Computers in Human Behavior Reports 4 (2021) 100141

Contents lists available at ScienceDirect

Computers in Human Behavior Reports

journal homepage: www.sciencedirect.com/journal/computers-in-human-behavior-reports

Acceptance of immersive head-mounted display virtual reality in stroke patients

Julian Specht [a,*], Helene Schroeder [b], Karsten Krakow, Prof. Dr. [c], Günter Meinhardt, Prof. Dr. [d], Barbara Stegmann [a], Bozana Meinhardt-Injac, Prof. Dr. [b]

[a] Department of Applied Psychology, SRH University of Applied Sciences Heidelberg, Germany
[b] Department of Psychology, Catholic University of Applied Sciences Berlin, Germany
[c] Asklepios Neurologische Klinik Falkenstein, Königstein, Germany
[d] Department of Psychology, Johannes Gutenberg-Universität Mainz, Germany

ARTICLE INFO

Keywords:
Virtual reality (VR)
Head mounted display-VR (HMD-VR)
Stroke
Technology acceptance
Attitudes
Cybersickness
User experience

ABSTRACT

Background: As virtual reality (VR) has become more accessible, it has increasingly come into focus for clinical application. Therapy with VR shows potential as an engaging, effective, and economic way to improve cognitive abilities following a stroke.
Objectives: While application of VR in clinical settings appears promising, its widespread use will crucially depend on acceptance of immersive head mounted display virtual reality (HMD-VR) systems when used in different patient groups. This study aimed to investigate acceptance of HMD-VR in stroke patients compared to a control group of healthy age-matched adults.
Methods: The attitude towards HMD-VR, as one of the most important predictors of technology acceptance within the technology acceptance model framework, was assessed in 20 stroke patients and 20 age-matched healthy adults. Further HMD-VR acceptance related measures concerned self-reported user experience, computer self-efficacy and cybersickness (see also Huygelier et al., 2019). Additionally, performance measures for memory span and speed were recorded in two VR-tasks.
Results: Both groups showed positive attitude above the neutral point of the scale and reported positive user experience in the VR-setting. Self-reported cybersickness was at modest levels and comparable in both groups. Controls had higher and more homogeneous scores in user experience, and performed notably faster in the VR-task while there were no significant differences in memory-span.
Conclusion: The study results suggest that treatment provided by immersive HMD-VR is tolerated by older adults, including those who have had strokes. This was the case without prior acquaintance with the VR-device or -software, and it was neither hindered by negative attitudes towards VR, nor cybersickness.

1. Introduction

Stroke is one of the most common neurological diseases leading to long-term disability in adults (Wilkins et al., 2017) and the second single leading cause of death in Europe. This includes an annual stroke-related death rate of 440 000 (OECD, 2016) amongst an incidence of 1.1 million people who suffer a stroke each year in Europe (Béjot et al., 2016). Stroke is not only a traumatic event for patients and their relatives, but also a significant economic burden; in 2017, the costs associated with stroke were estimated at €60 billion in the European Union (Luengo-Fernandez et al., 2019). This cost will further increase over the next decades as Europe faces the burden of demographic changes. The financial impact includes not only the direct costs of acute treatment, but also the subsequent burden on society of lost employment and the expense related to the need for extensive, multi-faceted rehabilitation and supportive care (Wilkins et al., 2017).

Current approaches to rehabilitation for different expressions of stroke injury vary widely (Pollock et al., 2014; Bunketorp-Käll et al., 2017). Recently, treatment strategies have been complemented by a new technology, virtual reality (VR). VR may be defined as images and sounds created by a computer that seem almost real to a user, one who can interact with them by using sensors (Oxford Advanced Learners

* Corresponding author.
E-mail address: julian.specht.extern@srh.de (J. Specht).

https://doi.org/10.1016/j.chbr.2021.100141
Received 27 January 2021; Received in revised form 9 August 2021; Accepted 7 September 2021
Available online 13 September 2021

Dictionary, 2020). It can be experienced via computers or screens (semi-immersive) or purpose-made VR-headsets, generating a fully immersive, three-dimensional and simulated world (immersive VR) by placing the device on the user's head in front of the eyes so that the field of vision is completely occupied by the VR device lenses (head-mounted display, HMD).

While some progress has been made in understanding the efficiency of the VR application in therapy and rehabilitation with diverse clinical groups (e.g., De Luca et al., 2017), adoption and use of VR technology in clinical and ambulatory settings is limited. Research on the acceptance of immersive VR with HMD-VR in clinical samples is sparse. According to the technology acceptance model (TAM) positive attitudes toward a new technology are highly predictive for technology adoption and use (Adams et al., 1992; Chen & Chan, 2011; Davis et al., 1989; Venkatesh & Morris, 2000; Venkatesh, 2000a, 2000b). However, different contextual (perceived support; social group processes) and individual factors such as personality (e.g., computer self-efficacy - control beliefs regarding individual ability to use a system), and demographics (e.g., age) have been identified as important variables that determine individuals' perception and beliefs (Venkatesh & Davis, 2000; Venkatesh, 2000a, 2000b). Huygelier et al. (2019) used the TAM model as a framework to study acceptance of HMD-VR in the population of older adults. In accordance with this model their results revealed that certain characteristics of older individuals such as age and computer proficiency predicted initial attitudes towards HMD-VR. The study also showed that older adults without prior experience with HMD-VR had a more positive attitude towards this new technology after a first exposure to HMD-VR. The finding that attitudes towards HMD-VR can be improved through exposure has important implications for the development of HMD-VR health applications.

At the time there is the need to expand the evidence concerning acceptance and user experience in VR by including the group of post-stroke patients. While a stroke can occur at any age, the mean age is 73 years in Europe, with an interquartile range of 62–81 years (Busch et al., 2013). This cumulation in older age makes it necessary to explore the factors contributing to technology acceptance. In a recent study by Spreij et al. (2020) it was shown that immersive VR user interfaces are feasible for use in stroke patients, and that immersion has positive effects on engagement, transportation, flow, and presence experience. Importantly, neither the severity of injury by stroke, nor cognitive functioning and time post stroke onset had an effect on the feasibility and user-experience. Similar findings were reported by Huygelier et al. (2020) in a small sample of stroke patients. Interestingly, neither cybersickness predicted user-experience in stroke patients, nor were symptoms of cybersickness stronger in this group as compared to healthy controls. This was revealed by several earlier studies. For example, in a study by Kang et al. (2008) it was shown that post-stroke patients are equally sensitive to cybersickness due to HMD-VR exposure as an age-matched control group. Simone et al. (2006) revealed that objective performance in an HMD-VR driver simulation in a group of post-stroke patients and controls was not associated with subjective comfort level. These studies addressed some important factors regarding employment of VR in post-stroke therapy, but they did not show how stroke patients experience VR, or reveal their acceptance of immersive HMD-VR.

The goal of the present study was to investigate acceptance of HMD-VR in stroke patients as compared to age-matched healthy subjects. To this end, we measured attitudes towards HMD-VR in both groups as the most important predictor of technology acceptance. To our best knowledge, attitudes toward HMD-VR in stroke patients were out of scope of the previous studies. Additionally, we assessed technical self-efficacy, cybersickness, user experience, as further potentially important predictors of technology acceptance within the TAM framework (see also Huygelier et al., 2019). Speed and memory span in VR-tasks were recorded as control measures for cognitive performance.

2. Methods

2.1. Participants

Patients were screened and recruited for enrollment from a neuro-rehabilitation clinic in the Frankfurt am Main (Germany) area. From a total of 48 patients, 26 were excluded due to low cognitive performance (MMSE <24), and two others were excluded due to hemianopsia (0 dropouts). The remaining 20 adult stroke patients (17 male, with a mean age of 68.3 years; SD = 14,5) comprised the experimental group. The mean duration between stroke and assessment was 17.6 weeks (SD = 38; range 2.5–175.5 weeks). Stroke locations with respect to hemisphere varied - right (10), left (7), or both (3) hemispheres. Residual symptoms in the 20 patients included neglect (n=7), aphasia (n=2) and ataxia (n=5). The MMSE test score was ≥26 (test cut-off for mild cognitive impairment) in 18 patients; one had a score of 24 and one of 25 indicating mild cognitive impairment. All recruited patients had already overcome the first early rehabilitation stage, including patients who were predominantly clear-minded, but still in need of nursing assistance for performing many activities of daily living. Patients had to be able to sit through the session (in a wheel-chair) and the functionality of one arm had to be given, to be able to operate the VR controller. Table 1 provides other characteristics of the patient sample regarding limb function and NIHSS.

Additionally, 20 age-matched healthy adults (8 male, with a mean age of 67.5; SD = 4.6) were recruited via a university database in the region of Mainz (Germany) (control group). The age of one participant was estimated since that person left one questionnaire incomplete. All participants in the control group lived an active and independent life. The educational level for this group was generally high: Eight participants had university degrees, eleven had high school degrees or equivalent and only one person had no degrees. All participants had normal or corrected to normal vision and none of them had acute health problems. Also, no one in the control group had any previous experience with VR-technology.

The study was conducted between February and July 2020. The study was conducted in accordance with the Declaration of Helsinki (2008), Good Clinical Practice guidelines, and local regulatory requirements. The study received approval by the Ethics Committee of the regional Medical Association (Hessen (Germany)). Participants gave their informed consent to the evaluation prior to enrolment.

2.2. Procedure

Patients were recruited from the neurological rehabilitation clinic Asklepios Neurologische Klinik Falkenstein by a neuropsychologist from the research team, healthy participants were assessed by trained university staff. They were asked to participate in the study, explaining its benefits, duration and demands on participants' time and commitment. During participant enrolment, patients completed the MMSE, NIHSS and MI. The intervention was conducted by a trained team of in-house therapists in the clinic and trained university staff for the control group.

Before the VR-session they completed computer self-efficacy and attitude towards HMD-VR questionnaires. After the VR-session they

Table 1
MI (motricity index) is an instrument measuring limb strength (range 1-100, 1 being the lowest score). NIHSS (NIH Stroke Scale) is a scale indicating the severity of a stroke (0–42: 1–4 = minor stroke, 5–15 moderate stroke, 16–20 = moderate to severe stroke, 21–42 = severe stroke).

	Mean	SD	Min	Max
MI left arm	73.5	35.9	1	100
MI right arm	78.25	32.8	26	100
MI left leg	84.7	26.65	26	100
MI right leg	85.2	23.7	34	100
NIHSS	5.2	3.33	0	15

completed user experience and self-reported cybersickness questionnaires. They were instructed how to use the VR setup and went through a tutorial. After the tutorial, participants were given the possibility to ask questions before starting two VR-tasks (memory task and recycling task – see Supplementary Material for more information). All questionnaires are provided in the supplementary material. The duration of the VR experience ranged from 20 min to 70 min (M=42, SD=14.93).

The VR experience was conceptualized and technically implemented by the neurorehabilitation company living brain (Heidelberg, Germany) together with clinical, neuropsychological and medical neurorehabilitation specialists.

2.3. Measures

The applied questionnaires for measuring HMD-VR acceptance related variables were used according to Huygelier et al., 2019, who investigated the acceptance towards immersive HMD-VR in healthy older adults. The items were translated into German. Items of existing questionnaires were chosen, combined, and extended. Four variables, (1) attitude towards HMD-VR, (2) computer self-efficacy, (3) user experience and (4) cybersickness, were measured. The scales for measuring the first 3 variables had several items consisting of 5-point Likert rating scales with 3 as the neutral point. Cybersickness was measured using the Simulator Sickness Questionnaire (SSQ, Kennedy et al., 1993). More information on the instruments is found in the Supplementary Materials.

The VR exposure consisted of a tutorial, the memory game (Fig. 1A) and the recycling game (see Fig. 1B). The tasks were devised to potentially train cognitive functions, such as short-term memory, attention, and executive functions. The VR application was developed for the HMD-VR device Oculus Quest. In the memory game, users were asked to encode the order of flashed symbols and reproduce the sequence by clicking the corresponding button in the right order (see Fig. 1A). The number of correctly recalled items was used as a measure of cognitive performance, i.e., of memory span. Additionally, we measured speed as a dependent measure in the recycling game. Here, participants were asked to sort different objects of daily life such as a glass bottle or carton of milk into the right receptacle (see Fig. 1B) as fast as possible. The recycling game stopped after 10 items were completed. Since several stroke patients had difficulty to reach 10 items, the game was stopped after 7 min. The time necessary to complete one item, indicated in minutes per item (min/item), was used as a speed measure. Note that speed as measured here reflects not only differences in information processing, but also in execution of action in VR, and is thus a more general measure of functioning, rather than a pure measure of cognitive performance. More details on the VR application and games are provided in the Supplementary Materials. For the HMD-VR acceptance related variables only one missing observation was registered, concerning cybersickness in the stroke group. This value was replaced by the group median. Further missing data concerned the two VR performance measures. Three missing values occurred in variable memory and five in variable speed. In the latter, speed data were missing if no item was completed. Except one missing age entry in the control group (see 2.1 Participants), missing data only concerned the stroke patient group.

2.4. Power and statistical testing

Comparing controls and stroke patients in performance and HMD-VR acceptance related measures involves group mean comparisons from two independent samples. Power calculations (power analysis module of

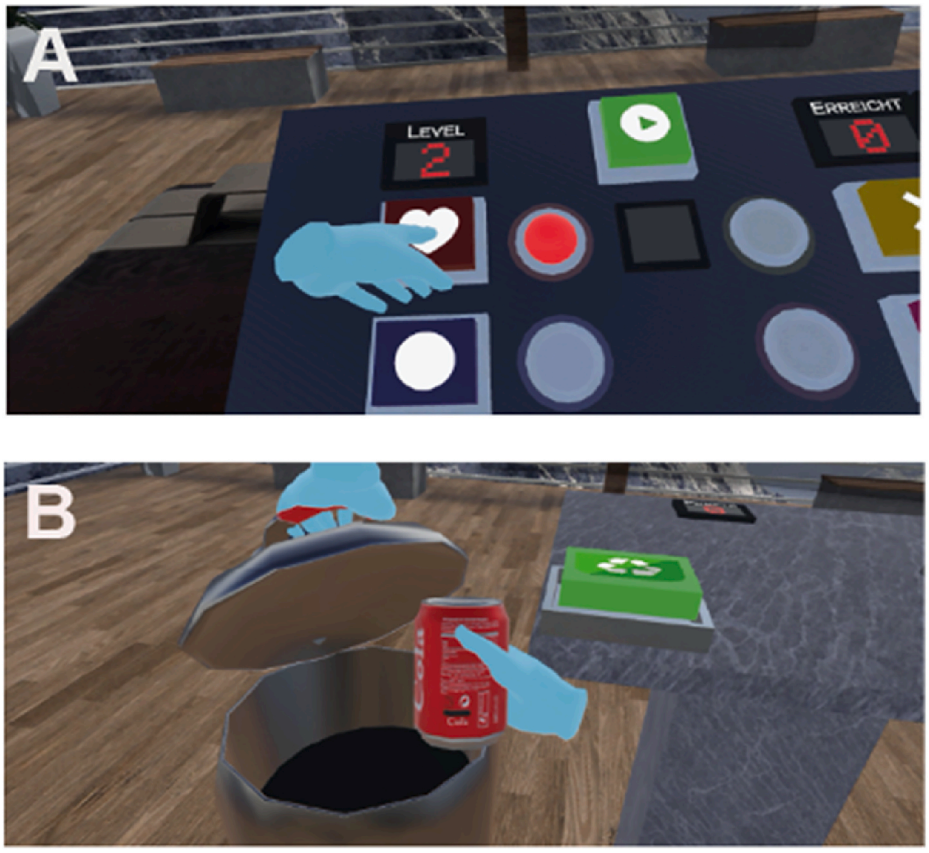

Fig. 1. A: Memory game: Lamps with heart and circle symbols were alternating lighting up in sequences. Participants encoded and repeated each sequence by sequential button presses. B: Recycling game: Cans, milk boxes or glass bottles were presented. Participants decided whether they had to be fed into the garbage press and then to be sorted into the correct bin.

Statistica 13.3, Tibco Soft Inc.) showed that our sample size of $N = 20$ participants per group was sufficient to detect a standardized estimated population mean difference of $d = 0.8$, which is considered as a large effect size (Cohen, 1988, p. 40), with a power of 1- $\beta = 0.7$ for $\alpha = 0.05$. A power goal of 1- $\beta > 0.8$ is reached for effect sizes $d > 0.91$. This shows that our quite small samples sizes constrain detection of between group effects at a conventional power goal of 0.8 (see, e.g., Fritz & MacKinnon, 2007) to large effects in the sense of Cohen. We report group mean effects together with effect size and achieved power level and also include Bayes factors (BF, Kass & Raftery, 1995) to aid judgement which of the hypotheses, H0 or H1, is more likely given the observed mean difference. To guide interpretation, we refer to the taxonomy of Kass and Raftery (1995), who consider BF in intervals of evidence strength: *weak* [1, 3.2], *substantial* (3.2,10], *strong* (10,100], and *decisive* (BF > 100). BFs are reported as the posterior odds in favor of H1 (BF_{10}), and in favor of H0 (BF_{01}), whereby equal a-priori probability of both alternatives is assumed. BFs were calculated with the Bayes Factors package of *R* (R Core Team, 2016).

To reveal the joint effect of all four HMD-VR acceptance related measures for separating control and stroke patient group we used discriminant function analysis, which is equivalent to a linear multiple regression model for predicting a dichotomous criterion variable if only 2 groups are to be separated (see Results section). In multiple regression, the effect size index $f^2 = \frac{R^2}{1-R^2}$ indicates a large effect for $f^2 > 0.35$, which corresponds to $R^2 > 0.26$ or $R > 0.51$. (see Cohen, 1988, pp. 413). To identify large effects at a power goal of 0.8, $N = 38$ is necessary for 3 predictors and $N = 42$ for 4 predictors in the equation, which means that our sample size was sufficient to substantiate the frequentist significance test for multiple correlations *R* beyond 0.5 with sufficient power.

3. Results

3.1. Descriptive statistics

First, we reviewed the data for outliers and assessed distribution properties. Basic descriptive statistics and results of checking distributions for normality are shown in Table 2. Box plots of the data are shown in Fig. 2. Outliers based on the Tukey criterion (data outside the interval [Q_{25} – 1.5 IQR, Q_{75} + 1.5 IQR], Tukey, 1977) were observed in the control variables memory span and speed (see Fig. 2). These values were kept in the sample, and testing across groups was done with nonparametric procedures. In the group of HMD-VR acceptance related measures, outliers were found in the variable cybersickness. Since these values concerned both the stroke and the control group, they were also kept in the sample. Normality was hurt for memory span in the control group and for speed in both groups.

In the four HMD-VR acceptance variables normality was hurt only for cybersickness, but for both the control and the stroke group. Since attitude, computer self-efficacy and user experience were measured on 5-point Likert scales we tested whether the scale means deviated from the neutral point, E_0 =3.0, with a one - sample *t*-test, including Bayes Factors for H1 over H0 (BF_{10}) and vice versa (BF_{01}) (see Table 4). *Attitude* and *user-experience* measures of both controls and stroke patients were significantly better than neutral (decisive evidence for H1).

Average *computer self-efficacy* complied with the neutral point of the scale for stroke patients (strong evidence for H0) while slightly better values than neutral were obtained from controls (significant, weak evidence for H1).

3.2. Group differences

3.2.1. Pairwise comparisons

The data shown in Fig. 2 indicate potentially unequal group variances in several variables. Since homogeneity of variances is one prerequisite of testing group-mean differences, we applied Levene test (see Table 3). Results indicated violations of variance homogeneity for age, memory span, and user experience, whereby the stroke group data showed larger variance compared to the controls (see SD listed in Table 2). We applied *t*-tests for independent groups and complemented results with nonparametric Mann-Whitney-*U*-test if *t*-test prerequisites were violated. Results are shown in Table 5.

Stroke patients and controls differed in average age about less than 1 year. Bayes factors showed weak at the very edge to substantial proof in favor of the Null hypothesis.

Since the stroke patients group showed notably larger age variance, we confirmed non-significant results with *U* test ($U = 184$, $z = -0.42$, $P = 0.675$). Further, there was evidence for equal performance of both groups in memory span (0.2 items difference, $BF_{01} \approx 3$, $U = 149.5$, $z = 0.61$, $P = 0.542$), while there was a pronounced difference in speed. The mean difference had large effect size, was strongly significant and at least substantial evidence in favor of H1, since BF_{10} was computable only after replacing all 5 missing values in the stroke group with the median, which might overestimate performance in places ($BF_{10} = 5.25$). Since speed distributions suffered from outliers, were not normal and variances inhomogeneous, we confirmed significant speed differences with nonparametric testing ($U = 36$, $z = -3.78$, $P < 0.001$). Comparing speed at the median (stroke patients: $Q_{50} = 0.78$ min/item; controls: $Q_{50} = 0.36$ min/item) showed that stroke patients needed about 25 s more time to complete one item.

In the four HMD-VR acceptance related measures we found weak evidence for same attitude and same self-efficacy in both groups. The control group showed significantly higher values in user-experience (large effect size, substantial evidence for H1), which was also confirmed in non-parametric testing ($U = 117.5$, $z = 2.22$, $P = 0.027$). Reported cybersickness was apparently not different in both groups (substantial evidence for H0; $U = 188.5$, $z = -0.30$, $P = 0.766$). To keep a family-wise α rate in multiple pairwise tests of the same difference hypothesis, a Bonferroni-corrected test alpha of $\alpha' = \frac{\alpha}{k}$, k the number of comparisons, can be chosen. Here, with $k = 6$ tests (2 performance tests and 4 acceptance tests), $\alpha' = 0.0083$ resulted, which means that significance in the frequentist test of H0 is save on a family-wise $\alpha = 0.05$ for speed but not for user-experience if the more conservative non-

Table 2
Valid observations (N), mean (M) and standard deviation (SD) in controls and stroke patients for age, the two VR task measures and the four HMD-VR acceptance related measures. Further, the Shapiro-Wilk (1965) W statistic (W_S) and its probability under the Null hypothesis (normal distribution) are listed. Violations of normality are indicated by an asterisk (*). Speed is indicated in minutes per item and memory span reflects the number of recalled items.

	Controls					Stroke Patients				
	N	M	SD	W_S	P	N	M	SD	W_S	P
Age	20	67.5	4.59	0.941	0.248	20	68.3	14.6	0.965	0.647
Memory-Span	20	3.55	1.36	0.601	<0.01*	17	3.35	2.15	0.927	0.197
Speed	20	0.42	0.22	0.664	<0.01*	15	1.17	1.20	0.652	<0.01*
Attitude	20	3.68	0.52	0.971	0.769	20	3.86	0.58	0.945	0.293
Self-Efficacy	20	3.25	0.49	0.970	0.755	20	3.02	0.57	0.960	0.547
User-Experience	20	4.10	0.28	0.933	0.173	20	3.66	0.61	0.929	0.147
Cybersickness	20	9.16	11.98	0.765	<0.01*	20	10.1	12.4	0.787	<0.01*

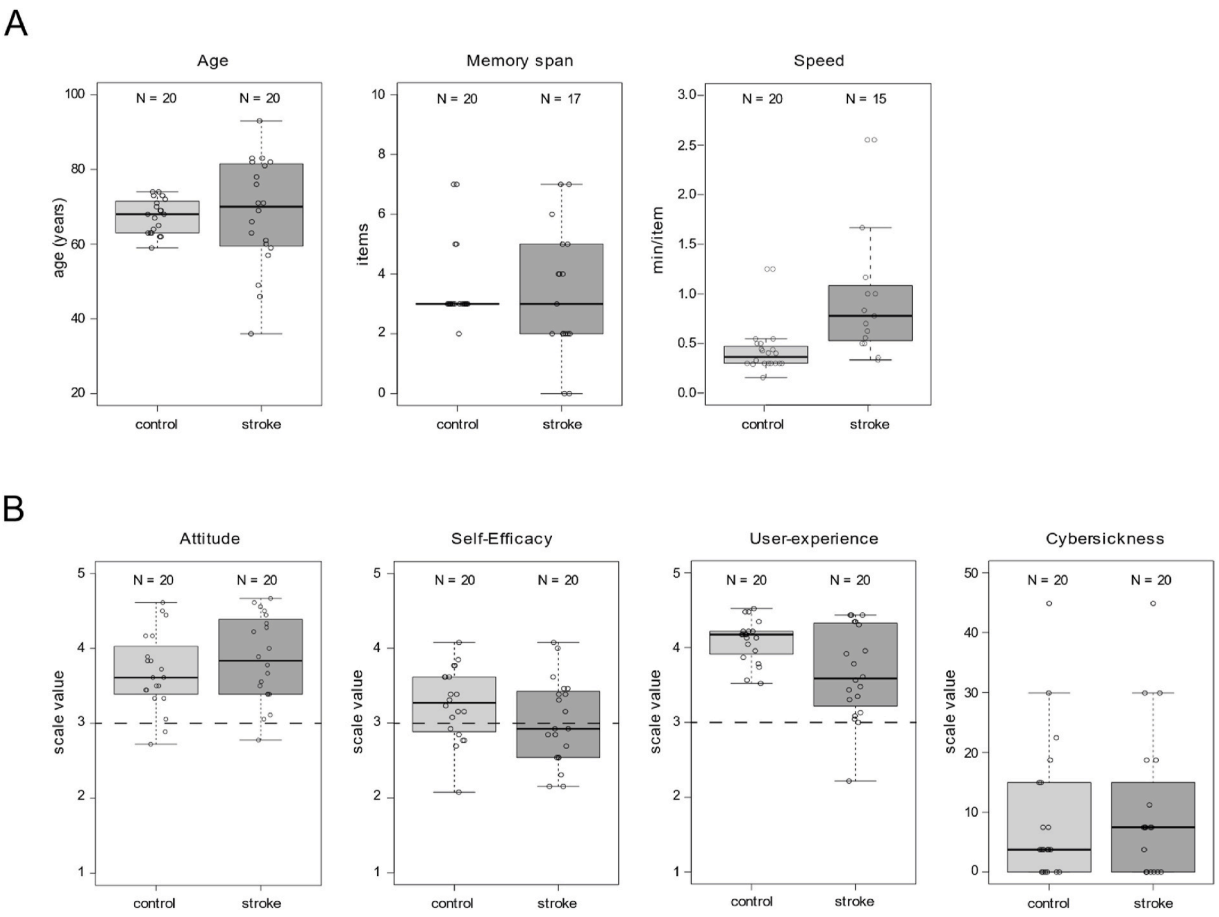

Fig. 2. Box plots for age, VR performance measures memory span and speed (A), and for the four *HMD-VR* acceptance related measures (B). Box plots indicate the 1st (Q_{25}) and 3rd (Q_{75}) quartiles (box), and the median (Q_{50}) with a solid black line, while the whiskers indicate real data points which fall within the interval [$Q_{25} - 1.5$ IQR, $Q_{75} + 1.5$ IQR]. Box plots are overlaid by the data with a jitter technique to illustrate their distribution across the scale. The dashed lines indicate the neutral point of the scales for attitude, self, efficacy and user-experience.

Table 3
Results of one-sample t-tests for proving compliance of sample means with the neutral point of the Likert scale for variables attitude, self-efficacy, and user-experience. The table lists deviation of sample mean and neutral point, t-statistic (all df = 19), probability of observed or larger difference given the Null hypothesis (P), and Bayes factor for proof in favor of H1 (BF_{10}) and H0 (BF_{01}).

	Controls					Stroke Patients				
	$M - E_0$	t(19)	P	BF_{10}	BF_{01}	$M - E_0$	t(19)	P	BF_{10}	BF_{01}
Attitude	0.68	5.90	<0.001	>100	<0.01	0.86	6.64	<0.001	>100	<0.01
Self-Efficacy	0.25	2.34	0.031	2.05	0.49	0.02	0.12	0.904	0.24	4.28
User-Experience	1.10	17.33	<0.001	>100	<0.01	0.66	4.80	<0.001	>100	<0.01

Table 4
Levene test results for the homogeneity of variances.

	MQ_A	MQ_e	F	P
Age	618	35.6	17.3	<0.001
Memory-Span	5.96	1.01	5.89	0.021
Speed	3.33	0.37	9.01	<0.01
Attitude	0.08	0.08	0.97	0.332
Self-Efficacy	0.07	0.08	0.89	0.356
User-Experience	0.81	0.07	11.54	0.002
Cybersickness	0.35	60.4	0.01	0.940

parametric testing results are considered.

3.2.2. Discriminant function analysis and case classification

Comparisons of group means indicated that only user experience differed among stroke patients and controls in the domain of HMD-VR acceptance related measures. However, statistical evidence in the frequentist test of H0 was weakened by multiple testing. To circumvent, we used discriminant function analysis (DFA) to decide on group separation with a single statistical test. Like in a multiple regression equation, the discriminant function is a linear combination of the variables while the variable weights are partial (i.e., corrected for the other predictors), indicating the unique contribution of each predictor to the discriminant function which maximizes the ratio γ of between-group to within-group variance. For the case of two groups and m variables, DFA is actually equivalent to a multiple regression on a dichotomous (0,1) criterion variable coding group assignment. Before the analysis, HMD-VR acceptance related measures were standardized across groups for convenience.

DFA returned a highly significant ratio of between-group to within-group variance with all four variables in the equation (Eigenvalue $\gamma = 0.435$, Wilks $\Lambda = 0.697$, $\chi^2 = 13.01$, df = 4, $P = 0.011$), with a ratio of

Table 5
Results of two-sample t-tests for comparison of stroke patients and control group. The table lists difference of means, its standard error, t-statistic, degrees of freedom, probability of observed or larger difference given the Null hypothesis (P), Cohen's effect size measure d, power 1-β at α = 0.05, and Bayes factor for proof in favor of H1 (BF_{10}) and H0 (BF_{01}).

	ΔM	$SE_{\Delta M}$	t	P	df	d	1-β	BF_{10}	BF_{01}
Age	−0.85	3.41	−0.25	0.805	38	0.08	0.06	0.32	3.16
Memory-Span	0.20	0.58	0.34	0.737	35	0.11	0.06	0.34	2.97
Speed	−0.75	0.27	−2.75	<0.01	33	0.93	0.76	5.25	0.19
Attitude	−0.18	0.17	−1.01	0.318	38	0.32	0.17	0.46	2.16
Self-Efficacy	0.24	0.17	1.43	0.161	38	0.45	0.29	0.69	1.45
User-Experience	0.44	0.15	2.89	0.006	38	0.91	0.80	7.09	0.14
Cybersickness	−0.94	3.86	−0.24	0.810	38	0.08	0.06	0.32	3.16

between-group to total variance of $R^2 = 1 - \Lambda = 0.303$. Evaluation of discriminant weights and utility ΔR^2, i.e., change in R^2 due to removing the variable, showed that cybersickness did not contribute to group separation at all, since removing it did practically not change R^2 ($\Delta R^2 = 0.03\%$), while the other variables contributed notably to group separation (see upper part of Table 6). Therefore, a three-variable solution (Eigenvalue $\gamma = 0.435$, Wilks $\Lambda = 0.697$, $\chi^2 = 13.17$, df = 3, $P = 0.004$) was better suited. Since we have a two groups case for which DFA yields a single discriminant function y, its values for each case i are obtained by $y_i = \sum_j b_j z_{ij}$, and group separation can be evaluated on y with a t-test. Results showed a highly significant difference of means ($\mu_1 = 0.64$, $\mu_0 = -0.64$, $t(38) = 4.06$, $P < 0.001$, $BF_{01} > 100$) with large effect size ($d = 1.28$) and power 1 - β = 0.98 at α = 0.05. Since DFA requires multivariate normally distributed measurement variables in each group we calculated the squared Mahalanobis distances from the group centroids, paired them with the corresponding quantiles from the χ^2 distribution with 3 degrees of freedom and performed the Q-Q plot correlation coefficient test for multivariate normality (Filliben, 1975; Looney & Gulledge, 1985). Results indicated no violations of multivariate normality with $r_{QQ} = 0.973$ for controls and $r_{QQ} = 0.987$ for stroke patients, which were both larger than critical Q-Q correlations that r_{QQ} should not fall below at a sample size of $N = 20$ ($\alpha = 0.01$, $r_{crit} = 0.927$; $\alpha = 0.05$, $r_{crit} = 0.951$; $\alpha = 0.1$, $r_{crit} = 0.960$; see Johnson & Wichern, 2002, p. 182). Checking normality of the y - scores also indicated no violations for controls or stroke patients (Shapiro-Wilk test, W_S (controls) = 0.968, $P = 0.719$; W_S (stroke patients) = 0.925, $P = 0.126$).

DFA allows us to check how well individual case classification works with the discriminant function y. We used the maximum likelihood ratio criterion for maximizing the proportion of correct classifications with the y - scores. This criterion is $y_0 = 0$ if the a-priori probabilities of both groups are equal, which was assumed (see Appendix). Classifying cases into the stroke group for $y > 0$ and into the control group otherwise led to 28 hits from 40 classifications. Inserting into the binomial distribution showed that the probability for 28 or more hits from 40 trials if the groups are equally likely ($p = q = 0.5$) is $P = 0.003$, which means that the DFA case classification based on attitude, computer self-efficacy and user-experience was significantly better than guessing.

Table 6
Results of discriminant function analysis for all four HMD-VR acceptance related variables (upper part) and for a 3 variables solution omitting cybersickness (lower part). The table lists the unstandardized discriminant weights b, Wilks Λ for a solution without the variable, F-value for the corresponding change in explained variance, probability of observed or larger change given the Null hypothesis (P) and utility (ΔR^2).

	b	*Λ*	F_{rem}	*P*	ΔR^2*(%)*
Attitude	0.652	0.777	4.024	0.052	8.01%
Self-Efficacy	−0.464	0.743	2.309	0.137	4.60%
User-Experience	−1.021	0.907	10.573	0.002	21.05%
Cybersickness	−0.038	0.697	0.013	0.908	0.03%
Attitude	0.663	0.787	4.659	0.038	9.02%
Self-Efficacy	−0.460	0.743	2.366	0.133	4.58%
User-Experience	−1.016	0.911	11.033	0.002	21.36%

4. Discussion

The primary objective of this study was to determine acceptance of immersive HMD-VR in a population of stroke patients compared to healthy age-matched controls. As a major result we found that measures of HMD-VR acceptance, as used by Huygelier et al. (2019), were positive above neutral in both the control and the stroke patient group, or complied well with neutral point of the scale (computer self-efficacy). Symptoms of cybersickness were of low intensity and rarely occurring. Albeit controls showed higher values in user experience, the individual values of the stroke patients were all above the neutral point of the scale, except one case (see Fig. 2). Hence, results for cybersickness and user experience support a generally positive experience with the VR setting for stroke patients as well. This is surprising, given the longer times to complete items and the fact that 5 stroke patients were unable to complete items in the given time in one specific exercise (recycling game). Generally, the feature characteristic in all four HMD-VR acceptance measures did not give rise to conclude negative cognitions about VR, or negative experiences made with the VR environment.

There were group differences, though, revealed by DFA and pairwise testing. Both procedures corroborated that controls had a more positive user experience in the VR setting than stroke patients, while there were no differences in cybersickness. Even with optimal weighting, the contribution of cybersickness to group separation was practically absent.

Within the framework of the TAM (Davis et al., 1989; Adams et al., 1992; Venkatesh & Morris, 2000; Venkatesh, 2000a, 2000b), attitude towards HMD-VR is the most important predictor of user acceptance. On the original scale we found no evidence for different HMD-VR related attitude among both groups. However, also evidence for same attitudes was weak in view of Bayes factor results (see Table 5). DFA results indicated that, if the unique contributions of each variable to group separation are considered, a specific portion of attitude variance can be isolated which significantly contributes to group separation. As indicated by the opposite sign of attitude and user experience weights, DFA exploits the slightly more positive attitude of stroke patients to improve the ratio of between to within group variance. In sum, results regarding attitude measured before VR exposure showed no evidence for more negative attitudes towards HMD-VR in stroke patients compared to age-matched healthy controls, but revealed a generally positive tuning toward VR technology.

Based on the findings, we conclude that the expected acceptance of immersive VR with HMD is high in both control and patient groups. This conclusion is in line with a recent study suggesting high acceptance of immersive HMD-VR. However, in regards specifically to an elderly population only (Huygelier et al., 2019), the present study provides support for acceptance in stroke patients.

Several recent studies provide evidence on feasibility of HMD-VR in stroke patients (Huygelier et al., 2020; Spreij et al., 2020, Salisbury et al., 2020; Hak Lee et al., 2019, Weber et al., 2019). For example, Spreij et al. (2020) evaluated the feasibility of HMD-VR for cognitive

assessment in stroke patients in contrast to traditional PC testing. In this study, the feasibility was operationalized as a complex measure of 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 trials), the total time needed to complete the VR-task, and the total number of products found on a shopping list (see also Spreij et al., 2020). The completion rate was notably higher in healthy controls (90.9%) than in patients (83.8%). Patients needed more time to complete the task and identified less products from the shopping list.

In this study, we did not have any dropout, but 24 patients were excluded from participation due to low cognitive status on MMSE. And yet, we still found that time needed to complete the task was significantly and notably longer in patients than in the control group. However, there was no evidence for differences in memory span (memory game in VR), while the variability among participants was considerably larger for stroke patients compared to controls. Together, these findings suggest that HMD-VR is generally feasible for use in stroke rehabilitation, even with patients that are more severely affected after stroke (see also Huygelier et al., 2020). Patient-tailored design of HMD-VR games could further increase feasibility of use in the clinical population (see also Huygelier et al., 2020).

In contrast to previous studies, the focus of this work was directed toward acceptance of HMD-VR. As a theoretical framework, we used the technology acceptance model (TAM), which has been shown to be highly predictive of technology adoption and use (Adams et al., 1992; Chen & Chan, 2011; Davis et al., 1989; Venkatesh & Morris, 2000; Venkatesh, 2000a, 2000b). Our study design aligns with that used by Huygelier et al. (2019), where the TAM model has been used as a framework to study acceptance of HMD-VR in the population of older adults. Initial attitudes of older adults, measured prior to VR-exposure, were more positive for individuals that were younger and had higher computer proficiency. It is important to note that attitudes are the most important predictor for the use of new technologies, according to the TAM model. However, it is also important to keep in mind that attitudes can be changed through experience and contextual influences (i.e., social dynamic in a group; support etc). The study by Huygelier et al. (2019) demonstrated that attitudes toward HMD-VR changed from initially neutral to positive after VR-exposure. We measured attitudes only before the VR-exposure, so it is noteworthy that we found above average positive attitudes in both groups (healthy controls and patients).

In accordance with the previous findings, the symptoms of cybersickness in the present study were weak, and no differences in cybersickness were found when compared to healthy controls (Huygelier et al., 2020; Kang et al., 2008; Simone et al., 2006; Spreij et al., 2020). This is of particular importance as immersive VR and the experienced content is often associated with a higher risk of simulator sickness due to the strong grade of immersion (Kemeny et al., 2020), which could be a limitation for the use of immersive VR in different settings. The immersive VR experience in this study was specifically designed to reduce symptoms of simulation sickness, such as no abrupt movement of objects in VR, complete control over movements by the user, and realistic scaling of interactable objects.

Given its good acceptance and low side effects, immersive VR in stroke patients offers good options for the implementation of new forms of therapy leading to new treatment options, e.g., applying more elaborated serious games in daily therapy (Bartolomé et al., 2011). Further research should be conducted on the acceptance of immersive VR with HMDs in patients suffering neurological diseases, particularly concerning the constraints imposed by anti-ease of use, costs, physical and cognitive impairments, the specific features of a device and the interface complexity of the device or technology (Klimova & Poulova, 2018).

The generalizability of our results might be limited as the implemented VR application was only used in this study. Another limitation is that there was no non-VR-control group, which offers no insights into whether symptoms of cybersickness to the same extent may have existed previously. It would have been beneficial to test participants concerning their attitudes after the VR-session as well.

5. Conclusion

The results of this study imply that both healthy older adults and older age stroke patients with MMSE $\geq$24 and sufficient motor functionality of at least one arm to operate the VR controller were able to interact with a HMD-VR device and the corresponding immersive VR software, and that both groups accepted the technology, supporting the notion that immersive VR meets the requirements of a technology for use in these important populations. The results of this study also indicate high potential for use of HMD-VR in stroke rehabilitation, which is in line with previous research about this topic, not only in terms of therapeutic benefits, but also as VR is a promising technology to foster homecare (Brennan et al., 2013, pp. 599–602; Salisbury et al., 2020) and to shorten hospitalization. Considering that prices of powerful mobile HMD-devices are continuously decreasing, the technology is becoming more and more accessible. This study serves as a profound base for further research with a thorough investigation of the factors leading to a successful use of immersive HMD-VR in stroke patients or patients in neurological rehabilitation as a whole.

Geolocation information

Frankfurt am Main Area (Germany).

Funding details

The study has been funded by a Catholic University of Applied Sciences Berlin Internal Research Funding Grant Zi./Leo. 14102019 Akzeptanz von immersive head-mounted (HMD) Virtual Reality bei Patienten nach Schlaganfall given to Prof. Dr. Bozana Meinhardt-Injac.

Disclosure statement

The authors report no conflict of interest.

Author note

This study was supported by a grant from the Catholic University of Applied Sciences Berlin given to Prof. Dr. Bozana Meinhardt-Injac. We thank Robert C. Knowlton, MD (UCSF) for valuable suggestions and Marilena Marino and Katharina Raab for continued support of this project.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chbr.2021.100141.

Appendix. Maximum likelihood ratio criterion z_β for maximizing the proportion of correct classifications

According to Bayes' theorem, the a-posteriori probability for hypothesis H_i is

$$P(H_i|x) = \frac{P(x \cap H_i)}{P(x)} = \frac{p(x|H_i)P(H_i)}{P(x)}$$

with x an observation (measurement) and $P(H_i)$ the a-priori probability of hypothesis H_i.

For two hypotheses H_0 and H_1, we decide in favor of the alternative with larger a-posteriori probability:

$$q_{10}(x) = \frac{\frac{p(x|H_1)P(H_1)}{P(x)}}{\frac{p(x|H_0)P(H_0)}{P(x)}} = \frac{p(x|H_1)P(H_1)}{p(x|H_0)P(H_0)} > 1$$

"If $q_{10} > 1$ choose H_1, and H_0 otherwise." Note that this rule is equivalent to the likelihood-ratio rule:

$$lr_{10}(x) = \frac{p(x|H_1)}{p(x|H_0)} > \frac{P(H_0)}{P(H_1)}$$

"If $lr_{10} > \beta$ choose H_1, and H_0 otherwise", with $\beta = P(H_0)/P(H_1)$. Note that the likelihood-ratio criterion $\beta = P(H_0)/P(H_1)$ maximizes the percentage of correct decisions (see Green & Swets, 1988, p. 23).

Assume the likelihood functions for H_0 and H_1 are normal distributions with same variance, z is the standard value of x, and d is the standardized difference of means. Then, the likelihood ratio becomes

$$lr_{10}(z) = \frac{p(z|H_1)}{p(z|H_0)} = \frac{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(z-\frac{d}{2}\right)^2\right)}{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(z+\frac{d}{2}\right)^2\right)} = \frac{\exp\left(-\frac{1}{2}\left(z-\frac{d}{2}\right)^2\right)}{\exp\left(-\frac{1}{2}\left(z+\frac{d}{2}\right)^2\right)}$$

Taking the logarithm

$$\begin{aligned}\ln(lr_{10}(z)) &= \ln\left(\exp\left(-\frac{1}{2}\left(z-\frac{d}{2}\right)^2\right)\right) - \ln\left(\exp\left(-\frac{1}{2}\left(z+\frac{d}{2}\right)^2\right)\right) \\ &= \frac{1}{2}\left(z+\frac{d}{2}\right)^2 - \frac{1}{2}\left(z-\frac{d}{2}\right)^2 = zd\end{aligned}$$

If $P(H_0) = P(H_1)$, then $\beta = 1$. Therefore $\ln(lr_{10}) = \ln(\beta) = 0$, from which follows:

$$0 = z_\beta d \Rightarrow z_\beta = 0$$

References

Adams, D., Nelson, R., & Todd, P. (1992). Perceived usefulness, ease of use, and usage of information technology: A replication. *MIS Quarterly, 16*(2), 227–247. https://doi.org/10.2307/249577.

Bartolomé, N. A., Zorrilla, A. M., & Zapirain, B. G. (2011). July). Can game-based therapies be trusted? Is game-based education effective? A systematic review of the serious games for health and education. In *2011 16th international conference on computer games (CGAMES)* (pp. 275–282). IEEE. https://doi.org/10.1109/CGAMES.2011.6000353.

Béjot, Y., Bailly, H., Durier, J., & Giroud, M. (2016). Epidemiology of stroke in Europe and trends for the 21st century. *Presse Med, 45*(12 pt 2), e391–e398. https://doi.org/10.1016/j.lpm.2016.10.003.

Brennan, P. F., Smith, C. A., Ponto, K., Radwin, R., & Kreutz, K. (2013). *Envisioning the future of home care: Applications of immersive virtual reality*. IOS Press, Studies in Health Technology and Informatics. https://doi.org/10.3233/978-1-61499-289-9-599.

Busch, M. A., Schienkiewitz, A., Nowossadeck, E., & Gößwald, A. (2013). Prevalence of stroke in adults aged 40-79 years in Germany. In *Results of the German health inverview and examination survey for adults (DEGS1)* (Vol. 56, pp. 656–660). Bundesgesundheitsblatt: Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/s00103-012-1659-0. *2013*.

Chen, K., & Chan, A. H. S. (2011). A review of technology acceptance by older adults. *Gerontechnology, 10*(1), 1–12. https://doi.org/10.4017/gt.2011.10.01.006.00.

Cohen, J. C. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ, USA: Lawrence Erlbaum Publishers.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science, 35*(8), 982–1003. https://doi.org/10.1287/mnsc.35.8.982.

De Luca, R., Russo, M., Naro, A., Tomasello, P., Leonardi, S., Santamaria, F., Desireè, L., Bramanti, A., Silvestri, G., Bramanti, P., & Calabrò, R. S. (2017). Effects of virtual reality-based training with BTs-nirvana on functional recovery in stroke patients: Preliminary considerations. *International Journal of Neuroscience, 17*. https://doi.org/10.1080/00207454. 2017.1403915.

Filliben, J. J. (1975). The probability plot correlation coefficient test for normality. *Technometrics, 17*, 111–117. https://doi.org/10.1080/00401706.1975.10489279.

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science, 18*, 233–239. https://doi.org/10.1111/j.1467-9280.2007.01882.x.

Green, D. M., & Swets, J. A. (1988). *Signal detection theory and psychophysics*. Los Altos, California: Peninsula publishing.

Huygelier, H., Schraepen, B., Lafosse, C., Vaes, N., Schillebeeckx, F., Michiels, K., Note, E., Abeele, V. V., van Ee, R., & Gillebert, C. R. (2020). An immersive virtual reality game to train spatial attention orientation after stroke: A feasibility study. *Applied Neuropsychology: Adult*. https://doi.org/10.1080/23279095.2020.1821030.

Huygelier, H., Schraepen, B., van Ee, R., Abeele, V. V., & Gillebert, C. R. (2019). Acceptance of immersive head-mounted virtual reality in older adults. *Scientific Reports, 9*, 4519. https://doi.org/10.1038/s41598-019-41200-6.

Johnson, R. A., & Wichern, D. W. (2002). *Applied multivariate statistical analysis*. Upper Saddle River: Prentice-Hall Inc.. NJ07458.

Kang, Y. J., Ku, J., Han, K., Kim, S. I., Yu, T. W., Lee, J. H., & Park, C. I. (2008). Development and clinical trial of virtual reality-based cognitive assessment in people with stroke: Preliminary study. *CyberPsychology and Behavior, 11*, 329–339. https://doi.org/10.1089/cpb.2007.0116.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*, 773–795. https://doi.org/10.1080/01621459.1995.10476572.

Kemeny, A., Chardonnet, J.-R., & Colombet, F. (2020). *Getting rid of cybersickness in virtual reality, augemented reality and simulators*. Springer Nature.

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*, 203–220. https://doi.org/10.1207/s15327108ijap0303_3.

Klimova, B., & Poulova, P. (2018). Older people and technology acceptance. In J. Zhou, & G. Salvendy (Eds.), *Human aspects of IT for the aged population. Acceptance, communication and participation. ITAP 2018. Lecture notes in computer science* (Vol. 10926). Cham: Springer. https://doi.org/10.1007/978-3-319-92034-4_7.

Looney, S. W., & Gulledge, T. R. (1985). Use of the correlation coefficient with normal probability plots. *The American Statistican, 39*, 75–79. https://doi.org/10.1080/00031305.1985.10479395.

Luengo-Fernandez, R., Violato, M., Candio, P., & Leal, J. (2019). Economic burden of stroke across Europe: A population-based cost analysis. *European Stroke Journal*. https://doi.org/10.1177/2396987319883160.

OECD. (2016). Mortality from heart disease and stroke. In *Health at a glance: Europe 2016: State of health in the EU cycle*. Paris: OECD Publishing.

Oxford Advanced Learners Dictionary. (2020), 19.12, retrieved from URL (https://www.oxfordlearnersdictionaries.com/us/definition/english/virtual-reality .

Pollock, A., Baer, G., Campbell, P., Choo, P. L., Forster, A., Morris, J., Pomeroy, V. M., & Langhorne, P. (2014). Physical rehabilitation approaches fort the recovery of function and mobility following stroke. *Cochrane Database of Systematic Reviews*, (4) https://doi.org/10.1002/14651858.CD001920.pub3. Art. No. CD001920.

R Core Team. (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://cran.microsoft.com/snapshot/2014-09-08/web/packages/dplR/vignettes/xdate-dplR.pdf.

Salisbury, J. P., Aronson, T. M., & Simon, T. J. (2020). At-home self-administration of an immersive virtual reality therapeutic game for post-stroke upper limb rehabilitation. In *Extended abstracts of the 2020 annual symposium on computer-human interaction in play* (pp. 114–121). https://doi.org/10.1145/3383668.3419935. Association for Computing Machinery.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality. *Biometrica*, 591–611.

Simone, L. K., Schultheis, M. T., Rebimbas, J., & Millis, S. R. (2006). Head-mounted displays for clinical virtual reality applications: Pitfalls in understanding user behavior while using technology. *CyberPsychology and Behavior, 9*, 591–602. https://doi.org/10.1089/cpb.2006.9.591.

Spreij, L. A., Visser-Meily, J. M. A., Sibbel, J., Gosselt, I. K., & Nijboer, T. C. W. (2020). Feasibility and user-experience of virtual reality in neuropsychological assessment following stroke. *Neuropsychological Rehabilitation*, 1–21. https://doi.org/10.1080/09602011.2020.1831935, 0.

Tukey, J. W. (1977). *Exploratory data analysis*. London, Amsterdam, Don Mills: Addison-Wesley Inc.

Venkatesh, V., & Davis, F. D.((2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(No. 2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926.

Venkatesh, V., & Morris, M. G. (2000). Age differences in technology adoption decisions: Implications for a changing workforce. *Personnel Psychology, 53*, 375–403. https://doi.org/10.1111/j.1744-6570.2000.tb00206.x.

Wilkins, E., Wilson, L., Wickramasinghe, K., Bhatnagar, P., Leal, J., Luengo-Fernandez, R., Burns, R., Rayner, M., & Townsend, N. (2017). *European cardiovascular disease statistics 2017*. Brussels: European Heart Network.

Bunketorp-Käll, L., Lundgren-Nilsson, A., Samuelsson, H., Pekny, T., Blomvé, K., Pekna, M., Pekny, M., Blomstrand, C., & Nilsson, M. (2017). Long-Term Improvements After Multimodal Rehabilitation in Late Phase After Stroke. A Randomized Controlled Trial. *Stroke, 2017*(48), 1916–1924. https://doi.org/10.1161/STROKEAHA.116.016433.