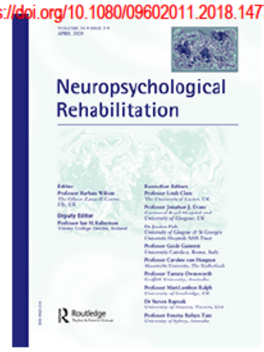

Routledge
Taylor & Francis Group

Neuropsychological Rehabilitation

An International Journal

ISSN: 0960-2011 (Print) 1464-0694 (Online) Journal homepage: https://www.tandfonline.com/loi/pnrh20

Use of immersive virtual reality to assess episodic memory: A validation study in older adults

Nick Corriveau Lecavalier, Émilie Ouellet, Benjamin Boller & Sylvie Belleville

To cite this article: Nick Corriveau Lecavalier, Émilie Ouellet, Benjamin Boller & Sylvie Belleville (2020) Use of immersive virtual reality to assess episodic memory: A validation study in older adults, Neuropsychological Rehabilitation, 30:3, 462-480, DOI: 10.1080/09602011.2018.1477684

To link to this article: https://doi.org/10.1080/09602011.2018.1477684

NEUROPSYCHOLOGICAL REHABILITATION
2020, VOL. 30, NO. 3, 462–480
https://doi.org/10.1080/09602011.2018.1477684

Use of immersive virtual reality to assess episodic memory: A validation study in older adults

Nick Corriveau Lecavalier[a,b], Émilie Ouellet[a,b], Benjamin Boller[a,b] and Sylvie Belleville[a,b]

[a]Research Centre, Institut universitaire de gériatrie de Montréal, Montreal, Québec, Canada; [b]Department of Psychology, University of Montreal, Montreal, Canada

ABSTRACT

Virtual reality (VR) allows for the creation of ecological environments that could be used for cognitive assessment and intervention. This study comprises two parts that describe and assess an immersive VR task, the *Virtual Shop*, which can be used to measure episodic memory. Part 1 addresses its applicability in healthy older adults by measuring presence, motivation, and cybersickness symptoms. Part 2 addresses its construct validity by investigating correlations between performance in the VR task and on a traditional experimental memory task, and by measuring whether the VR task is sensitive to age-related memory differences. Fifty-seven older and 20 younger adults were assessed in the *Virtual Shop*, in which they memorised and fetched 12 familiar items. Part 1 showed high levels of presence, higher levels of motivation for the VR than for the traditional task, and negligible cybersickness symptoms. Part 2 indicates that memory performance in the VR task is positively correlated with performance on a traditional memory task for both age groups, and age-related differences were found on the VR and traditional memory tasks. Thus, the use of VR is feasible in older adults and the *Virtual Shop* is a valid task to assess and train episodic memory in this population.

ARTICLE HISTORY Received 4 June 2017; Accepted 13 May 2018

KEYWORDS Virtual reality; Episodic memory; Aging; Validation study; Neuropsychological assessment

Memory is a complex function that relies on a range of interacting processes and systems. A variety of experimental and clinical tasks has been devised to measure memory and to attempt to tease apart these different processes. For instance, different tasks and testing conditions have been developed to distinguish familiarity from recollection, item from associative memory, or memory for verbal, visual, or spatial material. Memory tasks traditionally used in clinical practice or those used in experimental studies of aging are constructed to allow a fine control of the task parameters and testing conditions to reflect these fine-grained processes. However, these tasks generally lack ecological validity, as they fail to reflect the complexity and

CONTACT Sylvie Belleville sylvie.belleville@umontreal.ca Ph.D. Research Centre, Institut universitaire de gériatrie de Montréal, 4565, Queen-Mary, Montreal, Québec, Canada H3W 1W5

diversity of memory situations that older adults experience in their daily lives (Bowman, 1996; Chaytor & Schmitter-Edgecombe, 2003; Farias, Harrell, Neumann, & Houtz, 2003; Piolino, Desgranges, & Eustache, 2009; Sbordonne & Long, 1996; Schultheis, Himelstein, & Rizzo, 2002; Shuchat, Ouellet, Moffat, & Belleville, 2012). In real life, memorising often occurs in noisy environments with multi-dimensional material and often happens while completing other tasks, such as walking, talking, or problem solving. This is in marked contrast to the testing conditions that occur in experimental and clinical contexts, in which participants complete their tasks in quiet conditions, receive clear task instructions, encode unidimensional material most of the time and focus their attention on the task.

Virtual reality (VR) is a promising technology that could help increase the ecological validity of memory assessments and interventions. VR immerses the user in a dynamic virtual environment in which he/she carries out cognitive and sensorimotor activities while interacting with virtual stimuli (Fuchs, Moreau, & Berthoz, 2006). One major asset of VR is that it offers environments that reproduce the sensorial characteristics of the real world (e.g., visual scenes, audible conversations) and incorporate the cognitive and physical demands of situations that individuals face in their everyday lives. Thus, VR gives the opportunity to sample the integrity of cognitive functions in contexts that are more representative of everyday life.

These offer tremendous potential as measures of real-life cognition. Well-designed VR tasks might therefore better reflect real-life capacities than traditional neuropsychological tests (Rizzo, Schulteis, Kerns, & Mateer, 2004). Furthermore, VR has great potential to measure whether neuropsychological interventions transfer to daily life, which is a major challenge in rehabilitation studies (Adamovich et al., 2004; Lehmann et al., 2005; Ouellet, Boller, Corriveau-Lecavalier, Cloutier, & Belleville, 2018; Shuchat et al., 2012; Sveistrup, 2004).

However, VR is a recent technology, and so is its application to cognitive measurement. Many reasons justify measuring the applicability of VR in older adults. Studies using VR protocols with older adults are rare and hence many crucial questions regarding the applicability and validity of VR tasks among older adults remain to be investigated. Designing and testing tasks that reflect memory in real life is particularly interesting in the context of aging. Episodic memory declines with age, is frequently impaired by brain disease, and is one of the first signs of Alzheimer's disease. Thus, having access to a variety of sensitive and valid tools to assess and train episodic memory is crucial for clinical neuropsychologists. Furthermore, while there are many well-designed tasks to measure the fine processes involved in episodic memory, VR can provide tools that reproduce the complexity of memory in daily life. This is critical, as the impact of cognitive decline on autonomy is a major concern in the context of age-related cognitive decline, and VR could contribute to addressing these issues. Yet, the feasibility and applicability of VR in older adults is a potential problem because of differences in technological literacy. As older adults are less likely to make use of information and communication technologies than younger adults (Selwynn, 2004), it is critical to study factors that might contribute to their subjective experience when placed in a VR environment, as well as potential barriers to the use of this technology.

This study will address the applicability of VR technology in a population of older adults and its potential application to memory assessment by measuring presence, motivation, and cybersickness symptoms with a fully immersive episodic memory task (Part 1). It will also assess its construct validity (Part 2). The following section will briefly introduce these notions and how they have been addressed in the literature.

Presence is defined as the subjective experience of being in a place when one is in fact physically in another one (Witmer & Singer, 1998). Thus, in the context of VR, it refers to the subjective experience of actually being in the environment that is represented. It is calculated with questionnaires measuring the quality of the interaction with the environment, whether the experience in the environment was consistent with the real-world experience, and the quality and ease of the interface. A number of factors might determine the magnitude of presence experienced by the participants. Some are related to the software and hardware characteristics of the VR environment, for instance the interface quality, the type of interaction (e.g., joystick vs button response), or how participants navigate in the environment (e.g., active vs passive navigation). Psychological factors related to the user can also contribute to presence and performance in VR. These include the participant's perception regarding the degree of realism of the task, the level of control they have over the situation, the possibility they have to examine the elements of the environment, their subjective evaluation of their own performance, and their general motivation with respect to the task. Measuring presence is critical. Numerous studies have found that presence is positively related to performance in normal young individuals. In younger adults, larger presence has been associated with better sustained attention (Witmer & Singer, 1998), psychomotor performance (Stevens & Kincaid, 2015; Witmer & Singer, 1994), and spatial memory (Bailey & Witmer, 1994). It was also found that VR conditions that reduce presence, for instance those with limited user-environment interactions or which rely on less natural environments, also have deleterious effects on performance (for a review, see Nash, Edwards, Thompson, & Barfield, 2000). The association frequently reported between presence and cognitive performance implies that conditions that create a higher sense of presence will be more likely to provide an optimal assessment of cognition. We therefore examined if the VR environment elicits an appropriate sense of presence in older adults and if the expected correlation is found between presence and performance.

The degree of motivation towards the task might be particularly relevant when designing VR environments, as motivation optimises performance and is related to resource allocation in older adults (Hess, 1994; Hess, Germain, Swaim, & Osowski, 2009; Hess, Popham, Emery, & Elliott, 2012). Because older adults are generally less technologically experienced, one might expect them to be less motivated by VR than by non-VR tasks. Interestingly, some results suggest that this may not be the case. In a study led by Benoit et al. (2015), participants were presented with a photograph or an image-based VR representation of familiar locations in their home city or new locations and were asked to indicate whether they recognised the location. The motivation level of older adults, which was measured with a homemade questionnaire, was found to be larger for the VR than for the non-VR version of the task, although the difference was non-significant. Using the same motivation questionnaire, Manera et al. (2016) reported that older adults with mild cognitive impairment or Alzheimer's disease actually experienced higher levels of satisfaction and security, and lower levels of anxiety, discomfort, and fatigue, during a highly realistic image-based VR cancellation task than during its paper–pencil version (Manera et al., 2015). Thus, both studies reported that older adults experience a higher level of motivation for the VR rather than the non-VR version of the same tasks. Though there is clearly a need for more empirical data, these preliminary findings are interesting and suggest that VR has the potential to elicit positive motivation in older adults.

Cybersickness is potentially a major limitation for the use of VR. Indeed, nausea, headaches, and disorientation can occur while immersed in a virtual environment (Jaeger & Mourant, 2001; Slater, 1999). Their occurrence could seriously hamper the applicability of VR technology in populations that are sensitive to these symptoms. A few studies have found more frequent cybersickness symptoms in older adults relative to younger ones, although the reported difference appears to be of a relatively small magnitude (Arns & Cerney, 2005; Liu, Watson, & Miyazaki, 1999). However, a more recent study reported no increase in cybersickness symptoms following immersion in older adults (Benoit et al., 2015).

Additionally, it is critical to know whether cognitive VR tasks reflect the construct that they are intended to measure. Construct validity refers to the capacity of a test to accurately reproduce the attributes and characteristics of a given construct (Cronbach & Meehl, 1955). Convergent validity is a type of construct validity and is determined by measuring whether performance on the VR task is related to performance on tasks that measure similar theoretical concepts. A few studies have addressed the convergent validity of VR tasks by comparing them with traditional tasks that assess the same cognitive processes. Studies in younger adults have generally reported significant correlations between VR and traditional tasks of inhibition (Armstrong et al., 2013; Henry, Joyal, & Nolin, 2012), and VR and traditional tasks of attention (Parsons & Courtney, 2014). Parsons and Rizzo (2008) reported positive correlations between a traditional word memory task and memory performance in a VR task in which younger adults had to recall a list of 10 items (e.g., a blue car) encoded while navigating a virtual city. Plancher, Nicolas, and Piolino (2008) and Jebara, Orriols, Zaoui, Berthoz, and Piolino (2014) found that older adults' performance on the recognition of items seen in a 2D VR car ride was positively correlated with performance on traditional recognition and executive tests, suggesting that the VR memory task may also reflect other cognitive capacities, such as executive processes.

Construct validity can also be assessed by examining whether a VR task is sensitive to the differences in episodic memory typically associated with aging. Previous studies have found age-related differences on the free recall of spatiotemporal characteristics of a list of items encoded during a virtual car ride (e.g., where and when the items were seen during the car ride; Plancher et al., 2008) and on the free recall of items presented in a virtual apartment (Sauzéon et al., 2016). These results are broadly consistent with the literature, indicating that age is associated with a reduction of associative memory, defined as the capacity to bind pieces of information into a cohesive unit, and of episodic memory, defined as the memory for items encoded with their spatiotemporal context (Chalfonte, 1996; Johnson, 1996; Naveh-Benjamin, 1990, 2000).

In summary, VR has tremendous potential to measure memory in conditions that reflect cognition in everyday life. The use of VR might contribute greatly to how neuropsychologists assess cognition and provide interventions. Furthermore, the technology is becoming cheaper and more accessible, making its use with clinical populations likely to increase in the near future. However, there is a need for applicability and validation data to support VR as a useable technology in older adults and to ensure that VR variants of memory tasks reflect the constructs that they are meant to assess. The present study addresses these issues. The VR task developed here was meant to reflect a situation that is close to a real-life situation and likely reflects memory in action. Participants encode visually presented items and are then asked to find them in a small convenience store. As is the case in real life, their performance is probably based on a combination of active retrieval (for instance, "I need to go get the broom") and recognition, because it is likely that some items are recognised as they walk in the virtual environment. Thus, the task is quite unique relative to traditional memory tasks because the objects are present in the environment, yet the task involves active search and interference. The task requires a conscious mental representation of the items to fetch, which is to some extent close to the process of free recall: participants probably evoke their list while walking around in the store. Although the objects were present and could be used as cues, the subjective experience is clearly more complex than a typical recognition task because participants move in the environment to search for the memorised objects rather than being passively presented with lists of potential items. Another major innovative aspect of the study is to rely on a fully immersive 3D VR technology. Relying on 3D VR technology differs markedly from computerised flat screen VR tasks in that it provides a more immersive experience and more natural interaction with the surrounding environment. However, the technology might be more challenging to use by older adults or clinical populations than 2D technology. To our knowledge, no study has investigated the VR feasibility using a fully immersive technology.

Part 1 addresses the applicability of a fully immersive 3D VR episodic memory task in which participants had to memorise and fetch a series of items in a *Virtual Shop*. This is addressed in younger and older adults by measuring presence, motivation, and cybersickness symptoms. We hypothesise that the task will show strong feasibility in both younger and older adults. We also anticipate that older and younger adults will show a comparable level of presence and that presence will be related to performance in the VR task (Bailey & Witmer, 1994; Witmer & Singer, 1994). We also expected that the task would be motivating for participants, irrespective of their age. Finally, we did not expect cybersickness symptoms to interfere with the task completion, as a number of studies have reported that older adults have relatively few symptoms of cybersickness with tasks of short duration.

Part 2 measures construct validity of the immersive VR episodic memory task by comparing performance on the VR task with that obtained from traditional paper–pencil memory tasks and measuring whether the VR task was sensitive to the age difference typically found in episodic memory. Given the results from prior work, we hypothesise that the task will be a valid representation of episodic memory capacities. We expect that the task will have appropriate construct validity. This will be supported by finding a positive correlation between memory performance on the VR task and performance on a traditional task measuring immediate and delayed free recall of a list of visually presented words. Construct validity will also be supported by findings of a lower VR memory performance in older adults compared to young adults.

General methods, common to Parts 1 and 2

Participants

The study included 57 cognitively healthy older adults and 20 younger adults. It included a larger number of older than younger adults, as older adults were the main focus of our study. The goal was to assess the applicability and validity of VR in this population, and younger adults were included as a group of comparison. Furthermore, inter-individual variability increases with age (Hultsch & MacDonald, 2004; Hultsch, Strauss, Hunter, & MacDonald, 2008) and thus including a larger number of older adults increased the power to detect a group difference. The same participants were used for both parts to increase power and because this facilitates the comparison of the results obtained for feasibility and validity. Participants were recruited from the local community and were all native French speakers. Exclusion criteria included the following: presence or history of a neurodegenerative disease, life-threatening disease (e.g., cancer), stroke, uncontrolled sleep apnoea, major psychiatric disorders (i.e., depression, schizophrenia, etc.), excessive drinking (>25 drinks per week; for equivalence, see http://www.ccdus.ca/Resource%20Library/2012-Canada-Low-Risk-Alcohol-Drinking-Guidelines-Brochure-en.pdf), substance abuse, general anaesthesia during the past 6 months, balance difficulties, uncorrected visual impairment, and important hearing loss (corrected or not). We also used the *Montreal Cognitive Assessment* (MoCA) (Nasreddine et al., 2005), a short cognitive assessment battery, to exclude older adults with impaired cognition (score ≤26). The RL/RI-16 word recall task (Van der Linden et al., 2004) was included to characterise verbal memory in the sample of older adults and facilitate comparison between this sample and other samples that will be used in future studies.

The participants' demographic and clinical characteristics are presented in Table 1. Older and younger adults were equivalent on demographic characteristics. Furthermore, the scores of the older adults on the neuropsychological tests were within the normal range when considering their age and education level. This was expected, given that cognitively impaired participants were excluded.

VR task

The virtual environment of the *Virtual Shop* (*La boutique virtuelle*) was developed and rendered using the 3DVIA *Virtools 5* 3D engine and was run on a Dell Precision T3600 PC with an Inter(R) Xeon (R) CPU ES-1620 0 (3.60 Ghz, 10 Gbytes in RAM) processor and a NVIDIA GeForce GTX600 Ti graphics card. It was designed in collaboration with *Cliniques et développement in virtuo* (www.invirtuo.com). The virtual environment was in 3D and the immersion was produced by an *Nvisor ST50* audio-visual headgear and by a Worldviz PPT-X studio tracking system that allowed the participant to rotate his/her head in a 360-degree view around the room, as well as look up and down, and interact and walk freely in the virtual environment. The participant was asked to stand in the empty assessment room while the assistant installed the headgear and hand device. He/she was then presented with the virtual environment. The environment was a small convenience store built using the same dimensions as the assessment room (3.5 m × 6.5 m). Participants were told that they were free to move around the environment, explore and

Table 1. Demographic and clinical characteristic of participants.

	Younger (*n* = 20)	Older (*n* = 57)
Age (years)	21.65 (2.46)	67.77 (7.03)
Education (years)	13.90 (2.05)	14.86 (3.23)
Gender (f, m)	13, 7	47, 10
Montreal Cognitive Assessment (MoCA) (/30)	–	27.58 (1.74)
RL/RI-16 word recall test (3rd free recall)		11.46 (2.31)
RL/RI-16 word recall test (delayed free recall) (/16)	–	12.11 (2.08)
Geriatric Depression Scale (GDS)	–	1.77 (2.73)

fetch items. Participants used a hand remote control to select and retrieve items. The remote control allowed them to display a target sign that they could move in the virtual environment in order to point to the items they wanted to select.

Participants began the task in front of a cashier working behind a countertop and were presented with a list of 12 familiar virtual images of common items (e.g., belt, milk) that they were asked to memorise and then fetch in the store (see Figure 1). Each item was visually presented for 5 s on a notepad situated on the countertop with the name of the item written below the image to ensure that the item was properly encoded. During encoding, irrelevant conversations were presented via the headgear in order to mimic a noisy environment. Following the presentation of the last item, the programme initiated a 20-s conversation between the cashier and the participant (e.g., *Could you tell me the time, which is displayed on your right?*) as a filled interference delay. At the end of the delay period, the cashier instructed the participant to fetch the items in the store he/she had previously seen. The participant could then walk freely in the room to find and select the items that were shown on the learning list. There were 24 items displayed in the shop: 12 target items and 12 distractors. We chose distractors that matched the targets by taxonomic category. This is relevant when designing episodic memory tasks, as memory errors most often preserve the category in free recall conditions and more errors are made when distractors share the same semantic category as the target in recognition conditions. Furthermore, this feature reduces the likelihood of simply guessing the correct item based on having encoded its semantic category. It therefore makes the task more sensitive to memory failures. An ancillary benefit to using semantic distractors is that it enhances the ecological value of the test, as real-life shopping often requires selecting items among other ones from the same category. For instance, if one has to buy a particular vegetable to make a soup, it will be found in the "fruit and vegetable section" of the supermarket and the person will therefore be faced with distractors from the same category. The items were located on the shelves, on the floor, or hung on the walls.

Participants used a hand remote control to select and retrieve the items and were given unlimited time to find them. Prior to the testing, they were familiarised with the virtual devices using a different version of the convenience store in a condition in which they were simply asked to walk in the virtual environment and select an item that was not used in the memory test. During this familiarisation phase, additional information and practice trials were given to participants who were unsure about the procedure of the task or with the operation of the remote control. Familiarisation was continued until the participant was comfortable with the manipulation of the material in order to reduce the likelihood that problems would occur during the VR task.

Correct performance was measured as the number of targets that were correctly retrieved. A false recognition error response was recorded when the participant selected an incorrect item; that is, one that was not part of the encoded list. Note that all foils are related to one of the presented items and, therefore, false recognition errors are semantic errors by design. There were other built-in parameters that were measured by the VR task; for instance, time before the first item was selected and total time to complete the task. Although they were not used in the present study, they could be useful for researchers interested in a more extensive characterisation of the participant's behaviour in the task (see, for instance, Ouellet et al., 2018).

Figure 1. The *Virtual Shop*. Image A shows the notepad on the countertop, on which items appeared during the encoding phase, after which the cashier would talk to the participant as a filled interference delay. Image B shows a version of the *Virtual Shop*, with the items placed on the shelves and hung to the walls. Finally, Image C shows an item that has been selected by the participant with the remote control.

Traditional episodic memory task measures

A traditional experimental memory task was used to test convergent validity and compare motivation in a virtual vs non-virtual variant of a memory task. The task was adapted from a validated free recall word list test (Belleville et al., 2002). Two lists of 12 concrete words were visually presented on a laptop using *e-prime*. Two lists were used to increase the number of trials and, hence, reduce the impact of extraneous variables on performance. The lists were matched for word length (1–4 syllables), word frequency, and concreteness. Participants were presented with the words at a rate of one item every 5 s (4 s of presentation and 1 s of cross fixation). The lists were encoded and recalled with irrelevant verbal noise, similar to the noise used in the VR task, and were presented through a Plantronix Audio 550 headset. Participants were instructed to remember as many words as possible. Immediately after the presentation of the list, participants were asked to write down the words they remembered in the order in which they came to mind. Thus, participants did not have to retrieve the order in which items were presented, similar to what is typically done in clinical memory measures and as was the case for the VR task. Free recall was repeated 4 minutes after participants had completed a short-term memory task (a digit span task).

The use of an experimental measure was preferred over that of a clinical measure (for example, the logical memory test or the California Verbal Learning Test) as a test of convergent validity. This was done to allow for more flexibility and control over testing parameters. Using a computerised presentation facilitated strict control over presentation parameters (e.g., presentation rate, recall delay). Designing the task allowed us to control the frequency and concreteness of the items using data collected from the French-Canadian population. It also enabled us to match some of its parameters with those of the VR task (e.g., use of similar encoding rates, interfering noise during encoding, and a visual presentation of the stimuli). Note that the VR task is more complex by design and, hence, we did not expect a perfect match with the traditional task. One major difference relates to the retrieval phase. As is typical when shopping, items are retrieved while the participant walks in the VR environment and it is therefore more akin to a recognition than recall procedure. Thus, retrieval in the VR environment involves a mix of active search and recognition, while our traditional memory task involves recall.

Design

Health and demographic questionnaires were completed during a 30-min telephone interview. Eligible participants were tested with the cognitive and VR measures, and administered the presence, motivation, and cybersickness questionnaires at the CRIUGM during a single 2-hour session. Participants first received the traditional memory task, followed by the *Motivation Questionnaire* related to the traditional memory task and the cybersickness questionnaire. They then completed the VR task, followed by the cybersickness questionnaire for a second time, the motivation questionnaire related to the VR task, and the presence questionnaire. This study was approved by the Regroupement Neuroimagerie/Québec (RNQ) Comité mixte d'éthique de la recherche.

Part I: Applicability of the virtual shop

Presence, motivation, and cybersickness questionnaires

The French version of the *Presence Questionnaire* (Witmer & Singer, 1994), which was adapted by the Cyberpsychology Laboratory of the Université du Québec en Outaouais (UQO) (Robillard, Bouchard, Renaud, & Cournoyer, 2002), included 19 items (e.g., to what degree did your interactions with the environment seem natural?) divided into five subscales: realism, possibility to act, interface quality, possibility to examine, and self-evaluation of the performance. In this questionnaire, participants were asked to rate their VR experience on a Likert scale ranging from 1 (not at all) to 7 (completely). The questionnaire is constructed so that responses in the low range of the scale (1–3) indicate a negative experience, whereas responses in the high range of the scale (5–7) indicate a positive experience. This questionnaire was shown to have good internal consistency (Cronbach's alpha of 0.88), as well as content and construct validity (Witmer & Singer, 1994).

For the purpose of this study, we also constructed an experimental *Motivation Questionnaire* to assess motivation evoked by the tasks. The questionnaire was constructed based on a literature review regarding the different components of motivation according to the concept of flow (Csikszentmihalyi, 2000) in relation to video games (Klasen, Weber, Kircher, Mathiak, & Mathiak, 2012), and media enjoyment (Weber, Tamborini, Westcott-Baker, & Kantor, 2009). One version was used for the VR task and another for the traditional memory task. Each version comprised 7 items, against which participants rated their level of motivation and interest regarding the task they completed (e.g., I felt engaged during the task on the computer/Virtual environment) on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Lower scores indicated a low level of motivation. The questionnaire showed an appropriate internal consistency when tested in this sample (Cronbach's alpha = 0.79).

The French version of the *Simulator Sickness Questionnaire* (Kennedy, Lane, Berbaum, & Lilienthal, 1993), which was adapted by the UQO Cyberpsychology Laboratory (Bouchard, Robillard, & Renaud, 2007), was used to assess the occurrence, nature, and severity of cybersickness symptoms when immersed in a virtual environment. Two subscales (nausea and oculomotor difficulties) included 16 items (e.g., headaches), against which participants had to rate their symptoms on a scale of 0 (not at all) to 3 (severely). Here, lower ratings correspond to low levels of symptoms. This questionnaire was found to have good internal consistency (Cronbach's alpha of 0.81), and to be a valid measure of motion-induced sickness symptoms (Kennedy et al., 1993).

Statistical analyses and results

The data were analysed using *Statistical Package for Social Sciences* (SPSS) version 21.0. Three older adults did not complete the entire protocol: two of them reported severe cybersickness symptoms and thus could not complete the VR task, and one participant withdrew from the study before the VR task. Participants who withdrew for cybersickness symptoms completed the *Sickness Simulator Questionnaire* after the VR task and were thus included in the analysis of cybersickness symptoms. As there were no significant differences in terms of education and gender distribution between the younger and older adults, it was not necessary to use them as nuisance covariates (Miller & Chapman, 2001).

In order to compare younger and older adults on different *Presence Questionnaire* subscales, independent *t*-tests (two-tailed) with significance levels set to $p < .01$ according to the Bonferonni correction for multiple comparisons were conducted, and *Pearson* bivariate correlations (two-tailed) were performed to investigate the relationship between the different *Presence Questionnaire* subscales and memory performance on the VR task. Comparisons were adjusted according to Levene's tests for homogeneity of variance when needed. A 2 (Type of task) × 2 (Group) ANOVA was used to assess the degree of motivation evoked by the different memory tasks in younger and older adults. To assess the magnitude of cybersickness symptoms, a 2 (Immersion) × 2 (Group) ANOVA was conducted in order to compare younger and older adults before and after the immersion.

Presence

Results (means and standards deviations) for the *Presence Questionnaire* subscales, the *Motivation Questionnaire* for the VR and the traditional memory tasks, the *Sickness Simulator Questionnaire*, and performance on the VR and traditional memory task measures are listed in Table 2. Overall, participants rated the aspects related to the different *Presence Questionnaire* subscales in the positive range. No differences were found between the two groups on any of the subscales (*realism*, $t(1, 72) = 0.59$, $p = .60$; *possibility to act*, $t(1, 72) = 0.31$, $p = .76$; *interface quality*, $t(1, 72) = 0.40$, $p = .69$; *possibility to examine*, $t(1, 72) = 0.06$, $p = .95$; *self-evaluation of performance*, $t(1, 72) = 2.32$, $p = .29$).

In older adults, there were significant positive correlations between VR performance and all subscales of the *Presence Questionnaire* (*realism*, $r = .36$, $p < .01$; *possibility to act*, $r = .41$, $p < .01$; *interface quality*, $r = .42$, $p < .001$; *possibility to examine*, $r = .28$, $p < .05$; *self-evaluation of performance*, $r = .37$, $p < .01$). In younger adults, VR performance only correlated positively with the *interface quality* subscale, $r = .53$, $p < .05$, but there were no significant correlations with the other subscales (*realism*, $r = -.27$, $p = .25$; *possibility to*

Table 2. Mean scores for the questionnaires and time to complete the VR task.

	Younger ($n = 20$)	Older ($n = 57$)
Presence Questionnaire		
Realism (/49)	34.40 (6.18)	33.77 (7.50)
Possibility to act (/28)	19.40 (3.42)	19.32 (5.34)
Interface quality (/28)	14.50 (2.82)	15.17 (3.43)
Possibility to examine (/14)	13.85 (2.52)	14.17 (3.72)
Self-evaluation of performance (/14)	11.55 (2.06)	10.06 (2.82)
Motivation Questionnaires		
Virtual Reality (/35)	20.35 (7.50)	29.36 (5.16)
Traditional task (/35)	16.45 (7.97)	26.46 (5.14)
Sickness Simulator Questionnaire		
Pre-immersion		
Nausea (/27)	1.35 (3.76)	0.93 (1.33)
Oculo-motor difficulties (/21)	2.55 (2.46)	3.57 (2.39)
Total score (/48)	3.90 (5.42)	4.46 (3.24)
Post-immersion		
Nausea (/27)	2.75 (5.32)	2.73 (3.94)
Oculo-motor difficulties (/21)	3.05 (2.72)	2.95 (3.27)
Total score (/48)	5.80 (7.37)	5.67 (6.44)
Time to complete Task (seconds)	312.65 (99.17)	482 (153.30)

act, $r = -.01$, $p = .98$; *possibility to examine*, $r = -.21$, $p = .37$; *self-evaluation of performance*, $r = -.31$, $p = .18$).

Motivation

The ANOVA indicated a significant effect for the Type of task, as the motivation scores for the *Virtual Shop* were higher than for the traditional memory task for both groups, $F(1, 72) = 23.65$, $p < .001$, $\eta^2 = 0.25$. Results also revealed a Group effect, as older adults had higher motivation scores than younger adults overall, $F(1, 72) = 1132.37$, $p < .001$, $\eta^2 = 0.40$. There was no Group × Type of task interaction, $F(1, 72) \leq 1$, $p = .54$.

Cybersickness symptoms

There were slightly more cybersickness symptoms following than prior to immersion but the effect just missed significance, $F(1, 74) = 3.71$, $p = .06$. There was neither a Group nor interaction effect, $F < 1$, in both cases. Scores on the *Sickness Simulator Questionnaire* indicated that both groups experienced a low level of cybersickness symptoms, even after immersion.

Part II: Validation of the virtual shop

Method and results

The second part of this study assessed construct validity by computing *Pearson* bivariate correlations (two-tailed) to assess the relationship between performance on the VR task and on free immediate and delayed recall, measured with traditional word recall tasks in both younger and older adults (convergent validity). It also assesses construct validity by measuring whether the VR task was sensitive to the age-related effect typically observed on episodic memory tasks. Groups were compared on their performance on the *Virtual Shop* with independent *t*-tests (two-tailed), using the number of accuracies (correctly retrieved items) and the number of false recognitions as dependent variables. Their performance on the traditional memory task was compared with a 2 (Group) × 2 (Delay) ANOVA using correct word recall as the dependent variable.

Convergent validity

Performance on the VR task correlated with the immediate and delayed free recall scores of the traditional verbal memory task in both younger ($r = .57$, $p < .01$ and $r = .46$, $p < .05$, respectively) and older adults ($r = .28$, $p < .05$ and $r = .30$, $p < .05$, respectively) (see Table 3 for performance on the VR and traditional memory tasks).

Table 3. Mean scores for the memory tasks.

	Younger ($n = 20$)	Older ($n = 20$)
Memory tasks		
The *Virtual Shop*		
Items correctly identified	9.10 (2.13)	7.66 (2.37)
False recognitions	1.14 (0.25)	0.66 (0.09)
Episodic immediate recall (/12)	7.68 (2.08)	5.28 (1.42)
Episodic delayed recall (/12)	6.73 (2.59)	3.32 (1.75)

Construct validity

As expected, results on the traditional memory task showed a Group effect, $F(1, 75) = 41.28$, $p < .001$, $\eta^2 = 0.36$, a Delay effect, $F(1, 75) = 62.52$, $p < .001$, $\eta^2 = 0.60$, and a Group × Delay interaction, $F(1, 75) = 13.37$, $p < .001$, $\eta^2 = 0.15$. Tukey post-hoc analysis revealed a stronger effect of delay in older than in younger adults, $p < .001$. The VR task was sensitive to age, as younger adults performed significantly better than older adults when using the number of correct answers as a dependent variable, $t(1, 73) = 2.38$, $p < .05$, $d = 0.30$. However, there was no group effect on the number of false recognitions, $t(1, 75) = 0.79$, $p = .937$.

Discussion

To our knowledge, this study is the first to address the applicability and validity of a fully immersive episodic memory VR task in younger and older individuals. Overall, results indicate that the VR technology is a useable tool in aging and that the *Virtual Shop* has adequate validity properties to reflect episodic memory in a virtual context. This indicates that the technology is suitable to assess, and eventually train, episodic memory in older adults. These aspects of the study are discussed below.

The first goal of the study was to assess the feasibility of the VR task. Our hypothesis was that the *Virtual Shop* would show strong feasibility in younger as well as older adults. All three indicators suggest good feasibility. The five subscales of the *Presence questionnaire* were positively rated by both age groups, and younger and older adults were comparable on all subscales. We found higher levels of motivation for the *Virtual Shop* than for the traditional memory task in both age groups. Finally, negligible cybersickness symptoms were found following immersion for both younger and older adults.

Hence, the fact that older adults might have been less exposed to technology and electronic devices does not seem to impact their capacity to feel comfortable in virtual environments, to experience similar feelings and reactions as in real-life situations, and to enjoy realising cognitive tasks in that sort of setting. Our environment was fully immersive, and the high degree of interaction could account for the high sense of presence (Slobounov, Ray, Johnson, Slobounov, & Newell, 2015) and to making it a more inviting and interesting experience than the traditional task. Furthermore, the resemblance of the *Virtual Shop* with everyday situations may have provided a more meaningful environment to older adults and contributed to our finding that younger and older adults experienced an equivalent level of presence and motivation. This is consistent with the literature reporting that presence is optimised by meaningful environments and by conditions that allow user-environment interactions (for a review, see Nash et al., 2000). The fact that presence was high and that all subscales of the presence questionnaire were positively related to performance in older adults indicates that the age effect in memory is not accounted for by a poor sense of presence. This result supports the use of VR as a valid measure of memory in older adults. Interestingly, though, performance in younger and older adults is influenced by different characteristics from the presence experience within VR. The performance of younger adults appears to be influenced by the task's interface quality, whereas that of older adults appears to be influenced by the task's interface quality but also by the content of the task as well as their personal appraisal and confidence with respect to the task. Thus, many dimensions related to presence appear to influence cognition measurement when using VR with older adults, and these should be taken into account when designing tasks adapted to this population. Importantly, this may not be particular to VR, as many of these characteristics have been shown to influence performance when testing older adults with traditional tools as well.

One frequently reported drawback regarding the use of VR in older adults is the fact that it elicits cybersickness and that these effects might be more frequent and/or severe in older adults. The number of cybersickness symptoms did not significantly increase following immersion. Nevertheless, the two age groups reported slightly more cybersickness symptoms following immersion and that the pre–post immersion effect was close to significance. Importantly, however, the Group × Immersion interaction was far from significance, which indicates that older adults were not more prone to cybersickness symptoms than younger adults. Furthermore, the reported level of cybersickness symptoms was relatively low. Thus, cybersickness symptoms slightly increased with immersion but did not hamper the completion of the VR task. Of note: our task was of a relatively short duration and hence whether participants would have experienced more cybersickness symptoms with longer durations remains to be determined.

The second part of the study measured the construct validity of the *Virtual Shop*. Our hypotheses regarding construct validity were that the memory performance in the *Virtual Shop* would be correlated to that obtained from a traditional verbal episodic memory task and that the task would be sensitive to age-related differences. Both hypotheses were confirmed. Performance in the VR task was positively correlated with performance in a traditional word-recall test, suggesting that both reflect verbal memory processes. This concurs with previous studies indicating that VR can measure similar constructs as those measured by clinical or experimental measures (Armstrong et al., 2013; Henry et al., 2012; Parsons & Courtney, 2014; Parsons & Rizzo, 2008; Plancher et al., 2008). This is an important finding. VR measures memory performance in complex conditions in the presence of auditory and visual distractions, and while the participant navigates the environment and manipulates new devices. Furthermore, older adults showed the typical age-related memory decrement when examining the number of correctly retrieved items as a dependent variable. This indicates that the VR task is sensitive to typical memory impairment.

However, it is important to stress that even if the traditional and VR tasks share some common variance, they also differ in many other ways. Importantly, retrieval phases differed markedly between the traditional tasks and the VR tasks. Indeed, the former was a recall task whereas the latter involved recognition, which are subtended by different cognitive and brain processes. Furthermore, the VR task involves retrieval in challenging conditions because the environment is noisy, and participants have to retrieve information while walking and selecting the objects. Other studies have shown that recognition is impaired in older adults when performed in attention-demanding conditions (Anderson, Craik, & Naveh-Benjamin, 1998; Li, Lindenberger, Freund, & Baltes, 2001). Hence, the *Virtual Shop* has the potential to unravel cognitive difficulties encountered by older adults in real life, when conditions are more distracting or more demanding.

The fact that our VR and traditional tasks share common variance but are not totally equivalent indicates that VR can have a unique contribution to memory assessment. Traditional memory tasks can be used to isolate individual memory processes. In turn, VR tasks can provide a reflection of how memory works in real-life situations in which different processes interact to result in complex behaviour. VR memory measures such as the *Virtual Shop* are designed to reflect real-life memory and hence are quite complex tasks. Future studies should include a larger number of measures to address convergent criterion validity. It is indeed likely that other cognitive and sensorimotor processes other than memory contribute to performance in such complex immersive VR tasks.

Interestingly, the two age groups showed an equivalent number of false recognition errors. Thus, older adults omit more items than younger ones, but do not select erroneous semantic foils. This is contrary to some studies indicating that older adults have a more liberal response bias than younger ones (Dodson, Bawa, & Krueger, 2007; Huh, Kramer, Gazzaley, & Delis, 2006), but consistent with others reporting no age differences in response bias (e.g., see Bastin & Van der Linden, 2003). Interestingly, a yes–no recognition format similar to the one used here was found to yield a lower level of commission when compared to a forced-choice recognition format (Bastin & Van der Linden, 2003). In forced-choice recognition, the presence of concurrent alternatives would increase the likelihood of false-recognition errors. Thus, the finding of a similar number of false recognition in older and younger adults might be due to the fact that our recognition format is not one that favours this type of error.

Results must be interpreted within the context of some limitations. As stated above, it remains unclear whether other cognitive processes are implicated in the VR task, aside from episodic memory. Also, we have created a life-like task involving shopping in an immersive environment with demanding retrieval conditions and background noise. Our goal was to construct a task that would approximate real-life conditions, keeping in balance the limitations inherent to simulation procedures and to the technology involved in VR. Thus, it was not possible to make the VR and traditional memory tasks entirely equivalent. This was the case for the retrieval condition. Shopping in real life typically involves a relatively long delay between encoding (when one determines a list of items to buy) and retrieval (when one has arrived at the place where those items will be purchased). A simulation does not allow such long delays. We also reduced the time in the VR environment to lower the risk of cybersickness. Cybersickness is not found in the real environment and, therefore, eliciting these symptoms would have reduced rather than increased the ecological validity of our task. Thus, there were some differences with real-life shopping, but our goal was to keep a balance between feasibility and real-life resemblance. Virtual reality remains a tool and, while VR scenarios mimic those found in real life, we cannot claim that they are entirely akin to real-life conditions. However, different components were included to increase the similarity between the VR task and real-world situations, for instance including background conversations, providing complete immersion, and having individuals physically walk in the VR environment to retrieve their items. Altogether, these characteristics helped increase the similarity with real-world contexts, even if not totally similar to everyday life situations, and represent assets compared to other, less-elaborated protocols which rely on flat screens, joystick navigation, or passive exploration. Another limitation is that the same sample was used for the two parts of the study. Hence, the possibility remains that performance in the VR task might have influenced the participants' responses on the motivation and/or presence questionnaire. However, as older adults showed lower performance than younger ones but higher motivation and an equivalent sense of presence, it does not seem to be the case. We did not include measures of exposure to video games and technology, cognitive training, or use of cognitive games, and we did not measure test-retest reliability and comparability of parallel version. Finally, the sample size for the group of younger adults was small, which might have reduced statistical power, particularly for correlations.

In conclusion, our results indicate that the use of a fully immersive *Virtual Shop* task is feasible in older adults: it elicits presence, is engaging, and provokes limited symptoms of cybersickness within the conditions that were used here. Furthermore, it has appropriate construct validity to measure episodic memory: performance in the VR task is positively related with performance on a traditional memory task and is sensitive to age-related differences. The finding that the VR task is a feasible, valid, and sensitive measure of memory makes it a promising tool to contribute to the clinicians' knowledge regarding an individual's daily functioning and the impact that memory impairment may have on his/her daily life. Thus, VR memory tasks could become useful instruments to reflect real-life memory and provide complementary information relative to more traditional measures. VR tasks might also contribute to enriching cognitive interventions with environments that are realistic and engaging, thus addressing some of the challenges encountered in geriatric rehabilitation, such as lack of motivation and engagement, for instance (Choi & Twamley, 2013). Finally, given that VR is feasible and that VR-based tasks are valid measures, relying on the technology to devise real-life tasks represents an interesting avenue to assess whether interventions or rehabilitation strategies provided in clinical contexts generalise to more complex environments of daily life (for illustrations see Bier, Ouellet, & Belleville, in press; Zelinski, 2009).

Acknowledgements

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), Quebec Network for Research on Aging (RQRV), and Fondation canadienne pour l'innovation grants (FCI) to SB. NCL was supported by NSERC summer and FRQS scholarships. EO was supported by the Canadian Institutes of Health Research doctoral research award scholarship. BB was supported by a post-doctoral fellowship from the Fondation de la famille Lemaire and Fondation Institut de gériatrie de Montréal. We would like to thank Nadia Jaffer and the CRIUGM *banque de Participants* for their help with recruitment, Émilie Lepage and Christel Cornelis for their help in testing the participants, and Gabrielle Ciquier for English editing. We would also like to thank Stéphane Bouchard from the Université du Québec en Outaouais and his team for programming the VR task.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), Quebec Network for Research on Aging (RQRV), and Fondation canadienne pour l'innovation grants (FCI; grant no. FCI-30063) to SB. NCL was supported by NSERC (number: BRPC-467787-2014) summer scholarship and Fonds de Rercherche du Québec – Santé (FRQS) master's scholarship.

References

Adamovich, S. V., Merians, A. S., Boian, R., Tremaine, M., Burdea, G. S., Recce, M., & Poizner, H. (2004). A virtual reality based exercise system for hand rehabilitation post-stroke: Transfer to function. In *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE* (Vol. 2, pp. 4936–4939). IEEE.

Anderson, N. D., Craik, F. I., & Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval in younger and older adults: I. Evidence from divided attention costs. *Psychology and Aging*, *13*(3), 405–423.

Armstrong, C. M., Reger, G. M., Edwards, J., Rizzo, A. A., Courtney, C. G., & Parsons, T. D. (2013). Validity of the virtual reality stroop task (VRST) in active duty military. *Journal of Clinical and Experimental Neuropsychology*, *35*(2), 113–123. doi:10.1080/13803395.2012.740002

Arns, L. L., & Cerney, M. M. (2005). The relationship between age and incidence of cybersickness among immersive environment users. *IEEE Proceedings. VR 2005. Virtual Reality*, *2005*, 267–268. doi:10.1109/VR.2005.1492788

Bailey, J. H., & Witmer, B. G. (1994). Learning and transfer of spatial knowledge in a virtual environment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *38*(18), 1158–1162. doi:10.1177/154193129403801803

Bastin, C., & Van der Linden, M. (2003). The contribution of recollection and familiarity to recognition memory: A study of the effects of test format and aging. *Neuropsychology*, *17*(1), 14–24.

Belleville, S., Chatelois, J., Fontaine, F., Peretz, I., Renaseau-Leclerc, C., & Van der Linden, M. (2002). *Mémoria: Batterie informatisée d'évaluation de la mémoire pour Mac et PC*. Montréal: Institut universitaire de gériatrie de Montréal.

Benoit, M., Guerchouche, R., Petit, P.-D., Chapoulie, E., Manera, V., Chaurasia, G., & Robert, P. (2015). Is it possible to use highly realistic virtual reality in the elderly? A feasibility study with image-based rendering. *Neuropsychiatric Disease and Treatment*, *11*, 557–563. doi:10.2147/NDT.S73179

Bier, B., Ouellet, É., & Belleville, S. (under press). Computerized attentional training and transfer with virtual reality: Effect of age and training type. *Neuropsychology*.

Bouchard, S., Robillard, G., & Renaud, P. (2007). Revising the factor structure of the simulator sickness questionnaire. *Annual Review of CyberTherapy and Telemedicine*, *5*, 128–137. Retrieved from http://s3.amazonaws.com/academia.edu.documents/30756632/ARCTT2007.pdf?AWSAccessKeyId=AKIAJ56TQJRTWSMTNPEA&Expires=1466519574&Signature=Ln4uFk2F9TVhmoGBvMZVKhjUZD9o3D&response-content-disposition=inline3B20filename%3DA_Virtual_Human_Agent_for_Training_Novic.pdf#page=117

Bowman, D. (1996). Conceptual design space—Beyond walkthrough to immersive design. In *Designing digital space* (pp. 225–236). Retrieved from http://cs.vt.edu/node/6196

Chalfonte, B. L. (1996). Spatial location memory in amnesia: Binding item and location information under incidental and intentional encoding conditions. *Memory (Hove, England)*, *4*(6), 591–614. doi:10.1080/741940998

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review*, *13*(4), 181–197. doi:10.1023/B:NERV.0000009483.91468.fb

Choi, J., & Twamley, E. W. (2013). Cognitive rehabilitation therapies for Alzheimer's disease: A review of methods to improve treatment engagement and self-efficacy. *Neuropsychology Reviews*, *23*(1), 48–62.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. doi:http://doi.org/10.1037/h0040957

Csikszentmihalyi, M. (2000). *Beyond boredom and anxiety*. San Francisco: Jossey-Bass.

Dodson, C. S., Bawa, S., & Krueger, L. E. (2007). Aging, metamemory, and high-confidence errors: A misrecollection account. *Psychology and Aging*, *22*(1), 122–133.

Farias, S. T., Harrell, E., Neumann, C., & Houtz, A. (2003). The relationship between neuropsychological performance and daily functioning in individuals with Alzheimer's disease: Ecological validity of neuropsychological tests. *Archives of Clinical Neuropsychology*, *18*(6), 655–672. doi:10.1016/S0887-6177(02)00159-2

Fuchs, P., Moreau, G., & Berthoz, A. (2006). *Le traité de la réalité virtuelle volume 1: L'Homme et l'environnement virtuel*. Paris: Presse des Mines.

Henry, M., Joyal, C. C., & Nolin, P. (2012). Development and initial assessment of a new paradigm for assessing cognitive and motor inhibition: The bimodal virtual-reality stroop. *Journal of Neuroscience Methods*, *210*(2), 125–131. doi:10.1016/j.jneumeth.2012.07.025

Hess, T. M. (1994). Social cognition in adulthood: Aging-related changes in knowledge and processing mechanisms. *Developmental Review*, *14*(4), 373–412. doi:10.1006/drev.1994.1015

Hess, T. M., Germain, C. M., Swaim, E. L., & Osowski, N. L. (2009). Aging and selective engagement: The moderating impact of motivation on older adults' resource utilization. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *64B*(4), 447–456. doi:10.1093/geronb/gbp020

Hess, T. M., Popham, L. E., Emery, L., & Elliott, T. (2012). Mood, motivation, and misinformation: Aging and affective state influences on memory. *Aging, Neuropsychology, and Cognition*, *19*(1-2), 13–34. doi:10.1080/13825585.2011.622740

Huh, T. J., Kramer, J. H., Gazzaley, A., & Delis, D. C. (2006). Response bias and aging on a recognition memory task. *Journal of the International Neuropsychological Society*, *12*(1), 1–7.

Hultsch, D. F., & MacDonald, W. S. (2004). Intraindividual variability in performance as a theoretical window onto cognitive aging. In R. A. Dixon, L. Bäckman, & L.-G. Nilsson (Eds.), *New frontiers in cognitive aging* (pp. 65–68).

Hultsch, D. F., Strauss, E., Hunter, M. A., MacDonald, S. W. (2008). Intraindividual variability, cognition, and aging. In F. I. Craik, F. I. & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (3rd ed., pp. 491–556). New York: Psychology Press.

Jaeger, B. K., & Mourant, R. R. (2001). Comparison of simulator sickness using static and dynamic walking simulators. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *45*(27), 1896–1900. doi:10.1177/154193120104502709

Jebara, N., Orriols, E., Zaoui, M., Berthoz, A., & Piolino, P. (2014). Effects of enactment in episodic memory: A pilot virtual reality study with young and elderly adults. *Frontiers in Aging Neuroscience*, *6*, 314. http://doi.org/10.3389/fnagi.2014.00338

Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, *24*(4), 403–416. doi:10.3758/BF03200930

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, *3*(3), 203–220. doi:10.1207/s15327108ijap0303_3

Klasen, M., Weber, R., Kircher, T. T., Mathiak, K. A., & Mathiak, K. (2012). Neural contributions to flow experience during video game playing. *Social Cognitive and Affective Neuroscience*, *7*(4), 485–495. doi:10.1093/scan/nsr021

Lehmann, K. S., Ritz, J. P., Maass, H., Çakmak, H. K., Kuehnapfel, U. G., Germer, C. T., & Buhr, H. J. (2005). A prospective randomized study to test the transfer of basic psychomotor skills from virtual reality to physical reality in a comparable training setting. *Annals of Surgery*, *241*(3), 442–449.

Li, K. Z., Lindenberger, U., Freund, A. M., & Baltes, P. B. (2001). Walking while memorizing: Age-related differences in compensatory behavior. *Psychological Science*, *12*(3), 230–237.

Liu, L., Watson, B., & Miyazaki, M. (1999). VR for the elderly: Quantitative and qualitative differences in performance with a driving simulator. *CyberPsychology & Behavior*, *2*(6), 567–576. doi:10.1089/cpb.1999.2.567

Manera, V., Chapoulie, E., Bourgeois, J., Guerchouche, R., David, R., Ondrej, J., & Robert, P. (2016). A feasibility study with image-based rendered virtual reality in patients with mild cognitive impairment and dementia. *PloS one*, *11*(3), e0151487. http://doi.org/10.1371/journal.pone.0151487

Manera, V., Petit, P.-D., Derreumaux, A., Orvieto, I., Romagnoli, M., Lyttle, G., … Robert, P. H. (2015). "Kitchen and cooking", a serious game for mild cognitive impairment and Alzheimer's disease: A pilot study. *Frontiers in Aging Neurosciences*, *7*(24), doi:10.3389/fnagi201500024 PMID: 25852542

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*(1), 40–48.

Nash, E. B., Edwards, G. W., Thompson, J. A., & Barfield, W. (2000). A review of presence and performance in virtual environments. *International Journal of Human-Computer Interaction*, *12*(1), 1–41.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., & Chertkow, H. (2005). The Montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. doi:10.1111/j.1532-5415.2005.53221.x

Naveh-Benjamin, M. (1990). Coding of temporal order information: An automatic process? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(1), 117–126. doi:http://doi.org/10.1037/0278-7393.16.1.117

Naveh-Benjamin, M. (2000). Adult Age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1170–1187.

Ouellet, É, Boller, B., Corriveau-Lecavalier, N., Cloutier, S., & Belleville, S. (2018). The Virtual Shop: A new immersive virtual reality environment and scenario for the assessment of episodic memory. *Journal of Neurosciences Methods*, *303*, 126–135.

Parsons, T. D., & Courtney, C. G. (2014). An initial validation of the virtual reality paced auditory serial addition test in a college sample. *Journal of Neuroscience Methods*, *222*, 15–23. doi:10.1016/j.jneumeth.2013.10.006

Parsons, T. D., & Rizzo, A. A. (2008). Initial validation of a virtual environment for assessment of memory functioning: Virtual reality cognitive performance assessment test. *CyberPsychology & Behavior*, *11*(1), 17–25. doi:10.1089/cpb.2007.9934

Piolino, P., Desgranges, B., & Eustache, F. (2009). Episodic autobiographical memories over the course of time: Cognitive, neuropsychological and neuroimaging findings. *Neuropsychologia*, *47*(11), 2314–2329. doi:10.1016/j.neuropsychologia.2009.01.020

Plancher, G., Nicolas, S., & Piolino, P. (2008). Apport de la réalité virtuelle en neuropsychologie de la mémoire: étude dans le vieillissement. *Psychologie & NeuroPsychiatrie du Vieillissement*, *6*(1), 7–22. doi:10.1684/pnv.2008.0119

Rizzo, A., Schulteis, M., Kerns, K., & Mateer, C. (2004). Analysis of assets for virtual reality applications in neuropsychology. *Neuropsychological Rehabilitation*, *14*(1–2), 207–239. doi:10.1080/09602010343000183

Robillard, G., Bouchard, S., Renaud, P., & Cournoyer, L. (2002). *Validation canadiennefrançaise de deux mesures importantes en réalité virtuelle: l'immersive tendencies questionnaire et le presence questionnaire.* Poster presented at the 25e congrès annuel de la Société Québécoise pour la Recherche en Psychologie (SQRP), Trois-Rivières.

Sauzéon, H., K'Kaoua, B., Pala, P. A., Taillade, M., Auriacombe, S., & Guitton, P. (2016). Everyday-like memory for objects in ageing and Alzheimer's disease assessed in a visually complex environment: The role of executive functioning and episodic memory. *Journal of Neuropsychology*, *10*(1), 33–58.

Sbordonne, R., & Long, C. (1996). *The ecological validity of neuropsychological tests.* Orlando, FL: Deutch.

Schultheis, M. T., Himelstein, J., & Rizzo, A. A. (2002). Virtual reality and neuropsychology: Upgrading the current tools. *The Journal of Head Trauma Rehabilitation*, *17*(5), 378–394. Retrieved from http://journals.lww.com/headtraumarehab/Abstract/2002/10000/Virtual_Reality_and_Neuropsychology__Upgrading_the.2.aspx

Selwynn, N. (2004). The information aged: A qualitative study of older adult's use of information and communications technology. *Journal of Aging Studies*, *18*(4), 369–384. doi:https://doi.org/10.1016/j.jaging.2004.06.008

Shuchat, J., Ouellet, E., Moffat, N., & Belleville, S. (2012). Opportunities for virtual reality in cognitive training with persons with mild cognitive impairment or Alzheimer's disease. *Non-Pharmacological Therapies in Dementia*, *3*(1), 35. Retrieved from http://search.proquest.com/openview/a2896d865f593322ea54775b951cd41e/1?pq-origsite=gscholar

Slater, M. (1999). Measuring presence: A response to the Witmer and Singer presence questionnaire. *Presence: Teleoperators and Virtual Environments*, *8*(5), 560–565. doi:10.1162/105474699566477

Slobounov, S. M., Ray, W., Johnson, B., Slobounov, E., & Newell, K. M. (2015). Modulation of cortical activity in 2D versus 3D virtual reality environments: An EEG study. *International Journal of Psychophysiology*, *95*(3), 254–260. doi:10.1016/j.ijpsycho.2014.11.003

Stevens, J. A., & Kincaid, J. P. (2015). The relationship between presence and performance in virtual simulation training. *Open Journal of Modelling and Simulation*, *3*, 41–48.

Sveistrup, H. (2004). Motor rehabilitation using virtual reality. *Journal of Neuroengineering and Rehabilitation*, *1*(1), 10.

Van der Linden, M., Adam, S., Agniel, A., Baisset-Mouly, C., Bardet, F., Coyette, F., et al. (2004). *L'Évaluation des troubles de la mémoire: Présentation de quatre tests de mémoire épisodique (avec leur étalonnage).* Marseille: Solal Editeurs.

Weber, R., Tamborini, R., Westcott-Baker, A., & Kantor, B. (2009). Theorizing flow and media enjoyment as cognitive synchronization of attentional and reward networks. *Communication Theory*, *19*(4), 397–422. doi:10.1111/j.1468-2885.2009.01352.x

Witmer, B. G., & Singer, M. F. (1994). *Measuring presence in virtual environments (No. ARI-TR-1014).* Alexandria, VA: Army research institute for the behavioral and social sciences.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225–240.

Zelinski, E. M. (2009). Far transfer in cognitive training of older adults. *Restorative Neurology and Neurosciences*, *27*(5), 455–471.