Feasibility and user-experience of virtual reality in neuropsychological assessment following stroke

Lauriane A. Spreij, Johanna M.A. Visser-Meily, Jacqueline Sibbel, Isabel K. Gosselt & Tanja C.W. Nijboer

Routledge
Taylor & Francis Group

Feasibility and user-experience of virtual reality in neuropsychological assessment following stroke

Lauriane A. Spreij [a], Johanna M.A. Visser-Meily [a,b], Jacqueline Sibbel[b], Isabel K. Gosselt [a] and Tanja C.W. Nijboer [a,c]

[a]Center of Excellence for Rehabilitation Medicine, UMC Utrecht Brain Center, University Medical Center Utrecht, De Hoogstraat Rehabilitation, Utrecht, Netherlands; [b]Department of Rehabilitation, Physical Therapy Science & Sports, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht, Netherlands; [c]Department of Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, Netherlands

ABSTRACT
Virtual Reality (VR) offers the possibility to assess cognitive functioning in a dynamic environment resembling daily life. In this cross-sectional study, we used two user interfaces, namely non-immersive VR by using a computer monitor (CM) and immersive VR by using a head-mounted display (HMD). We investigated (1) potential differences in feasibility, user-experience, and a potential preference for one user interface over another between stroke patients and healthy controls; (2) potential differences in feasibility, user-experience, and preference between patients referred for *inpatient* rehabilitation care and patients referred for *outpatient* rehabilitation care; and (3) potential demographic and clinical characteristics that were related to patients' preference for one user interface over another. Stroke patients ($n = 88$) and healthy controls ($n = 66$) performed a VR-task with a CM and HMD. Both user interfaces were feasible to use, irrespective of clinical referral (in- or outpatient rehabilitation care). Patients reported an enhanced feeling of engagement, transportation, flow, and presence, but more negative side effects when tested with a HMD, compared to a CM. The majority of stroke patients had no preference for one user interface over the other, yet younger patients tended to prefer a HMD. VR seems highly feasible in stroke patients.

Introduction

Cognitive rehabilitation refers to a set of interventions that focus on improving cognitive functioning to promote functional independence during activities of daily living (ADL) and social participation (Cicerone et al., 2000). Cognitive

CONTACT Tanja C.W. Nijboer ✉ t.c.w.nijboer@uu.nl 🖂 Center of Excellence for Rehabilitation Medicine, UMC Utrecht Brain Center, University Medical Center Utrecht, De Hoogstraat Rehabilitation, P.O. Box 85500, Heidelberglaan 100, 3584 CX Utrecht, Netherlands

rehabilitation typically begins with a thorough neuropsychological assessment to identify cognitive strengths and weaknesses. The conclusions of the assessment are used to formulate an appropriate treatment plan. Nowadays, neuropsychological assessment usually consists of paper-and-pencil tests that are conducted in a quiet room where distractions are minimized. Although these tests are convenient to purely measure underlying cognitive functions, research has often reported a lack of ecological validity (Chaytor & Schmitter-Edgecombe, 2003; Dawson & Marcotte, 2017). Performances on paper-and-pencil tests do not translate easily to daily life functioning, which results in a poor understanding of the difficulties patients encounter in daily life (Donovan et al., 2008).

Ecologically valid assessment has evolved as an area of focus within clinical neuropsychology (Chaytor & Schmitter-Edgecombe, 2003). Several standardized tests have been developed with an improved ecological validity, such as the Test of Everyday Attention (TEA; Robertson et al., 1996), the Rivermead Behavioural Memory Test (RBMT; Wilson et al., 1985), and the Behavioural Assessment of the Dysexecutive Syndrome (BADS; Wilson et al., 1996). However, even if most researchers agree that these tests seem similar to everyday tasks, participants remain well aware of the laboratory setting. For this reason, ecologically valid tests have been developed that are conducted in the real-world, such as the Multiple Errands Test (Shallice & Burgess, 1991) or the Executive Secretarial Task (Lamberts et al., 2010). A limitation, however, is the lack of a standardized and controlled setting, which results in an inconsistent degree of distractions within and between assessments.

Virtual Reality (VR) offers a novel possibility to assess cognitive functioning in simulated environments resembling daily life (Bohil et al., 2011; Maggio et al., 2019a; Parsey & Schmitter-Edgecombe, 2013; Parsons, 2011, 2015; Parsons et al., 2017; Parsons et al., 2014; Rizzo et al., 2004; Schultheis et al., 2002). VR allows the development of ecologically valid environments without losing control over stimulus presentation, while capturing precise and detailed performance measures due to a continuous data acquisition. In this study, we used two primary user interfaces, namely non-immersive VR by using a computer monitor (CM) and immersive VR by using a head-mounted display (HMD). CMs are considered the least interactive implementation of VR, but are often already accessible and are therefore a low-cost implementation of VR. HMDs are considered the highest interactive implementation of VR and allow patients to be fully immersed and to interact naturally with the virtual environment. Although the use of VR in neuropsychological assessment has been promising, is it feasible to use in stroke patients? How do stroke patients experience non-immersive and immersive VR?

Feasibility studies are used to determine whether an intervention is appropriate for further testing, in other words, whether or not the ideas and findings can be shaped to be relevant (Bowen et al., 2009). The objective of this study was to determine the feasibility (as measured with objective parameters, such as

completion rate), user-experience and preference (as measured with subjective parameters) of VR in stroke patients. We asked stroke patients to perform a VR-task in a *virtual supermarket* twice, one time by using a CM and one time by using a HMD. We investigated (1) potential differences in feasibility, user-experience, and a potential preference for one user interface over another between stroke patients and healthy controls; (2) potential differences in feasibility, user-experience, and preference between patients referred for *inpatient* rehabilitation care and patients referred for *outpatient* rehabilitation care; and (3) potential demographic and clinical characteristics that were related to patients' preference for one user-interface over another.

Methods

Participants

In the Netherlands, stroke patients are referred for *inpatient* rehabilitation care when: (a) a safe discharge from hospital to home is not achievable within 5 days; (b) the patient is physically and cognitively capable to participate in therapy; (c) a multidisciplinary approach is essential to reach complex rehabilitation goals; and (d) discharge from inpatient rehabilitation to home is expected in view of the prognosis and availability of the caregivers within 3 months. Stroke patients are referred for *outpatient* rehabilitation care when: (a) a safe discharge from hospital to home is achievable; and (b) a multidisciplinary approach is essential to reach rehabilitation goals.

We recruited participants between June 2016 and July 2019. We recruited stroke patients who were referred for inpatient or outpatient rehabilitation care at *De Hoogstraat Rehabilitation Center*, and stroke patients who were referred for outpatient rehabilitation care at the *University Medical Center Utrecht,* the Netherlands. Outpatients referred for rehabilitation care are a very specific group of stroke patients that have a relative good outcome (so-called "walk and talk group"). Inclusion criteria for all patients were (1) clinically diagnosed with stroke (confirmed by an MRI or CT scan); (2) aged $\geq$ 18 years; (3); physically and cognitively able to perform two VR-tasks as evaluated by the multidisciplinary team (clinicians who were actively engaged in the treatment, such as rehabilitation physicians, occupational therapists, neuropsychologists) and substantiated with objective measurements. When the opinion of the team was that motor or communication problems were so severe that patients could either not work with the joystick or controllers, or would not be able to understand task instructions or fill out the questionnaires, they would not be included in the study. The exclusion criteria were the diagnosis of (1) epilepsy (as the changing images could potentially trigger a seizure in patients with photosensitive epilepsy), and (2) severe visuo-spatial neglect based on a screening that was administered within the first two weeks of admission (care as

usual). Patients who would largely ignore one side of space and were not able to compensate for this were excluded. Inpatients who met the inclusion criteria received more information about the study and participation was discussed. Outpatients were invited by an information letter handed out by a clinician or sent by post. Participation was discussed by phone. When patients were willing to participate, an appointment was scheduled that was appropriate given their individual rehabilitation schedule.

We recruited healthy controls among acquaintances and colleagues, and by using advertisements in newsletters in (elderly, sports) associations. We aimed to match age, sex, and level of education as best as possible. The inclusion criteria for healthy controls were: (1) no history of neurological and/or psychiatric disorders for which treatment was needed; and (2) aged $\geq$ 18 years. All participants gave written informed consent. The experiment was performed in accordance with the Declaration of Helsinki (The World Medical Association, 2008). The research protocol was approved by the Medical Ethical Committee of the University Medical Center Utrecht (METC protocol number 15-751/C).

Apparatus

A virtual supermarket was developed with the software Unity by Atoms2Bits for commercial purposes, and was adapted for research and potential clinical purposes in close collaboration with the *University Medical Centre Utrecht*, *De Hoogstraat Rehabilitation Centre*, and *Utrecht University*. It was designed to be used on a regular computer in combination with two user interfaces: a CM and a HMD. The virtual supermarket was modelled according to a regular Dutch supermarket and contained 18 shelves, eight cash registers, several product displays (e.g., bread, fruit, vegetables) and freezing compartments (Figure 1). Approximately 20,000 products were designed referring to real brands and packages from common products in Dutch supermarkets. The surface was 50×30 virtual metres. Participants navigated at a maximum speed of .5 m/s.

The CM was a 24 inch monitor with a resolution of 1920×1200 pixels. A wired controller was used to navigate (Xbox 360©). Participants were seated on an office chair in front of the CM, which was placed at approximately 90 cm from their eyes. Two types of HMDs were used in this study. Participants included in between June 2016 and February 2017 were tested with the Oculus Rift DK2© with a 100° field of view, a resolution of 960×1080 per eye, and a refresh rate of 75 Hz. A wired controller (Xbox 360©) was used to navigate. Participants included in between January 2018 and July 2019 were tested with the HTC Vive© with a 110° field of view, a resolution of 1080×1200 per eye, and a refresh rate of 90 Hz. The HTC Vive contained two controllers to navigate and two base stations with a tracking system for participants to navigate through

Figure 1. Impression of the virtual supermarket used in this study (reprinted with permission of Atoms2Bits).

real-time movement in the virtual environment (maximum space of 3×3 metres). Since balance deficits are common in stroke patients (Geurts et al., 2005), participants (healthy controls also) were seated on an office chair for safety reasons.

Procedure

Participants provided written informed consent before initiation of the experiment. Participants were asked to perform a VR-task twice, one time by using a CM and one time by using a HMD. To avoid a possible bias on the results due to learning or boredom, the order in which the conditions were administered was randomized: with half of the participants starting with the CM and the other half starting with the HMD (Figure 2). Participants received a practice trial to get familiar with the VR apparatus and environment (i.e., virtual supermarket with empty shelves to prevent a learning effect). After the practice trial, participants were instructed to (1) start the VR-task by passing through the entry gates, (2) find three products from a shopping list, and (3) pass the cash registers to finish. A grocery list was presented over three trials, and participants were asked to recall the products. There were two different shopping lists: (1) salt, matches, sprinkles; (2) hair wax, cookies, socks. The shopping lists were semi-randomized across conditions. Task duration was limited to 15 minutes per condition. After the VR-task, a questionnaire was administered to assess the user-experience. The procedure was then repeated with the other user interface. Finally, a questionnaire was administered to assess the preference for one of the two user interfaces. The total duration of the experiment was approximately one hour.
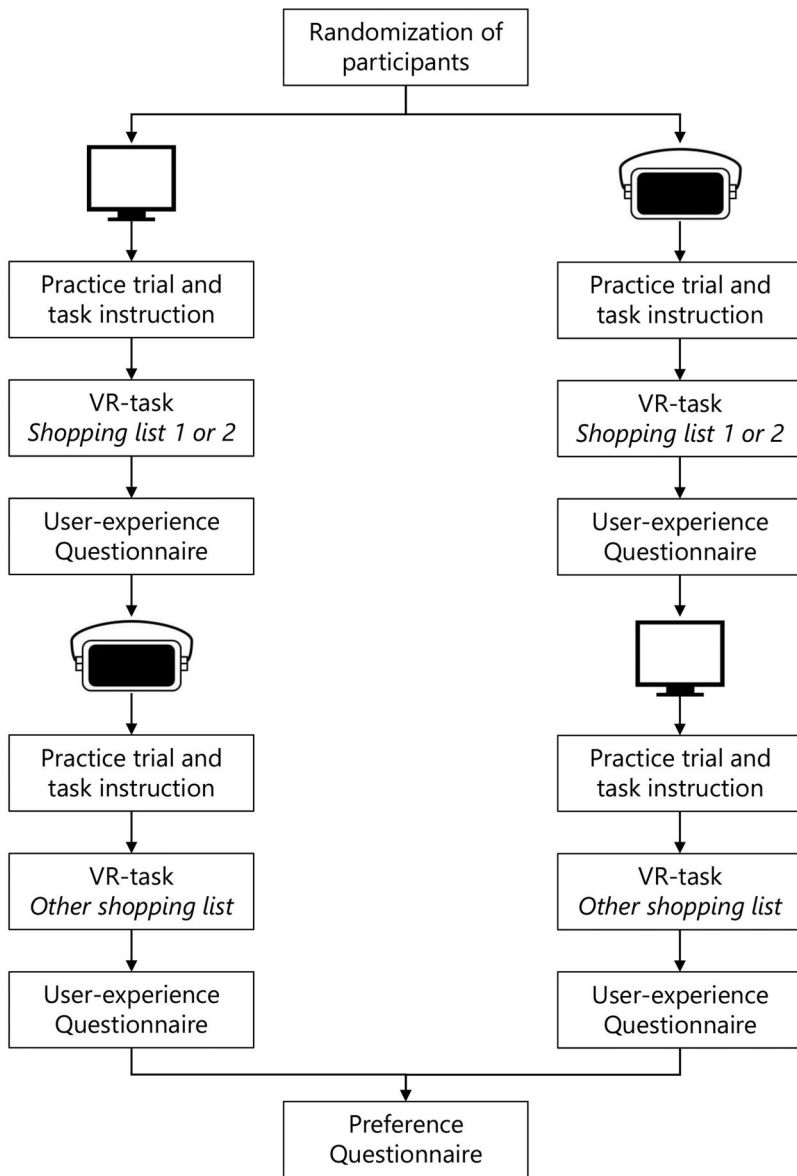
Figure 2. Procedure of the within subject design of this study.

Outcome measures

Feasibility measures

To investigate the feasibility, we calculated the completion rate (i.e., number of participants who completed the VR-task, who aborted the VR-task, and who did not start the VR task because of negative side effects during the practice trial), the total time needed to complete the VR-task (in minutes), and the total number of products found that were presented on the shopping list (range 0–3).

Questionnaire regarding user-experience

We developed a questionnaire to measure the user-experience based on previous cross-media research (Lessiter et al., 2001; Schuemie et al., 2001; Weibel et al., 2008). The questionnaire consisted of 15 items divided over five scales (three items per scale): (1) "engagement" defined as the feeling of involvement and enjoyment of the content; (2) "transportation" defined as the feeling of arriving in another world than the real world; (3) "flow" defined as a mental state in which a person is fully immersed in an activity with utmost concentration and distorted sense of time; (4) "presence" defined as the feeling of being physically present inside a virtual environment; (5) "negative effects" defined as adverse physiological reactions such as nausea. Response options were based on a 6-point Likert scale ranging from negative (--- [0]) to positive (+++ [5]). We summed the three items-scores belonging to a scale, resulting in a scale-score ranging from 0-15. An English translation of the questionnaire is presented in Appendix A (note that the results in this study are obtained with a Dutch version).

The face validity of the questionnaire was explored in an additional sample of 55 healthy controls (20% male, 89% high-educated, average age of 29.14 years [SD 9.78]). Those healthy controls did not participate in the main study. We asked the participants to cluster the items into five scales. A high percentage of participants clustered the right items into the scales engagement (86%), transportation (69%), and negative effects (96%), which indicated a valid face validity. A lower percentage of participants clustered the right items into the scales flow (40%) and presence (51%), which indicated a weaker face validity.

Questionnaire regarding preference

Participants were asked to indicate their preference for one of the two user interface in regard to five statements. The response options were: CM, HMD, or both. We quantified how many times a participant preferred the CM, HMD, or indicated to have no preference. An English translation of the questionnaire is presented in Appendix B.

Demographic and clinical characteristics

We collected data on sex, age and level of education. Level of education was assessed using a Dutch classification system (Verhage, 1965), consisting of 7 levels ([1] less than primary education; [2] primary education; [3] primary education and less than 2 years of low-level secondary education; [4] low-level secondary education; [5] average-level secondary education; [6] high-level secondary education; [7] academic degree). These levels were converted into three categories for analysis: low (level 1-4), average (level 5), and high (level 6-7). We asked six questions about the participant's familiarity with (1) 2D games; (2) 3D games; (3) 3D games with "first persons view"; (4) keyboards/touchscreens; (5) controllers; and (6) VR. Response

options were based on a 3-point Likert scale (- [0]; ± [1]; + [2]). The sum of the item scores was used as an indication of gaming experience, resulting in a score ranging from 0 (no gaming experience) to 12 (a great deal of gaming experience).

For all patients, we extracted time since stroke, stroke type (ischaemic, haemorrhage or subarachnoid haemorrhage) and lesion side (left, right or both) from the medical files. For inpatients, we extracted the scores on several clinical variables that were administered at admission as care as usual: communication skills as measured with the Stichting Afasie Nederland test (Deelman et al., 1981), independence during ADL as measured with the Barthel Index (Collin et al., 1988), and motor strength of upper and lower extremities as measured with the Motricity Index (Collin & Wade, 1990). For outpatients, these clinical variables were not administered since these patients would have had a maximum score to support their clinical referral to outpatient rehabilitation care. Global cognitive functioning was measured with the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005). This score was extracted from the medical files for inpatients and administered at the beginning of the test session for outpatients.

Statistical analyses

Demographic and clinical characteristics
We compared demographic and clinical characteristics between stroke patients and healthy controls by using non-parametric tests (Chi-square test for categorical variables and Mann-Whitney U test for continuous variables).

Differences in feasibility, user-experience and preference between user interfaces (CM vs. HMD)
The development of VR HMDs is ongoing. Important differences may be seen between old generation HMDs and new generation HMDs (Kourtesis et al., 2019). Halfway through this study, we switched from the Oculus Rift DK2 to the more sophisticated HTC Vive. To avoid a possible bias on the results, we first compared the feasibility, user-experience and preference between healthy controls who conducted the VR-task with the Oculus Rift DK2 ($n = 33$) and healthy controls who conducted the VR task with the HTC Vive ($n = 33$). We used non-parametric tests (Chi-square test for categorical variables and Mann-Whitney U test for continuous variables).

Furthermore, we compared the feasibility, user-experience and preference between the user interfaces (CM vs. HMD) and between patients and healthy controls. We used Chi-square tests (2×2), Fisher's exact tests (between-subject), and McNemar tests (within-subject) for categorical variables (i.e., completion rate, preference). We conducted a mixed analysis of variance

(ANOVA) with "user interface" as a within-subject factor (CM vs. HMD) and "group" as a between-subject factor (stroke patients vs. healthy controls) for continuous variables (i.e., total time, total products, scale-scores). A Benjamini-Hochberg correction was applied to counteract the problem of multiple comparisons (Benjamini & Hochberg, 1995). False discovery rate was set at .1.

Potential effect of clinical referral on feasibility, user-experience and preference

We compared the feasibility, user-experience and a potential preference for one user interface over another between patients who were referred for *inpatient* rehabilitation care (moderate to severe impaired patients) and patients who were referred for *outpatient* rehabilitation care (mild impaired patients). We used Chi-square tests, Fisher's exact tests, or McNemar tests for categorical variables and a mixed ANOVA for continuous variables.

Relations between demographic and clinical characteristics and patients' preference

Spearman correlations were computed between demographic (age, gaming experience) and clinical characteristics (MoCA score and time post-stroke onset) and preference for the CM, HMD, or both. An *r* of .1 was considered a small, .3 a moderate, and .5 a large relation (Field, 2009). The level of significance was set at $p = .05$.

Results

Demographic and clinical characteristics

From 249 stroke patients who were evaluated at the rehabilitation centre, 68 stroke patients were unable to participate as evaluated by clinicians ($n = 52$) or due to early discharge ($n = 16$). These numbers were not systematically recorded at the medical centre and no estimation can be given. A total of 181 patients were invited to participate, and 93 patients did not respond or declined due to various reasons (e.g., no time/interest). In total, 88 stroke patients, from the rehabilitation and medical centre combined, were included in this study. In addition, 66 healthy controls were included. See Table 1 for the demographic and clinical characteristics. There were more men in the patient sample than in the healthy controls sample ($\chi^2(2) = 5.43$, $p = .020$). Healthy controls were younger ($U = 1847.50$, $z = -3.86$, $p < .001$), higher educated ($\chi^2(2) = 18.82$, $p < .001$), and had more gaming experience ($U = 1961.00$, $z = -3.47$, $p = .001$), when compared to stroke patients.

Table 1. Demographic characteristics (split for stroke patients and healthy controls) and clinical characteristics (split for stroke patients who were referred for inpatient and outpatient rehabilitation care).

Demographic characteristics	Stroke patients (n=88)	n	Healthy controls (n=66)	n
Sex (% male)*	61.4	88	42.4	66
Age in years (mean, SD)*	55.32 (13.72)	88	46.24 (15.60)	66
Level of education (%)*		88		66
Low	21.6		3.0	
Moderate	27.3		13.6	
High	51.1		83.3	
Gaming experience 0–12 (mean, SD)*	4.00 (2.80)	88	5.45 (2.72)	66
Clinical characteristics	Inpatients (n=43)	n	Outpatients (n=45)	n
Stroke type (%)		43		45
Ischemic	69.8		71.1	
Haemorrhage	25.6		4.4	
Subarachnoid haemorrhage	4.7		20	
Rare causes of stroke[a]	0		4.4	
Lesion side (%)		43		45
Left	46.5		28.9	
Right	41.2		40	
Bilateral	0		15.6	
Unknown	2.3		15.6	
Days post stroke (mean, range)	153.74 (245.69)	43	490.76 (634.68)	45
Montreal Cognitive Assessment 0–30 (mean, SD)	22.92 (4.84)	39	24.86 (3.97)	35
MoCA Score 26 or below (%)	62.8		44.4	
Stichting Afasie Nederland 1–7 (mean, SD)	6.08 (4.97)	36	Not applicable[b]	
Barthel Index 0-20 (mean, SD)	13.39 (5.37)	41	Not applicable[b]	
Motricity Index upper 0–100 (mean, SD)	62.59 (38.34)	39	Not applicable[b]	
Motricity Index lower 0–100 (mean, SD)	74.58 (31.95)	38	Not applicable[b]	

Note: *Significantly different between stroke patients and healthy controls, p-value <.05.
[a]One patient diagnosed with cerebral venous sinus thrombosis and one patient with cerebral fat embolism, which are less common causes of stroke.
[b]These clinical variables were not administered since patients would have had a maximum score to support the clinical referral for outpatient rehabilitation care.

Differences in feasibility, user-experience and preference between user interfaces (CM vs. HMD)

Healthy controls who were tested with the Oculus Rift DK2 completed the VR-task less often ($\chi 2(2) = 8.41$, $p < .001$), reported less transportation ($U = 358.50$, $z = -2.24$, $p = .025$), less flow ($U = 379.00$, $z = -1.97$, $p = .049$), less presence ($U = 330.50$, $z = -2.46$, $p = .014$), more negative effects ($U = 313.50$, $z = -2.83$, $p = .005$), and had a distinct preference for the CM ($\chi 2(2) = 42.62$, $p < .001$), when compared to healthy controls who were tested with the HTC Vive. To avoid a possible bias of the type of HMD on the results, we only used the data of patients ($n = 74$) and healthy controls ($n = 33$) who were tested with the more sophisticated HTC Vive in further analyses. The HTC Vive offered a better quality and further VR HMD development would only make the devices better suited.

Stroke patients did not abort the VR-task more often than healthy controls with the CM (Fisher's exact, $p = .592$) nor with the HMD (Fisher's exact, $p = .732$). Stroke patients did complete the VR-task more often with the CM than with the HMD (McNemar test, $p = .039$). Based on a mixed ANOVA, we found a

Table 2. Feasibility in stroke patients and healthy controls, split for user interface (CM vs. HMD).

	Stroke patients ($n = 74$)		Healthy controls ($n = 33$)	
	Computer monitor	Head-mounted display	Computer monitor	Head-mounted display
Completion rate (%)	93.2	83.8	100	90.9
Aborted during VR-task (%)	4.1	10.8	0	6.1
Aborted after practice trial (%)	2.7	5.4	0	3.0
Total time to complete VR-task[a] (minutes)	10.46 (4.56)	10.15 (4.29)	7.47 (3.94)	7.31 (3.50)
Total number of found products[a] (0–3)	2.49 (.77)	2.43 (.96)	2.93 (.25)	2.87 (.43)

Note: [a]Participants who did not started (6 patients; 1 healthy control) or completed (7 patients; 2 healthy controls) one of the two conditions were excluded from these analyses (included: 61 patients; 30 healthy controls).

main effect of group, where stroke patients needed more time ($F(1, 88) = 18.97, p < .001$) and found less products ($F(1, 89) = 10.13, p = .002$), compared to healthy controls (Table 2). We found a main effect of user interface, where both patients and healthy controls reported an enhanced feeling of engagement ($F(1, 96) = 21.99, p < .001$), transportation ($F(1, 98) = 132.10, p < .001$), flow ($F(1, 98) = 29.60, p < .001$), and presence ($F(1, 94) = 109.75, p < .001$), but more negative effects ($F(1, 98) = 47.92, p < .001$) when tested with the HMD, compared to the CM. There was no significant difference in preference for one user interface over another between stroke patients and healthy controls ($\chi2(2) = 4.88, p = .088$), with the majority reporting to have no preference (Figure 3).
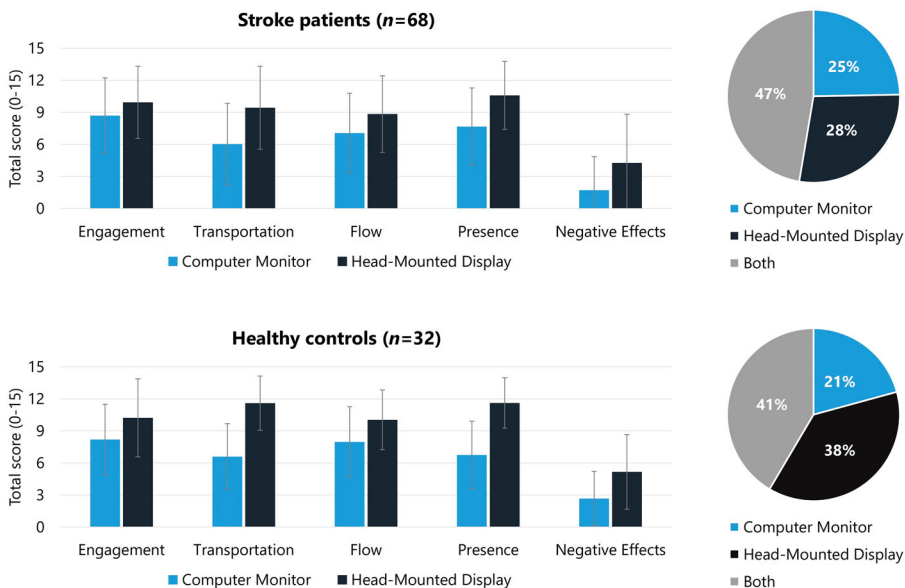
Figure 3. The user-experience of both user interface (CM vs. HMD) and the preference (for CM, HMD, or both) is depicted, split for stroke patients and healthy controls. Note that patients ($n = 6$) and healthy controls ($n = 1$) who did not started one of the two conditions were excluded from these analyses (included patients $n = 68$; healthy controls $n = 32$). Number of participants varies per variable since data was missing on one question within a scale for 6 participants.

Potential effect of clinical referral on feasibility, user-experience and preference

Patients who were referred for *inpatient* rehabilitation care did not abort the VR-task more often than patients who were referred for *outpatient* rehabilitation care with the CM (Fisher's exact = 1.33, p = .632), nor with the HMD (Fisher's exact = 612, p = .797). Outpatients reported significantly more negative effects compared to inpatients (F (1, 66) = 7.22, p = .009). We only found an interaction effect (group x user-interface) on the feeling of engagement (F (1, 59) = 8.66, p = .005). Outpatients, in comparison to inpatients, reported a significant improved feeling of engagement when conducting the VR-task with the HMD, compared to the CM. There was no significant difference in preference for one user interface over another between the patient groups ($\chi2(2)$ = 1.41, p = .494), with the majority reporting to have no preference (Table 3).

Table 3. Feasibility, user-experience and preference, split for stroke patients who were referred for inpatient rehabilitation care and patients who were referred for outpatient rehabilitation care.

	Inpatients (*n*=43)		Outpatients (*n*=31)	
	Computer monitor	Head-mounted display	Computer monitor	Head-mounted display
Feasibility				
Completion rate (%)	90.7	86	96.8	80.6
Aborted during VR-task (%)	4.7	9.3	3.2	12.9
Aborted after practice trial (%)	4.7	4.7	0	6.5
Completion time VR-task[a] (minutes)	11.14 (4.34)	10.20 (4.63)	9.44 (4.78)	10.07 (3.82)
Total number of found products[a] (0-3)	2.56 (.67)	2.31 (1.06)	2.40 (.87)	2.60 (.76)
User-experience[b]				
Engagement (0-15)	9.55 (3.24)	9.87 (3.40)	7.54 (3.63)	10.07 (3.39)
Transportation (0-15)	5.87 (3.89)	9.56 (3.82)	6.28 (3.75)	9.28 (4.02)
Flow (0-15)	7.64 (3.88)	9.41 (2.91)	6.31 (3.40)	8.10 (4.10)
Presence (0-15)	8.00 (3.58)	10.59 (3.73)	7.19 (3.69)	10.62 (2.16)
Negative effects (0-15)	.87 (1.87)	3.28 (4.44)	2.86 (4.06)	5.62 (4.42)
Preference[b]	Inpatients (*n*=43)		Outpatients (*n*=31)	
Computer Monitor	22%		28%	
Head-mounted Display	30%		25%	
Both	48%		47%	

Note. [a]Participants who did not started (4 inpatients; 2 outpatients) or completed (3 inpatients; 4 outpatients) one of the two condition were excluded from these analyses (included: 36 inpatients; 25 outpatients).
[b]Participtants who did not started one of the two conditions were excluded from these analyses (Included: 39 inpatients; 29 outpatients). Number of participants varies per variable since data was missing on one question within a scale for 5 participants.

Relations between demographic and clinical characteristics and patients' preference

Age and preference for the HMD were negatively related among stroke patients. Patients who were younger tended to prefer the HMD over the CM more often. There was no relation between preference and gaming experience, general cognitive functioning and days post stroke onset (Table 4).

Table 4. Relation between demographic and clinical characteristics and preference for one user interface in stroke patients ($n = 68$).

	Age	Gaming experience	MoCA score	Days post-stroke
Preference for head-mounted display	−.25 ($p = .043$)*	.18 ($p = .135$)	.09 ($p = .474$)	−.13 ($p = .279$)
Preference for computer monitor	−.01 ($p = .928$)	−.09 ($p = .489$)	.06 ($p = .629$)	.13 ($p = .304$)
No preference	.19 ($p = .116$)	−.076 ($p = .54$)	−.048 ($p = .707$)	.06 ($p = .650$)

Note. MoCA scores were extracted from medical files for inpatients and administered at the beginning of the test session for outpatients. Results should be interpreted with caution.
*$p < .05$.

Discussion

In this study, stroke patients performed a VR-task in a *virtual supermarket* twice, one time by using a CM and once time by using a HMD. We investigated (1) potential differences in feasibility, user-experience, and preference for one user interface over another between stroke patients and healthy controls; (2) potential differences in feasibility, user-experience, and preference between patients referred for *inpatient* rehabilitation care and patients referred for *outpatient* rehabilitation care; and (3) potential demographic and clinical characteristics that were related to patients' preference for one user-interface over another. A high percentage of patients completed the VR-task when tested with a CM (93%) and a HMD (84%). This suggests that it is feasible to use both non-immersive and immersive VR user interfaces (CM and HMD) in stroke patients. Patients and healthy controls reported an enhanced feeling of engagement, transportation, flow, and presence when tested with the HMD, when compared to the CM. Hence, the use of a HMD evokes an enhanced user-experience, which is expected to lead to a more natural behaviour and interaction with the virtual environment (Parsons, 2015). However, more adverse physiological reactions were reported by both stroke patients and healthy controls when tested with a HMD, when compared to a CM. Negative effects are expected to decrease with further VR HMD development (Kourtesis et al., 2019), which was also shown in this study where healthy controls experienced more negative effects when tested with the older Oculus Rift DK2 compared to the more sophisticated HTC Vive. Current best practice guides for VR development focus on alleviating negative effects by using several approaches to reduce sensory mismatch, such as display factors (e.g., higher refresh rates) and intuitiveness of interaction and navigation (Kourtesis et al., 2019; Kourtesis et al., 2020; Oculus, 2017; Weech et al., 2019). Furthermore, negative effects decrease over repeated exposures, which emphasizes the importance of practice trials to help a user become more familiar with a particular device (Germine et al., 2019; Kennedy et al., 2000). Importantly, stroke patients reported no preference for one user interface (CM vs. HMD), which increases the usability of VR in clinical practice as patients are willing to work with both user interfaces with their own set of strengths and limitations. This allows for a tailor-made application, dependent on the aim of the assessment and the willingness of a

patient. The use of a HMD seems preferable in neuropsychological assessment since it induces more natural behaviour, but a CM remains a valid alternative when a HMD is not accessible or not feasible with a particular patient.

We did not find an effect of clinical referral to *inpatient* or *outpatient* rehabilitation care on the feasibility, user-experience and preference, which indicates that VR (when using a CM or HMD) is feasible in patients who are more severely injured by stroke. Indeed, general cognitive functioning did not affect the preference for one user interface, nor did the time post stroke onset. We only found a small negative relation between age and preference for the HMD, indicating that younger patients tended to prefer a HMD over a CM more often. Gaming experience did not affect the preference in our sample. One should be cautious however, as only limited research on this topic has been performed in stroke patients and opposite effects have also been reported in healthy controls (Weech et al., 2019). A next step on this topic, would be the investigation whether gaming experience would affect cognitive performance in a virtual environment. Previous research shows that individuals with more computer experience tend to demonstrate a better cognitive performance on computer-based assessment, than individuals with less computer experience (Iverson et al., 2009; Tun & Lachman, 2010). This might also be the case with VR-based assessments (Iverson et al., 2009; Tun & Lachman, 2010). This would mainly mean that we might have to facilitate longer practice trials for patients with less gaming experience, to help them get more familiar with the devices and virtual environment.

Strengths and limitations

A strength of this study was the inclusion of a large number of stroke patients ($n = 88$) and the recruitment in both a rehabilitation centre and medical centre, which increases the representativeness of our sample, at least for the way rehabilitation care is organized in the Netherlands. The sample of stroke patients in this study encompasses a wide range of severity of stroke and severity of *consequences* of stroke. Furthermore, research emphasizes the importance of including patients in the evaluation of new medical technological devices (Lee, 2019), so incorporating the user-experience and preference of stroke patients provides useful insights into the use of VR in clinical practice. An unknown factor, however, is the feasibility, user-experience and preference of the more severely hampered stroke patients. Clinicians evaluated whether participation would be made possible, and we excluded patients with interfering impairments (e.g., severe motor problems hindering the use of a controller, severe communication problems preventing them to understand the instruction, severe fatigue). The feasibility of VR in severely injured patients remains therefore unknown. This may be considered as a limitation, as using VR could make the whole testing experience less

threatening and more enjoyable, which in turn could motivate patients to undergo assessment more often and monitor their cognitive functioning more closely (Zygouris & Tsolaki, 2015b).

A limitation that should be considered, is that a part of the stroke patients and healthy controls was excluded from further analyses, due to significant differences in user-experience when comparing the Oculus Rift DK2 and the HTC Vive. This resulted in smaller samples of especially healthy controls for the subsequent analyses (stroke patients from $n = 88$ to $n = 74$; healthy controls from $n = 66$ to $n = 33$). For the feasibility, user-experience and preference of stroke patients this most likely did not have a large effect, yet for comparisons of these results with those of the healthy controls, we need to be cautious. Another potential limitation is the difference in demographic characteristics and gaming experience of the healthy controls and patients. Even though we tried to match age, sex, and level of education of both groups, the healthy controls were younger and higher educated. Also, there were more men among stroke patients. For the aims on feasibility and preference among the stroke patients, this does not have a large impact, but when comparing the outcomes of the stroke patients to those of the healthy controls, we cannot be sure that the current results (i.e., feasibility, user-experience and preference) would be comparable with an older and lower educated sample of healthy controls. For the feasibility, user-experience and preference of stroke patients, this does not change the conclusion.

We used questionnaires as subjective parameters to assess user-experience and preference. The questionnaire regarding user-experience has given important insights into the five scales (engagement, transportation, flow, presence, negative effects). With regard to the CM condition, few negative effects were reported, namely nausea (7% of patients; 12% of healthy controls), feeling warm (13% of patients, 19% of healthy controls) and having a headache (11% of patients; 5% of healthy controls). With regard to the HMD condition, more negative effects were reported, namely nausea (32% of patients; 54% of healthy controls), feeling warm (47% of patients, 65% of healthy controls) and having a headache (20% of patients; 14% of healthy controls). However, one might argue that it was not elaborated enough. For example, "dizziness" was not part of the scale "negative effects." Certain patients (e.g., patients with cerebellar and/or midbrain stroke) might experience dizziness in daily life situation, and thus also while performing a VR-task. As "dizziness" has been a commonly reported effect of VR (Szpak et al., 2019), it seems crucial to incorporate "dizziness" as item in future studies. Furthermore, the preference questionnaire did not have the response option "none/neither", so based on the questionnaire, we would not be able to know whether there were patients that would not like to use VR at all. Informal feedback, however, suggests that patients do see the potential of such technology, as virtual environments resemble real-life environments and replicate the challenges found in daily

life situations. This feedback is very valuable as it stresses the relevance of VR-tasks in neuropsychological assessment.

Future research

VR offers the opportunity to gain valuable information which cannot be obtained through paper-and-pencil tests, such as wayfinding features (e.g., crossing one's own pathway, location and duration stops), eye movement features (e.g., fixation duration, number of re-fixations, pupillometry), and time-based measures (e.g., reaction time, fluctuations in pace) (Lutz et al., 2017; Parsey & Schmitter-Edgecombe, 2013). A first step should be the development of outcome measures, on which patients with impaired cognitive functions perform significantly different than cognitively healthy controls. For instance, based on the relatively simple outcome measures in this study, we found that stroke patients needed more time and found less products, when compared to healthy controls. Numerous studies describe VR-tasks in different populations and report significant differences in performance between patients and healthy controls. For instance, a VR office and meeting room is used to test patients with frontal lobe lesions on multitasking abilities by evaluating the quality of the performance on different tasks (failure, partial or satisfactory completion) (Denmark et al., 2019), and a VR shopping task is used to test patients with traumatic brain injury on prospective memory by evaluating event-based measures (press button when sale-announcement is heard) and time-based measures (send text message at three different time points) (Canty et al., 2014). A next important step is the development of outcome measures to accurately discriminate patients from healthy controls. For instance, a virtual supermarket task showed a correct classification of 87% when discriminating patients with mild cognitive impairment from healthy controls with outcome measures, such as test duration and correctly bought products (Zygouris et al., 2015a). A combination of outcome measures, may be used to identify distinct patterns of scores discriminating patients and healthy controls more accurately. In a large sample, data-driven machine learning analyses might reveal which behavioural patterns occur from the data. Data-driven analyses may enable a shift towards developing more sophisticated models of behaviour to further improve the sensitivity of neuropsychological assessment by using VR-tasks. Finally, a next step would be the validation and the reliability of the VR-task, followed with the derivation of normative data to help clinicians interpret the complementing information (Iriarte et al., 2012).

Finally, previous research has reported promising results in using VR in cognitive rehabilitation (Larson et al., 2014; Laver et al., 2015; Maggio et al., 2019b; Moreno et al., 2019; Rizzo et al., 2004; T. Rose et al., 2018; Shin & Kim, 2015). The use of VR in cognitive rehabilitation may have important benefits, such as the

opportunity to train skills and compensation strategies in a safe environment while interacting with people and/or objects (F. D. Rose et al., 2005).

Conclusion

In this study, we found that the use of both non-immersive and immersive VR user interfaces (CM and HMD) is feasible in stroke patients, irrespective of clinical referral for inpatient or outpatient rehabilitation care. Patients reported an enhanced feeling of engagement, transportation, flow, and presence, but more negative effects when tested with a HMD, when compared to a CM. Negative effects are likely to decrease with more sophisticated HMD, which is a lead focus in best practice guides for VR development. The majority of stroke patients had no preference for one user interface, yet younger patients tended to prefer HMDs more often. Future research should focus on novel outcome measures, the validation and reliability, and the development of normative data.

Acknowledgements

Disclosure statement

Funding

ORCID

Lauriane A. Spreij http://orcid.org/0000-0001-5558-9746
Johanna M.A. Visser-Meily http://orcid.org/0000-0002-5955-8012
Isabel K. Gosselt http://orcid.org/0000-0001-8086-8843
Tanja C.W. Nijboer http://orcid.org/0000-0001-6683-0267

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False discovery rate – a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methological)*, *57*(1), 289–300. https://www.jstor.org/stable/2346101

Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews Neuroscience*, *12*(12), https://doi.org/10.1038/nrn3122

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., Bakken, S., Kaplan, C. P., Squiers, L., Fabrizio, C., & Fernandez, M. (2009). How we design feasibility studies. *American Journal of Preventive Medicine*, *36*(5), 452–457. https://doi.org/10.1016/j.amepre.2009.02.002

Canty, A. L., Fleming, J., Patterson, F., Green, H. J., Man, D., & Shum, D. H. K. (2014). Evaluation of a virtual reality prospective memory task for use with individuals with severe traumatic brain injury. *Neuropsychological Rehabilitation*, *24*(2), 238–265. https://doi.org/10.1080/09602011.2014.881746

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. https://doi.org/10.1023/B:NERV.0000009483.91468.fb

Cicerone, K. D., Dahlberg, C., Kalmar, K., Langenbahn, D. M., Malec, J. F., Bergquist, T. F., Felicetti, T., Giacino, J. T., Harley, J. P., Harrington, D. E., Herzog, J., Kneipp, S., Laatsch, L., & Morse, P. A. (2000). Evidence-based cognitive rehabilitation: Recommendations for clinical practice. *Archives of Physical Medicine and Rehabilitation*, *81*(12), 1596–1615. https://doi.org/10.1053/apmr.2000.19240

Collin, C., & Wade, D. (1990). Assessing motor impairment after stroke : A pilot reliability study. *Journal of Neurology, Neurosurgery, and Psychiatry*, *53*(7), 576–579. https://doi.org/10.1136/jnnp.53.7.576 Collin,

Collin, C., Wade, D. T., Davies, S., & Horne, V. (1988). The Barthel ADL Index: A reliability study. *International Disability Studies*, *10*(2), 61–63. https://doi.org/10.3109/09638288809164103

Dawson, D. R., & Marcotte, T. D. (2017). Special issue on ecological validity and cognitive assessment. *Neuropsychological Rehabilitation*, *27*(5), 599–602. https://doi.org/10.1080/09602011.2017.1313379

Deelman, B., Koning-Haanstra, M., Liebrand, W., & Van den Burg, W. (1981). *Stichting Afasie Nederland - de SAN-test*. Swets & Zeitlinger.

Denmark, T., Fish, J., Jansari, A., Tailor, J., Ashkan, K., & Morris, R. (2019). Using virtual Reality to investigate multitasking ability in individuals with frontal lobe lesions. *Neuropsychological Rehabilitation*, *29*(5), 767–788. https://doi.org/10.1080/09602011.2017.1330695

Donovan, N. J., Kendall, D. L., Heaton, S. C., Kwon, S., Velozo, C. A., & Duncan, P. W. (2008). Conceptualizing functional cognition in stroke. *Neurorehabilitation and Neural Repair*, *22* (2), 122–135. https://doi.org/10.1177/1545968307306239

Field, A. (2009). *Discovering statistics using SPSS*. https://doi.org/10.1234/12345678

Germine, L., Reinecke, K., & Chaytor, N. S. (2019). Digital neuropsychology: Challenges and opportunities at the intersection of science and software. *The Clinical Neuropsychologist*, *33*(2), 271–286. https://doi.org/10.1080/13854046.2018.1535662

Geurts, A. C. H., De Haart, M., Van Nes, I. J. W., & Duysens, J. (2005). A review of standing balance recovery from stroke. *Gait and Posture*, *22*(3), 267–281. https://doi.org/10.1016/j.gaitpost.2004.10.002

Iriarte, Y., Diaz-Orueta, U., Cueto, E., Irazustabarrena, P., Banterla, F., & Climent, G. (2012). AULA —advanced virtual reality tool for the assessment of attention: Normative study in Spain. *Journal of Attention Disorders*, *20*(6), 542–568. https://doi.org/10.1177/1087054712465335

Iverson, G. L., Brooks, B. L., Ashton, V. L., Johnson, L. G., & Gualtieri, C. T. (2009). Does familiarity with computers affect computerized neuropsychological test performance? *Journal of Clinical and Experimental Neuropsychology*, *31*(5), 594–604. https://doi.org/10.1080/13803390802372125

Kennedy, R. S., Stannney, K. M., & Dunlap, W. P. (2000). Duration and exposure to virtual environments: Sickness curves during and across sessions. *Presence: Teleoperators and Virtual Environments*, *9*(5), 463–472. https://doi.org/10.1162/105474600566952

Kourtesis, P., Collina, S., Doumas, L. A. A., & MacPherson, S. E. (2019). Technological compe-tence is a pre-condition for effective implementation of virtual reality head mounted dis-plays in human neuroscience: A technological review and meta-analysis. *Frontiers in Human Neuroscience*, *13*. https://doi.org/10.3389/fnhum.2019.00342

Kourtesis, P., Korre, D., Collina, S., Doumas, L. A. A., & MacPherson, S. E. (2020). Guidelines for the development of immersive virtual reality software for cognitive Neuroscience and neu-ropsychology: The development of virtual Reality everyday assessment Lab (VR-EAL), a neuropsychological test Battery in immersive virtual Reality. *Frontiers in Computer Science*, *1*(12), 1–24. https://doi.org/10.3389/fcomp.2019.00012

Lamberts, K. F., Evans, J. J., & Spikman, J. M. (2010). A real-life, ecologically valid test of execu-tive functioning: The executive secretarial task. *Journal of Clinical and Experimental Neuropsychology*, *32*(1), 56–65. https://doi.org/10.1080/13803390902806550

Larson, E. B., Feigon, M., Gagliardo, P., & Dvorkin, A. Y. (2014). Virtual reality and cognitive rehabilitation: A review of current outcome research. *NeuroRehabilitation*, *34*(4), 759–772. https://doi.org/10.3233/NRE-141078

Laver, K. E., George, S., Thomas, S., Deutsch, J. E., & Crotty, M. (2015). Virtual reality for stroke rehabilitation. *Cochrane Database of Systematic Reviews*, (2), 1–17. https://doi.org/10.1002/14651858.CD008349.pub3

Lee, D. H. (2019). A model for designing healthcare service based on the patient experience. *International Journal of Healthcare Management*, *12*(3), 180–188. https://doi.org/10.1080/20479700.2017.1359956

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence: Teleoperators and Virtual Environments*, *10*(3), 282–297. https://doi.org/10.1162/105474601300343612

Lutz, O. H.-M., Burmeister, C., Santos, L. F. d., Morkisch, N., Dohle, C., & Krüger, J. (2017). Application of head-mounted devices with eye-tracking in virtual reality therapy. *Current Directions in Biomedical Engineering*, *3*(1), 53–56. https://doi.org/10.1515/cdbme-2017-0012

Maggio, M. G., De Luca, R., Molonia, F., Porcari, B., Destro, M., Casella, C., Salvati, R., Bramanti, P., & Calabro, R. S. (2019a). Cognitive rehabilitation in patients with traumatic brain injury: A narrative review on the emerging use of virtual reality. *Journal of Clinical Neuroscience*, *61*, 1–4. https://doi.org/10.1016/j.jocn.2018.12.020

Maggio, M. G., Latella, D., Maresca, G., Sciarrone, F., Manuli, A., Naro, A., De Luca, R., & Calabrò, R. S. (2019b). Virtual reality and cognitive rehabilitation in people with stroke: An overview. *Journal of Neuroscience Nursing*, *51*(2), 101–105. https://doi.org/10.1097/JNN.0000000000000423

Moreno, A., Wall, K. J., Thangavelu, K., Craven, L., Ward, E., & Dissanayaka, N. N. (2019). A sys-tematic review of the use of virtual reality and its effects on cognition in individuals with neurocognitive disorders. *Alzheimer's and Dementia: Translational Research and Clinical Interventions*, *5*(1), 834–850. https://doi.org/10.1016/j.trci.2019.09.016

Nasreddine, Z. S., Philips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & 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. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Oculus, V. L. (2017). Best practices guide. In *Version 310-30000-02*. https://doi.org/10.2106/JBJS.H.00657

Parsey, C. M., & Schmitter-Edgecombe, M. (2013). Applications of technology in neuropsycho-logical assessment. *The Clinical Neuropsychologist*, *27*(8), 1328–1361. https://doi.org/10.1080/13854046.2013.834971

Parsons, T. D. (2011). Neuropsychological assessment using virtual environments: Enhanced assessment technology for improved ecological validity. In S. Brahnam & L. C. Jain (Eds.),

Advamced computer intelligence paradigms in healthcare (pp. 271–289). Berlin: Springer-Verlag.

Parsons, T. D. (2015). Ecological validity in virtual reality - based neuropsychological assessment. In M. Khosrow-Pour (Ed.), *Encyclopedia of information Science and technology, Third Edition* (pp. 1006–1015). Information Science Reference.

Parsons, T. D., Carlew, A. R., Magtoto, J., & Stonecipher, K. (2017). The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology. *Neuropsychological Rehabilitation*, 27(5), 777–807. https://doi.org/10.1080/09602011.2015.1109524

Parsons, T. D., McPherson, S., & Interrante, V. (2014). Enhancing neurocognitive assessment using immersive virtual reality. *IEEE Virtual Reality*, 27–34. https://doi.org/10.1109/VAAT.2013.6786190

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

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1996). The structure of normal human attention: The test of everyday attention. *Journal of the International Neuropsychological Society*, 2(6), 525–534. https://doi.org/10.1017/s1355617700001697

Rose, F. D., Brooks, B. M., & Rizzo, A. A. (2005). Virtual reality in brain damage rehabilitation: Review. *CyberPsychology & Behavior*, 8(3), 241–262. https://doi.org/10.1089/cpb.2005.8.241

Rose, T., Nam, C. S., & Chen, K. B. (2018). Immersion of virtual reality for rehabilitation – review. *Applied Ergonomics*, 69(February 2017), 153–161. https://doi.org/10.1016/j.apergo.2018.01.009

Schuemie, M. J., Van der Straaten, P., Krijn, M., & Van der Mast, C. A. P. G. (2001). Research on presence in virtual reality: A survey. *Cyberpsychology and Behavior*, 4(2), 183–201. https://doi.org/10.1089/109493101300117884

Schultheis, M. T., Himelstein, J., & Rizzo, A. A. (2002). Virtual reality and neuropsychology: Upgrading the current tools. *Journal of Head Trauma Rehabilitation*, 17(5), 378–394. https://doi.org/10.1097/00001199-200210000-00002

Shallice, T., & Burgess, P. (1991). Deficits in strategy application following frontal lobe damage in. *Brain*, 114(November 2014), 727–741. https://doi.org/10.1093/brain/114.2.727

Shin, H., & Kim, K. (2015). Virtual reality for cognitive rehabilitation after brain injury: A systematic review. *Journal of Physical Therapy Science*, 27(9), 2999–3002. https://doi.org/10.1589/jpts.27.2999

Szpak, A., Michalski, S., Saredakis, D., & Loetscher, T. (2019). Beyond feeling Sick: The Visual and cognitive Aftereffects of virtual Reality. *PsyArχiv*(July), https://doi.org/10.31234/osf.io/xerwz

Tun, P. A., & Lachman, M. E. (2010). The association between computer use and cognition across adulthood: Use it so You Won't Lose it? *Psychology and Aging*, 25(3), 560–568. https://doi.org/10.1037/a0019543

Verhage, F. (1965). Intelligence and age in a Dutch sample. *Human Development*, 8(4), 238–245. https://doi.org/10.1159/000270308

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: A review. *Frontiers in Psychology*, 10(FEB), 1–19. https://doi.org/10.3389/fpsyg.2019.00158

Weibel, D., Wissmath, B., Habegger, S., Steiner, Y., & Groner, R. (2008). Playing online games against computer- vs. Human-controlled opponents: Effects on presence, flow, and enjoyment. *Computers in Human Behavior*, 24(5), 2274–2291. https://doi.org/10.1016/j.chb.2007.11.002

Wilson, B. A., Alderman, N., Burgess, P., Emslie, H., & Evans, J. J. (1996). *Behavioural assessment of the Dysexecutive Syndrome (BADS)*. Thames Valley Test Company.

Wilson, B. A., Cockburn, J., & Baddeley, A. D. (1985). *The Rivermead behavioural memory test manual*. Thames Valley Test Company.

The World Medical Association. (2008). *Declaration of Helsinki: Ethical principles for medical research involving human subjects.* The World Medical Association, INC.

Zygouris, S., Giakoumis, D., Votis, K., Doumpoulakis, S., Ntovas, K., Segkouli, S., Segkouli, S., Karagiannidis, C., Tzovaras D., & Tsolaki, M. (2015a). Can a virtual reality cognitive training application fulfill a dual role? Using the virtual supermarket cognitive training application as a screening tool for mild cognitive impairment. *Journal of Alzheimer's Disease*, *44*(4), 1333–1347. https://doi.org/10.3233/JAD-141260

Zygouris, S., & Tsolaki, M. (2015b). Computerized cognitive testing for older adults: A review. *American Journal of Alzheimer's Disease and Other Dementias*, *30*(1), 13–28. https://doi.org/10.1177/1533317514522852

Appendices

Appendix A. Questionnaire regarding user-experience translated in English. Note that the results of this study were obtained with the original Dutch version.

Through this questionnaire, we want to ask about your experience with the virtual environment. Please take your time and choose the most suitable answer for each question.	− − − + + + − − + + − +
1. I was hardly aware of the real world.	
2. I had the feeling to be transported into another world.	
3. I only paid attention to the virtual environment and less to second thoughts.	
4. I wanted to explore the virtual environment.	
5. I felt nauseous.	
6. I was part of the virtual environment.	
7. I felt present in the virtual environment.	
8. I felt warm.	
9. I was curious about the virtual environment.	
10. At the end of the task, I felt I was coming back from another world.	
11. I had a headache.	
12. I forgot the time.	
13. The virtual environment drew my attention.	
14. I felt immersed in the virtual environment.	
15. At the start of the task, I felt like stepping into another world.	

Note: The questionnaire consisted of five scales: engagement (items 4,9,13); transportation (items 2,10,15); flow (items 1,3,12); presence (items 6,7,14); negative effects (items 5,8,11).

Appendix B. Questionnaire regarding preference for one user interface over another translated in English. Note that the results of this study were obtained with the original Dutch version.

Through this questionnaire, we want to ask you which user interface you prefer. Please take your time and choose the most suitable answer for each question.	Head-mounted display	Computer monitor	Both
1. I was motivated to perform the VR-task			
2. I enjoyed the VR-task.			
3. I would like to do the VR-task again.			
4. I would like to do the VR-task at home.			
5. I would like to do the VR-task regularly.			