Presence and cognitive performances in virtual environment for cognitive training: an exploratory study in healthy aging and mild cognitive impairment

Pierre-Alexandre Andrieu-Devilly¹, Marc Gandit², Didier Schwab³, Christelle Nahas² ⁴ ⁵, Lisa Quillon-Dupr² ⁴, Emmanuel Monfort⁵

¹Univ. Grenoble Alpes, UFR SHS, 38000 Grenoble, France
²Univ Grenoble Alpes, LIP/PC2S, EA 4145, 38000 Grenoble, France
³Univ. Grenoble Alpes, CNRS, UMR 5217, Grenoble INP, LIG, 38000 Grenoble, France
⁴ COVIRTUA Healthcare, 31770 Colomiers, France
⁵ Univ. Grenoble Alpes, CNRS, UMR 5525, VetAgro Sup, Grenoble INP, TIMC, 38000 Grenoble, France
emmanuel.monfort@univ-grenoble-alpes.fr

Abstract

Background: The growing field of artificial intelligence has opened new opportunities in computerized cognitive-training design (CCT), closer to everyday situations, specifically in the prevention of neurocognitive disorders. The aim of our exploratory study was to characterize the importance of social and cognitive facilitators in avatar-mediated CCT with reference to aging and minor neurocognitive disorders and to specify its relationship with cognitive difficulties.

Methods: We conducted a quasi-experimental study using a virtual environment simulating a supermarket. We compared performance and sense of spatial and social presence for 16 young adults, 16 healthy older adults, and 17 older adults with mild neurocognitive disorders.

Results: Older adults (with and without neurocognitive impairment) showed significantly poorer performance on the social presence dimension, but not on the spatial presence and self-dimensions. In contrast, healthy adults (young or old) make significantly fewer errors of execution on the travelling task in a virtual environment that included verbal support mediated by an avatar.

Conclusion: Considering that two types of facilitators can be considered for the use of CCT (social facilitators that can take the form of avatars, and cognitive facilitators in the form of structured cognitive support), then the results obtained suggest that only cognitive support would be decisive.

Keywords: Virtual environment, social presence, computerized cognitive training, aging

I. INTRODUCTION

Age-related neurocognitive disorders are one of the most disabling health problems and a challenge in the context of an aging world population [1]. Their prevalence and incidence may increase in the coming decades [2]. A significant consequence of these disorders is the decline of high-level instrumental activities of daily living (IADLs) right from the prodromal stage [3]. A qualitative and targeted assessment of daily life difficulties can therefore open up avenues for targeted therapeutic management to delay the progression of cognitive impairment [4]. A promising approach is to implement support strategies as early as possible, using compensatory structures, particularly through computerized cognitive training (CCT) [5].

A. Computerized Cognitive Training

A benefit of CCT is admitted for older persons with MCI on cognitive performance, but no transfer to daily life skills is recognized [6]. This is also the case following more focused strokes [7]. This implies taking better account of veridicality [8], the degree of an empirical relationship between the measure used and those that can be applied to daily life functioning, with more subtle measures of functional activities [9]. In addition, a limitation of CCT lies in the lack of characterization of the commitment, motivation, and meaning that the participants may associate with this type of training [10]. A key factor in the meaning that users give is narrative transport, a process by which an individual is mentally "transported" from the physical world to an imaginary world, through the story told by the environment [11]. Verisimilitude [8], the need for the cognitive demands of training to be similar to the cognitive demands in an everyday environment, is now made possible through immersion and interactivity. These appear to be particularly important for the experience associated with training, as they would help people with MCI to better engage in real-life experiences [12].

B. Ecological Training in Virtual Environments
It is possible to design ecological tasks based on their realization within virtual environments (VE), and not just based on the transposition of decontextualized tasks onto a computer medium [13]. Immersion then becomes an essential objective sought by VE, to properly account for the user’s movements in digitally composed surroundings [14]. Chirico and colleagues [15] have recently shown the feasibility of functional assessment in an immersive kitchen environment. In this context, taking subtle errors into account could then allow for the development of more targeted strategies for performing simulated activities in VE.

An approach to assessing the validity of ecological tasks in VE is to develop an ad hoc model of participants’ activity. This approach is supported by the association between different types of errors in IADLs with deficits on classic neuropsychological decontextualized tasks and on alteration of the cerebral functioning [16]. In particular, cognitive control could be a critical process, including for simple and automated actions, as automated in VE. Giovannetti and colleagues [17] proposed to use a generic model of goal control to account for errors in everyday action, which distinguishes between two types of errors: deficits in competing object and goal control (which include errors of commission, i.e. inaccurate performance of task steps) and deficits in goal activation (which include omissions, i.e. failure to complete task steps). Control operates in a top-down activation, which allows for a shift from a general goal to an adapted behavioral sequence, and in a bottom-up activation, which allows for real-time adaptation to the context [18]. Thus, age-related cognitive decline may manifest itself in subtle errors that do not necessarily result in failure to complete the task, whereas disease-related cognitive decline may manifest as higher-level errors.

C. Social Interaction in Virtual Environments

Cognitive training requires positive facilitation strategies to create a safe and comfortable environment related to supportive facilitation without inducing dilemmas [19]. This is particularly the case for supporting the use of self-correction skills that are relevant during supportive interactions with older adults using a technological device [20]. An approach that assumes that confrontation with errors should be kept to a minimum to support learning for people with neurocognitive impairments [21] can help structure support tailored to the level of understanding of older adults with cognitive impairments. It has shown relevance in virtual tasks [22]. Thus, CCT using VE should certainly appeal to these principles.

VE also implies taking better account of the social dimension of the interactions that take place within it. This seems particularly relevant because immersion can be obtained through the means of characters who play a driving role or describe a story (avatars), or through a similarity effect between the characters in the game and the players themselves, to promote immersion [23]. Improving the sense of realism in a VE is notably based on the sense of presence [24], that is to say by the feeling of ‘being there’ which can refer to (1) spatial presence, (2) social presence, or (3) self-presence [25]. Enhancing this sense of presence may improve interactions with the digital interface and could optimize the therapeutic benefit of CCTs [26]. However, to our knowledge, the role of social interactions in CCTs is poorly understood, especially when mediated by an avatar.

The study aimed to investigate the social and cognitive aspects of interactions in a virtual ecological task with verbal support among older adults with and without MCI. It was hypothesized that older adults without MCI would perform better than those with MCI on decontextualized tasks and that older adults with MCI would make more execution errors and require more support. Additionally, the aim was to explore the relationship between presence and goal achievement during a contextualized task in non-immersive VE.

II. Method

The study applied a quasi-experimental design. The experimental protocol was approved by UGA’s ethical committee (agreement CERGA-Avis-2021-12). Methods were carried out in accordance with relevant guidelines and regulations, and informed consent was obtained from all participants.

D. Participants

A convenience sample was used to recruit participants who were over 65 years old, and under 30 years old. The test group consisted of older adults with MCI older than 65 years (N = 17, M = 85.12 years, SD = 7.87 years, age range = 68–99, 11 females – 6 males). Participants were assigned to the MCI group based on the diagnostic criteria for MCI as defined by Petersen and colleagues [27]. The two healthy control groups consisted of young adults between 20 and 30 years of age (N = 16, M = 23.75 years, SD = 2.14 years, age range = 20–29, 7 females – 9 males), and those older than 65 years (N = 17, M = 70.06 years, SD = 4.39 years, age range = 65–78, 7 females – 11 males), respectively.

E. Materials

1) Computerized Tasks
Participants performed a series of computerized tasks developed for the CCT, on a 26.4 cm diagonal Samsung Galaxy Tab S6 Lite tablet. These were decontextualized tasks, followed by a contextualized task. They performed using the COVIRTUA software (See https://www.youtube.com/watch?v=_Sw-Qrh_MjA for an introduction of the COVIRTUA software). Four different tasks had to be completed obtaining performance indicators: three decontextualized tasks (a selective attention task, a visual semantic-categorization task, and a working-memory updating task), and a contextualized task (presented as a game in a non-immersive virtual environment). The VE represented a supermarket, arranged in thematic departments. The participant moved around in this environment through a first-person view (Figure 1). It was possible to move forward in successive increments and to orient the camera at 360° in both dimensions of the screen. The experimenter could manipulate the avatar of an in-game supermarket employee to interact with the participant. The participants could move around the supermarket aisles and interact verbally with the experimenter’s avatar through a microphone. The test phase was a moving target tracking task.

![Figure 1. Virtual environment: the simulated supermarket (contextualized task)](image)

2) Coding and Indicators

Two categories of indicators were collected for the VE task. First, we considered goal achievement (i.e. success and speed of task execution) and second three indicators for the control of gestures needed to achieve the goals, with reference to the goal-control model [17]: wandering errors (i.e. an unnecessary round trip in an aisle), corresponding to a deficit in maintaining goal activation, and orientation errors (i.e. turning in the wrong shelf), corresponding to a deficit of control of the information in competition. Orientation errors could be uncorrected (i.e. turning in the wrong shelf without turning back), or self-corrected (i.e. turn in the wrong shelf but turn back). They were operationalized by analyzing the participant’s in-game travel pattern, in reference to 55 possible positions (figure 2), to create an index for each type of error [(number of errors / number of in game moves) * 100].

![Figure 2. 55 possible positions in the simulated supermarket (in green: right pathway, in red: wrong pathway)](image)

During a test phase, the participants were invited to follow the sales assistant through the aisles, with the aim of gaining ease and familiarity with the VE. According to Anselme and colleagues [28], we defined four types of support provided to participants. The first three categories involve verbal interactions, ordered by the level of accuracy of the information provided to the participant: general supports (incomplete information with respect to the expected action, e.g. "be careful"), specific supports (specific information for the realization of the expected action, without explaining the procedure necessary for its realization, e.g. "you started on the wrong side"), and total supports (specific information for the realization of the expected action, and explaining the procedure necessary for its realization, e.g. "I'm on your left"). The last type of support, the physical supports, concerns the direct intervention on the tablet (e.g. the participant continues to confuse the direction and movement gestures, and requires a new demonstration). We calculated a support index for each type, in the same way as for the error indices [(number of support interactions / number of in game moves) * 100].

3) Sense of presence

To assess the three dimensions of presence, we used a French adaptation of the Multimodal Presence Scale for virtual reality environments (MPS) [29].
III. RESULTS

A. Contextualized Task

Considering goal achieving of the contextualized task, 100% of the young adults ($N = 16$) and healthy elderly people ($N = 17$) reached the finish area, compared to 18.8% of the elders with MCI ($N = 16$). A KW test showed that travel speeds – measured by the number of movements per second (m.s$^{-1}$) – are significantly affected by groups, $H(2) = 41.6$, $p < .001$. The DSCF post-hoc test tells us that this difference is between young ($Mdn = 0.687$ m.s$^{-1}$) and healthy older participants ($Mdn = 0.362$ m.s$^{-1}$), $p < .001$, as well as between older participants without and with MCI ($Mdn = 0.126$ m.s$^{-1}$), $p < .001$.

Considering execution errors during the contextualized task, KW tests indicated a statistically significant difference between groups for wandering errors, $H(2) = 19.01$, $p < .001$, and uncorrected orientation errors, $H(2) = 24.86$, $p < .001$, but not for self-corrected orientation errors, $H(2) = 4.02$, $p = .134$. Bonferroni-adjusted Dunn post-hoc tests revealed no difference between young participants and healthy older participants for either wandering errors, $p = 1$, or uncorrected orientation errors, $p = 1$. Differences between healthy older participants and older participants with MCI were statistically significant for wandering errors, $p < .001$, and for uncorrected orientation errors, $p < .001$ (Figure 3).

![Figure 3. Error frequency depending on groups and on error types, for young adults, healthy older adults, and MCI older adults. Asterisks denote significant differences (p < .05)](image)

Considering verbal support during the contextualized task, KW tests indicate a statistically significant difference between groups for general supports $H(2) = 35.4$, $p < .001$, specific supports, $H(2) = 33.6$, $p < .001$, and total supports, $H(2) = 32.3$, $p < .001$.

Bonferroni-adjusted Dunn post-hoc tests revealed no difference between healthy young and older participants for general support, $p = .062$, specific support, $p = .073$, and total support, $p = .269$. Differences between older participants without MCI and older participants with MCI were statistically significant for general supports, $p < .001$, specific supports, $p < .001$, and total supports, $p < .001$.

With regard to errors of execution, correlations using Bonferroni-adjusted alpha threshold of .016 (0.05/3) between errors and finish area reaching show a negative relationship for wandering errors, $tb = -.354$, $p = .004$, and uncorrected orientation errors, $tb = -.529$, $p = .001$, but no significant correlations were found between self-corrected errors and finish area reaching, $tb = -.074$, $p = .551$.

With regard to verbal support, the stronger the participant’s need for support, the lower the finish area is reached. At the Bonferroni-adjusted alpha threshold of .16 (.05/3), general verbal supports, $tb = -.398$, $p < .001$, and specific verbal supports, $tb = -.333$, $p = .007$, are negatively correlated with finish area reaching. Similarly, total verbal supports are trendily correlated with finish area reaching, $tb = -.298$, $p = .017$.

Considering correlation between decontextualized and contextualized tasks performance, it is notable that finish area reaching was significantly correlated with selective attention performance, $tb = -.325$, $p = .009$, and working memory update performance, $tb = -.405$, $p = .001$. No significant correlation were found between finish area reaching and visual semantic categorisation performance, $tb = -.120$, $p = .342$.

B. Sense of presence
Considering all participants, a high level of spatial presence is reported \((M = 3.63, SD = 1.20)\), a slightly lower level of social presence was reported \((M = 3.10, SD = 1.14)\), and a near absence of self-presence was reported \((M = 0.35, SD = 0.58)\) (Figure 4).

Under the assumption of a response difference by group, a Fisher's exact test applied to the levels measured for each dimension of presence \((N = 49)\) reveals a significant intergroup difference only for the feeling of social presence, \(p = .002, \text{ FET}\). Moreover, Fisher’s exact tests allow us to reject null hypothesis of an age-related difference, \(p = .007, \text{ FET}\), but not the null hypothesis of an impairment-related difference, \(p = .800, \text{ FET}\).

### IV. Discussion

The management of neurocognitive aging can be considered on the basis of detailed and contextualized evaluations, and then by training adapted to the characteristics resulting from these evaluations. CCT by means of complex tasks closer to daily life is today a promising opportunity to limit the consequences of pathological aging, in particular for early disorders [5]. VE is therefore particularly interesting, as it potentially provides a way of better engaging the patient who is concerned [30]. Although little is known about the determinants of engagement in these simulated activities, especially when training is mediated by an avatar, our results provide notable information.

An important perspective is that ecological tasks performed in VE could be closer to cognitive functioning in daily life. An approach centred on functional activities offers the possibility of early detection of cognitive impairment, insofar as functions that are very sensitive to decline are solicited by IADLs, such as executive functions or prospective memory [13].

The goal control model [17], allows the interpretation of interactions with a VE that involves movement and thus sets precise guidelines to design adapted aids: an increase in the number of errors, reflects difficulty in activating correct goal representations, and an increase in the number uncorrected orientation errors reflects the degradation of goal representations. Our results suggest a continuum of severity, whereby the greater the impairment, the greater the number of errors in general and of low-level errors in particular. Older adults with MCI have significantly less successful goal reaching and more wandering and orientation errors. Yet, orientation errors would be high-level difficulties, which occur more frequently, but are also more easily controllable by older people without neurocognitive impairment. The wanderings would be lower-level errors, which first characterize older people with neurocognitive disorders. These interpretations must, however, be qualified by the possibility of clumsy errors [15], even though the touch screen can improve accuracy and psychomotor speed compared to using a mouse [31].

In contrast to actual errors, it appears that the greater the impairment, the less self-corrected the errors were. In particular, wanderings are not spontaneously corrected, but orientation errors could potentially be corrected by older people, and much less so when they had a neurocognitive disorder. This is certainly a key point in contributing to task completion, insofar as self-correction abilities are essential resources in cognitive support [20]. An adapted EV should certainly consider encouraging these self-corrections to increase its effectiveness, but also collect this indicator to provide fine-grained information on the skill level of the older people.

Our results show that young adults require less verbal support than healthy older adults, who in turn require less verbal support than older adults with MCI. Older adults with MCI would require a more specific level of verbal support than healthy older adults. An important finding was that the verbal support provided by the avatar according to the hierarchical approach proposed by Thivierge and colleagues [21] would not significantly limit errors, but would be sufficient to enable participants to complete the task. This is therefore a potentially useful approach, which could be automated in CCT.

Regarding the sense of presence, the results show a significant difference in the level of social presence between adults and older adults, as well as no significant difference between healthy older adults and older adults with MCI. Yet, this is a critical dimension of sense of presence [32]. Thus, this result suggests a preservation of a sense of social presence in older adults with MCI, even though verbal support is less effective than in healthy older adults. Consequently, verbal support through avatar mediation may not disrupt engagement in the narrative framework of a VE task.

### V. Conclusion and Perspectives
In view of our results, while artificial intelligence may now improve the realism of VEs and sense of presence, it is crucial to integrate an accurate cognitive model of errors related to cognitive decline into CCTs, in order to increase participants' engagement and the effectiveness of the help provided. Future developments will involve incorporating automated graded aids that promote sustained cognitive training, first on orientation, then on wandering. At the same time, it is essential to incorporate an accurate cognitive framework of errors related to cognitive decline into CCTs to increase participant engagement and detection of neurocognitive disorder-specific error patterns.

ACKNOWLEDGMENT
This research has received funding from the Pôle Grenoble Cognition (PGC) under grant agreement and has been financially supported by MIAI@Grenoble-Alpes, (ANR-19-P3IA-0003).

REFERENCES
[29] G. Makransky, L. Lilleholt and A. Aaby. Development and validation of the Multimodal Presence Scale for virtual rea...

