hänninen, A. Jula, T. Laatikainen, *et al.*, "A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (finger): a randomised controlled trial," *The Lancet*, vol. 385, no. 9984, pp. 2255–2263, 2015.

[96] A. Rosenberg, T. Ngandu, M. Rusanen, R. Antikainen, L. Bäckman, S. Havulinna, T. Hänninen, T. Laatikainen, J. Lehtisalo, E. Levälahti, *et al.*, "Multidomain lifestyle intervention benefits a large elderly population at risk for cognitive decline and dementia regardless of baseline characteristics: The finger trial," *Alzheimer's & Dementia*, vol. 14, no. 3, pp. 263–270, 2018.

[97] E. P. M. van Charante, E. Richard, L. S. Eurelings, J.-W. van Dalen, S. A. Ligthart, E. F. Van Bussel, M. P. Hoevenaar-Blom, M. Vermeulen, and W. A. van Gool, "Effectiveness of a 6-year multidomain vascular care intervention to prevent dementia (prediva): a cluster-randomised controlled trial," *The Lancet*, vol. 388, no. 10046, pp. 797–805, 2016.

[98] S. Andrieu, S. Guyonnet, N. Coley, C. Cantet, M. Bonnefoy, S. Bordes, L. Bories, M.-N. Cufi, T. Dantoine, J.-F. Dartigues, *et al.*, "Effect of long-term omega 3 polyunsaturated fatty acid supplementation with or without multidomain intervention on cognitive function in elderly adults with memory complaints (mapt): a randomised, placebo-controlled trial," *The Lancet Neurology*, vol. 16, no. 5, pp. 377–389, 2017.

[99] M. O. Onyesolu and F. U. Eze, "Understanding virtual reality technology: advances and applications," *Adv. Comput. Sci. Eng*, pp. 53–70, 2011.

[100] J. M. Mittelstaedt, "Individual predictors of the susceptibility for motion-related sickness: A systematic review," *Journal of Vestibular Research*, no. Preprint, pp. 1–29.

[101] M. Zahabi and A. M. A. Razak, "Adaptive virtual reality-based training: a systematic literature review and framework," *Virtual Reality*, pp. 1–28, 2020.

[102] S. Mottura, L. Fontana, S. Arlati, A. Zangiacomi, C. Redaelli, and M. Sacco, "A virtual reality system for strengthening awareness and participation in rehabilitation for post-stroke patients," *Journal on Multimodal User Interfaces*, vol. 9, no. 4, pp. 341–351, 2015.

[103] M. Csikszentmihalyi, *Finding flow.* Hachette Audio, 2017.

[104] R. I. Garcia-Betances, V. Jiménez-Mixco, M. T. Arredondo, and M. F. Cabrera-Umpiérrez, "Using virtual reality for cognitive training of the elderly," *American Journal of Alzheimer's Disease & Other Dementias®*, vol. 30, no. 1, pp. 49–54, 2015.

[105] R. Herne, S. Rai, M. Shiratuddin, H. Laga, and M. Byrnes, "Using a driving simulator to improve driving awareness in stroke survivors: a pilot study," *Journal of Fundamental and Applied Sciences*, vol. 10, no. 2S, pp. 201–214, 2018.

[106] M. C. Ashe, C. L. Ekegren, A. M. Chudyk, L. Fleig, T. K. Gill, D. Langford, L. Martin-Martin, and P. Ariza-Vega, "Telerehabilitation for community-dwelling middle-aged and older adults after musculoskeletal trauma: A systematic review," *AIMS Medical Science*, vol. 5, no. 4, pp. 316–336, 2018.

[107] V. Venkatesh and F. D. Davis, "A theoretical extension of the technology acceptance model: Four longitudinal field studies," *Management science*, vol. 46, no. 2, pp. 186–204, 2000.

[108] S. V. Cobb, S. Nichols, A. Ramsey, and J. R. Wilson, "Virtual reality-induced symptoms and effects (vrise)," *Presence: Teleoperators & Virtual Environments*, vol. 8, no. 2, pp. 169–186, 1999.

[109] J. M. Mittelstaedt, J. Wacker, and D. Stelling, "Vr aftereffect and the relation of cybersickness and cognitive performance," *Virtual Reality*, vol. 23, no. 2, pp. 143–154, 2019.

[110] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson, "Virtual reality induced symptoms and effects (vrise): Comparison of head mounted display (hmd), desktop and projection display systems," *Displays*, vol. 29, no. 2, pp. 58–69, 2008.

[111] H. Huygelier, B. Schraepen, R. van Ee, V. V. Abeele, and C. R. Gillebert, "Acceptance of immersive head-mounted virtual reality in older adults," *Scientific reports*, vol. 9, no. 1, pp. 1–12, 2019.

[112] G. Optale, C. Urgesi, V. Busato, S. Marin, L. Piron, K. Priftis, L. Gamberini, S. Capodieci, and A. Bordin, "Controlling memory impairment in elderly adults using virtual reality memory training: a randomized controlled pilot study," *Neurorehabilitation and neural repair*, vol. 24, no. 4, pp. 348–357, 2010.

[113] D. W. Man, J. C. Chung, and G. Y. Lee, "Evaluation of a virtual reality-based memory training programme for hong kong chinese older adults with questionable dementia: a pilot study," *International journal of geriatric psychiatry*, vol. 27, no. 5, pp. 513–520, 2012.

[114] N. Jebara, E. Orriols, M. Zaoui, A. Berthoz, and P. Piolino, "Effects of enactment in episodic memory: a pilot virtual reality study with young and elderly adults," *Frontiers in aging neuroscience*, vol. 6, p. 338, 2014.

[115] V. Manera, E. Chapoulie, J. Bourgeois, R. Guerchouche, R. David, J. Ondrej, G. Drettakis, and P. Robert, "A feasibility study with image-based rendered virtual reality in patients with mild cognitive impairment and dementia," *PloS one*, vol. 11, no. 3, p. e0151487, 2016.

[116] S.-C. Yeh, Y.-C. Chen, C.-F. Tsai, and A. Rizzo, "An innovative virtual reality system for mild cognitive impairment: diagnosis and evaluation," in *2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences*, pp. 23–27, IEEE, 2012.

[117] P. J. White and Z. Moussavi, "Neurocognitive treatment for a patient with alzheimer's disease using a virtual reality navigational environment," *Journal of experimental neuroscience*, vol. 10, pp. JEN–S40827, 2016.

[118] I. Tarnanas, A. Tsolakis, and M. Tsolaki, "Assessing virtual reality environments as cognitive stimulation method for patients with mci," in *Technologies of Inclusive Well-Being*, pp. 39–74, Springer, 2014.

[119] I. Tarnanas, C. Mouzakidis, and W. Schlee, "Functional impairment in virtual-reality-daily-living-activities as a defining feature of amnestic mci: Cognitive and psychomotor correlates," in *2013 International Conference on Virtual Rehabilitation (ICVR)*, pp. 27–34, IEEE, 2013.

[120] P. Gamito, J. Oliveira, C. Coelho, D. Morais, P. Lopes, J. Pacheco, R. Brito, F. Soares, N. Santos, and A. F. Barata, "Cognitive training on stroke patients via virtual reality-based serious games," *Disability and rehabilitation*, vol. 39, no. 4, pp. 385–388, 2017.

[121] E. R. Zanier, T. Zoerle, D. Di Lernia, and G. Riva, "Virtual reality for traumatic brain injury," *Frontiers in neurology*, vol. 9, p. 345, 2018.

[122] K. P. Padala, P. R. Padala, S. Y. Lensing, R. A. Dennis, M. M. Bopp, C. M. Parkes, M. K. Garrison, P. M. Dubbert, P. K. Roberson, and D. H. Sullivan, "Efficacy of wii-fit on static and dynamic balance in community dwelling older veterans: a randomized controlled pilot trial," *Journal of aging research*, vol. 2017, 2017.

[123] K. P. Padala, P. R. Padala, S. Y. Lensing, R. A. Dennis, M. M. Bopp, P. K. Roberson, and D. H. Sullivan, "Home-based exercise program improves balance and fear of falling in community-dwelling older adults with mild alzheimer's disease: a pilot study," *Journal of Alzheimer's disease*, vol. 59, no. 2, pp. 565–574, 2017.

[124] M. Colombo, E. Marelli, R. Vaccaro, E. Valle, S. Colombani, E. Polesel, S. Garolfi, S. Fossi, and A. Guaita, "Virtual reality for persons with dementia: an exergaming experience," in *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, vol. 29, p. 1, IAARC Publications, 2012.

[125] C. Anderson-Hanley, P. J. Arciero, A. M. Brickman, J. P. Nimon, N. Okuma, S. C. Westen, M. E. Merz, B. D. Pence, J. A. Woods, A. F. Kramer, *et al.*, "Exergaming and older adult cognition: a cluster randomized clinical trial," *American journal of preventive medicine*, vol. 42, no. 2, pp. 109–119, 2012.

[126] F. Clay, D. Howett, J. FitzGerald, P. Fletcher, D. Chan, and A. Price, "Use of immersive virtual reality in the assessment and treatment of alzheimer's disease: A systematic review," *Journal of Alzheimer's Disease*, no. Preprint, pp. 1–21, 2020.

[127] A. E. Kazdin, "Acceptability of alternative treatments for deviant child behavior," *Journal of Applied Behavior Analysis*, vol. 13, no. 2, pp. 259–273, 1980.

[128] P. Astrand, "Measurement of maximal aerobic capacity.," *Canadian Medical Association Journal*, vol. 96, no. 12, p. 732, 1967.

[129] C.-W. Lee and G.-H. Cho, "Effect of stationary cycle exercise on gait and balance of elderly women," *Journal of physical therapy science*, vol. 26, no. 3, pp. 431–433, 2014.

[130] N.-G. Kim, Y.-Y. Kim, and T.-K. Kwon, "Development of a virtual reality bicycle simulator for rehabilitation training of postural balance," in *International Conference on Computational Science and Its Applications*, pp. 241–250, Springer, 2006.

[131] D. A. Brown and S. Kautz, "Increased workload enhances force output during pedaling exercise in persons with poststroke hemiplegia," *Stroke*, vol. 29, no. 3, pp. 598–606, 1998.

[132] Z. Ren, J. Meng, J. Yuan, and Z. Zhang, "Robust hand gesture recognition with kinect sensor," in *MM '11 Proceedings of the 19th ACM international conference on Multimedia*, pp. 759–760, ACM, November 2011.

[133] "Kinect v1 touchless touch, available at: http://www.touchlesstouch.com/."

[134] "Leap motion, available at: https://www.leapmotion.com/."

[135] G. Coldham and D. M. Cook, "Vr usability from elderly cohorts: Preparatory challenges in overcoming technology rejection," in *2017 National Information Technology Conference (NITC)*, pp. 131–135, IEEE, 2017.

[136] K. M. Stanney, R. S. Kennedy, and J. M. Drexler, "Cybersickness is not simulator sickness," in *Proceedings of the Human Factors and Ergonomics Society annual meeting*, vol. 41, pp. 1138–1142, SAGE Publications Sage CA: Los Angeles, CA, 1997.

[137] M. E. Nelson, W. J. Rejeski, S. N. Blair, P. W. Duncan, J. O. Judge, A. C. King, C. A. Macera, and C. Castaneda-Sceppa, "Physical activity and public health in older adults: recommendation from the american college of sports medicine and the american heart association," *Circulation*, vol. 116, no. 9, p. 1094, 2007.

[138] G. Carlesimo, C. Caltagirone, G. Gainotti, L. Fadda, R. Gallassi, S. Lorusso, G. Marfia, C. Marra, U. Nocentini, and L. Parnetti, "The mental deterioration battery: normative data, diagnostic reliability and qualitative analyses of cognitive impairment," *European neurology*, vol. 36, no. 6, pp. 378–384, 1996.

[139] P. Caffarra, G. Vezzadini, F. Dieci, F. Zonato, and A. Venneri, "Rey-osterrieth complex figure: normative values in an italian population sample," *Neurological Sciences*, vol. 22, no. 6, pp. 443–447, 2002.

[140] H. Spinnler and G. Tognoni, "Italian group on the neuropsychological study of ageing: Italian standardization and classification of neuropsychological tests," *The Italian Journal of Neurological Sciences*, vol. 6, no. 8, pp. 1–120, 1987.

[141] A. R. Giovagnoli, M. Del Pesce, S. Mascheroni, M. Simoncelli, M. Laiacona, and E. Capitani, "Trail making test: normative values from 287 normal adult controls," *The Italian journal of neurological sciences*, vol. 17, no. 4, pp. 305–309, 1996.

[142] I. Appollonio, M. Leone, V. Isella, F. Piamarta, T. Consoli, M. Villa, E. Forapani, A. Russo, and P. Nichelli, "The frontal assessment battery (fab): normative values in an italian population sample," *Neurological Sciences*, vol. 26, no. 2, pp. 108–116, 2005.

[143] G. Novelli, C. Papagno, E. Capitani, M. Laiacona, *et al.*, "Tre test clinici di ricerca e produzione lessicale. taratura su sogetti normali.," *Archivio di psicologia, neurologia e psichiatria*, 1986.

[144] R. I. Pfeffer, T. Kurosaki, C. Harrah Jr, J. Chance, and S. Filos, "Measurement of functional activities in older adults in the community," *Journal of gerontology*, vol. 37, no. 3, pp. 323–329, 1982.

[145] S. Mrakic-Sposta, M. Gussoni, M. Montorsi, S. Porcelli, and A. Vezzoli, “Assessment of a standardized ros production profile in humans by electron paramagnetic resonance,” *Oxidative Medicine and Cellular Longevity*, vol. 2012, 2012.

[146] J. Nielsen and J. Levy, “Measuring usability: preference vs. performance,” *Communications of the ACM*, vol. 37, no. 4, pp. 66–76, 1994.

[147] M. M. Wolf, “Social validity: The case of subjective measurement or how applied behavior analysis is finding its heart,” *Journal of applied behavior analysis*, vol. 11, no. 2, pp. 203–214, 1978.

[148] F. D. Davis, “Perceived usefulness, perceived ease of use, and user acceptance of information technology,” *MIS quarterly*, pp. 319–340, 1989.

[149] V. Braun and V. Clarke, *Thematic analysis.* American Psychological Association, 2012.

[150] N. T. Hill, L. Mowszowski, S. L. Naismith, V. L. Chadwick, M. Valenzuela, and A. Lampit, “Computerized cognitive training in older adults with mild cognitive impairment or dementia: a systematic review and meta-analysis,” *American Journal of Psychiatry*, vol. 174, no. 4, pp. 329–340, 2016.

[151] A. Nunomura, R. J. Castellani, X. Zhu, P. I. Moreira, G. Perry, and M. A. Smith, “Involvement of oxidative stress in alzheimer disease,” *Journal of Neuropathology & Experimental Neurology*, vol. 65, no. 7, pp. 631–641, 2006.

[152] D. Grembowski, D. Patrick, P. Diehr, M. Durham, S. Beresford, E. Kay, and J. Hecht, “Self-efficacy and health behavior among older adults,” *Journal of health and social behavior*, pp. 89–104, 1993.

[153] E. M. Lexell, U.-B. Flansbjer, and J. Lexell, “Self-perceived performance and satisfaction with performance of daily activities in persons with multiple sclerosis following interdisciplinary rehabilitation,” *Disability and rehabilitation*, vol. 36, no. 5, pp. 373–378, 2014.

[154] B. Dohnke, B. Knäuper, and W. Müller-Fahrnow, “Perceived self-efficacy gained from, and health effects of, a rehabilitation program after hip joint replacement,” *Arthritis Care & Research: Official Journal of the American College of Rheumatology*, vol. 53, no. 4, pp. 585–592, 2005.

[155] K. D. Cicerone, C. Dahlberg, K. Kalmar, D. M. Langenbahn, J. F. Malec, T. F. Bergquist, T. Felicetti, J. T. Giacino, J. P. Harley, D. E. Harrington, *et al.*, “Evidence-based cognitive rehabilitation: recommendations for clinical practice,” *Archives of physical medicine and rehabilitation*, vol. 81, no. 12, pp. 1596–1615, 2000.

[156] “Leantween, available at: https://assetstore.unity.com/packages/tools/animation/leantween-3595.”

[157] M. Monaco, A. Costa, C. Caltagirone, and G. A. Carlesimo, “Forward and backward span for verbal and visuo-spatial data: standardization and normative data from an italian adult population,” *Neurological Sciences*, vol. 34, no. 5, pp. 749–754, 2013.

[158] C. Marra, G. Gainotti, E. Scaricamazza, C. Piccininni, M. Ferraccioli, and D. Quaranta, “The multiple features target cancellation (mftc): an attentional visual conjunction search test. normative values for the italian population,” *Neurological Sciences*, vol. 34, no. 2, pp. 173–180, 2013.

[159] R. Capasso and G. Miceli, *Esame neuropsicologico per l’afasia enpa metodologie riabilitative in logopedia.* 2001.

[160] J. A. Yesavage, “Geriatric depression scale,” *Psychopharmacol Bull*, vol. 24, no. 4, pp. 709–711, 1988.

[161] R. S. Marin, R. C. Biedrzycki, and S. Firinciogullari, “Reliability and validity of the apathy evaluation scale,” *Psychiatry research*, vol. 38, no. 2, pp. 143–162, 1991.

[162] F. Franchini, *Fattori di rischio modificabili e declino cognitivo nell’anziano: uno studio trasversale e un trial randomizzato e controllato.* PhD thesis, Italy, 2020.

[163] S. K. Nam, H. J. Chu, M. K. Lee, J. H. Lee, N. Kim, and S. M. Lee, “A meta-analysis of gender differences in attitudes toward seeking professional psychological help,” *Journal of American College Health*, vol. 59, no. 2, pp. 110–116, 2010.

[164] C. S. Mackenzie, W. Gekoski, and V. Knox, “Age, gender, and the underutilization of mental health services: The influence of help-seeking attitudes,” *Aging and mental health*, vol. 10, no. 6, pp. 574–582, 2006.

[165] S. Elena, N. Georgeta, G. Cecilia, and L. Elena, “The attitude of the elderly persons towards health related physical activities,” *Procedia-Social and Behavioral Sciences*, vol. 30, pp. 1913–1919, 2011.

[166] “Physical activity and older adults, available at: https://www.who.int/dietphysicalactivity/factsheet-olderadults/en/.”

[167] H. Yamaguchi, Y. Maki, and T. Yamagami, “Overview of non-pharmacological intervention for dementia and principles of brain-activating rehabilitation,” *Psychogeriatrics*, vol. 10, no. 4, pp. 206–213, 2010.

[168] O. Lazarov, J. Robinson, Y.-P. Tang, I. S. Hairston, Z. Korade-Mirnics, V. M.-Y. Lee, L. B. Hersh, R. M. Sapolsky, K. Mirnics, and S. S. Sisodia, “Environmental enrichment reduces aβ levels and amyloid deposition in transgenic mice,” *Cell*, vol. 120, no. 5, pp. 701–713, 2005.

[169] S. Arlati, V. Colombo, D. Spoladore, L. Greci, E. Pedroli, S. Serino, P. Cipresso, K. Goulene, M. Stramba-Badiale, G. Riva, *et al.*, "A social virtual reality-based application for the physical and cognitive training of the elderly at home," *Sensors*, vol. 19, no. 2, p. 261, 2019.

[170] J. J. Gibson, *The ecological approach to visual perception: classic edition.* Psychology Press, 2014.

[171] D. Novak, A. Nagle, U. Keller, and R. Riener, "Increasing motivation in robot-aided arm rehabilitation with competitive and cooperative gameplay," *Journal of neuroengineering and rehabilitation*, vol. 11, no. 1, p. 64, 2014.

[172] R. M. Baños, C. Botella, I. Rubió, S. Quero, A. García-Palacios, and M. Alcañiz, "Presence and emotions in virtual environments: The influence of stereoscopy," *CyberPsychology & Behavior*, vol. 11, no. 1, pp. 1–8, 2008.

[173] J. Freeman, J. Lessiter, K. Pugh, and E. Keogh, "When presence and emotion are related, and when they are not," in *Proceedings of the 8th annual international workshop on presence (PRESENCE 2005)*, pp. 213–219, International Society for Presence Research, 2005.

[174] P. J. Rosa, D. Morais, P. Gamito, J. Oliveira, and T. Saraiva, "The immersive virtual reality experience: a typology of users revealed through multiple correspondence analysis combined with cluster analysis technique," *Cyberpsychology, Behavior, and Social Networking*, vol. 19, no. 3, pp. 209–216, 2016.

[175] M. Mondellini, S. Arlati, L. Greci, G. Ferrigno, and M. Sacco, "Sense of presence and cybersickness while cycling in virtual environments: Their contribution to subjective experience," in *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pp. 3–20, Springer, 2018.

[176] J. Hakkinen, T. Vuori, and M. Paakka, "Postural stability and sickness symptoms after hmd use," in *IEEE International Conference on Systems, Man and Cybernetics*, vol. 1, pp. 147–152, IEEE, 2002.

[177] T. Schubert, F. Friedmann, and H. Regenbrecht, "The experience of presence: Factor analytic insights," *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 3, pp. 266–281, 2001.

[178] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *The international journal of aviation psychology*, vol. 3, no. 3, pp. 203–220, 1993.

[179] R. M. Ryan, "Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory.," *Journal of personality and social psychology*, vol. 43, no. 3, p. 450, 1982.

[180] M. M. North and S. M. North, "A comparative study of sense of presence of traditional virtual reality and immersive environments," *Australasian Journal of Information Systems*, vol. 20, 2016.

[181] M. Slater and A. Steed, "A virtual presence counter," *Presence: Teleoperators & Virtual Environments*, vol. 9, no. 5, pp. 413–434, 2000.

[182] J. Chung, H.-J. Yoon, and H. J. Gardner, "Analysis of break in presence during game play using a linear mixed model," *ETRI journal*, vol. 32, no. 5, pp. 687–694, 2010.

[183] M. Slater, "Presence and the sixth sense," *Presence: Teleoperators & Virtual Environments*, vol. 11, no. 4, pp. 435–439, 2002.

[184] X. Dong, K. Yoshida, and T. A. Stoffregen, "Control of a virtual vehicle influences postural activity and motion sickness.," *Journal of Experimental Psychology: Applied*, vol. 17, no. 2, p. 128, 2011.

[185] C. B. Lourenco, L. Azeff, H. Sveistrup, and M. F. Levin, "Effect of environment on motivation and sense of presence in healthy subjects performing reaching tasks," in *2008 Virtual Rehabilitation*, pp. 93–98, IEEE, 2008.

[186] W. IJsselsteijn, Y. De Kort, R. Bonants, M. De Jager, and J. Westerink, "Virtual cycling: Effects of immersion and a virtual coach on motivation and presence in a home fitness application," in *Proceedings Virtual Reality Design and Evaluation Workshop*, pp. 22–23, 2004.

[187] M. Slater, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3549–3557, 2009.

[188] J. R. Bruun-Pedersen, S. Serafin, and L. B. Kofoed, "Going outside while staying inside-exercise motivation with immersive vs. non-immersive recreational virtual environment augmentation for older adult nursing home residents," in *2016 IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 216–226, IEEE, 2016.

[189] Y. Ling, H. T. Nefs, W.-P. Brinkman, C. Qu, and I. Heynderickx, "The relationship between individual characteristics and experienced presence," *Computers in Human Behavior*, vol. 29, no. 4, pp. 1519–1530, 2013.

[190] G. De Leo, L. A. Diggs, E. Radici, and T. W. Mastaglio, "Measuring sense of presence and user characteristics to predict effective training in an online simulated virtual environment," *Simulation in Healthcare*, vol. 9, no. 1, pp. 1–6, 2014.

[191] P. Gamito, J. Oliveira, D. Morais, A. Baptista, N. Santos, F. Soares, T. Saraiva, and P. Rosa, "Training presence: the importance of virtual reality experience on the "sense of being there"," *Annual Review of Cybertherapy and Telemedicine 2010*, pp. 128–133, 2010.

[192] T. M. Porcino, E. Clua, D. Trevisan, C. N. Vasconcelos, and L. Valente, "Minimizing cyber sickness in head mounted display systems: design guidelines and applications," in *2017 IEEE 5th international conference on serious games and applications for health (SeGAH)*, pp. 1–6, IEEE, 2017.

[193] J. J. LaViola Jr, “A discussion of cybersickness in virtual environments,” *ACM Sigchi Bulletin*, vol. 32, no. 1, pp. 47–56, 2000.

[194] A. Gaggioli, L. Greci, S. Arlati, M. Stramba-Badiale, E. Pedroli, D. Colombo, S. Serino, P. Cipresso, and G. Riva, “Positive bike - an immersive biking experience for combined physical and cognitive training of elderly patients,” 2017.

[195] M. Montero-Odasso, S. W. Muir, and M. Speechley, “Dual-task complexity affects gait in people with mild cognitive impairment: the interplay between gait variability, dual tasking, and risk of falls,” *Archives of physical medicine and rehabilitation*, vol. 93, no. 2, pp. 293–299, 2012.

[196] E. Al-Yahya, H. Dawes, L. Smith, A. Dennis, K. Howells, and J. Cockburn, “Cognitive motor interference while walking: a systematic review and meta-analysis,” *Neuroscience & Biobehavioral Reviews*, vol. 35, no. 3, pp. 715–728, 2011.

[197] X. Wang, Y. Pi, P. Chen, Y. Liu, R. Wang, and C. Chan, “Cognitive motor interference for preventing falls in older adults: a systematic review and meta-analysis of randomised controlled trials,” *Age and ageing*, vol. 44, no. 2, pp. 205–212, 2014.

[198] P. Plummer, G. Eskes, S. Wallace, C. Giuffrida, M. Fraas, G. Campbell, K. Clifton, E. R. Skidmore, *et al.*, “Cognitive-motor interference during functional mobility after stroke: state of the science and implications for future research,” *Archives of physical medicine and rehabilitation*, vol. 94, no. 12, pp. 2565–2574, 2013.

[199] V. E. Kelly, A. J. Eusterbrock, and A. Shumway-Cook, “A review of dual-task walking deficits in people with parkinson’s disease: motor and cognitive contributions, mechanisms, and clinical implications,” *Parkinson’s Disease*, vol. 2012, 2012.

[200] A. Mirelman, L. Rochester, I. Maidan, S. Del Din, L. Alcock, F. Nieuwhof, M. O. Rikkert, B. R. Bloem, E. Pelosin, L. Avanzino, *et al.*, “Addition of a non-immersive virtual reality component to treadmill training to reduce fall risk in older adults (v-time): a randomised controlled trial,” *The Lancet*, vol. 388, no. 10050, pp. 1170–1182, 2016.

[201] D. Schoene, S. R. Lord, K. Delbaere, C. Severino, T. A. Davies, and S. T. Smith, “A randomized controlled pilot study of home-based step training in older people using videogame technology,” *PloS one*, vol. 8, no. 3, p. e57734, 2013.

[202] R. Mazzocchio, S. Meunier, S. Ferrante, and F. Molteni, “Cycling, a tool for locomotor recovery after,” *NeuroRehabilitation*, vol. 23, pp. 67–80, 2008.

[203] J. Brooke *et al.*, “Sus-a quick and dirty usability scale,” *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.

[204] "Middlevr unity plug-in. available at: https://www.middlevr.com/middlevr-for-unity/."

[205] "Vicon tracking system. available at: https://www.vicon.com/."

[206] S. A. Jackson and H. W. Marsh, "Development and validation of a scale to measure optimal experience: The flow state scale," *Journal of sport and exercise psychology*, vol. 18, no. 1, pp. 17–35, 1996.

[207] A. Bangor, P. Kortum, and J. Miller, "Determining what individual sus scores mean: Adding an adjective rating scale," *Journal of usability studies*, vol. 4, no. 3, pp. 114–123, 2009.

[208] A. Bangor, P. T. Kortum, and J. T. Miller, "An empirical evaluation of the system usability scale," *Intl. Journal of Human-Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008.

[209] H. R. Marston, M. Kroll, D. Fink, and Y. J. Gschwind, "Flow experience of older adults using the istoppfalls exergame," *Games and Culture*, vol. 11, no. 1-2, pp. 201–222, 2016.

[210] B. Galna, D. Jackson, G. Schofield, R. McNaney, M. Webster, G. Barry, D. Mhiripiri, M. Balaam, P. Olivier, and L. Rochester, "Retraining function in people with parkinson's disease using the microsoft kinect: game design and pilot testing," *Journal of neuroengineering and rehabilitation*, vol. 11, no. 1, p. 60, 2014.

[211] A. E. Holland, C. J. Hill, P. Rochford, J. Fiore, D. J. Berlowitz, and C. F. McDonald, "Telerehabilitation for people with chronic obstructive pulmonary disease: feasibility of a simple, real time model of supervised exercise training," *Journal of telemedicine and telecare*, vol. 19, no. 4, pp. 222–226, 2013.

[212] D. Anderson *et al.*, "A social virtual reality system with 3d animation, spoken interaction, and runtime modifiability," tech. rep., Technical Report at Mitsubishi Electronic Research Laboratories, Cambridge, 1996.

[213] E. Pedroli, P. Cipresso, L. Greci, S. Arlati, L. Boilini, L. Stefanelli, M. Rossi, K. Goulene, M. Sacco, M. Stramba-Badiale, *et al.*, "An immersive motor protocol for frailty rehabilitation," *Frontiers in neurology*, vol. 10, 2019.

[214] "Htc vive, available at: https://www.vive.com/eu/."

[215] "Virtual reality toolkit (vrtk), available at: https://vrtoolkit.readme.io/."

[216] "Steamvr, available at: https://store.steampowered.com/steamvr."

[217] J. Lessiter, J. Freeman, E. Keogh, and J. Davidoff, "A cross-media presence questionnaire: The itc-sense of presence inventory," *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 3, pp. 282–297, 2001.

[218] J. D. Moss and E. R. Muth, "Characteristics of head-mounted displays and their effects on simulator sickness," *Human factors*, vol. 53, no. 3, pp. 308–319, 2011.

[219] S. Nisenfeld, *Using reality to Evaluate the ITC Presence Questionnaire.* PhD thesis, Department of Computer Science, Brown Unversity, 2003.

[220] J. Brade, M. Lorenz, M. Busch, N. Hammer, M. Tscheligi, and P. Klimant, "Being there again–presence in real and virtual environments and its relation to usability and user experience using a mobile navigation task," *International Journal of Human-Computer Studies*, vol. 101, pp. 76–87, 2017.

[221] M. Busch, M. Lorenz, M. Tscheligi, C. Hochleitner, and T. Schulz, "Being there for real: presence in real and virtual environments and its relation to usability," in *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, pp. 117–126, ACM, 2014.

[222] M. Mousavi, Y. H. Jen, and S. N. B. Musa, "A review on cybersickness and usability in virtual environments," in *Advanced Engineering Forum*, vol. 10, pp. 34–39, Trans Tech Publ, 2013.

[223] L. Rebenitsch and C. Owen, "Review on cybersickness in applications and visual displays," *Virtual Reality*, vol. 20, no. 2, pp. 101–125, 2016.

[224] R. Pals, L. Steg, J. Dontje, F. Siero, and K. van Der Zee, "Physical features, coherence and positive outcomes of person–environment interactions: A virtual reality study," *Journal of environmental psychology*, vol. 40, pp. 108–116, 2014.

[225] S. Arlati, D. Spoladore, D. Baldassini, M. Sacco, and L. Greci, "Virtualcruisetour: An ar/vr application to promote shore excursions on cruise ships," in *Augmented Reality, Virtual Reality, and Computer Graphics* (L. T. De Paolis and P. Bourdot, eds.), (Cham), pp. 133–147, Springer International Publishing, 2018.

[226] W. Powell, B. Stevens, and M. Simmonds, "Treadmill interface for virtual reality vs. overground walking: A comparison of gait in individuals with and without pain.," *Studies in health technology and informatics*, vol. 144, p. 198, 2009.

[227] E. Medina, R. Fruland, and S. Weghorst, "Virtusphere: Walking in a human size vr "hamster ball"," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 52, pp. 2102–2106, SAGE Publications Sage CA: Los Angeles, CA, 2008.

[228] S. Razzaque, Z. Kohn, and M. C. Whitton, *Redirected walking.* Citeseer, 2005.

[229] M. Olbrich, H. Graf, J. Keil, R. Gad, S. Bamfaste, and F. Nicolini, "Virtual reality based space operations: a study of esa's potential for vr based training

and simulation," in *International Conference on Virtual, Augmented and Mixed Reality*, pp. 438–451, Springer, 2018.

[230] W. E. Marsh, J. W. Kelly, V. J. Dark, and J. H. Oliver, "Assessing the use of cognitive resources in virtual reality," in *International Conference on Human-Computer Interaction*, pp. 120–124, Springer, 2011.

[231] D. G. Liebermann, S. Berman, P. L. Weiss, and M. F. Levin, "Kinematics of reaching movements in a 2-D virtual environment in adults with and without stroke," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 6, pp. 778–787, 2012.

[232] A. Viau, A. G. Feldman, B. J. McFadyen, and M. F. Levin, "Reaching in reality and virtual reality: A comparison of movement kinematics in healthy subjects and in adults with hemiparesis," *Journal of NeuroEngineering and Rehabilitation*, vol. 1, 2004.

[233] M. P. Furmanek, L. F. Schettino, M. Yarossi, S. Kirkman, S. V. Adamovich, and E. Tunik, "Coordination of reach-to-grasp in physical and haptic-free virtual environments," *Journal of NeuroEngineering and Rehabilitation*, vol. 16, no. 1, p. 78, 2019.

[234] "Oculus rift, available at: https://www.oculus.com/."

[235] "Vicon datastream sdk, available at: https://www.vicon.com/software/datastream-sdk/."

[236] "Vicon vr alignment tool, available at: https://www.vicon.com/software/utilities-and-sdk/vr-alignment-tool/."

[237] "Full body modeling with plug-in gait, available at: https://docs.vicon.com/display/nexus210/plug-in+gait+reference+guide."

[238] M. Cirstea, A. Mitnitski, A. Feldman, and M. Levin, "Interjoint coordination dynamics during reaching in stroke," *Experimental Brain Research*, vol. 151, no. 3, pp. 289–300, 2003.

[239] F. El Jamiy and R. Marsh, "Distance estimation in virtual reality and augmented reality: A survey," in *2019 IEEE International Conference on Electro Information Technology (EIT)*, pp. 063–068, IEEE, 2019.

[240] A. Loftus, S. Murphy, I. McKenna, and M. Mon-Williams, "Reduced fields of view are neither necessary nor sufficient for distance underestimation but reduce precision and may cause calibration problems," *Experimental brain research*, vol. 158, no. 3, pp. 328–335, 2004.

[241] C. González-Alvarez, A. Subramanian, and S. Pardhan, "Reaching and grasping with restricted peripheral vision," *Ophthalmic and Physiological Optics*, vol. 27, no. 3, pp. 265–274, 2007.

[242] S. Murphy, C. Maraj, and J. Hurter, "Field of view: Evaluation of oculus rift and htc vive," 2018.

[243] V. Venkatesh and H. Bala, "Technology acceptance model 3 and a research agenda on interventions," *Decision sciences*, vol. 39, no. 2, pp. 273–315, 2008.

[244] E. M. Kolasinski, *Simulator sickness in virtual environments*, vol. 1027. US Army Research Institute for the Behavioral and Social Sciences, 1995.

[245] R. C. Allen, M. J. Singer, D. P. McDonald, and J. E. Cotton, "Age differences in a virtual reality entertainment environment: A field study," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 44, pp. 542–545, SAGE Publications Sage CA: Los Angeles, CA, 2000.

[246] W. Lo and R. H. So, "Cybersickness in the presence of scene rotational movements along different axes," *Applied ergonomics*, vol. 32, no. 1, pp. 1–14, 2001.

[247] S. Davis, K. Nesbitt, and E. Nalivaiko, "Comparing the onset of cybersickness using the oculus rift and two virtual roller coasters," in *Proceedings of the 11th Australasian Conference on Interactive Entertainment (IE 2015)*, vol. 27, p. 30, 2015.

[248] M. Melo, J. Vasconcelos-Raposo, and M. Bessa, "Presence and cybersickness in immersive content: Effects of content type, exposure time and gender," *Computers & Graphics*, vol. 71, pp. 159–165, 2018.

[249] O. D. Kothgassner, A. Goreis, J. X. Kafka, H. Hlavacs, L. Beutl, I. Kryspin-Exner, and A. Felnhofer, "Agency and gender influence older adults' presence-related experiences in an interactive virtual environment," *Cyberpsychology, Behavior, and Social Networking*, vol. 21, no. 5, pp. 318–324, 2018.

[250] A. R. Roberts, B. De Schutter, K. Franks, and M. E. Radina, "Older adults' experiences with audiovisual virtual reality: perceived usefulness and other factors influencing technology acceptance," *Clinical gerontologist*, vol. 42, no. 1, pp. 27–33, 2019.

[251] S. Syed-Abdul, S. Malwade, A. A. Nursetyo, M. Sood, M. Bhatia, D. Barsasella, M. F. Liu, C.-C. Chang, K. Srinivasan, M. Raja, *et al.*, "Virtual reality among the elderly: a usefulness and acceptance study from taiwan," *BMC geriatrics*, vol. 19, no. 1, p. 223, 2019.

[252] A.-S. Melenhorst, W. A. Rogers, and D. G. Bouwhuis, "Older adults' motivated choice for technological innovation: Evidence for benefit-driven selectivity.," *Psychology and aging*, vol. 21, no. 1, p. 190, 2006.

[253] M. J. Brosnan, "The impact of computer anxiety and self-efficacy upon performance," *Journal of computer assisted learning*, vol. 14, no. 3, pp. 223–234, 1998.

[254] L. Berger and K. Wolf, "Wim: Fast locomotion in virtual reality with spatial orientation gain & without motion sickness," in *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia*, pp. 19–24, 2018.

[255] M. Csikszentmihalyi, *Finding flow: The psychology of engagement with everyday life.* Basic Books, 1997.

[256] N. Hauk, J. Hüffmeier, and S. Krumm, "Ready to be a silver surfer? a meta-analysis on the relationship between chronological age and technology acceptance," *Computers in Human Behavior*, vol. 84, pp. 304–319, 2018.

[257] N. Wenk, J. Penalver-Andres, R. Palma, K. A. Buetler, R. Müri, T. Nef, and L. Marchal-Crespo, "Reaching in several realities: Motor and cognitive benefits of different visualization technologies," in *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, pp. 1037–1042, IEEE, 2019.

[258] E. B. Nash, G. W. Edwards, J. A. Thompson, and W. Barfield, "A review of presence and performance in virtual environments," *International Journal of human-computer Interaction*, vol. 12, no. 1, pp. 1–41, 2000.

[259] M. T. Robert, L. Ballaz, and M. Lemay, "The effect of viewing a virtual environment through a head-mounted display on balance," *Gait & posture*, vol. 48, pp. 261–266, 2016.

[260] H. Akizuki, A. Uno, K. Arai, S. Morioka, S. Ohyama, S. Nishiike, K. Tamura, and N. Takeda, "Effects of immersion in virtual reality on postural control," *Neuroscience letters*, vol. 379, no. 1, pp. 23–26, 2005.

[261] M. Lövdén, P. Ghisletta, and U. Lindenberger, "Social participation attenuates decline in perceptual speed in old and very old age.," *Psychology and aging*, vol. 20, no. 3, p. 423, 2005.

[262] H. Hikichi, K. Kondo, T. Takeda, and I. Kawachi, "Social interaction and cognitive decline: Results of a 7-year community intervention," *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, vol. 3, no. 1, pp. 23–32, 2017.

[263] W. Peng and J. Crouse, "Playing in parallel: The effects of multiplayer modes in active video game on motivation and physical exertion," *Cyberpsychology, behavior, and social networking*, vol. 16, no. 6, pp. 423–427, 2013.

[264] A. L. Snyder, C. Anderson-Hanley, and P. J. Arciero, "Virtual and live social facilitation while exergaming: competitiveness moderates exercise intensity," *Journal of Sport and Exercise Psychology*, vol. 34, no. 2, pp. 252–259, 2012.

[265] A. E. Staiano, A. A. Abraham, and S. L. Calvert, "Adolescent exergame play for weight loss and psychosocial improvement: a controlled physical activity intervention," *Obesity*, vol. 21, no. 3, pp. 598–601, 2013.

[266] M. Tousignant, P. Boissy, H. Corriveau, and H. Moffet, “In home telerehabilitation for older adults after discharge from an acute hospital or rehabilitation unit: A proof-of-concept study and costs estimation,” *Disability and Rehabilitation: Assistive Technology*, vol. 1, no. 4, pp. 209–216, 2006.

[267] M. Cattan, M. White, J. Bond, and A. Learmouth, “Preventing social isolation and loneliness among older people: a systematic review of health promotion interventions,” *Ageing & Society*, vol. 25, no. 1, pp. 41–67, 2005.

[268] N. Krause, “Neighborhood deterioration and social isolation in later life,” *The International Journal of Aging and Human Development*, vol. 36, no. 1, pp. 9–38, 1993.

[269] A. L. Blanchard and M. L. Markus, “The experienced” sense” of a virtual community: characteristics and processes,” *ACM Sigmis Database: the database for advances in information systems*, vol. 35, no. 1, pp. 64–79, 2004.

Appendix A

List of abbreviations

The following list the significance of various acronyms and abbreviations used throughout the thesis. Abbreviations are alphabetically sorted.

AD	Alzheimer's Disease
ADL	Activities of Daily Living
AES	Apathy Evaluation Scale
AM	Attention Matrices
BI	Behavioural Intention
CAIDE	Cardiovascular Risk Factors, Aging, and Incidence of Dementia
CANX	Computer Anxiety
CAVE	Cave Automatic Virtual Environment
CCT	Computerized Cognitive Training
CE	Cash Error
CT	Cognitive Training
CTR	Control group
DB	Digit-span Backward
DE	Dropped item Error
DF	Digit-span Forward
DT	Dual Task
ENJ	Perceived Enjoyment
EXP	Experimental group

FAB	Frontal Assessment Battery
FAQ	Functional Assessment Questionnaire
FE	Fallen item Error
fov	field-of-view
GDS	Global Deterioration Score
HMD	Head Mounted Display
HR	Heart rate
IADL	Instrumental Activities of Daily Living
IJC	Inter-Joint Coordination
IPQ	Igroup Presence Questionnaire
ITC-SOPI	ITC-Sense of Presence Inventory
LCT	Line Cancellation Test
MCI	Mild Cognitive Impairment
MFTC	Multiple Features Target Cancellation
MMSE	Mini-Mental State Examination
MoCA	Montreal Cognitive Assessment
MRI	Magnetic Resonance Imaging
ND	Naming from Description
OUT	Output quality
PE	Physical Exercise
PEC	Perception of External Control
PEOU	Perceived Ease-Of-Use
PET	Positron Emission Tomography
PF	Phonological Fluency
PU	Perceived Usefulness
QoL	Quality of Life
RAVLT_D	Delayed Recall of Rey Auditory Test
RAVLT_I	Immediate Recall of Rey Auditory Test

RCT	Randomized Controlled Trial
ROCFT	Rey-Osterreith Complex Figure Test
RoM	Range of Motion
ROS	Reactive Oxygen Species
RPM	Revolutions Per Minute
RW	Real World
RWC	Real World while holding a Controller
TMT	Trail Making Test
SD	Standard Deviation
SF	Semantic Fluency
SoP	Sense of Presence
SPECT	Single-Photon Emission Computed Tomography
SSD	Short Story Delayed recall
SSI	Short Story Immediate recall
SSQ	Simulator Sickness Questionnaire
SFS	Short Flow Scale
SUS	System Usability Scale
TAM3	Technology Acceptance Model questionnaire, version 3
UI	User Interface
VE	Virtual Environment
VF	Verbal Fluency
VR	Virtual Reality
WE	Wrong item Error
XML	eXtensible Markup Language

Appendix B

Levels of difficulty of the 2D cognitive training

For the performance of the randomized controlled study described in Chapter 2, the virtual environment dedicated to the CS was improved from different points of view §2.9. One of the improvements was related to the increase of the levels of difficulty that passed from 5 to 52. In Table B.1 and Table B.2 and , the features of all these levels of difficulty are described for the *aisle task* and for the *shelf task*, respectively.

Table B.1: Levels of difficulty of the aisle task. Item number represents the number of names displayed on each aisle sign. Y = yes, N = no.

Level	**Aisle number**	**Item number**	**Similar word**
1	2	1	N
2	2	1	N
3	2	1	N
4	2	2	N
5	2	2	N
6	3	1	N
7	3	1	N
8	3	2	N
9	3	2	N
10	3	2	N
11	3	3	N
12	3	3	N
13	3	4	N
14	3	4	N
15	4	1	N
16	4	1	N
17	4	2	N
18	4	2	N
19	4	2	N
20	4	3	N
21	4	3	N

22	4	4	N
23	4	4	N
24	2	1	Y
25	2	1	Y
26	2	2	Y
27	2	2	Y
28	2	2	Y
29	3	2	Y
30	3	2	Y
31	3	3	Y
32	3	3	Y
33	3	3	Y
34	3	3	Y
35	3	4	Y
36	3	4	Y
37	3	4	Y
38	4	2	Y
39	4	2	Y
40	4	2	Y
41	4	3	Y
42	4	3	Y
43	4	3	Y
44	4	3	Y
45	4	3	Y
46	4	4	Y
47	4	4	Y
48	4	4	Y
49	4	4	Y
50	4	4	Y
51	4	4	Y
52	4	4	Y
53	4	4	Y
54	4	4	Y
55	4	4	Y
56	4	Y	Y

Table B.2: Levels of difficulty of the shelf task. Similarity determines how similar the products not randomly placed are; VarFactor is the percentage of the products on the shelves placed picking completely random objects. Y = yes, N = no.

Level	**List length**	**Max items**	**Similarity**	**Small items**	**Discount items**	**Var Factor**	**Hidden** textbflist
1	3	2	1	N	N	0.6	N
2	3	5	1	N	N	0.6	N
3	3	10	1	N	N	0.6	N
4	3	25	1	N	N	0.6	N
5	3	40	1	N	N	0.6	N
6	5	10	2	N	N	0.6	N
7	5	25	2	N	N	0.6	N
8	5	25	2	N	N	0.2	N
9	5	40	2	N	N	0.6	N
10	5	40	2	N	N	0.2	N
11	5	10	2	Y	N	0.6	N
12	5	25	2	Y	N	0.6	N
13	5	25	2	Y	N	0.2	N
14	5	40	2	Y	N	0.6	N
15	5	40	2	Y	N	0.2	N
16	5	10	2	Y	Y	0.6	N
17	5	25	2	Y	Y	0.6	N
18	5	25	2	Y	Y	0.2	N
19	5	40	2	Y	Y	0.6	N
20	5	40	2	Y	Y	0.2	N
21	1	10	2	N	N	0.4	Y
22	1	25	2	N	N	0.4	Y
23	1	40	2	N	N	0.4	Y
24	1	10	2	Y	N	0.4	Y
25	1	25	2	Y	N	0.4	Y
26	1	40	2	Y	N	0.4	Y
27	1	10	2	Y	Y	0.4	Y
28	1	25	2	Y	Y	0.4	Y
29	1	40	2	Y	Y	0.4	Y
30	2	10	2	Y	Y	0.4	Y
31	2	25	2	Y	Y	0.4	Y
32	2	40	2	Y	Y	0.4	Y
33	3	10	2	Y	Y	0.4	Y
34	3	25	2	Y	Y	0.4	Y
35	3	40	2	Y	Y	0.4	Y
36	4	10	2	Y	Y	0.4	Y
37	4	25	2	Y	Y	0.4	Y
38	4	40	2	Y	Y	0.4	Y
39	5	10	2	Y	Y	0.4	Y
40	5	25	2	Y	Y	0.4	Y
41	5	40	2	Y	Y	0.4	Y

42	6	10	2	Y	Y	0.4	Y
43	6	25	2	Y	Y	0.4	Y
44	6	40	2	Y	Y	0.4	Y
45	7	10	2	Y	Y	0.4	Y
46	7	25	2	Y	Y	0.4	Y
47	7	40	2	Y	Y	0.4	Y
48	8	10	2	Y	Y	0.4	Y
49	8	25	2	Y	Y	0.4	Y
50	8	40	2	Y	Y	0.4	Y
51	9	10	2	Y	Y	0.4	Y
52	9	25	2	Y	Y	0.4	Y
53	9	40	2	Y	Y	0.4	Y
54	10	10	2	Y	Y	0.4	Y
55	10	25	2	Y	Y	0.4	Y
56	10	40	2	Y	Y	0.4	Y

Appendix C

Reach and transfer gesture data

Table C.1: Mean (SD) values for the young adults' group. MT_r=movement time during reaching, $MTMT_t$= movemebt time during transport, V_r=reaching velocity, V_t=transport velocity, C=curvature, Should=shoulder, Adb=abduction, Flex=flexion, Elb=elbow, Rot=rotation [28].

		CT	MT	IT	CC	MC	IC	CB	MB	IB
MT$_r$ [s]	VR	1.59(0.37)	1.41(0.63)	1.44(0.34)	1.67(0.40)	1.38(0.42)	1.70(0.82)	1.54(0.41)	1.55(0.34)	1.24(0.33)
	RW	1.02(0.18)	1.07(0.18)	0.98(0.13)	1.01(0.19)	0.92(0.17)	1.01(0.18)	0.97(0.14)	1.05(0.24)	0.87(0.16)
	RWC	1.11(0.24)	1.11(0.27)	1.01(0.22)	1.08(0.28)	1.04(0.26)	1.01(0.25)	1.06(0.21)	1.05(0.19)	1.00(0.19)
MT$_t$ [s]	VR	1.58(0.45)	1.45(0.81)	1.44(0.38)	1.73(0.55)	1.31(0.32)	1.65(0.76)	1.46(0.37)	1.53(0.34)	1.14(0.23)
	RW	0.99(0.17)	1.07(0.19)	0.98(0.13)	1.02(0.19)	0.91(0.18)	0.99(0.18)	0.97(0.15)	1.06(0.27)	0.87(0.14)
	RWC	1.10(0.24)	1.10(0.27)	1.02(0.26)	1.09(0.29)	1.01(0.24)	1.02(0.25)	1.06(0.22)	1.03(0.19)	0.99(0.18)
V$_r$ [m/s]	VR	1.14(0.47)	1.16(0.41)	1.27(0.38)	1.17(0.43)	1.18(0.39)	1.22(0.28)	1.32(0.52)	1.17(0.40)	1.33(0.45)
	RW	1.98(0.36)	1.80(0.42)	1.94(0.37)	2.29(0.44)	1.96(0.52)	1.77(0.33)	2.24(0.37)	2.14(0.38)	1.79(0.35)
	RWC	1.50(0.37)	1.39(0.34)	1.47(0.32)	1.50(0.32)	1.45(0.32)	1.30(0.27)	1.56(0.38)	1.41(0.27)	1.18(0.32)
V$_t$ [m/s]	VR	1.68(0.33)	1.51(0.32)	1.72(0.30)	1.75(0.33)	1.67(0.28)	1.56(0.32)	1.71(0.41)	1.66(0.32)	1.43(0.28)
	RW	1.73(0.31)	1.66(0.37)	1.73(0.40)	1.82(0.23)	1.69(0.38)	1.71(0.25)	1.91(0.31)	1.76(0.29)	1.24(0.32)
	RWC	1.44(0.29)	1.44(0.26)	1.56(0.28)	1.50(0.34)	1.46(0.29)	1.41(0.21)	1.54(0.32)	1.47(0.31)	1.20(0.21)
C [-]	VR	1.65(2.07)	1.04(0.30)	1.31(0.94)	1.72(2.37)	0.97(0.28)	1.69(1.99)	1.43(1.12)	1.54(1.23)	1.70(2.09)
	RW	1.12(0.09)	1.11(0.23)	1.10(0.05)	1.09(0.03)	1.04(0.12)	1.27(0.47)	1.10(0.06)	1.25(0.52)	1.21(0.19)
	RWC	1.28(0.62)	1.12(0.15)	1.13(0.22)	1.08(0.07)	1.30(0.71)	1.06(0.05)	1.41(0.75)	1.67(1.42)	1.41(0.55)
Should Adb [°]	VR	57.11(14.98)	57.12(15.65)	64.51(21.97)	61.78(17.81)	53.15(15.95)	50.88(19.91)	66.91(29.49)	57.31(12.43)	38.91(24.17)
	RW	67.41(18.93)	68.58(15.16)	79.29(19.94)	75.05(27.11)	57.40(17.42)	61.33(18.50)	68.69(30.68)	57.89(15.43)	45.62(21.20)
	RWC	43.81(22.05)	39.49(22.96)	47.51(29.86)	39.57(20.38)	29.16(12.65)	25.61(13.96)	55.32(36.96)	39.57(19.87)	21.67(14.15)
Should Flex [°]	VR	63.91(16.61)	61.69(18.85)	65.43(12.60)	67.24(14.73)	63.40(13.81)	62.08(16.82)	66.65(18.53)	63.17(11.82)	48.24(12.77)
	RW	72.44(6.92)	71.87(11.76)	71.18(9.44)	79.75(27.67)	69.52(13.71)	69.33(15.24)	73.07(8.81)	72.19(8.90)	58.07(10.33)
	RWC	72.38(13.83)	71.20(15.65)	69.16(15.94)	71.20(15.28)	66.29(5.74)	68.61(17.95)	76.30(17.34)	69.18(11.27)	54.25(16.35)
Elb Flex [°]	VR	76.23(13.15)	71.09(11.14)	77.67(12.55)	79.41(9.55)	74.73(10.36)	74.81(11.45)	76.01(11.78)	74.82(10.95)	60.16(14.33)
	RW	71.95(6.39)	71.87(11.76)	71.67(9.84)	77.51(28.05)	69.52(13.71)	71.57(15.86)	73.78(8.45)	72.19(8.90)	57.36(9.39)
	RWC	72.20(11.74)	66.56(13.41)	69.80(11.89)	67.33(14.75)	70.85(15.32)	66.00(14.76)	68.92(9.98)	71.76(14.43)	62.67(15.28)
Trunk Tilt [°]	VR	13.88(11.19)	13.41(10.26)	14.48(10.94)	13.59(10.13)	15.24(12.66)	13.72(9.75)	14.63(10.89)	13.69(10.35)	12.40(8.96)
	RW	8.34(3.85)	10.06(5.31)	9.83(4.85)	8.72(3.29)	10.13(7.02)	10.86(5.77)	10.39(4.09)	8.59(2.94)	11.18(6.15)
	RWC	5.98(2.57)	7.35(3.68)	6.70(2.44)	6.55(2.46)	7.56(4.64)	6.74(2.65)	7.02(2.19)	7.10(2.96)	7.12(5.17)
Trunk Rot [°]	VR	66.21(16.52)	63.35(9.84)	69.59(15.93)	67.88(11.10)	66.45(11.51)	67.08(19.59)	72.60(12.66)	68.52(11.30)	52.70(13.92)
	RW	44.59(6.03)	42.95(11.04)	39.83(7.32)	46.99(8.59)	42.36(7.01)	45.76(10.06)	50.25(7.24)	45.45(8.81)	37.97(14.45)
	RWC	30.77(8.72)	27.40(7.17)	30.30(8.62)	35.31(9.55)	30.43(7.58)	30.00(7.41)	35.52(8.63)	33.20(10.05)	29.01(8.29)

Table C.2: Mean (SD) values for the older adults' group. MT_r=movement time during reaching, $MTMT_t$= movemebt time during transport, V_r=reaching velocity, V_t=transport velocity, C=curvature, Should=shoulder, Adb=abduction, Flex=flexion, Elb=elbow, Rot=rotation [28].

		CT	MT	IT	CC	MC	IC	CB	MB	IB
MT$_r$ [s]	VR	2.04(0.62)	2.48(0.82)	2.06(0.14)	2.16(0.62)	1.82(0.17)	2.27(0.70)	2.49(0.77)	1.66(0.46)	2.15(0.64)
	RW	0.94(0.14)	1.13(0.04)	1.03(0.26)	1.06(0.43)	1.09(0.18)	1.21(0.17)	0.95(0.18)	1.00(0.34)	0.85(0.09)
	RWC	1.26(0.21)	1.09(0.19)	1.17(0.18)	1.08(0.41)	1.39(0.40)	1.35(0.34)	1.23(0.14)	1.15(0.31)	1.14(0.35)
MT$_t$ [s]	VR	2.04(0.55)	2.36(0.79)	2.12(0.16)	2.14(0.62)	2.03(0.34)	2.19(0.63)	2.28(0.64)	1.65(0.46)	2.13(0.58)
	RW	0.93(0.14)	1.13(0.05)	1.09(0.31)	1.00(0.42)	1.04(0.16)	1.17(0.19)	0.93(0.17)	1.03(0.28)	0.85(0.07)
	RWC	1.27(0.22)	1.09(0.24)	1.14(0.22)	1.08(0.43)	1.36(0.27)	1.32(0.27)	1.20(0.19)	1.11(0.26)	1.04(0.25)
V$_r$ [m/s]	VR	1.47(0.31)	1.30(0.14)	1.27(0.27)	1.28(0.43)	1.20(0.35)	1.26(0.42)	1.38(0.22)	1.19(0.10)	1.28(0.35)
	RW	1.83(0.55)	1.75(0.36)	1.83(0.41)	2.07(0.41)	1.77(0.66)	1.49(0.16)	1.92(0.33)	1.79(0.47)	1.51(0.09)
	RWC	1.51(0.56)	1.71(0.49)	1.19(0.54)	1.57(0.56)	1.47(0.68)	1.39(0.48)	1.46(0.62)	1.51(0.58)	1.10(0.51)
V$_t$ [m/s]	VR	1.37(0.33)	1.63(0.21)	1.45(0.21)	1.58(0.11)	1.52(0.11)	1.49(0.20)	1.57(0.24)	1.37(0.22)	1.20(0.14)
	RW	1.65(0.20)	1.87(0.39)	1.71(0.43)	1.80(0.23)	1.87(0.03)	1.54(0.16)	1.83(0.18)	1.54(0.25)	1.11(0.14)
	RWC	1.83(0.21)	1.95(0.36)	1.82(0.30)	1.93(0.35)	1.78(0.30)	1.57(0.19)	1.89(0.40)	1.72(0.39)	1.26(0.40)
C [-]	VR	1.12(0.01)	1.35(0.06)	1.18(0.06)	1.24(0.09)	1.18(0.11)	1.20(0.05)	1.12(0.04)	1.11(0.04)	1.36(0.35)
	RW	1.10(0.04)	1.14(0.07)	1.14(0.03)	1.09(0.04)	1.16(0.08)	1.16(0.05)	1.07(0.03)	1.11(0.02)	1.26(0.05)
	RWC	1.09(0.03)	1.06(0.03)	1.13(0.07)	1.06(0.03)	0.95(0.36)	1.12(0.06)	1.12(0.08)	1.08(0.03)	1.30(0.25)
Should Adb [°]	VR	63.00(7.94)	60.20(18.68)	53.38(7.66)	73.82(13.48)	42.83(16.09)	57.10(10.75)	72.76(25.51)	53.11(6.06)	43.03(25.97)
	RW	71.87(3.56)	70.00(10.89)	65.37(8.43)	79.62(5.55)	80.04(51.23)	46.47(23.76)	78.01(19.34)	54.72(15.82)	35.09(13.67)
	RWC	35.97(8.59)	36.05(6.15)	28.94(11.52)	38.13(5.13)	25.87(4.86)	31.24(5.07)	37.15(4.18)	36.83(12.44)	19.39(6.72)
Should Flex [°]	VR	65.26(16.70)	61.69(18.85)	64.08(12.49)	68.77(14.96)	63.40(13.81)	60.54(15.91)	66.60(18.56)	63.17(11.82)	48.29(12.80)
	RW	72.44(6.92)	71.87(11.76)	71.18(9.44)	79.75(27.67)	69.52(13.71)	69.33(15.24)	73.07(8.81)	72.19(8.90)	58.07(10.33)
	RWC	73.66(15.76)	71.20(15.65)	67.87(13.56)	72.39(16.13)	66.29(5.74)	67.42(16.90)	76.57(17.41)	69.18(11.27)	53.97(15.85)
Elb Flex [°]	VR	62.50(12.33)	71.85(2.54)	70.51(9.25)	73.18(7.86)	64.27(5.05)	59.63(8.61)	72.44(11.42)	65.79(5.19)	49.87(0.39)
	RW	67.65(4.18)	75.89(4.18)	66.89(11.35)	75.29(3.61)	72.09(9.85)	67.34(15.17)	68.57(9.39)	68.13(9.08)	52.78(9.60)
	RWC	72.32(11.55)	63.12(11.35)	60.61(19.69)	74.01(12.73)	74.93(12.29)	61.67(9.87)	66.08(16.92)	62.84(17.24)	43.46(11.95)
Trunk Tilt [°]	VR	11.55(7.02)	9.98(6.08)	8.99(2.55)	7.73(4.30)	12.52(5.56)	10.54(4.82)	12.64(7.06)	11.25(5.01)	12.29(4.74)
	RW	11.10(3.98)	8.54(1.61)	10.73(4.73)	8.77(4.57)	10.62(3.79)	10.44(3.21)	10.29(2.55)	8.04(3.53)	8.86(4.44)
	RWC	5.91(4.83)	5.42(1.32)	4.88(0.62)	8.13(7.65)	6.11(3.26)	5.78(3.38)	7.81(7.28)	8.21(5.88)	6.96(7.15)
Trunk Rot [°]	VR	54.20(23.20)	56.45(20.22)	58.64(17.30)	57.44(19.25)	54.53(16.40)	59.39(23.51)	63.30(21.19)	53.45(27.28)	49.66(25.89)
	RW	32.73(12.38)	38.00(13.05)	35.30(10.81)	41.06(12.96)	36.07(9.78)	34.24(10.07)	37.11(12.31)	35.90(18.74)	29.50(6.71)
	RWC	36.25(24.87)	39.40(24.08)	39.24(20.09)	37.70(17.47)	36.70(14.88)	33.91(21.73)	37.99(25.07)	38.51(30.34)	33.13(20.89)