Treatment and assessment tools for stroke patients with and without visuo-spatial neglect

Bernadette C. Tobler-Ammann
Treatment and assessment tools for stroke patients with and without visuo-spatial neglect

DISSERTATION

to obtain the degree of Doctor at Maastricht University, on the authority of the Rector Magnificus, Prof. dr. Rianne M. Letschert, in accordance with the decision of the Committee of Deans, to be defended in public on

Wednesday, June 27, 2018, at 12.00h

by

Bernadette C. Tobler-Ammann
**Supervisor**
Prof. dr. R. A. de Bie, professor of Physiotherapy Research, Maastricht University, the Netherlands

**Co-supervisors**
Dr. R. H. W. Knols, Physiotherapy and Occupational Therapy Research Center, University Hospital Zurich, Switzerland
Prof. dr. E. D. de Bruin, department of Health Sciences and Technology, ETH Zurich, Switzerland

**Degree Committee**
Prof. dr. C. G. Faber, professor of Neuromuscular Disorders, Maastricht University, the Netherlands (chair)
Dr. P. Kubben, department of Neurosurgery, Maastricht UMC+, the Netherlands
Prof. dr. G. Kwakkel, professor of Neurorehabilitation, VUmc Amsterdam, the Netherlands
Dr. T. Satink, self-management in post-stroke, HAN University of Applied Sciences, Nijmegen, the Netherlands
Prof. dr. A.M.C.F. Verbunt, professor of Clinical Epidemiology in Rehabilitation Medicine, with a focus on Chronic Pain, Maastricht University, the Netherlands

The research presented in this dissertation was conducted at various occupational therapy (OT) outpatient practices that are registered as specialized in neurorehabilitation in the canton of Bern, Switzerland, and at two rehabilitation clinics, namely the “Klinik Bethesda Neurorehabilitation, Parkinson-Zentrum, Epileptologie”, Tschugg, Bern, and “Zürcher RehaZentrum Wald”, Faltigberg-Wald, Zurich, Switzerland. This work was partially supported by the REWIRE project (www.rewire-project.eu) and funded by the European Commission under the FP7 framework with contract 287713 and the Occupational Therapy Association of Switzerland.
“Such a thing [stroke event] happens quite unexpectedly, you know. In one fell swoop, your whole life is completely different. Nothing is the same anymore. This is really sad. Suddenly you are dependent on outside help, because you can’t walk anymore and you are only able to do things with one arm. And everything you did before you can’t do anymore. This is really, really difficult. [...] I was always doing something from morning till evening. I’ve been working 12 to 14 hours a day and additionally doing lots of things during my leisure time. I always used to be active – and now all of a sudden you are a sack of potatoes.”

Patient’s description of experiencing a stroke event
Contents

Chapter 1 .......................................................................................................................... 1
General introduction

Chapter 2 .......................................................................................................................... 41
Application of principles of exercise training in sub-acute and chronic stroke survivors: a systematic review

Chapter 3 .......................................................................................................................... 79
Concurrent validity and test-retest reliability of the Virtual Peg Insertion Test to quantify upper limb function in patients with chronic stroke

Chapter 4 .......................................................................................................................... 109
The Zürich Maxi Mental Status Inventory (ZüMAX): Test-retest reliability and discriminant validity in stroke survivors

Chapter 5 .......................................................................................................................... 137
Exergames encouraging exploration of hemineglected space in stroke patients with visuospatial neglect: A feasibility study

Chapter 6 .......................................................................................................................... 177
User perspectives on exergames designed to explore the hemineglected space for stroke patients with visuospatial neglect: Usability study

Chapter 7 .......................................................................................................................... 213
From unawareness to emergent awareness: experiences of stroke patients with unilateral spatial neglect near the end of inpatient rehabilitation

Chapter 8 .......................................................................................................................... 241
General discussion

Summary ........................................................................................................................... 271
Zusammenfassung ........................................................................................................... 277
Valorisation Addendum ................................................................................................. 285
Acknowledgment .......................................................................................................... 289
Curriculum vitae .......................................................................................................... 293
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADL</td>
<td>activities of daily living</td>
</tr>
<tr>
<td>BBT</td>
<td>Box and Block Test</td>
</tr>
<tr>
<td>BIT</td>
<td>Behavioral Inattention Test</td>
</tr>
<tr>
<td>CBS</td>
<td>Catherine Bergego Scale</td>
</tr>
<tr>
<td>CIMT</td>
<td>constraint-induced movement therapy</td>
</tr>
<tr>
<td>CO-OP</td>
<td>cognitive orientation to daily occupational performance</td>
</tr>
<tr>
<td>CT</td>
<td>computed tomography</td>
</tr>
<tr>
<td>DALYs</td>
<td>disability-adjusted life-years</td>
</tr>
<tr>
<td>EEG</td>
<td>electroencephalography</td>
</tr>
<tr>
<td>ETNT</td>
<td>Eye Tracker Neglect Test</td>
</tr>
<tr>
<td>FITT</td>
<td>Frequency, Intensity, Time, Type</td>
</tr>
<tr>
<td>ICF</td>
<td>international classification of functioning</td>
</tr>
<tr>
<td>IGER</td>
<td>Intelligent Game Engine for Rehabilitation</td>
</tr>
<tr>
<td>LE</td>
<td>lower extremity</td>
</tr>
<tr>
<td>MCI</td>
<td>mild cognitive impairments</td>
</tr>
<tr>
<td>(m)CIMT</td>
<td>(modified) constraint-induced movement therapy</td>
</tr>
<tr>
<td>MMSE</td>
<td>Mini Mental State Examination</td>
</tr>
<tr>
<td>MoCA</td>
<td>Montreal Cognitive Assessment</td>
</tr>
<tr>
<td>MRI</td>
<td>magnetic resonance imaging</td>
</tr>
<tr>
<td>MV</td>
<td>movement visualization</td>
</tr>
<tr>
<td>NET</td>
<td>Neglect Test</td>
</tr>
<tr>
<td>NFT</td>
<td>neurofeedback therapy</td>
</tr>
<tr>
<td>NHPT</td>
<td>Nine Hole Peg Test</td>
</tr>
<tr>
<td>NICE</td>
<td>National Institute for Health and Care Excellence</td>
</tr>
<tr>
<td>NIHSS</td>
<td>National Institutes of Health Stroke Scale</td>
</tr>
<tr>
<td>NMV</td>
<td>neck muscle vibrations</td>
</tr>
<tr>
<td>OKS</td>
<td>optokinetic stimulation</td>
</tr>
<tr>
<td>OT(s)</td>
<td>occupational therapist(s)</td>
</tr>
<tr>
<td>PC</td>
<td>personal computer</td>
</tr>
<tr>
<td>PNF</td>
<td>proprioception neuromuscular facilitation</td>
</tr>
<tr>
<td>PSCI</td>
<td>post-stroke cognitive impairments</td>
</tr>
<tr>
<td>PSD</td>
<td>post-stroke depression</td>
</tr>
<tr>
<td>PSDs</td>
<td>psychosocial difficulties</td>
</tr>
</tbody>
</table>
Chapter 1
General introduction

Pre exergames intervention results of a figure copying task of the Neglect Test (NET) of P5, male, 53 years old, 6 weeks onset right-hemispheric stroke and visuo-spatial neglect
Stroke

Definition

Humankind has always had to live with stroke and its consequences. An English physician named William Cole was likely the first to introduce the word “stroke” into medicine in 1689 [1, 2]. Until then, the term “apoplexy” (from the term in ancient Greek ἀποπληξία, meaning “a striking away”) had been used to describe acute non-traumatic brain injuries [3]. With increasing medical knowledge about the nature of stroke and technological advances in its diagnosis, the definition of stroke has changed. Examples of such advances are the imaging of the brain through computed tomography (CT) scans since the 1970s [4], and magnetic resonance imaging (MRI) since the 1990s [5]. To date, stroke is defined as “an acute episode of focal dysfunction of the brain, retina, or spinal cord lasting longer than 24h, or of any duration if imaging (CT or MRI) or autopsy show focal infarction or hemorrhage relevant to the symptoms” [2]. Stroke is distinguished from transient ischemic attack (TIA), being defined as “focal dysfunction of less than 24h duration and with no imaging evidence or infarction” [6].

Global prevalence, incidence rates and risk factors

With almost 35.7 million survivors in 2013, 6.5 million deceased persons, 113 million disability-adjusted life-years (DALYs; numbers of years lost because of disability) and 10.3 million new incidents, stroke is the third most common cause of disability and the second most common cause of death worldwide [7, 8]. Despite the trends of declining mortality and stroke incidence rates, the global burden of stroke is increasing [6, 9, 10]. The increasing numbers of stroke cases are likely due to major advances in the treatment of stroke and the therewith-related increasing number of people who remain disabled [11]. Aging, growth of the population and increased prevalence of modifiable risk factors are further reasons that stroke remains a major global health problem [12, 13]. Physical inactivity [14], obesity [15], sedentary behavior [16], excessive use of alcohol [17, 18], unhealthy diet [19-22], tobacco use [23], plus long working hours and work stress [24, 25] are all examples of modifiable lifestyle risk factors to prevent stroke [6, 7, 13, 26, 27]. Although age is still one of the strongest risk factors for stroke, being 25 times more prevalent in people aged 75-84 years than for those aged 45-54 years [26], stroke should no longer be regarded as a disease of the elderly [28]: two thirds of all stroke events occur among people younger than 70 years of age [7]. Some of the
above listed risk factors – such as being overweight – are likely to be the main drivers in this trend in lifestyle, because adolescents are also affected.

**Prevalence, incidence rates and costs in Switzerland**

In Switzerland in 2017, the number of estimated stroke events was 21’928 with 8,449 million inhabitants [29]. Men presented higher age-specific stroke events than women (179/100’000 versus 120/100’000), with an overall stroke event rate of 296/100’000 population in 2004 [30], and 260/100’000 in 2017 [29]. Those rates are comparable with global incidence rates of 258/100’000/year in 2010 and an age-adjusted incidence 1.5 times higher in men than in women [28]. Specifically, there were about 133/100’000 person-years in men and 99/100’000 person-years in women global ischemic stroke events in 2013 [7, 27], and 82/100’000 person-years for hemorrhagic stroke in 2010 worldwide [31]. DALYs due to hemorrhagic stroke was estimated at 178/100’000 in Switzerland in 2010 [31]. The cost per year of inpatient treatment in Switzerland was estimated to CHF 691 million in 2003, corresponding to 2.9% of total inpatients costs and 1.4% of total health care costs [32]. For comparison with other industrial countries, the total health-care costs caused by stroke were around 3% [28, 33]. The mean total costs per patient, who initially received inpatient treatment, was CHF 62’139 in 2014, the first year of treatment accounting for about 76% of the costs [32]. This considerable proportion of follow-up costs after acute care is because of rehabilitation thanks to the availability of highly specialized stroke centers in high income countries [34, 35]. Further costs originate from long-term care, nursing and loss of productivity [36]. For example, only 14.5% of patients of working age were able to fully return to work post-stroke in Switzerland in 2014 [32]. About 50% of those patients who survive a stroke event make incomplete recovery and remain dependent in activities of daily living (ADL) [37, 38].

**Subtypes**

The pathophysiological classification of stroke distinguishes between cerebral ischemia and cerebral hemorrhage with differing disease mechanisms [26]. An ischemic stroke is “an episode of neurological dysfunction caused by focal cerebral, spinal, or retinal infarction” [2]. The hemorrhagic stroke is subdivided in intracerebral and subarachnoid hemorrhagic stroke, depending on the localization of the bleeding. The former is “rapidly developing clinical signs of neurological dysfunction attributable
to a focal collection of blood within the brain parenchyma or ventricular system that is not caused by trauma” [2]. The latter is “rapidly developing signs of neurological dysfunction and/or headache because of bleeding into the subarachnoid space, which is not caused by trauma” [2]. About 80% of strokes are caused by primary cerebral ischemia, and 20% by cerebral hemorrhage [26], from which approximately 15% are of intracerebral and 5% of subarachnoid origin [39]. Mortality within the first 30 days post-stroke is approximately 25-48% for hemorrhagic and 13-23% for ischemic stroke [40].

Clinical features
Initial assessment and triage of the stroke patient is crucial: the more time that passes without accurate medical help, the more severe becomes the brain damage and its consequences [41]. Therefore, knowledge of cerebral arterial anatomy and location of different functions of the brain are required for a swift and exact localization, allowing precise medical treatment [26]. Some clinical stroke features are common in the anterior and posterior circulation territory of the brain, such as hemiparesis, hemianopia or hemisensory loss. Other symptoms, for example ataxia or dysphagia, are typical for a posterior circulation stroke, while dysphasia, apraxia or neglect are usually caused by an anterior circulation stroke [26, 41]. Many lacunar infarct strokes are asymptomatic, while some result in a purely motor or sensory stroke [26]. As lacunar strokes spare the cortex, symptoms such as neglect or dysphasia are no clinical signs, while cognitive impairment seems to be common [42, 43]. The most commonly used standard assessment tool for determining baseline stroke severity is the National Institutes of Health Stroke Scale (NIHSS) [44]. An accurate diagnosis of stroke and its neurological consequences is important for the subsequent therapeutic management of the syndrome.

Sensory-motor impairments
A hemiparesis is a “muscular weakness or partial paralysis restricted to one side of the body”, while a hemiplegia is defined as “total or partial paralysis of one side of the body that results from disease of or injury to the motor centers of the brain” [45]. Post-stroke hemiplegia can affect the upper extremity (UE), the lower extremity (LE) or both, usually affecting the limbs contralesional to the brain damage. A significant number of stroke patients will not fully recover the use of their affected limb. Incidence of
spasticity, contracture and pain affect 33-90% of upper-limb hemiplegia, symptoms of pain and spasticity appearing within one week post-stroke, and contracture within 2 weeks [46]. There is evidence that recovery of UE function is achieved by 79% of patients with mild hemiparesis but by only 18% of patients with severe hemiparesis [47]. Recovery of walking is usually achieved in 65-85% of stroke patients within 6 months post-stroke [48]. The presence of hemisensory loss – the loss of sensation on one side of the body – negatively influences recovery from hemiplegia. There is evidence that functional recovery post-stroke is completed within 12.5 weeks in 95% of stroke survivors, while 80% of those patients had reached their best ADL function within 6 weeks [49].

**Cognitive impairment**

Besides sensory-motor impairments, the prevalence of post-stroke cognitive impairments (PSCI) is high, affecting almost 50% of stroke survivors at 3 month onset stroke [50]. PSCI – defined as “cognitive disorders with or without loss of self-sufficiency” – ranges from mild cognitive impairments (MCI) to post-stroke dementia [50]. PSCI may affect all major cognitive domains, such as executive function, attention, visuospatial abilities (classified as perceptual rather than cognitive skills; see chapter “Spatial neglect”), language and memory/learning [42, 51]. The prevalence of post-stroke dementia ranges from 7-41% within the first year post-stroke [52], while the prevalence of MCI ranges from 17-92% at 3 months post-stroke [53]. The wide ranges result from different study designs (e.g. population- vs hospital-based studies), from different criteria used and from included stroke subtypes. PSCI constitutes a substantial economic world health burden with estimated total costs of dementia, including post-stroke dementia, of 105 billion Euros in Europe in 2010 [54]. Therefore, accurate diagnosis and early treatment are essential. To date, the Montreal Cognitive Assessment (MoCA) and Mini Mental State Examination (MMSE) are the most valid and clinically feasible screening tools for accurate diagnosis [55, 56].

**Spatial neglect**

Post-stroke motor impairment or somatosensory deficits, however, might have other causes than muscular weakness or post-stroke dementia. Spatial neglect, a “failure to attend to the (in most instances) contralesional side of space” might be the cause of such symptoms [57]. Spatial neglect should be distinguished from hemianopia, a
“blindness in one half of the visual field of one or both eyes” [45], and optic ataxia, a “specific deficit of the visual control of movement unrelated to motor, somatosensory, visual acuity or visual field deficits” [58]. It is a cognitive disorder having an attentional and perceptual basis, since it may be overcome by cuing or increasing motivation [59]. Its prevalence varies widely from 13 to 81% of patients who had right hemispheric stroke [57], being most common and persistent with lesions of the right temporoparietal cortex [60]. This large variation in prevalence might be because previous studies have examined patients at different times post-stroke, and have used different tests to detect spatial neglect [61]. Spatial neglect is classified in two subtypes: i) motor origin, where patients exhibit directional hypokinesia for actions in the contralesional space, and ii) perceptual origin, where patients fail to respond to stimuli on the left side of their eyes, head or body irrespective of the motor response [57]. Neglect can involve several modalities, including the visual, auditory, olfactory, tactile and gustatory senses [62] (see Figure 1 for clinical examples). It further may affect a) the contralesional body (personal neglect), where patients are, for example, unaware of their left arm, b) the contralesional near space within reaching distance (peripersonal neglect), where patients, for example, leave one side of a plate of food uneaten, or c) the space beyond reaching distance (extrapersonal neglect), where patients are unable to find items placed on their affected side [57].

To measure the severity of neglect, a standardized battery of assessments rather than a single one is recommended [62]. Two assessments have shown to be most sensitive in detecting behavioral and visual neglect: the Behavioral Inattention Test (BIT) [63, 64], and the Catherine Bergego Scale (CBS) [65]. Both assess the presence of neglect in everyday situations, where symptoms usually persist longer than in paper-and-pencil tasks, which makes them useful for rehabilitation, as such impairments may restrict the patient after discharge [66].

Neglect has been associated with impaired functional recovery following stroke [59]. Between 9 to 22% of neglect patients additionally suffer from anosognosia, being a negative predictor of functional outcome in stroke rehabilitation [67, 68]. Anosognosia is a “lack of awareness (insight) or underestimation of a specific loss in sensory, perceptual, motor, affective or cognitive functions caused by a brain lesion” [68]. Spatial neglect resolves in most individuals, for example from initially 43% of right brain-lesioned stroke patients to 17% at 3 months post-stroke [60]. The larger the size
of the infarct, however, the more likely neglect symptoms persist [60]. Nevertheless, many neglect patients remain impaired in ADL, as visually guided behavior dominates the performance of ADL [57, 69]. Interestingly, stroke patients with neglect and anosognosia usually perceive their quality of life (QOL) as being better than neglect patients without anosognosia, due to their unawareness of their deficits [68].

**Psychosocial consequences**

In addition to all the previously identified bodily and functional impairments, the psychosocial burden of a stroke event and its consequences on daily life for the affected person is huge. Stroke survivors often have to cope with psychosocial difficulties (PSDs) [71], productivity loss [72], and/or post-stroke depression (PSD) [73]. PSDs, such as sleep disturbances or emotional instability, are prevalent with approximately 37% of stroke-survivors and commonly used to estimate disease severity [71]. They can cause (i) functional problems related to ADL, (ii) leisure or mobility restrictions (e.g. walking or driving), or (iii) a reduction in motivation and independence due to cognitive issues or over-solicitous families [74]. Environmental determinants such as the help, assistance and attitude of the family and health professionals can, thus, positively or negatively influence PSDs [71]. Furthermore, affected persons described experiencing lack of control, stress on the family, uncertainty about the future and work limitations due to their PSDs. Therefore, being enabled to get out of the house to meet people and to get back to work may reduce PSDs [74]. For return to work, the most robust predictor is initial stroke severity, followed by support from family and coworkers, employer flexibility and social benefit [75]. Organizational, personal and social factors can, thus, facilitate or impede return to work [75].

PSD occurs in approximately one third of stroke survivors at any time post-stroke, adversely affecting QOL and functional outcome by limiting participation in rehabilitation [76]. Caregiving relatives are also at a particular risk for depression, with depression rates even exceeding those of stroke survivors [77]. Therefore, brief psychosocial interventions and information provision for both the patient and spouse may be useful and effective in preventing and treating PSD [73].
**Figure 1.** Conventional test results of the standardized Behavioral Inattention Test from a person with right-sided intracerebral hemorrhagic stroke in the anterior circulation and severe visuo-spatial neglect 10 weeks onset stroke [70].

**Figure copying task:** “Please copy this figure.”

**Representational drawing task:** “Please draw a picture of a clock face, together with the numbers and a setting of the hands.”

**Visual search task:** “Detect and cross out all the target lines on the page.”
Stroke rehabilitation

There are two options to reduce global stroke burden: prevention and cure. Prevention strategies, such as dietary modifications [78-81], exercise [82, 83], or medication such as aspirin, antihypertensive agents [84] or statin therapy are effective to reduce the risk of a first-ever or recurrent stroke event [13, 85]. Other promising approaches in this respect are the development of “stroke Apps”, such as the “Stroke RiskometerTM App” [86]. This free mobile-phone application aims to spread stroke awareness by calculating the individual risk and by educating people on warning symptoms and signs.

Although prevention is always better than cure, effective therapies in the acute phase [41], as well as rehabilitation and long-term follow-up efforts, are all needed to prevent stroke recurrence and improve functional outcomes [7]. Therefore, rehabilitation starts as soon as the patient’s general health status allows it [39, 41]. Applying rehabilitation guidelines for recovery may facilitate the provision of the best treatment available, as they reflect best clinical practice and the evidence base available. Careful individualization is, however, needed at the time point of care [87]. To achieve optimal recovery from the stroke event, rehabilitation additionally requires efforts from a multidisciplinary team, including the patient and his/her family, plus physicians, occupational, speech and physical therapists, social workers, and others. Within this team, communication and coordination are essential to maximize the recovery potential of the affected person [87].

Conventional treatment methods

Rehabilitation of sensory-motor impairments

Notwithstanding the inputs of varying professional groups, UE sensory-motor rehabilitation is often a primary focus for occupational therapists (OTs) and includes the training of UE-ADL, as well as of touch, proprioception and range of motion of the affected arm [87]. Therapy of the affected arm is complex, including techniques such as splinting [88], joint mobilization, positioning [89] and stretching [46, 90]. The aim is to reduce difficulties due to the hemiparesis of the affected arm, which are pain, spasticity or contractures [89]. However, it is uncertain whether stretching has clinically important short-term effects on pain and activity limitations, and there is high-quality evidence that stretching did not have clinically important effects on joint mobility [90]. There is evidence that splinting the wrist in either a neutral or extended position does
not reduce wrist contracture after stroke [88]. The addition of botulinum toxin yielded only slightly better results in reducing wrist spasticity than splinting alone [91]. However, a strong evidence base exists for the use of botulinum toxin to reduce UE spasticity post-stroke [92]. Positioning the hemiplegic arm in a contracture preventive position slows down the development of shoulder abduction contracture, but does not show additional value on other outcome measures [89]. Modified constraint-induced movement therapy (mCIMT) and task-specific training show significant positive effects on UE function [93, 94]. Supplementary use of biofeedback and electrical stimulation further enhance UE function, while use of Bobath therapy is not supported [94]. Further beneficial interventions for improving UE function after stroke include mirror therapy [95], interventions for sensory impairment (e.g., sensory re-education, repetitive sensory practice or desensitization) [96] and repetitive task practice [97].

LE sensory-motor rehabilitation is more often a primary focus for physical therapists (PTs) and includes mobility (e.g., locomotion) and balance/vestibular training [38, 87, 98, 99]. Established rehabilitation strategies are, for example, the use of concepts such as Bobath, or the application of proprioception neuromuscular facilitation (PNF), motor learning, strengthening and functional principles [38]. Although these approaches are mostly based on empirical results rather than scientific evidence, several studies have shown the positive effects of these interventions on the recovery of motor function post-stroke [100-102]. To date, they are complemented with more advanced therapy methods, such as electric or thermal stimulation [103], mirror therapy [104], partial body weight supported treadmill training [105] or mCIMT [38]. Furthermore, there is strong evidence for high intensive task-oriented and task-specific training in all phases post-stroke to improve gait and mobility related functions and activities [99]. Interventions that include behavior change techniques appear more effective at improving real-world walking habits than exercise alone [106]. A recent systematic review, however, found that no one approach to physical rehabilitation is any more (or less) effective in promoting recovery of function and mobility after stroke, claiming for evidence-based physical therapy regardless of historical or philosophical origin [98].

Rehabilitation of cognitive impairments
Cognitive rehabilitation is more often a primary focus for neuropsychologists and OTs for executive dysfunction, attention deficits and memory/learning difficulties, and by speech therapists for language problems. Interventions aim either to restore lost skills
or to teach compensatory techniques by means of paper-and-pencil tasks, computer-based therapy or real-world activities. However, the evidence base is weak. For example, there is no evidence that cognitive rehabilitation interventions (interventions to restore, compensate for, or adapt to the functional loss) are helpful to restore executive function [107]. The effectiveness of occupational therapy for the treatment of cognitive impairment post-stroke (e.g., the use of a cognitive skills remediation program) is unclear [108]. There is evidence that computer-based treatment of attention deficits improves divided attention, but does not improve functional abilities, mood, or QOL [109]. Although evidence to support or refute the effectiveness of memory rehabilitation is limited [110], the use of an electronic paging reminder system seems to be promising to improve memory [111]. There is preliminary evidence that two promising interventions, the “cognitive orientation to daily occupational performance” (CO-OP) approach and a computer-based executive function training, have a positive impact on real-world outcomes [112]. Another promising approach for cognitive rehabilitation post-stroke is neurofeedback therapy (NFT), showing improvements in cognitive deficits [113]. However, limitations in study quality and strength of evidence restricts generalizability to the larger stroke population [114]. Furthermore, as NFT requires electroencephalography (EEG) modalities and a specific software to display the visual or auditory feedback to the user, its application is limited to laboratory settings.

**Rehabilitation of spatial neglect**

Treatment of spatial neglect can be split predominantly into “top-down” and “bottom-up” interventions [62, 115]. Top-down interventions encourage awareness of the disability and potential compensatory strategies by utilizing remaining intact brain function (compensation) [116]. Examples are visual scanning training (VST; involving progressive exercises and feedback to the patient while training the patient to voluntarily alter their gaze) and visual/mental imagery (e.g., the Lighthouse Strategy [117], where the patients are asked to imagine being a lighthouse, sweeping the light to the right and left to guide ships in to safety during therapy training tasks). The aim is to train the patients to compensate for their deficits, thus requiring their full cooperation, which is not always possible due to lack of awareness or anosognosia [62].
Bottom-up interventions use sensory stimulation to increase the patient’s perception of the affected space or limb. The aim is to retrain impaired function and to modulate internalized perceptions of the outer world (restitution) [116]. Examples are CIMT, limb activation therapy, prism adaptation, eye patching, mirror therapy, feedback training and sensory stimulation (e.g., repeated neck muscle vibrations (NMV), or optokinetic stimulation (OKS)) [62, 118]. These types of interventions are directed at the impairment, but do not address the underlying lack of self-awareness of the deficits, resulting in only short-lived improvement [116]. Adaptation by modifying the environment to the patient’s impairment is a third approach to treat neglect (substitution). Reducing environmental distractors to minimize the required visual attention during navigation is one example for such modifications. This method is promising, as the improvement of spatial neglect symptoms relies on both adaptation and subconscious active learning [116]. Furthermore, severely affected patients can benefit from this approach, as they cannot appreciate their underlying deficits, which is, for example, a prerequisite for learning compensatory strategies.

Despite the many interventions used to treat spatial neglect symptoms, the effectiveness of cognitive rehabilitation for this syndrome remains inconclusive. There is very limited evidence that cognitive rehabilitation may have an immediate beneficial effect on standardized neglect assessments, but there is no evidence for immediate or persisting effects on ADL and for persisting effects on neglect tests [109, 115]. Limb activation therapy (consists of spatio-motor cueing, where the patient volitionally moves the affected limb in the left hemi-space) can improve functional outcome, but it is not clear if this improvement translates into ADL [119]. Eye patching shows some amelioration of symptoms in their spatial bias [120] and improvement in reading tasks [121], but shows no amelioration in the functional level as measured with the CBS [121]. Repeated NMV using transcutaneous electrical nerve stimulation (TENS) and feedback training (verbal cueing by a therapist or video-feedback to increase awareness of the disability) have shown only transient remission of neglect symptoms [62, 118]. Furthermore, the impact of intervention was even found to be significantly different when rehabilitation was compared with “no treatment” versus “attention control”, suggesting that the presence of a therapist may be the crucial point for improvements rather than therapy content per se [109]. Although no treatment has been definitively shown to be of benefit, there is strong evidence that prism adaptation and VST help rehabilitate neglect [116, 122]. The
former causes a lateral shift in the gaze so that visual targets are altered so that the patient can identify objects in the previously neglected space [123]. This adaptation can improve daily functioning, as measured by reading/writing and ADL direct tests [123]. The latter uses visual anchors involving cues such as a red line, which are located on the left part of, for example, a book page and the patient is asked to start reading there [124]. Accordingly, there is evidence that VST improves functional skills such as reading and writing [125], and there is also evidence for the generalization of such training to other ADL that were not specifically trained [126]. Furthermore, repeated small-field OKS training seems to be an effective treatment for neglect to improve, for example, neglect dyslexia [118, 127]. OKS training uses moving visual personal computer (PC) displays to evoke an optokinetic nystagmus without the subjective impression of body rotation when using large visual displays that fill the field of vision [127]. Although this technique shows some enduring recovery from neglect [128], it requires active patient-participation [129]. Furthermore, evidence shows that if OKS combined with hemi-field eye patching was applied in acute stroke patients with neglect, it had no additive effect to the spontaneous recovery of the disorder [130]. This implies that the time-point of admission seems to be crucial for successful rehabilitation. Besides the provision of evidence-based neglect treatments, however, dedicating time for the education and training of family to understand the nature of neglect and how to best assist the affected person on a daily basis are equally important for successful rehabilitation.

Rehabilitation of activity performance and participation in daily life
As a result of the above described stroke-related physical and mental impairments, many stroke survivors have limitations in the performance of ADL, thus having difficulties in self-care (e.g. dressing), leisure (e.g. reading), domestic (e.g. cleaning), mobility (e.g. driving), social (e.g. meeting friends) and/or vocational (e.g. office tasks) activities. As a consequence, many remain dependent on others. Therefore, the restoration of occupational performance and the therewith-related ability to participate in daily life are important stroke rehabilitation goals. There is some evidence that occupational therapy targeted towards ADL can improve its performance in stroke survivors and, therefore, reduce the risk of deterioration in these abilities [37, 131]. Despite the variety of interventions used to improve occupational performance post-stroke, including but not limited to repetitive task practice, CIMT and mirror therapy
[132], there seem to be some commonalities, including being performance focused, using compensatory approaches, being goal-directed, individualizing tasks, and involving strategy training [133]. Other interventions aiming to improve motor function post-stroke, such as for example, using repetitive peripheral magnetic stimulation (rPMS) to create a painless, deep muscle stimulation, show no clear effect on ADL improvements post-treatment [134]. There is some evidence that leisure therapy attains short-term improvements in QOL and mood, together with increasing participation in and satisfaction with leisure activities [135]. Interventions used by OTs to improve on-road-fitness-to-drive, for example, computer-based driving simulator training or off-road skill-specific training, showed limited evidence to improve driving skills [136]. Other promising therapeutic strategies increasing participation in daily life are, for example, the use of the CO-OP approach [137], or the provision of post-acute self-management programs for stroke survivors [138].

**Adherence to stroke treatment**

There is strong evidence that it is not only individual treatment approaches which increase independence in ADL, but also organized, inter-professional stroke care [87]. In order to achieve an optimal rehabilitation outcome, however, the patient must adhere well to the therapy program. Successful adherence - being “the act of doing what is required” [45] - depends on both the therapist and the patient. The therapist has to be mindful in the individual planning and adjustment of the therapy program according to the patient’s progress. Dose and timing of the intervention are crucial for successful adherence, although its evidence is limited by substantial heterogeneity [139]. For example, the National Institute for Health and Care Excellence (NICE) guidelines 2013 for adult stroke rehabilitation recommend that patients receive 45 minutes of each relevant therapy five times a week, each therapy being provided at an intensity that will produce functional change [140]. A Cochrane systematic review concludes that a dose of 30 to 60 minutes per day delivered five to seven days per week is effective to recover function and mobility following stroke [98]. Paying attention to the principles of exercise training can facilitate clinical decision making regarding intensity and dose of the planned therapy and, therefore, increase adherence to the therapy schedule [141]. Considering patient-centered goal setting can further improve the patient’s motivation to participate in treatment [142]. There is consensus that a higher dose of repetitive task practice and extra practice during rehabilitation are
beneficial for post-stroke motor recovery [99, 143, 144]. However, the latter is often unfeasible due to organizational issues, or due to the patient’s reduced health status. Occurrence of post-stroke fatigue [145], emotional impairment [146] or anosognosia [147] can further hinder the patient’s motivation and adherence to treatment and, thus, reduce rehabilitation outcome.

**Choosing the best-practice method for the treatment of stroke**

Stroke rehabilitation is a dynamic process with the overall aim of reducing stroke-related disability [148]. The complex nature of stroke symptoms and the diversity of the population require health care practitioners to have a strong knowledge base of best-practice methods to support affected persons [131]. While the current evidence base for the sensory-motor treatment of UE impairments is moderate [97], interventions targeting improvements of gait and mobility related functions and activities have proven effectiveness [98]. There is currently insufficient evidence of good quality for the effectiveness of rehabilitation for cognitive impairments post-stroke to support clear recommendations for clinical practice [109]. Furthermore, the effectiveness of rehabilitation interventions for reducing neglect symptoms remains unproven [115]. Besides TENS, OKS, somatosensory electrostimulation or mirror therapy, VR training seems to be one of the most effective and promising treatment methods to rehabilitate neglect [149]. VR environments also show promise in the rehabilitation of ADL after stroke, thanks but not limited to their ecological validity [150, 151].

**Virtual reality as treatment and assessment tool for stroke rehabilitation**

In recent years, novel treatments and assessment tools have evolved to rehabilitate stroke symptoms, complementing conventional methods. These approaches, for example, include non-invasive brain stimulation, pharmacological augmentation and robot-assisted therapy [148]. Since the 1990s, many joint-specific robotic devices have been developed and tested in the rehabilitation of different neurological conditions [152]. They have the potential to deliver reproducible, high-intensity and repetitive therapy, thus potentially assisting therapists in rehabilitation programs [148]. Examples are the Bi-Manu-Track [153], the Arm robot (ARMin) [154, 155], and the Amadeo [156] to improve arm function and ADL for people after stroke. (For further UE examples see Brewer et al. 2007 [157], Merholz et al. 2012 [158], and Zhang et al. 2017 [159].) Robot-assisted arm training has the potential to improve ADL, arm function and arm muscle
strength [158, 159]. Results of the meta-analysis of Zhang et al. [159] further suggest that the UE function significantly improved when the robot-assisted arm training was used with standard care compared with standard care alone. Several robot-assisted devices have also been developed for the training of gait and walking (stationary and overground walking systems) [160]. Examples are the Lokomat [161, 162], and the GT I [163]. Those devices have proven to be effective complements to conventional therapy to improve walking distance [164], but not balance [165]. Moreover, some of the robotic-devices for the LE require two therapists to support the patient, and are restricted to the use in a clinic due to their size and weight.

Another proposed novel solution is virtual reality (VR) technologies including interactive video-gaming [166-172]. VR is defined as “an advanced form of human-computer interface that allows the user to ‘interact’ with and become ‘immersed’ in a computer-generated environment in a naturalistic fashion” [173]. Its use occupies many advantages not only for the patient, but also for the therapist: VR allows the therapist to exercise more control over the environment compared with real-life settings [174]. VR further provides a safe and supportive environment for practice, for example, the training of driving [175]. VR allows a broader range of activities to be offered in hospital and rehabilitation settings and provides a better real-world task than simulated ADL (e.g., reaching for cones the therapist holds in his hand as if taking an object from a shelf). Furthermore, VR may be a more cost-effective way of providing assessment and treatment by training the patient to use the VR device on their own, additionally increasing amount of daily therapy time [176]. This approach may also be useful for rural settings or nursing homes, where access to therapy is limited, and for patients without available transport options [174]. VR provides an enriched, challenging and appealing environment together with the opportunity of multimodal feedback on performance and progress [177]. In this way, VR shares similarities with interactive video-games, which are also used in therapeutic settings with promising results (e.g., the Nintendo Wii [178]) [179].

Barriers to the implementation of VR and video-gaming in clinical practice are mentioned in terms of application, education and research [174, 179]. The use of such devices requires education and computer skills from the therapist, together with the difficulty of keeping abreast of the rapid technological advances. Furthermore, many devices are unaffordable and occupy too much space for many therapeutic practices [179]. Creating a user-friendly program appears to be one of the greatest challenges
at present [174]. Similarly, a number of aspects of the application of VR technology remains unclear, for example, which patients are most likely to benefit from its use, or when, and for which issues a VR approach is most useful. Evidence for the effectiveness of VR technologies and video-games is rapidly growing [167-172]. However, the quality of that evidence remains low due to small sample sizes and often poorly reported methods. Furthermore, the majority of trials focused on motor rather than cognitive function, or on activity performance in relatively young patients who are more than one year post-stroke [167].

**Virtual reality for the treatment and assessment of motor function**

There is evidence to support the use of VR for the promotion of walking [168], for balance and mobility [180], and for the VR and interactive video-gaming for improving UE function [167]. The evidence whether VR elicits greater benefits when used as an adjunct to conventional therapy or when compared with the same dose of conventional therapy, however, is inconsistent. One systematic review, for example, concluded that if VR technology replaces some or all of the standard rehabilitation, then walking speed, balance and mobility significantly improved [181]. Another review reported better results when a multimodal approach is used, thus combining VR and conventional therapy, with respect to significant improvements in gait speed, balance and LE motor function post-stroke [182]. A meta-analysis exploring custom built virtual environments and commercial games in therapy used the international classification of functioning (ICF) framework as outcome measures [171]. They found a significant benefit of VR therapy compared to conventional therapy for body function and activity outcomes, but no significant difference between VR technology and commercial games. Therefore, custom built VR therapy demonstrated moderate advantages in body function and activity outcomes when compared to commercial games. These results are in line with another meta-analysis evaluating the effect of VR on arm motor function improvements after stroke [170]. They found a benefit in favour of VR in comparison to conventional rehabilitation alone, accounting for 15% improvement in motor impairment and 20% improvement in motor function after VR. There are preliminary but promising results for VR tests to assess sensorimotor functions of the UE [183, 184] and LE [185]. However, further research is necessary to improve quality criteria of such VR assessments and to evaluate their benefit in comparison to standardized stroke assessments [186]. Furthermore, evidence for the effectiveness
of VR to improve grip strength, gait speed and global motor function is currently insufficient [167]. At present, it is unclear which characteristics of VR are most important and it is unknown whether effects are sustained in the longer term.

**Characteristics and types of virtual reality applications**

Given the beneficial effects of the use of VR technologies to rehabilitate motor function after stroke, a recent systematic review investigated the characteristics of VR applications, specifically the visual information on the user’s movement in the virtual environment, also called movement visualization (MV) [187]. The broad variety of MV methods used across studies (e.g., an avatar or augmented reality), however, made a precise conclusion impossible. The authors could only conclude that individuals’ movements during VR-based motor training can be displayed in different ways. Another review examined the underlying training mechanisms involved in each VR intervention [166]. The results suggest that VR applications used for LE rehabilitation predominantly mediate learning through providing a task-oriented and graduated learning under a variable and unpredictable practice.

Henderson et al. [188] examined immersive and non-immersive types of virtual environments for UE motor recovery. In an immersive VR system, the users view parts of themselves (e.g. their hands or feet) or an avatar (icon or figure representing the user) in a video game on the computer screen [189]. To deepen the sense of immersion, the environment can be projected onto a concave surface, a large screen projection, or on a head mounted display. In a non-immersive environment, users interact to different degrees with the virtual environment without viewing themselves on the computer screen. This interaction takes place with different interface devices (e.g., a computer mouse, haptic device, force sensors, or a joystick). Henderson et al. [188] found an advantage to training in immersive VR environments versus conventional or no therapy. Conflicting results were found for the training in non-immersive environments versus no therapy and some evidence when compared to conventional therapy, concluding that additional trials with this form of VR environment are justified.

**Virtual reality for the treatment and assessment of cognition**

To date, the majority of studies have evaluated VR interventions and VR assessments to improve motor impairments post-stroke. Only a few randomized controlled trials
(RCTs) have addressed cognitive rehabilitation and assessment [167]. Si Hyun et al. [190], for example, assessed the effectiveness of a computerized visual perception rehabilitation program with an interactive patient-computer interface using motion-tracking technology for visual perception impairment post-stroke. Aside from the higher interest in the novel program, they found no differences between groups and the existing computer-based rehabilitation program. Kim et al. [191] compared a VR intervention (IREX system®) with a computer-based cognitive rehabilitation program (ComCog®), investigating the effects on the recovery of cognitive impairment in stroke patients. Results showed significant improvements in both groups, with the VR group having greater improvements in the attention domain. A recent RCT developed a VR based serious games application for cognitive training post-stroke, with attention and memory tasks consisting of ADL [151]. Results showed significant improvements in both attention and memory function in the intervention group, but not in the control group that underwent conventional cognitive training. These results are in line with another RCT comparing a VR-based cognitive rehabilitation program (Reh@City) to conventional cognitive therapy [150]. The experimental stroke group showed significantly greater improvements in global cognitive functioning, attention and executive functions than the control group.

Virtual reality for the treatment and assessment of neglect symptoms
Despite those promising results in favour of VR cognitive rehabilitation programs, the most prevalent assessment for measuring cognition post-stroke in clinical practice is still pen-and-paper (73%), followed by VR with 11%, computer 6%, observational functional performance 5%, informant 3% and telephone 3% [192]. The application of VR in the assessment of spatial neglect, however, seems to grow steadily. Two recent systematic reviews aiming to provide an overview of the most recent VR applications for the assessment and rehabilitation of neglect resulted in 13 [193] and 22 studies [194], respectively. Devices for identifying and measuring neglect include, for example, the Virtual Reality Lateralized Attention Test (VRLAT), a computerized measure that mimics the attentional demands of real-world tasks [195]. Results outperformed traditional paper-and-pencil tests in the prediction of real-world collisions. Another example is the Virtual Reality Diagnostic Test Battery (VRDiSTRO) developed to create an interactive and immersive 3D experience to assess spatial neglect [196]. The sensitivity was 100% and the specificity 82% for the VRDiSTRO to correctly identify
neglect. (See Pedroli et al. [193] and Ogourtsova et al. [194] for further examples of such devices.) Thus, VR assessments have the potential to identify neglect symptoms that were previously not identified with conventional methods [197]. However, given the lack of analysis of psychometric properties of several assessment studies, the quality of such VR assessments is debatable [194]. The evidence for the rehabilitation of neglect using VR methods is still limited due to the lack of high-quality studies substantiating conclusions with strong evidence [194]. The presence of significant within-group improvements in all VR treatment groups included in the systematic review of Ogourtsova et al. [194], however, remains promising, suggesting that VR interventions have the potential to reduce neglect symptoms. The RCTs by Akinwuntan et al. [198] and Van Kessel et al. [199], for example, both found significant within-group differences but no differences between groups. The former compared a simulator-based driving-related intervention with a non-computer-based cognitive training in stroke patients with and without neglect [198]. The latter compared visual scanning training alone with visual scanning training plus a VR driving simulator task [199]. Both stroke groups improved their driving-related visual attention skills after training, with benefits lasting up to 6 months. Another RCT, however, found significantly higher changes (for example in the CBS scores) in the VR group using the IREX system® compared to conventional neglect therapy alone [200]. This system consists of a monitor, a video camera, computer-recognizing gloves and virtual objects, stimulating the left side of the body. Thus, although there is limited evidence that VR is more effective than conventional therapy in improving neglect symptoms, VR has the capacity to enhance and complement current methods of diagnosis and treatment of neglect [201, 202]. To date, more VR systems are being developed and tested, for example the Bts-Nirvana System [203].

**Virtual reality for the treatment of activity performance**

VR and interactive video-gaming have proven to be beneficial for improving ADL performance when used as an adjunct to conventional therapy or when compared with the same dose of conventional therapy [167]. They provide the opportunity to practice cognitive and motor activities that cannot always be easily performed within the clinical environment. Examples are the training of attention abilities to safely cross the street [204, 205], of driving-related visual attention skills [175, 206], or of executive functions and multitasking by visiting a supermarket [207, 208]. Furthermore, VR systems can
provide ecologically valid environments for the training of ADL, thus represent a better real-world transfer rehabilitation than conventional functional therapy or paper-and-pencil tasks [150]. A recent study, for example, paid attention to designing an ecologically valid VR system for the training of ADL, the Reh@City, a three-dimensional environment with streets, sidewalks, commercial buildings, parks and moving cars [150]. Although the results are in favor of the experimental group regarding improvements in global cognitive functioning, attention and executive functions, there is insufficient evidence of if and how the use of VR technologies impacts everyday functioning [174].

Potential of Virtual Reality for stroke rehabilitation

Overall, the evidence for the use of VR applications for stroke rehabilitation is promising, with the body of knowledge growing fast. VR applications can be used either as assessment or intervention tools to complement – or in the (near) future – replace some of the more conventional therapy methods. They have the potential to overcome limitations of conventional rehabilitation, such as being time-consuming or labor-and resource-intensive [170], or language specific [209]. Furthermore, VR systems may allow earlier training of impaired functions than conventional therapy, for example, when the hemiparesis of the dominant arm interferes with performing paper-and-pencil tasks to restore cognition [150]. With most VR systems providing different difficulty levels and immediate feedback on performance and progress, paying attention to the application of exercise training principles becomes easier [87]. VR and video game applications, however, may be used to increase the intensity of treatment and to promote motor and cognitive recovery after stroke not only during rehabilitation [167], but also after discharge. Thanks to advances in VR systems - for example, through the development of transportable, light-weight and low-cost devices – self-directed home-based VR therapy became feasible. Such devices range from VR gaming platforms (e.g., Nintendo® Wii [178] or Microsoft® Kinect [210]), to dedicated instrumented devices (e.g., the ArmeoSenso system [211]), but also including combinations of these (e.g. the intelligent game engine for rehabilitation (IGER) together with the Falcon from Novint Technologies [212]). Thus, VR applications can serve as a perfect option for daily training on an individual level. However, not all patients would be eligible for this technology. Although VR tools may revolutionize stroke rehabilitation, they may still need to be provided in a supervised manner, allowing monitoring and guidance of the
patient to avoid adverse events [174]. Therapists need to assess the suitability of the VR tasks for their patients, grade the tasks and evaluate progress as part of a holistic and goal-oriented rehabilitation plan. Therefore, close partnerships between clinicians, engineers and patients are necessary to optimally advance the field of VR-based stroke rehabilitation [186].

While VR offers exciting prospect for rehabilitation, it is still a new and evolving field. Thus, while many studies are promising, it is not surprising that their power and design are still limited and the support they can offer remains moderate in quality [172]. Overall, further research is needed to advance this promising field, specifically, to provide further psychometric properties analysis of VR-based stroke assessments and to develop and test novel and innovative VR-interventions for their feasibility to treat stroke symptoms, such as neglect. Furthermore, assessing the users’ perspectives on this novel therapy method is important for maximizing acceptability of such novel rehabilitation methods. Gaining knowledge from personal experiences of patients having a stroke is deemed crucial for designing custom-made VR-interventions.
Rationale and objectives of the thesis

Despite huge advances in the field of stroke rehabilitation, the global burden of stroke remains substantial and is increasing as populations age. With increasing demand for rehabilitation services, it is vital to explore new ways of service delivery. It seems likely that VR will be important in the future, encouraging therapists to engage with engineering and gaming groups to explore innovative approaches to the delivery of rehabilitation programs.

In this respect, this thesis grew out of a collaboration with an engineering group and other (health) professionals working together in the European research consortium “rehabilitative wayout in responsive home environments” (REWIRE). REWIRE developed a non-immersive, low-cost and transportable VR system for stroke patients after discharge from rehabilitation to improve gait, posture and visuo-spatial neglect (VSN) symptoms. The main goal of REWIRE was to create a personalized health system that can be deployed massively within the patients’ homes, to enable effective home-based rehabilitation to improve disabilities and functions. REWIRE’s further goal was to target at patients recently discharged from hospital, who needed to continue rehabilitation. The idea was to provide them with continuing home-based rehabilitation under remote monitoring by the hospital (see www.rewire-project.eu for further information).

The author of this thesis was mainly involved in the development of a series of exergames to treat neglect symptoms post-stroke. As an OT, she provided clinical knowledge important to address the target group’s needs, while the game engineers designed and programmed the REWIRE neglect exergames accordingly. This series of exergames developed within the REWIRE research consortium aimed at improving exploration of the neglected space for patients with VSN post-stroke.
The primary objectives of this thesis were

1) to determine the feasibility of usage of exergames, with minimal therapist supervision, in terms of implementation of the intervention, including adherence, attrition, and safety, and limited efficacy testing, also aiming to evaluate possible effects on VSN symptoms in patient users (feasibility study).

2) to assess users’ perspectives (patients and therapists) when using the REWIRE exergames for rehabilitation of VSN symptoms due to a stroke. Specifically, we wanted to evaluate the perceived user-friendliness of the exergames, attitude towards using the exergames, and intention to use the exergames in the future (usability study).

Secondary objectives of this thesis were

3) to gain knowledge about the quality and level of reporting of exercise training principles used in the rehabilitation of the patient group, plus patient adherence to the prescribed intervention (systematic review). Subsequently, this knowledge was integrated in the development of the exergames.

4) to test psychometric properties of a novel VR-based stroke assessment to quantify UE function in patients with chronic stroke (the Virtual Peg Insertion Test (VPIT)) and of a novel cognitive paper-and-pencil tool to screen for disturbances in neuropsychological function (the Zürich maxi mental status inventory (ZüMAX)) aiming to evaluate their clinical utility (methodological studies). After having accomplished the methodological studies of this thesis, we tested the REWIRE exergames for their feasibility in the target group. As those games aimed at exploring the neglected space, the VPIT with its focus on restoring UE motor function was not used any further despite promising results. The ZüMAX, however, was used as one of the outcome measurements in the feasibility study of this thesis.

5) to explore everyday life experiences of persons with right hemispheric stroke and spatial neglect symptoms near the end of inpatient rehabilitation (qualitative study). The purpose was to provide insider knowledge important for health care professionals and families working or living with those persons to better meet their needs post-stroke. Specifically, we aimed to investigate how they experienced their performance of daily activities, their body and self.
Outline of the thesis

This thesis focuses on the testing of novel VR treatment and assessment tools in stroke patients to contribute to development and advancements in the field of stroke rehabilitation (Figure 2).

Figure 2. Flow chart of the outline of the thesis.

In CHAPTER 2, a systematic review of principles of exercise training in sub-acute and chronic stroke survivors is presented. The 37 RCTs included in this systematic review showed incomplete and inconsistent reporting of underlying exercise training principles and components, patient adherence and a broad variation in the methodological quality. CHAPTERS 3 and 4 comprise analyses of psychometric properties of novel stroke assessments. CHAPTER 3 presents a VR assessment tool to quantify UE motor function of the affected limb post-stroke in a sample of 31 chronic stroke patients, the VPIT. Results showed that the VPIT is a feasible but demanding test in this target group, requiring other components of UE motor function than the conventional gold standard tests. CHAPTER 4 includes the examination of a novel paper-and-pencil cognitive assessment tool, the ZüMAX, in 33 community-dwelling stroke survivors and 35 healthy controls. It showed moderate-to-good test-retest reliability and
discriminative abilities between healthy persons and chronic stroke survivors in 3/5 subscales. **CHAPTERS 5 and 6** investigate the feasibility, potential effects and usability of a novel series of nine exergames aimed at improving exploration of the neglected space for patients with VSN symptoms in the sub-acute stage post-stroke. The high adherence rate showed that these exergames were feasible for the participants (**CHAPTER 5**). The results further suggest that these exergames have the potential to reduce VSN symptoms. Additionally, **CHAPTER 5** shows promising preliminary results of a newly developed assessment tool to quantify spatial exploration skills, the Eye Tracker Neglect Test (ETNT). **CHAPTER 6** comprises an examination of users’ perspectives of previously appraised VSN exergames, evaluating the perceived-user friendliness, attitude towards using them and intention for future use. Seven patients and 12 therapists participated in this usability study, the latter being generally more critical towards this novel treatment method than the former. In **CHAPTER 7**, the qualitative findings of how persons with right hemispheric stroke and spatial neglect symptoms experience their everyday life near the end of inpatient rehabilitation are presented. The seven participants experienced a stepwise adaptation process post-stroke, expressed in three over-arching themes. Finally, **CHAPTER 8** provides a general discussion of the thesis outcomes, including implications for clinical practice and recommendations for further research relevant for this topic.
References


Chapter 2
Application of principles of exercise training in sub-acute and chronic stroke survivors: a systematic review

Post exergames intervention results of a figure copying task of the Neglect Test (NET) of P5, male, 53 years old, 10 weeks onset right-hemispheric stroke and visuo-spatial neglect

This chapter has been published as: Ammann BC, Knols RH, Baschung P, de Bie RA, de Bruin ED. Application of principles of exercise training in sub-acute and chronic stroke survivors: a systematic review. *BMC Neurology* 2014,14:167.
Abstract

Background: There is increasing evidence for the beneficial effects of exercise training in stroke survivors. In order to reach the desired training effects, exercise training principles must be considered as this ensures the prescription of adequate exercises at an adequate dose. Moreover, exercise training interventions must be designed in a way that maximizes patients’ adherence to the prescribed exercise regimen. The objectives of this systematic review were (1) to investigate whether training principles for physical exercise interventions are reported in RCTs for sub-acute and chronic stroke survivors, (2) to evaluate whether the RCTs reported the prescription of the FITT components of the exercise interventions as well as (3) patients’ adherence to this prescription, and (4) to assess the risk of bias of the included studies.

Methods: We performed a systematic review of RCTs with exercise training as the primary intervention and muscular strength and/or endurance as primary outcomes. The Cochrane library’s risk of bias (ROB) tool was used to judge the methodological quality of RCTs.

Results: Thirty-seven RCTs were included in this systematic review. Eighteen studies (48.7%) focused on aerobic, 8 (21.6%) on resistance and 11 (29.7%) on combined interventions of aerobic and resistive strength exercise. Twenty-nine studies (78.4%) included only chronic stroke survivors, 5 studies (13.5%) only sub-acute stroke survivors whilst 3 studies (8.1%) included both. In terms of principle of exercise training, 89% reported specificity, 75.7% progression, 48.7% overload, 37.8% initial values, 32.4% reversibility and 13.5% diminishing returns. One RCT described all principles of physical exercise training and 19 (51.4%) all FITT components. Patients’ adherence to exercise prescription was accounted for in 3 studies (8.1%). Failure to report blinding in patients and participants and failure to report allocation concealment were the most prevalent methodological shortcomings.

Conclusions: Incomplete and inconsistent reporting of (1) training components, (2) underlying exercise training principles and (3) patient adherence together with (4) a broad variation in the methodological quality of the included RCTs limit both the utility and reproducibility of physical exercise programs in stroke patients.

Keywords: Stroke, Cerebrovascular accident, Aerobic exercise, Resistance exercise, Training principles, Training components
Background
Stroke is one of the leading causes of disability and death worldwide [1]. Due to demographic shifts in the global population, the number of affected people will increase, even with stable stroke incidence rates from approximately 1.1 million per year in 2000 to 1.5 million per year by 2025 [2,3]. The difference by 2025 will be ±150 000 stroke events when compared with stable rates [3]. Approximately 50-70% of persons with stroke regain functional independence, but 15-30% of the stroke survivors are left with permanent disability [4]. Disability – manifested by impairment of body function or body structure, activity limitation and/or participation restriction [5] – results in poor physical fitness, defined as “the ability to carry out daily tasks with vigor and alertness, without undue fatigue, and with ample energy to enjoy leisure-time pursuits and respond to emergencies” [6].
Physical fitness includes health-related (cardiorespiratory endurance, muscular endurance, muscular strength, flexibility and body composition) and skill-related components (agility, coordination, balance, speed, reaction time and power) [6]. All of these render the development of exercise training programs for stroke rehabilitation a complex enterprise. Although several exercise recommendations have been published [7-9], the complex interactions present in stroke rehabilitation preclude the definition of specific, detailed exercise prescriptions. Still, several systematic reviews and meta-analyses provide evidence that aerobic exercise and resistive strength training are beneficial to improve aerobic capacity, walking distance, muscular strength and physical function in stroke survivors without increasing pain or tone in the paretic limbs [10-17]. Thus, the available literature suggests that impaired physical fitness is partly responsible for the disability evident in stroke survivors.
When reporting the results of an exercise intervention, it is important that the precise principles of the exercise training are consistently and accurately reported [18,19]. Recognised principles are specificity, overload, progression, initial values, reversibility and diminishing returns (Table 1) [18-20]. Their application in the design of an exercise intervention ensures that the dose and type of exercise is planned such that benefits for the recipient are maximized. Furthermore, it seems fair to assume that when principles of exercise training are applied to the development of exercise protocols, clinicians in practical settings can have greater confidence that non-significant research findings reflect deficiencies in exercise efficacy rather than deficiencies in exercise prescription [18].
Table 1 Exercise training principles

<table>
<thead>
<tr>
<th>Principle</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specificity</td>
<td>Exercising a certain body part or component of the body primarily develops that part. To become better at a particular exercise or skill, you must perform that exercise or skill.</td>
</tr>
<tr>
<td>Overload</td>
<td>A greater than normal stress or load on the body is required for training adaptation to take place. The body will adapt to this stimulus.</td>
</tr>
<tr>
<td>Progression</td>
<td>There is an optimal level of overload that should be achieved, and an optimal time frame for this overload to occur. A gradual and systematic increase of the workload over a period of time will result in improvements in fitness without risk of injury.</td>
</tr>
<tr>
<td>Initial values</td>
<td>Improvement in the outcome of interest will be greatest in those with lower initial values. In other words, those with lowest level of fitness have greatest room for improvement.</td>
</tr>
<tr>
<td>Reversibility</td>
<td>Once a training stimulus is removed, fitness levels will eventually return to baseline (‘use it or lose it!’).</td>
</tr>
<tr>
<td>Diminishing returns</td>
<td>Refers to the decreasing expected degree of improvement in fitness as individuals become fit, thereby increasing the effort required for further improvements.</td>
</tr>
</tbody>
</table>

However, a perfectly planned intervention adopting all the principles of exercise training is almost useless when it is not reported in sufficient detail to permit intervention replication and results interpretation. Therefore, the prescription of the components of the exercise program and participants’ adherence to that exercise prescription should ideally be reported according to the so-called FITT components (Frequency, Intensity, Time and Type of exercise) (Table 2) [21]. FITT represents components of physical conditioning programs that determine effect on cardiorespiratory endurance, muscular strength and/or endurance and flexibility. Furthermore, participants’ adherence to each of the prescribed FITT components should be reported. Detailed reporting of the dose of exercise prescribed (and received) allows for an adequate interpretation of results – including possible dose-response effects – and provides information about the tolerability and safety of the intervention. Without detailed information on both the type and dose of exercise that is prescribed and actually received, developing optimally designed and dosed exercise prescriptions for a desired level of benefit (i.e., response) remains difficult, thus limiting the implementation of evidence-based training programs.
Table 2 FITT components applied to physical conditioning programs

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Intensity</th>
<th>Time</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>The number of times an exercise or activity is performed generally expressed in sessions, episodes or bouts per week.</td>
<td>Refers to how much work is being performed or the magnitude of the effort required performing an activity or exercise.</td>
<td>The length or duration in which an activity or exercise is performed, usually expressed in minutes.</td>
<td>E.g. running/swimming for cardiorespiratory endurance; free weights/resistance machines for muscular strength or endurance.</td>
</tr>
</tbody>
</table>

Well-designed randomized controlled trials (RCTs) provide the best evidence regarding the effectiveness of health care interventions. Trials with inadequate methodological approaches may overstate treatment effects and bias results. Critical appraisal of the quality of clinical trials is possible only if the design, execution and analysis of RCTs are described thoroughly and accurately [22,23]. Thus, in order to properly interpret the results of an RCT, it is important to know the principles underlying the prescribed exercises, the FITT components of the intervention, the adherence to these components and the methodological quality of a trial.

The objectives of this systematic review were (1) to investigate whether training principles for physical exercise interventions are reported in RCTs for sub-acute and chronic stroke survivors, (2) to evaluate whether the RCTs reported the prescription of the FITT components of the exercise interventions as well as (3) patients’ adherence to this prescription, and (4) to assess the risk of bias of the included studies.

**Methods**

**Search strategy**

An electronic search strategy was developed and performed by a librarian of Zurich University, using the databases Medline, OvidSP, EMBASE, CINAHL, PsycInfo and the Cochrane Library (Figure 1). The search was restricted to English language literature from inception of the databases up to February 2014. Combinations of the following medical subject headings (MeSH) and free text words were used: stroke (cerebral stroke, vascular accident, brain vascular accident, apoplexy, cerebrovascular apoplexy, cerebrovascular stroke, cerebrovascular accident) and exercise (cardiovascular training, cardiopulmonary training, cardiorespiratory training, aerobic training, endurance training, exercise, endurance exercise, ergometry, resistive
strength training, physical exercise principles, specificity, overload, progression, initial values, reversibility, diminishing returns).

**Eligibility criteria**

Studies were eligible for inclusion if they: (1) were RCTs (2) evaluated the effects of aerobic physical exercise training or resistive strength training alone or in combination, (3) included adult (> 18 years) sub-acute or chronic stroke survivors, and (4) included aerobic capacity and/or aerobic endurance and/or muscle strength as primary outcome measurement. The sub-acute phase was defined as the range between 15 and 180 days after initial stroke, thus all RCTs with a mean or median within this time range were considered a sub-acute population. Chronic stroke was defined as the open-ended time period starting 180 days after initial stroke and characterised by generally slow or no clinical progress [24].
Figure 1 Search history. Example of search history from OVID MEDLINE(R).

Study selection
After removal of duplicates, the search results were screened for eligibility by two teams of two reviewers (BA/PB and RK/EDB), sharing the retrieved citations. In case of disagreement within one team, a reviewer from the other team served as referee.

Data collection process
Using existing systematic reviews that evaluated the effects of exercise interventions in cancer patients [18,19], BA created a purpose-designed data collection sheet to include information about each included RCT regarding (1) the reporting of exercise training principles, (2) the description of the exercise training components, (3) participants’ adherence to the training plan according the FITT components, and (4) further quality characteristics of the included studies (risk of bias (ROB)). Corresponding authors of selected trials were contacted by email to retrieve unpublished data. Three reviewers (BA, PB and RK) independently collected and rated these data.

Rating of the reporting of exercise training principles
For each exercise principle, three rating categories were used: reported (+), not reported (-) and unclear or inconsistently reported (?).

Rating of the description of exercise training components and participants’ adherence to the training plan
Corresponding to the rating of the training principles, we judged the description with (+) if the component of the exercise prescription was reported; (-) if the component was not reported and (?) if it was unclearly or inconsistently reported. Participants’ adherence to the exercise prescription was judged with (+) if the adherence of each component to the exercise prescription was reported; (-) if the adherence was not reported and (?) if it was unclearly or inconsistently reported.
Rating of risk of bias in individual studies
Two reviewers (BA and PB) independently applied the Cochrane Collaboration’s tool for assessing ROB to assess the risk of over- or under-estimating the effects of an intervention [25]. The five bias domains (selection, performance, detection, attrition and reporting bias) were judged as: low ROB (+), unclear ROB (?) or high ROB (-). Rating (+) is unlikely to alter the results seriously, (?) raises some doubt about the results and (-) seriously weakens confidence in the results. With unclear information on an item, the score given was (?). The arbitration of a third reviewer (RK) was used in the event of any disagreement between the rating reviewers (BA and PB).

Data synthesis
All extracted data from the RCTs were transformed into percentages. The percentage agreement and Cohen’s kappa were calculated and interpreted in accordance with Landis and Koch’s benchmarks for assessing the agreement between raters: poor (≤0), slight (between 0.0 and 0.20), fair (between 0.21 and 0.40), moderate (between 0.41 and 0.60), substantial (between 0.61 and 0.80), and almost perfect (between 0.81 and 1) [26]. The PRISMA statement was followed for reporting items of this systematic review [27,28].

Results
Study selection
The systematic search provided a total of 1599 citations. After automatic de-duplication, 1265 hits remained. From this total, 1158 titles and abstracts failed to meet the inclusion criteria and were thus excluded. The remaining 107 abstracts were retrieved as full texts and screened together with their references. Sixteen additional studies were considered after manually scanning the reference lists of identified studies. After excluding another 86 studies based on in- and exclusion criteria, 37 studies were included in the final review (Figure 2).
Study characteristics

From the 37 studies with a grand total of 2135 participants, 18 studies (48.7%) focused on aerobic exercises [29-46], 8 (21.6%) on resistance exercises [47-54] and 11 (29.7%) on combined interventions of aerobic and resistive strength exercise [55-65] (Additional file 1). Most of the included studies compared an experimental exercise intervention to usual care, or to an alternative intervention. Four RCTs included 3 or 4 training groups in their design [38,53,58,59]. Twenty-nine studies (78.4%) included chronic stroke survivors [32-41,43, 44,46-54,58-65], 5 (13.5%) included sub-acute stroke survivors [29-31, 55, 56] whilst 3 studies (8.1%) included both types of patients [42,45,57]. In the 18 studies evaluating aerobic exercise only (n=1278) [29-46], treadmill training with or without body weight support was the most frequently used exercise intervention. In the 8 studies that focused on resistance training (n=235) [47-54], circuit classes with progressive strength training constituted the most frequently used intervention. Of the 11 studies focusing on combined interventions of aerobic and resistive strength exercise (n=622) [55-65], most focused on improvement of the lower extremity or both upper and lower extremities. No study focused on improvement of
the upper extremity in the intervention group. The length of the exercise interventions ranged from 3 weeks [38] to 12 months [42,45]. Sixteen studies (43.2%) described follow-up measurements [29-34,36-38,41,47,51,54,58,61,62] in which the follow-up period ranged from 2 months [51] to one year [32,62].

**Application of exercise principles**

One study described all six principles of exercise training [32], one described 5 out of 6 [47], 11 described 4 out of 6 [34,36,37,40,41,43,44,46,53,58,60] and 12 described three training principles [30,35,38,39,51,59,61-64] (Additional file 2: Table S1a-c). The remaining 12 studies, representing 32.4% of the included studies, described just one [42,48,50,57] or two [29,31,33,45,52,54-56,65] of these. With respect to the specific training principles, 33 (89.2%) of the RCTs reported **specificity** [30,32,34-56,58-65], 28 (75.7%) described **progression** [32-41,43,44,46,47,49,51-53,55,56,58-65], 18 (48.7%) described **overload** [32,34-37,39,40,43,44,46,47,49,53,58-60,63,64], 14 (37.8%) described **initial values** [29-32,40,41,43-47,53,57,60], 12 (32.4%) described **reversibility** [29,31-34,38,41,47,51,54,61,62] and 5 (13.5%) described **diminishing returns** [30,32,36,37,58]. Nineteen studies (51.4%) did not describe the **initial values** of their participants [34-37,39,42,48-51,55,56,58,59,61-65] and 14 (37.8%) did not describe the principle of **overload** [29-31,33,42,48,50,51,54-56,61,65]. Seven studies (18.9%) were unclear in describing **progression** [29-31,42,45,48,57], whilst 5 (13.5%) were unclear with respect to **overload** [38,41,52,57,62].

**Prescription of the exercise training**

The reporting of each FITT component of the exercise prescription is summarized in Figure 3. Details of the FITT component reporting of each study are presented in the Additional file 1. In general, **Frequency** was described in 35 studies (94.6%) [29-32,34-41,43-65], leaving 2 studies (5.4%) that were unclear in their reporting [33,42]. Twenty-two studies (59.5%) reported training **Intensity** [32-37,39-41,43-47,49,50,53,58,59,61,64,65], 6 studies (16.2%) did not report this item [42,51,54-57] whilst 9 studies (24.3%) reported it unclearly [29-31,38,48,52,60,62,63]. Thirty-four studies (91.9%) reported the component **Time** [29-32,34-48,50-52,54-65] and 35 RCTs (94.6%) reported **Type** [29-32,34-56,58-65] of physical exercise. Two studies (5.4%) did not report the **Time** dedicated to exercises [33,49] whilst two studies (5.4%) unclearly reported the **Type** of each exercise bout [33,57]. Nineteen studies (51.4%) 50
described all FITT components [32,34-37,39-41,43-47,50,58,59,61,64,65], 15 (40.5%) reported 3 of 4 FITT components [29-31,38,48,49,51-56,60,62,63], two (5.4%) reported 2 out of 4 items [42,57], and one (2.7%) reported one of the 4 items [33].

**Figure 3 Reporting of the FITT components.** Review authors' judgments about the prescription of the training program according to the FITT format presented as percentages across all included studies.

**Adherence to the exercise training**

Adherence to each FITT component of the exercise prescription is summarized in figure 4. Twenty-one studies (56.8%) reported the item *Frequency* [30, 32, 36, 37, 40-42, 46-52, 54-57, 61, 63, 65], 5 studies (13.5%) reported the item *Intensity* of the activity [32, 37, 46, 55, 58], 7 studies (18.9%) reported the item *Time* duration of each exercise bout [32, 37, 40-42, 55, 56] and 9 studies (24.3%) reported the item *Type* of exercise [32, 37, 40, 41, 46, 50, 52, 55, 61]. Three studies (8.1%) unclearly reported adherence to the *Type* of exercise [48, 54, 63] whilst 3 studies (each representing 2.7%) unclearly reported the items *Frequency* [60], *Intensity* [63] and *Time* [46] of the intervention. Fifteen (40.5%) studies did not report the item *Frequency* [29, 31, 33-35, 38, 39, 43-45, 53, 58, 59, 62, 64], 29 studies did not report the item *Time* (78.4%) [29-31, 33-36, 38, 39, 43-45, 47-54, 57-65], 25 studies did not report the *Type* of exercises completed (67.6%) [29-31, 33-36, 38, 39, 42-45, 47, 49, 51, 53, 56-60, 62, 64, 65] and the item *Intensity* was not reported in 31 studies (83.8%) [29-31, 33-36, 38-45, 47-54, 56, 57, 59-62, 64, 65]. Fifteen studies (40.5%) did not report any of the FITT components [29, 31, 33-35, 38, 39, 43-45, 53, 59, 60, 62, 64], 11 studies (29.7%) reported 1 out of 4 components [30, 36, 47-49, 51, 54, 57, 58, 63, 65] and 5 studies
(13.5%) reported 2 out of 4 items [42, 50, 52, 56, 61]. Three studies (8.1%) reported all four FITT items [32, 37, 55].

Figure 4 Adherence to the FITT components. Review authors’ judgments about the adherence of the study sample to the exercise prescription according to the FITT format presented as percentages across all included studies.

Methodological risk of bias
Methodological quality varied substantially across the studies (Additional file 3: Table S2a-c). All included studies had a low risk of selective reporting (reporting bias) [29-65], 19 (51.4%) had a low risk concerning the blinding of their outcome assessments (detection bias) [29,31,32,34-36,40,42,45,46,48-50,52,55,57,60,61,63], 21 studies (56.8%) had a low risk of incomplete outcome data (attrition bias) [32,34,35,40-42,46-50,52, 54-57,59,61-63,65], 19 studies (51.4%) had a low risk of selection bias concerning allocation concealment [32,33,35,38,40-46,48,50,51,56,57,61-63] and 22 studies (59.5%) showed a low risk of selection bias regarding random sequence generation [32,33,35-38,40-42,45,46,48,51,55-63]. With the exception of one study [40], the blinding of participants and personnel (performance bias) was a weak point in all studies: 30 studies (81.1%) had an unclear risk [29-39,43,44,46,48-60,62,64,65] whilst the remaining 6 studies (16.2%) had a high risk of performance bias [41,42,45,47,61,63].

The inter-rater agreement of the 3 rating categories “low, unclear, high ROB” of the quality assessment was substantial (Kappa: 0.71, SE of kappa: 0.04, 95%CI: 0.630-0.796). The number of observed agreements between the two raters was 83.3% of the judgments.

52
Discussion

This systematic review evaluated the explicit reporting of the principles of (aerobic and/or resistance) exercise training in sub-acute and chronic stroke survivors. The results showed that these principles were inconsistently reported. This result impacts on the clinical reproducibility of trials, as clinicians cannot be confident whether non-significant findings are due to lack of efficacy or occur through limitations in treatment prescription. The risk of bias in the 37 studies depended on the bias domain being judged.

In this review, specificity and progression were the most frequently applied (i.e., explicitly reported) training principles, in 89.2% and 75.7% of reviewed studies respectively. Accordingly, most exercise trials clearly outlined training progression and designed their intervention to be specific to the target population, thus warranting reproducibility with respect to these training principles. In contrast, initial values and overload principles were not described in 37.8% and 48.7% of the reviewed RCTs respectively. Without knowing the baseline fitness levels of studied participants, it is difficult to generalize the findings to a clinical setting. Moreover, it is impossible to verify whether the provided exercise program was of adequate intensity, which hinders the interpretation of results. This situation is further aggravated by the fact that only 59.5% of the included studies reported the intensity of their exercise training. Unfortunately, it is not always feasible to accurately measure the intensity of an exercise due to lack of equipment to monitor energy, work and/or power. In addition, a lack of knowledge of the exact effort (in terms of energy, work and/or power) that healthy subjects require to perform certain exercises [66] further complicates the dosing of exercises in patient populations. Only in studies where the mechanical output of physical activity can be controlled, such as by using cycle ergometry [35, 39] or leg press machines [49], is the required effort known. Yet even in light of these difficulties, it should always be possible to report the number of repetitions required in a certain exercise or the total time dedicated to exercise training. The latter was successfully described in 91.9% of the exercise trials included in this review. Furthermore, energy spent performing certain exercises might be assessed via proxy measures such as the Borg’s Rating of Perceived Exertion and Pain Scales [67]. Although perceived exertion is a subjective measure, it may nonetheless provide a fairly good estimate of the actual heart rate during physical activity and hence the intensity of that activity. Indeed, practitioners
generally agree that perceived exertion ratings of 12 to 14 on the Borg Scale suggest a moderate level of intensity of physical activity [67].

There is evidence of a positive relationship between the time dedicated for therapy and therapy outcomes, indicating a positive relationship between dose and response [68, 69]. Lohse and colleagues [68] and Kwakkel et al. [69] both reported that the benefit of large increases in therapy volume is similar across a range of post-stroke times. That is, patients will benefit from an increase in therapy volume, regardless of whether their stroke occurred several months or several years ago. This finding also highlights the importance of exercise even after discharge from rehabilitation and reflects the training principle of reversibility: i.e. ‘use it or lose it’. Following the principle of diminishing returns – reported in only 13.5% of the studies – the effort required to achieve further improvements should increase over time. In line with the findings of the present systematic review, the principles of diminishing returns and reversibility were the least reported exercise principles in RCTs on physical training interventions in cancer survivors according to two recently published systematic reviews [18, 19]. A possible explanation is related to the fact that the assessment of these training principles requires follow-up measurements, which would increase research expenses and heighten the burden on patients participating in the study. However, assessment and reporting of both principles are crucial to determine the volume, frequency and intensity required in an exercise intervention to achieve durable long-term benefits for stroke patients [20].

Perhaps the most striking finding in this review was the discrepancy between the reporting of the FITT components in the exercise intervention (Figure 3) and the adherence to those components (Figure 4). To illustrate: 94.6% of the studies reported the FITT component Type yet only 24.3% of the studies reported adherence to this component. Similarly, Time, Frequency and Intensity was reported by 91.9%, 94.6% and 59.5% of the studies, respectively (Figure 3), yet adherence to these components was reported in only 18.9%, 56.8% and 13.5% of them (Figure 4). This general failure of reporting adherence with respect to the FITT components all but obscures this crucial aspect of an intervention [21] and hence prohibits important considerations that must be made in before replicating an intervention in a clinical setting. There are several reasons why exercise programs might not be performed as prescribed, including patient-related factors (e.g. fatigue or depression [70, 71]; lack of motivation; stroke impairments [72, 73]), environmental factors (e.g. lack of transportation) or
health concerns (e.g. fear of falls [74]). For clinicians and researchers alike, addressing these perceived barriers to exercise training is crucial both for successful rehabilitation and for provision of replicable exercise training programs. Promising first steps in breaking down these perceived barriers to training might be to remove costs for transportation or to integrate patients’ relatives in the rehabilitation process [75]. In support of the latter, it has been found that social support is an important motivator in achieving and maintaining physical fitness [72].

The most prevalent methodological shortcoming of the included studies was the absence of blinding of participants and personnel in 97.3% of the included RCTs. This is in line with findings of other systematic reviews of stroke exercise training [10,16]. Such lack of blinding can cause overestimations of the treatment effect and therefore bias the study results. For example in a meta-epidemiological study by Wood et al. [76], estimates of treatment effects were exaggerated by 7% in non-blinded compared to blinded trials. Although blinding of participants and personnel may not always be feasible, it may still be possible to blind the outcome assessments. In this review, 19 RCTs (51.4%) had a low risk concerning this form of blinding (detection bias). Objectively assessed outcomes are less susceptible to bias than subjectively assessed ones [76]. Therefore, efforts to minimize other forms of bias are particularly important when objective measurements are not feasible. Allocation concealment was ambiguous in 48.6% of studies whilst 15 studies (40.5%) had a high or unclear risk of selection bias regarding random sequence generation. These findings are in line with those of other reviews [77]. Because concealment of allocation can lead to exaggeration of treatment effects [78], details of both sequence generation and concealment of allocation should always be clearly reported [22,23].

Through improving cardiovascular fitness and muscle strength, disability after stroke may be reduced [16]. This is an important aspect of training since regaining function and independence are important goals for patients. The benefits of aerobic exercise might even be broader: Converging evidence suggests that aerobic exercise is a valuable intervention for improving brain function [79, 80], promoting neuroplasticity by upregulation of neurotrophins [81]. Aerobic exercise programs after stroke have also been shown to improve blood pressure [82] and arterial function [83], as well as enhancing glucose regulation [84]. It is also highly plausible that exercise could be an effective treatment for fatigue [85], especially in combination with the treatment of the associated depressive symptoms of post-stroke fatigue (PSF) [86], even though it must
be noted that there is insufficient evidence of an association between PSF and physical fitness to date [70]. Taken together, the total body of evidence clearly highlights the importance of maintaining physical fitness after stroke, as it greatly reduces the effort required to carry out daily tasks and therefore contributes to a more active and independent lifestyle [87]. Moreover, an optimal level of physical fitness decreases the risk of subsequent stroke, which is particularly significant in view of the fact that around 30% of stroke survivors will have recurrent stroke within their lifetimes, of which 18% will prove fatal [75]. Given the great restorative potential of achieving and maintaining physical fitness after stroke, the need for RCTs to properly report both exercise prescription and adherence to exercise prescription – which would allow full replication of positive findings in clinical settings – becomes particularly pressing.

Call for transparency to facilitate evidence-based practice

In contemporary clinical practice, clinicians must ensure that the time allotted for therapy is used effectively and efficiently. To be successful, it is imperative that the goals of an exercise program be both reasonable and attainable [20,88]. These aims are best achieved with a custom-made and individually tailored training program with variables that can be manipulated from workout to workout [21]. Such variables might be the choice, volume, intensity and order of exercise as well as the frequency and length of training and the length of rest periods. In order to achieve an optimum training effect, programs used in research should comply with and clearly report the exercise training principles [20] and FITT components [21]. This is expected to facilitate application of effective programs in clinical practice. However, the current state of the evidence still renders it difficult for practitioners to choose the optimal evidence-based training program for their individual patients. Only by reporting sufficient detail about volume and type of exercise actually performed by trial participants will more accurate interpretations of study outcomes and more appropriate translations of programs into non-research settings be permitted. A good starting point for future trials would be the American Heart Association (AHA) recommendation for stroke exercise training [89] (modified by Billinger et al. [75]). Such a detailed description of adherence, in combination with an equally detailed description of the developed program and the target population, would greatly facilitate reproduction of trials in clinical settings.
Limitations of the review
To the best of our knowledge, this systematic review is the first to investigate the application of exercise training principles for sub-acute and chronic stroke patients. The review focuses on the reporting of intervention content rather than on the actual intervention outcomes. In striving to achieve a robust systematic review, we developed and documented the methods (e.g. a systematic search strategy and several worksheets for collecting and synthesizing the data) in advance. Due to the large number of existing trials on stroke rehabilitation, we decided to focus exclusively on RCTs to ensure high external validity. Nevertheless, some limitations remain. Firstly, we limited our search to English language peer-reviewed journal literature. Hence clearly there is a possibility that important RCTs published in other languages were missed. Secondly, due to the scope of the review, we did not perform meta-analyses of RCT results and hence cannot make recommendations concerning preferable exercise interventions for sub-acute and chronic stroke survivors. For training recommendations based on best available evidence we refer to the literature [75,89]. Thirdly, although we included only RCTs in this review, the clinical and methodological diversity of the studies considered was still rather large. Finally, we exclusively focused on sub-acute and chronic stroke patients; thus our results cannot be generalized to acute patient groups. The reason for not including this patient population is twofold: First, there is no consensus in the literature as to how early physical activity should begin after a stroke [90] and second, information on how to influence and evaluate aerobic capacity in severely affected individuals is lacking [10].

Conclusions
The present systematic review showed that RCTs on aerobic and/or resistance rehabilitation training in sub-acute and chronic stroke survivors incompletely and inconsistently reported the prescription of, and especially the adherence to, exercise prescriptions in investigated training programs. Allowing practitioners to use established training principles to guide the therapy process will help them to plan and implement their stroke therapy more efficiently and more effectively. Therefore, it is of the utmost importance for researchers to plan and document investigated exercise interventions in as detailed a manner as possible. The consideration of all principles of exercise training and all FITT components in the development and reporting of exercise prescriptions is a course of action ideally fitted to this purpose.
Competing interests
The authors declare that they have no competing interests.

Authors' contributions
BA conceived the methodology and carried out data collection, quality assessment, data analysis and manuscript writing. RK supervised progress, participating in methodology conception as well as quality assessment, data analysis and manuscript writing & revision. PB helped with data collection, quality assessment and data analysis. RDB contributed in manuscript writing. EDB initiated the study, supervised progress, helped with methodology conception, data collection and manuscript writing & revision. All authors read and approved the final manuscript.

Acknowledgements
The authors thank Martina Gosteli for performing the search strategy. Further thanks go to Dr R van de Langenberg and to Dr Martin J Watson for proof reading the manuscript on English and structure. This work was partially supported by the REWIRE project (www.rewire-project.eu), funded by the European Commission under the FP7 framework with contract 287713.
### Additional file 1 Characteristics of included RCTs.

#### Aerobic exercise only (n=18)

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Participants (Int/Con)</th>
<th>Intervention (Int)</th>
<th>Control (Con)</th>
<th>Primary outcomes (instrument)</th>
<th>Other outcomes (instrument)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kosak (2000)</td>
<td>n=56 (22/34) (LE/LE)</td>
<td>F: 5d/w for duration of inpatient stay  I: ur  T: 45min/d  T: PBWSTT + PT of the same FIT</td>
<td>F: 5d/w for duration of inpatient stay  I: ur  T: 45min/d  T: ABAW+PT</td>
<td>Walking endurance (distance until fatigue) 0  Walking speed (patients' average speed during 2min) 0</td>
<td></td>
</tr>
<tr>
<td>Mayo (2013)</td>
<td>n=87 (43/44) (LE/LE)</td>
<td>F: daily for 12 months  I: 50-70% HRR  T: 15-30min/d  T: home ex. program with mobility ex. &amp; brisk walking involving stationary cycling</td>
<td>F: daily for 12 months  I: 11-15 on Borg Scale (6-20 ratings)  T: 15-30min/d  T: home ex. program with mobility ex. &amp; brisk walking</td>
<td>Walking capacity (6minWT) 0  HRQoL (RAND-36) 0  Balance (BBS) 0  Physical function (SIS) 0  Depression (GDS) 0</td>
<td></td>
</tr>
<tr>
<td>Dean (2012)</td>
<td>n=151 (76/75) (LE/UE)</td>
<td>F: at least 3x/w for 40w over 12 months  I: nr  T: 45-60min/d  T: home ex. program and weekly circuit-style group ex. class (progressive balance &amp; strengthening ex., walking, stair climbing)</td>
<td>F: at least 3x/w for 40w over 12 months  I: nr  T: 45-60min/d  T: home ex. program and weekly circuit-style group ex. class to improve UE and cognitive functions (task-related strength &amp; coordination training; sequencing tasks)</td>
<td>Walking capacity (6minWT) +  Walking speed (10MWT) +  Amount of falls (Fall calendars) 0  Fall risk (SF-PPA) 0  Physical activity (steps/d) (Digimax pedometer) 0  HRQoL (SF-12) 0  Community participation (AAP) +  Isometric knee muscle strength 0  Activity (TUG) 0</td>
<td></td>
</tr>
<tr>
<td>Bang (2013)</td>
<td>n=30 (15/15) (LE/LE)</td>
<td>F: 5x/w for 4 weeks  I: ½ of max overground walking speed, then increased  T: 30min/d plus 10min video  T: gait-related video, then treadmill training</td>
<td>F: 5x/w for 4 weeks  I: ½ of max overground walking speed, then increased  T: 30min/d plus 10min video  T: gait-unrelated video, then treadmill training</td>
<td>Walking endurance (6minWT) +  Dynamic balance test (TUG) +  Gait speed (10MWT) +  Walking capacity (max flexed knee angle in the swing phase (Dartfish motion analysis software)) +</td>
<td></td>
</tr>
<tr>
<td>Study (Year)</td>
<td>n</td>
<td>Study Group</td>
<td>Exercise Protocol</td>
<td>Outcome Measures</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>---</td>
<td>-------------</td>
<td>-------------------</td>
<td>------------------</td>
<td></td>
</tr>
</tbody>
</table>
| Carda (2013) [16] | 38 | (19/19) LE/LE | F: 5x/w for 6w  
I: 80% of 6minWT speed, then progressively increased  
T: 45min/d (PT) + 30min/d (treadmill training)  
T: PT + treadmill training with 5% descending slope | Walking endurance (6minWT) +  
Gait speed (10MWT) +  
Dynamic balance test (TUG) + |
| Jin (2013) [16] | 128 | (65/63) LE/LE | F: 5x/w for 12w  
I: 50-70%HRR  
T: 40min/d  
T: progressive aerobic cycling | Cardiovascular fitness (V0₂peak) +  
Knee extension muscle strength (isokinetic dynamometer) +  
Walking endurance (6minWT) +  
Balance (BBS) 0  
Spasticity of LE (MASc) 0 |
I: progressed from 40-80%HRR  
T: 60min/d  
T: high-intensity aerobic ex. (brisk level and inclined overground walking, cycle ergometry) | Cardiovascular fitness (V0₂peak) 0  
Arterial stiffness (pulse pressure in mm/Hg) +  
Ambulatory capacity (6minWT) 0  
Hemodynamics & cardiac function (echocardiography; 2D-Doppler) +  
Lipid, glucose & homocysteine levels 0 |
| Globas (2012) [17] | 38 | (20/18) LE/LE&UE | F: 3d/w for 13w  
I: 40-80% HRR  
T: 30-50min/d  
T: Progressive graded, high-intensity TAEX | Peak exercise capacity (V0₂peak) +  
Walking capacity (6minWT) +  
Gait velocity (10MWT) +  
Leg strength (5 chair-raise) 0  
Balance (BBS) 0  
Self-rated mobility (RMI) +  
HRQoL (SF-12) + |
| Jin (2012) [19] | 133 | (68/65) LE/LE | F: 5x/w for 8w  
I: 50-70%HRR  
T: 40min/d (plus 30min+10min)  
T: progressive aerobic cycling with lower limb weights on the paretic limb (plus balance exercise and stretching) | Cardiorespiratory fitness (V0₂peak) +  
Walking endurance (6minWT) +  
Mobility (RMI) 0  
Muscle strength (isokinetic dynamometer) +  
Balance (BBS) 0  
Spasticity of LE (MASc) 0 |
<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>Session</th>
<th>Training Protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moore (2010)</td>
<td>20</td>
<td>4w PT-4w LT</td>
<td>F: 2-5d/w for 4w for PT and LT I: 0-85% HRR or Borg 17/ LT session T: nr T: Progressive TAEX with BWS 4w PT-4w LT-4w delay period</td>
</tr>
<tr>
<td>Quaney (2009)</td>
<td>38</td>
<td>8w</td>
<td>F: 3d/w for 8w I: 40-70% HRR T: 45min/d T: Progressive, resistive aerobic ex. on a stationary bike</td>
</tr>
<tr>
<td>Lennon (2008)</td>
<td>48</td>
<td>10w</td>
<td>F: 2d/w for 10w I: 50-60% HRR T: 30min/d T: Cycle ergometer ex.</td>
</tr>
<tr>
<td>Luft (2008)</td>
<td>113</td>
<td>6 month</td>
<td>F: 3d/w for 6 month I: 40-60% HRR T: 40min/d T: Progressive TAEX Subset (N=15) for fMRI testing</td>
</tr>
<tr>
<td>Macko (2005)</td>
<td>61</td>
<td>6 month</td>
<td>F: 3d/w for 6 month I: Treadmill at 30-40% HRR T: 35min /5min T: Stretching /Treadmill walking</td>
</tr>
<tr>
<td>Peurala (2005)</td>
<td>45</td>
<td>3w</td>
<td>F: 5d/w for 3w &amp; additional PT 5d/w for 3w for all groups I: ur T: 20min/d &amp; additional PT 55min/d for all groups T: G1: Gait trainer ex. with FES G2: Gait trainer ex. without FES G3: Practicing walking</td>
</tr>
</tbody>
</table>

Self-selected 0 and fastest velocity + (Gait MAT II) Walking endurance (12MWT) 0 Physical fitness (O2 cost +; peak treadmill speed +; VO2 0) Mobility (TUG 0; StepWatch +) Balance (BBS) 0

V02 max (metabolic stress test) + Motor learning (SRTT) for response time + / for random sequence response time 0 Balance (BBS) 0 Sensorimotor function (FMMA) 0 Conditional learning ability (PGFM) + Rule learning (WCST) 0 Selective attention (Stroop task21) 0 Visual search ability & working memory (Trail-Making Task) 0

Cardiorespiratory fitness (V02peak) + Gait velocity (10MWT) + Walking ability & endurance (6minWT) 0 fMRI +

Cardiorespiratory fitness (V02peak L/min) + Economy of gait (V02 mL/kg/min) 0 Ambulatory performance measures (10MWT 0; 6minWT +) Functional Mobility (RMI 0; WIQ +)

Walking speed (10MWT) 0 & endurance (6minWT) 0 Postural sway (force plate) 0 Spasticity of LE (MASc) + in the ankle (only in G3) but 0 in the knee and hip (all groups) Muscle force (scale 0-5) + for the ankle dorsiflexors (only in G1) and + for the hip flexors (only in G2) but 0 for the knee extensors (all groups) Motor ability (MMAS) 0 Functional independence (FIM) 0
### Potempa (1995) [44]

n=42  
(19/23)  
(LE/LE&UE)

| F | 3d/w for 10w  
| I | 20-94 watts  
| T | 30min/d  
| T | Progressive cycle ergometer  
| F | 3d/w for 10w  
| I | nr  
| T | 30min/d  
| T | PROM ex.  

Physical fitness (HR at rest and during ex. (RR interval)  
0; blood pressure (sphygmonanometer) for the systolic  
+ but for the diastolic blood pressure during ex. 0;  
metabolic parameters VO$_2$ +; VCO$_2$ +; VE + and RER 0  
(MGCESTS); workload (ergometer) +; exercise time  
(ergometer) +  
Sensorimotor function (FMMA) 0

### Resistance exercise only (n=8)

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Participants (Int/Con)</th>
<th>Intervention (Int)</th>
<th>Control (Con)</th>
<th>Primary outcomes (instrument)</th>
<th>Other outcomes (instrument)</th>
</tr>
</thead>
</table>
| Clark (2013) [57] | n=35  
(18/17)  
(LE/LE) | F: 3x/w for 5w (plus 3w)  
I: 3-4x 10 reps (nr)  
T: ECC-PRT of the paretic leg (followed by PBWSTT)  
G: 1: PRT with CKC ex. on a leg press  
G2: PRT with OKC ex. on a leg press  
G3: routine activities, no regular exercise program  
F: 3x/w for 5w (plus 3w)  
I: 3-4x 10 reps (nr)  
T: ECC-PRT of the paretic leg (followed by PBWSTT)  
| CON 0 and ECC 0 muscle strength quadriiceps (isokinetic dynamometer)  
Muscle activity of the (RF+VM) +, (ST+BF) 0 (EMG)  
Walking speed (motion analysis system) 0  
LE muscle activity (in %) of the RF +, BF +, GC + + and TA + +  
(EMG)  
Balance (Good Balance System platform: (A-P) + + & (M-L) + +  
sway velocities (mm/s))  
+ + (more significant changes in favor of G1 in post hoc analysis)  
| LE muscle endurance (6minWT) 0  
LE movement (STREAM) +  
Gait evaluations (F-SCAN; in-shoe plantar pressure mapping system) +  
Balance ability (BBS) 0  
Plantar flexor spasticity (MASc) +  
| Waldman (2013) [59] | n=24  
(12/12)  
(LE/LE) | F: 3x/w for 6w  
I: 3 sets of 70% of 1RM  
T: ur  
G: 1: PRT with CKC ex. on a leg press  
G2: PRT with OKC ex. on a leg press  
G3: routine activities, no regular exercise program  
| Isometric ankle muscle strength (DF + & PF 0) / ankle AROM (DF 0  
& PF 0) and PROM (DF +) (robotic device)  
LE muscle endurance (6minWT) 0  
LE movement (STREAM) +  
Gait evaluations (F-SCAN; in-shoe plantar pressure mapping system) +  
Balance ability (BBS) 0  
Plantar flexor spasticity (MASc) +  
| Flansbjer (2008) [52] | n=24  
(15/9)  
(LE/LE&UE) | F: 2d/w for 10w  
I: 6-8 reps in 2 sets, 80% 1RM  
T: 90min/d with 6min PRT  
T: PRT of the knee muscles  
F: 3x/w for 6w  
I: nr  
T: 60min/d  
T: stretching plantar flexors and AROM exercises for ankle mobility and strength  
| Muscle strength (dynamic + & isokinetic +; MASc 0)  
Gait performance (TUG +; Fast gait speed 0; 6minWT 0)  
Perceived participation (SIS) +  
(24/24)  
(LE/LE&UE) | F: 3d/w for 4w  
I: ur  
T: 30min/d  
T: ex. for LE strengthening (circuit class)  
F: nr  
I: nr  
T: nr  
T: Usual care  
| LE muscle strength (dynamometer) +  
LE muscle endurance (6minWT) +  
Gait velocity, cadence & stride length (GAITRite) +  
Mobility (TUG) +  
Balance (step test) +  
<p>|</p>
<table>
<thead>
<tr>
<th>Study ID</th>
<th>Participants (Int/Con)</th>
<th>Intervention (Int)</th>
<th>Control (Con)</th>
<th>Primary outcomes (instrument)</th>
<th>Other outcomes (instrument)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ouellette (2004)</td>
<td>n=42 (21/21) ((LE/LE&amp;UE)</td>
<td>F: 3d/w for 12w&lt;br&gt;I: 3 sets, 8-10 reps/set, 70% 1RM&lt;br&gt;T: nr&lt;br&gt;T: Supervised high-intensity PRT with 4 ex.</td>
<td>F: 3d/w for 12w&lt;br&gt;I: nr&lt;br&gt;T: nr&lt;br&gt;T: Bilateral ROM ex. + UE stretching</td>
<td>LE muscle strength (1RM Leg Press +; knee extension ex. +; DF-ex. in the paretic + &amp; in the non-paretic limb 0; PF-ex. +; peak muscle power Leg Press 0; knee extension ex. +)&lt;br&gt;LE muscle endurance (6minWT) 0&lt;br&gt;Stair-climb time (10-stair flight) 0&lt;br&gt;Repeated chair-rise time 0&lt;br&gt;Gait velocity (10MWT) 0&lt;br&gt;Function &amp; disability (LLFDI) 0</td>
<td>Self-perceived ability for 5 functional tasks (ESES) 0&lt;br&gt;Self-reported functional outcome (PF10) 0&lt;br&gt;Depression (GDS) 0</td>
</tr>
<tr>
<td>Kim (2001)</td>
<td>n=20 (10/10) (LE/LE)</td>
<td>F: 3d/w for 6w&lt;br&gt;I: 3 sets, 10 reps/set at RM&lt;br&gt;T: 45min/d&lt;br&gt;T: max concentric isokinetic strength training</td>
<td>F: 3d/w for 6w&lt;br&gt;I: nr&lt;br&gt;T: 45min/d&lt;br&gt;T: PROM of paretic LE</td>
<td>Isokinetic strength of LE (dynamometer) 0&lt;br&gt;Walking speed (level-walking and stair-walking) 0&lt;br&gt;HRQoL (SF-36) 0&lt;br&gt;Mental health (MSC) 0</td>
<td></td>
</tr>
<tr>
<td>Dean (2000)</td>
<td>n=9 (5/4) (LE/UE)</td>
<td>F: 3d/w for 4w&lt;br&gt;I: nr&lt;br&gt;T: 60min/d&lt;br&gt;T: Task-related strength training (circuit of 10 workstations)</td>
<td>F: 3d/w for 4w&lt;br&gt;I: nr&lt;br&gt;T: 60min/d&lt;br&gt;T: Practiced UE tasks (circuit of workstations &amp; ex.)</td>
<td>Walking speed (10MWT) + and endurance (6minWT) +&lt;br&gt;LE function; force production (TUG) +&lt;br&gt;Balance (step test) +</td>
<td></td>
</tr>
<tr>
<td>Duncan (2003)</td>
<td>n=92 (44/48) (LE&amp;UE / LE&amp;UE)</td>
<td>F: 3d/w for 12w&lt;br&gt;I: nr&lt;br&gt;T: 90min/d&lt;br&gt;T: Supervised in-home strength, endurance and balance program</td>
<td>F: nr&lt;br&gt;I: nr&lt;br&gt;T: nr&lt;br&gt;T: Usual care + completing intervention log</td>
<td>Muscular strength (dynamometer and ankle &amp; knee isometric peak torque) 0&lt;br&gt;Cardiorespiratory endurance (peak aerobic capacity) +&lt;br&gt;UE-function (WMFT) 0&lt;br&gt;LE &amp; UE motor control (Fugl-Meyer) 0&lt;br&gt;Mobility (10MWT 0, 6minWT +)&lt;br&gt;Balance (BBS) +</td>
<td></td>
</tr>
<tr>
<td>Duncan (2003)</td>
<td>n=20 (10/10) (LE&amp;UE / LE&amp;UE)</td>
<td>F: 3d/w for 12w&lt;br&gt;I: nr&lt;br&gt;T: 90min/d&lt;br&gt;T: Supervised in-home strength, endurance and balance program</td>
<td>F: nr&lt;br&gt;I: nr&lt;br&gt;T: nr&lt;br&gt;T: Usual care + completing intervention log</td>
<td>Gait speed (10MWT) + &amp; walking distance (6minWT) 0&lt;br&gt;Motor recovery LE + and UE 0 (FMMA)&lt;br&gt;ADL (BI of ADL) 0 and (LS-IADL) 0&lt;br&gt;HRQoL (SF-36) 0&lt;br&gt;Balance (BBS) 0&lt;br&gt;Upper extremity function (Jebsen test) 0</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>n</td>
<td>Intervention Details</td>
<td>Outcome Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>---</td>
<td>--------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salbach (2004) [62]</td>
<td>91</td>
<td>F: 3d/w for 6w I: nr T: 60min/d T: 10 functional tasks to strengthen the LE &amp; improve walking balance, speed &amp; distance</td>
<td>Walking capacity (6minWT) + Comfortable and max pace (5MWMT) + Mobility (TUG) 0 Balance (BBS) 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee (2010) [63]</td>
<td>48</td>
<td>F: 3d/w for 12w for all groups I: PRT: 2 sets, 8 reps/set, 50-80% 1RM Cycling: 50-80% HRR T: 30min/d for all groups T: G1: aerobic cycling + sham PRT G2: sham cycling + PRT G3: aerobic cycling + PRT G4: sham cycling + sham PRT</td>
<td>Max muscle force (pneumatic resistance machines; fixed dynamometer) +§ 0† Muscle power (peak force, velocity) +§ 0† Muscle endurance (as many repetitions as possible) +§ 0†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee (2008) [64]</td>
<td>52</td>
<td>F: 3d/w for all groups I: PRT: 2 sets, 8 reps/set, 50-80% 1 RM Cycling: 50-80% HRR T: 30min/d for all groups T: G1: aerobic cycling + sham PRT G2: sham cycling + PRT G3: aerobic cycling + PRT G4: sham cycling + sham PRT</td>
<td>Mobility (10MWMT, 6minWT) 0§ 0† Stair climbing test &amp; power +§ 0† Muscle endurance (as many repetitions as possible) +§ 0† Cardiorespiratory fitness +† (V0peak, PO peak +, peak HR) HRQoL (SF-36) +§ 0†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pang (2008) [65]</td>
<td>63</td>
<td>F: 3d/w for 19w I: ur T: 60min/d T: FAME: cardiorespiratory fitness, mobility, balance, LE strengthening</td>
<td>Walking capacity (6minWT) + V0₂ max (Excalibur cycle ergometer) + Muscle strength (dynamometer) in the paretic + &amp; in the nonparetic knee 0 Balance (ABCS 0; BBS 0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mead (2007) [66]</td>
<td>66</td>
<td>F: 3d/w for 12w I: Endurance: at least Borg 13, PRT: 4-15 reps per ex. T: 75min/d T: Progressive endurance training (cycle ergometer) &amp; PRT</td>
<td>Physical fitness (V0₂; mL/kg +; NPR 0) Mobility (RMI 0; reach 0; sit-to-stand 0; elderly mobility score 0; TUG +) ADL (FIM 0; NE-ADL 0) HRQoL (SF-36) + Depression (HADS 0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ohey (2006) [67]</td>
<td>74</td>
<td>F: 3d/w for 10w I: ur T: 90min/d T: Supervised ex. sessions: graded walking and/or cycling program + strength training</td>
<td>Walking endurance (6minWT) 0 Strength of LE muscles (dynamometer) 0 Oxygen cost (PCI 0 Activity (HAP) + Physical and mental health (SF36) 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>n</td>
<td>Intervention Details</td>
<td>Sessions as Intervention Group</td>
<td>Outcomes/Measurements</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---</td>
<td>----------------------</td>
<td>--------------------------------</td>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td>Pang (2005)</td>
<td>n=63 (32/31) (LE/UE)</td>
<td>F: 3d/w for 19w&lt;br&gt;I: ur&lt;br&gt;T: 60min/d&lt;br&gt;T: FAME: mobility, cardiopulmonary fitness, balance, LE strengthening</td>
<td>VO2 max (Excalibur cycle ergometer) +&lt;br&gt;Walking endurance (6minWT) +&lt;br&gt;Muscle strength (dynamometer) in the paretic + &amp; in the non-paretic knee 0&lt;br&gt;Activity and participation (PASPD) 0&lt;br&gt;Femoral neck bone mineral density (dual-energy x-ray absorptiometry) in the paretic + &amp; in the non-paretic leg 0&lt;br&gt;Balance (BBSS) 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carr (2003)</td>
<td>n=40 (nr/nr) (LE&amp;UE/LE&amp;UE)</td>
<td>F: 3d/w for 16w&lt;br&gt;I: Endurance: 40-70% of max wattage, Strength: 2sets, 10 reps/set&lt;br&gt;T: 20-40min/d + strength ex.&lt;br&gt;T: Aerobic &amp; strength training (same aerobic program &amp; progressively increasing strength-training)</td>
<td>VO2 max (stress test) +&lt;br&gt;Muscle strength (Biodex strength assessment): knee flexion, knee extension 0, shoulder extension/extension +&lt;br&gt;Blood lipid panel (cholesterol +, glucose 0, lipids 0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teixeira-Salmela (1999)</td>
<td>n=13 (6/7) (LE/UE)</td>
<td>F: 3d/w for 10w&lt;br&gt;I: Aerobic ex.: 50-70% HRR, Strength: 3sets, 10 reps/set, 50-80% 1RM&lt;br&gt;T: 60-90min/d&lt;br&gt;T: Graded walking &amp; stepping or cycling &amp; LE muscle strengthening</td>
<td>Comfortable gait speed (middle 22meters of 30meters) +&lt;br&gt;Activity per day (AAS) +&lt;br&gt;Quality of life (NHP) +&lt;br&gt;Comfortable stair-climbing speed +&lt;br&gt;Physical activity (HAP) +&lt;br&gt;LE muscle strength (dynamometer) +&lt;br&gt;LE spasticity of quadriceps (pendulum test) 0 &amp; ankle PF (controlled resistance to passive stretch) 0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Legend**<br ofrece significant between-group difference / 0 = non-significant difference between groups / both in favor of intervention group<br>AAP = Adelaide Activities Profile, ABAW = aggressive bracing assisted walking, ABCS = Activity-Specific Balance Confidence Scale, ADL = Activities of Daily Living, (A-L) = antero-posterior, BBS = Berg Balance Scale, BF = biceps femoris, BI = Barthel Index, BMI = Body Mass Index, BWS = body weight support, BWSTT = body weight supported treadmill training, CKC = closed kinetic chain, CMSA = Chedoke-McMaster Stroke Assessment, CON = concentric, CRS = Cardiac Risk Score, d = days, DF = dorsiflexion, ECC = eccentric, EMG = electromyography, ESES = Ewart's Self Efficacy Scale, ex. = exercise, FAI = Frenchay Activities Index, FAME = Fitness and Mobility Exercises, FIM = Functional Independence Measure, FMMA = Fugl-Meyer Motor Assessment, fMRI = functional Magnetic Resonance Imaging, GC = gastrocnemius, GDS = Geriatric Depression Scale, HADS = Hospital Anxiety and Depression Scale, HAP = Human Activity Profile, Hfmax = maximal heart rate, HR = heart rate, HRR = Heart Rate Reserve, HRQoL = Health Related Quality of Life, LE = lower extremity, LLFDI = Late Life Function and Disability Instrument, LS-IADL = Lawton Scale of Instrumental Activities of Daily Living, LT= locomotor training, MAS = Motor Assessment Scale, MASC = Modified Ashworth Scale, max = maximum, min = minutes, minWT = minute walking test, (M-L) = medio-lateral, MSC = Mental Health Component Summary Score, MWT = meter walking test, NE-ADL = Nottingham Extended Activities of Daily Living, NPR = Nottingham Power Rig, nr = not reported, OKC = open kinetic chain, PASPD = Physical Activity and Disability Scale, PASIPD = Physical Activity for Individuals with Physical Disability, PBWSTT = Partial body weight supported treadmill training, PCI= Physiological Cost Index, PF = plantar flexion, PF10 = Physical Functioning Subscale, PGIC = Patient Global Impression of Change, PNF = Proprioceptive Neuromuscular Facilitation Patterns, PO = power output, PPA = Physiological Profile Assessment, PROM = Passive Range of Motion, PRF = Progressive resistance training, PT = Physical Therapy, reps = repetitions, RF = rectus femoris, (1/RM = (One) Repetition Maximum, RMI = Rivermead Mobility Index, RPE = Borg Rate of Perceived Exertion, SF = Short Form, SF36 = 36-item Short-Form Health Survey, SIS = Stroke Impact Scale, ST = semitendinosus, STREAM = Stroke Rehabilitation Assessment of Movement, TA = tibialis anterior, TAEX = Aerobic Treadmill Exercise, TUG = Timed up and go test, UE = upper extremity, ur = unclear reporting, VM = vastus medialis, VO2 = oxygen consumption, w = weeks, WIQ = Walking Impairment Questionnaire, WMFT = Wolf Motor Function Test, FES = Functional Electric Stimulation
Additional file 2 Application of principles of exercise training.

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Specificity</th>
<th>Progression</th>
<th>Overload</th>
<th>Initial values</th>
<th>Reversibility</th>
<th>Diminishing returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbeau (2003) [29]</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Visintin (1998) [31]</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Mayo (2013) [45]</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dean (2012) [42]</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bang (2013) [40]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Carda (2013) [41]</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Jin (2013) [43]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tang (2013) [46]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Globas (2012) [32]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Jin (2012) [44]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Moore (2010) [33]</td>
<td>?</td>
<td>+</td>
<td>-</td>
<td>?</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Quaney (2009) [34]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Lennon (2008) [35]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table S1a Aerobic exercise only (n=18).

sub-acute stroke survivors

both

chronic stroke survivors
<table>
<thead>
<tr>
<th>Study ID</th>
<th>Specificity</th>
<th>Progression</th>
<th>Overload</th>
<th>Initial values</th>
<th>Reversibility</th>
<th>Diminishing returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luft (2008) [36]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Macko (2005) [37]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Peurala (2005) [38]</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>-</td>
</tr>
<tr>
<td>Potempa (1995) [39]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lee (2013) [53]</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Waldman (2013) [54]</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Flansbjer (2008) [47]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Yang (2006) [48]</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ouellette (2004) [49]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kim (2001) [50]</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dean (2000) [51]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

Table S1b Resistance exercise only (n=8).
Table S1c Aerobic and resistance exercises (n=11).

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Specificity</th>
<th>Progression</th>
<th>Overload</th>
<th>Initial values</th>
<th>Reversibility</th>
<th>Diminishing returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duncan (2003) [55]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Duncan (1998) [56]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lee (2010) [58]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Lee (2008) [59]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pang (2008) [60]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mead (2007) [61]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Olney (2006) [62]</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Pang (2005) [63]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Carr (2003) [64]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Teixeira-Salmela (1999) [65]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Additional file 3 Rating of risk of bias in individual studies.

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Random sequence generation (selection bias)</th>
<th>Allocation concealment (selection bias)</th>
<th>Blinding of participants and personnel (performance bias)</th>
<th>Blinding of outcome assessment (detection bias)</th>
<th>Incomplete outcome data (attrition bias)</th>
<th>Selective reporting (reporting bias)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mayo (2013) [45]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Dean (2012) [42]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Bang (2013) [40]</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Carda (2013) [41]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Tang (2013) [46]</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Globas (2012) [32]</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Moore (2010) [33]</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Quaney (2009) [34]</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Table S2a Aerobic exercise only (n=18).
<table>
<thead>
<tr>
<th>Study ID</th>
<th>Random sequence generation (selection bias)</th>
<th>Allocation concealment (selection bias)</th>
<th>Blinding of participants and personnel (performance bias)</th>
<th>Blinding of outcome assessment (detection bias)</th>
<th>Incomplete outcome data (attrition bias)</th>
<th>Selective reporting (reporting bias)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lennon (2008) [35]</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Macko (2005) [37]</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Peurala (2005) [38]</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>-</td>
</tr>
</tbody>
</table>

**Table S2b Resistance exercise only (n=8).**
### Table S2c Aerobic and resistance exercises (n=11).

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Random sequence generation (selection bias)</th>
<th>Allocation concealment (selection bias)</th>
<th>Blinding of participants and personnel (performance bias)</th>
<th>Blinding of outcome assessment (detection bias)</th>
<th>Incomplete outcome data (attrition bias)</th>
<th>Selective reporting (reporting bias)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sub-acute stroke survivors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duncan (1998) [56]</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>both</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salbach (2004) [57]</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Lee (2010) [58]</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Lee (2008) [59]</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pang (2008) [60]</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Mead (2007) [61]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>chronic stroke survivors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olivey (2006) [62]</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pang (2005) [63]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
References


24. World Health Organization, health topics: chronic diseases [http://www.who.int/topics/chronic_diseases/en/]


52. Clark DJ, Patten C: Eccentric versus concentric resistance training to enhance neuromuscular activation and walking speed following stroke. Neurorehabil Neural Repair 2013, 27(4):335-344.


Chapter 3

Concurrent validity and test-retest reliability of the Virtual Peg Insertion Test to quantify upper limb function in patients with chronic stroke

Follow-up exergames intervention results of a figure copying task of the Neglect Test (NET) of P5, male, 53 years old, 14 weeks onset right-hemispheric stroke and visuo-spatial neglect

This chapter has been published as: Tobler-Ammann BC, de Bruin ED, Fluet MC, Lambercy O, de Bie RA, Knols RH. Concurrent validity and test-retest reliability of the Virtual Peg Insertion Test to quantify upper limb function in patients with chronic stroke. Journal of NeuroEngineering and Rehabilitation 2016;13(1):8.
Abstract

**Background:** Measuring arm and hand function of the affected side is vital in stroke rehabilitation. Therefore, the Virtual Peg Insertion Test (VPIT), an assessment combining virtual reality and haptic feedback during a goal-oriented task derived from the Nine Hole Peg Test (NHPT), was developed. This study aimed to evaluate (1) the concurrent validity of key outcome measures of the VPIT, namely the execution time and the number of dropped pegs, with the NHPT and Box and Block Test (BBT), and (2) the test-retest-reliability of these parameters together with the VPIT’s additional kinetic and kinematic parameters in patients with chronic stroke.

**Methods:** The three tests were administered on 31 chronic patients with stroke in one session (concurrent validity), and the VPIT was retested in a second session 3–7 days later (test-retest reliability). Spearman rank correlation coefficients (ρ) were calculated for assessing concurrent validity, and intraclass correlation coefficients (ICCs) were used to determine relative reliability. Bland-Altman plots were drawn and the smallest detectable difference (SDD) was calculated to examine absolute reliability.

**Results:** For the 31 included patients, 11 were able to perform the VPIT solely via use of their affected arm, whereas 20 patients also had to utilize support from their unaffected arm. For n = 31, the VPIT showed low correlations with the NHPT (ρ = 0.31 for time (T_ex[s]); ρ = 0.21 for number of dropped pegs (N_dp)) and BBT (ρ = -0.23 for number of transported cubes (N_tc); ρ = -0.12 for number of dropped cubes (N_dc)). The test-retest reliability for the parameters T_ex[s], mean grasping force (F_mg[N]), number of zero-crossings (N_zcgo/return) and mean collision force (F_cmean[N]) were good to high, with ICCs ranging from 0.83 to 0.94. Fair reliability could be found for F_mgreturn (ICC=0.75) and trajectory error (E_trajgo[cm]) (0.70). Poor reliability was measured for E_trajreturn[cm] (0.67) and N_dp (0.58). The SDDs were: T_ex = 70.2 seconds, N_dp = 0.4 pegs; F_mggo/return = 3.5/1.2 Newton; N_zcgo/return = 0.2/1.8 zero-crossings; E_trajgo/return = 0.5/0.8 cm; F_cmean = 0.7 Newton.

**Conclusions:** The VPIT is a promising upper limb function assessment for patients with stroke requiring other components of upper limb motor performance than the NHPT and BBT. The high intra-subject variation indicated that it is a demanding test for this stroke sample, which necessitates a thorough introduction to this assessment. Once familiar, the VPIT provides more objective and comprehensive measurements of upper limb function than conventional, non-computerized hand assessments.
**Keywords:** Virtual Peg Insertion Test, upper limb function, stroke, concurrent validity, test-retest reliability

**Background**

Upper limb function relies on the delicate interaction between hand and brain, defining our ability to perform activities of daily living (ADL) [1]. Brain injury caused by cerebrovascular accident may result in reduced cerebral hand representation over time [1]. The consequences for patients after a stroke are tremendous and ADL become challenging. However, building on brain plasticity and intensive training, hand representation in the brain can be increased again [2]. Novel upper limb training modalities include the use of virtual reality (VR), robotics and computer gaming, all of which can provide ecologically valid, intensive and task specific training [3]. It is essential to regularly measure motor function in order to document changes in upper limb performance during the course of rehabilitation. Such regular evaluation also informs adaptation of therapy settings, as needs arise. To date, there are several assessments measuring upper limb motor function, such as the Box and Block Test (BBT) [4, 5] and the Nine Hole Peg Test (NHPT) [6, 7] which measure gross and fine manual dexterity respectively. Additionally, the use of VR and robotic devices are not only an effective alternative to conventional therapy [8, 9], but can also be used for precise and objective concurrent assessment of attributes such as motor function, cognition and ADL [10–16].

A promising assessment tool for measuring upper limb function is the Virtual Peg Insertion Test (VPIT), a computer-assisted assessment. The task is that of the NHPT, but since there is no precision grip required, the movement has more similarities to the BBT [17]. The VPIT allows measurement of three-dimensional hand position and orientation as well as grasp force during the accomplishment of a goal-oriented task consisting of grasping, transporting and inserting nine virtual pegs into the nine holes of a virtual pegboard. This is achieved by positioning and controlling a grasping force applied to grasping a handle instrumented with force sensors mounted on a PHANTOM Omni haptic device (Geomagic, USA). The clinical practicability and measurement properties of the VPIT were tested in patients with Multiple Sclerosis (MS) [18] and pilot-tested in patients with Autosomal Recessive Spastic Ataxia of Charlevoix-Saguenay (ARSACS) [19]. Both patient groups were significantly less coordinated and slower than age-matched healthy subjects when evaluated using the VPIT. The
preliminary results of both studies illustrate the feasibility of using the VPIT in both MS and ARSACS patients, and underline the potential of this test to evaluate upper limb motor function. To date, there is only one pilot study reporting initial evaluation of the VPIT outcome measures in a group of four chronic patients with stroke, showing significant differences in grasping force control and upper limb movement patterns compared to healthy subjects [17]. However, the relation between the analyzed performance parameters during the VPIT and impaired function needs to be further established. A way to achieve this would be to evaluate the validity and reliability of the VPIT parameters for this population. As the VPIT combines the characteristics of the conventional NHPT and the BBT, one option could be to validate their mutual parameters, namely the execution time (Tex[s]) and the number of dropped pegs (Ndp). As the VPIT provides in total nine different outcome measures including kinematic and kinetic parameters quantifying movement coordination, smoothness, upper limb synergies and force control, we decided to evaluate all of these for their reliability when administered twice in a test and retest procedure. Therefore, this study aimed to evaluate (1) the concurrent validity of key outcome measures of the VPIT, namely the execution time and the number of dropped pegs, with the NHPT and BBT, and (2) the test-retest reliability of these parameters together with the VPIT’s additional kinetic and kinematic parameters, in patients with chronic stroke.

Methods

Apparatus
In the VPIT, pegs and holes are displayed in a virtual environment and can be felt through a haptic interface. Subjects are asked to move the handle of the haptic device to grasp, displace and release the pegs (Fig. 1). The aim is to grasp, transport and place all pegs in the holes as fast as possible. To evaluate upper limb function during the execution of the task, the VPIT software records and computes nine specific parameters. These are (Fig. 1):
The two primary outcome parameters evaluating the overall arm and hand functional ability of the patient are: (1) Tex[s]: time to execute the task from the approach to the first peg to the insertion of the last peg. (2) Ndp: number of times a peg is dropped during the transport. Secondary outcome parameters pertain to four kinematic and three kinetic parameters.
The four kinematic parameters are: (3–4) $E_{\text{traj}}[\text{cm}]_{\text{go}}/\text{return}$: trajectory error measured as the distance between the actual trajectory and the ideal straightest trajectory projected on the horizontal plane. This parameter is used to evaluate movement accuracy and upper limb motor synergies during a movement requiring the simultaneous control of shoulder, elbow and hand [20–22]. (5–6) $N_{zc}[1/\text{s}]_{\text{go}}/\text{return}$: the time-normalized number of zero-crossings (i.e., change of sign) of the acceleration during the gross movement from a peg to a hole, or from a hole to a peg. This parameter provides an estimate of the number of sub-movements a point-to-point movement is composed of, which is a commonly accepted measure of movement smoothness and upper limb coordination [19, 23, 24].

The three kinetic parameters are: (7–8) $F_g[N]_{\text{go}}/\text{return}$: average grasping force calculated during the transport of one peg and for the return trajectory. This parameter evaluates factors such as force control. (9) $F_{\text{cmean}}[N]$: mean collision force exerted against the virtual pegboard.

The “go” refers to the trajectory during the outbound trip when a peg is transported to the hole, and the “return” stands for the way back from the hole to approach a new peg. While performing the VPIT, the patient additionally receives visual feedback regarding the force applied on the peg being held: the cursor is yellow when no peg is held, turns orange to indicate that it is properly aligned with a peg, green when a peg is currently held and red when excessive grasping force is applied to the handle but no peg is held. Parameters are computed offline and no feedback related to the user’s performance is displayed during or after the test. For a more detailed description of the apparatus, we refer to Fluet et al. (2011) [17].
Fig. 1 VPIT setup and upper limb function parameters. Tex[s]: execution time in seconds; Ndp: number of dropped pegs during transport. Kinematic parameters: Etraj[cm]go/return: trajectory error; Nzc[1/s]go/return: number of zero-crossings of the acceleration. Kinetic parameters: Fg[N]go/return: mean grasping force; Fcmean[N]: mean collision force

**Conventional tests**

The NHPT version produced by Smith & Nephew Rehabilitation, Inc., Wisconsin, USA, was used. It consists of a plastic board with a shallow round dish to contain the pegs on one end of the board and the nine holes in a 3x3 grid on the opposite end. Initially, the participant had to use their affected hand to grasp, one by one, the nine pegs from the dish, inserting each one into a hole until all pegs are placed. After this the participant then has to replace each of the pegs back into the dish. All of this is carried out as fast as possible [7]. The test was timed, with a stopwatch, from the moment the participant touched the first peg until the moment when all pegs were removed from the holes. As the VPIT stops when all pegs are placed into the holes, we measured an intermediate time when all pegs were placed into the holes of the NHPT board. We labelled this parameter “time point 1” (TP1). All dropped pegs were counted and noted on the case report form (CRF).

The BBT consists of 50 wooden blocks (cubes of 2.54 cm on each side) placed in a wooden box that has 2 equal-sized compartments that are separated by a central wooden partition of 15.2 cm height [5]. The participant was instructed to use their affected hand to move blocks, one by one, from one compartment to the other,
evaluating the maximum number which could be transferred within 1 min. The transported blocks were counted during the test duration with the help of a counter clicker. All dropped blocks were noted on the CRF.

**Participants**

In total, 31 patients with stroke were consecutively recruited in this study. The occupational therapy (OT) outpatient practices that are registered as specialized in neurorehabilitation in the canton of Bern, Switzerland, were contacted to recruit patients with chronic stroke. The inclusion criteria were (1) a stroke diagnosis at least 6 months before study inclusion confirmed by a physician, (2) ≥ 18 years of age, (3) ability to communicate in German language, (4) capable of sitting in a (wheel-)chair with a backrest for up to 90 min, (5) able to lift and hold the arm in 90° of elbow flexion and 45° of shoulder abduction and (6) able to grasp a wooden block of 2.54 cm on each side as used in the BBT [25]. Exclusion criteria were (1) a diagnosis of a brain injury other than stroke, (2) a diagnosed neglect, aphasia or hemianopsia and, (3) non-controlled medical conditions (chronic pain, drug abuse). All included patients were screened for dementia and stereopsis and their handedness was assessed. A Mini Mental State Examination (MMSE) of at least 20 points (light dementia or better) was required to participate in the study [26]. The stereopsis of each patient was evaluated with the Lang Stereotest 1 (LST), which shows three objects differing in disparity and perceived distance: a cat, a star and a car [27, 28]. It has a high predictive value for stereopositivity in adults [28], which we assumed would be of importance in conducting the VPIT, which is represented in a 3D virtual environment on a computer screen. The study was approved by the Ethics Committee (KEK-Nr. 119/13) of the canton of Bern (Switzerland). All subjects gave their informed consent prior to study entry.

**Procedures**

The measurements for the patients with stroke took place at the patients' outpatient OT practice or in their home environment. The setting and test instructions for the BBT and NHPT corresponded to the standards set by Mathiowetz et al. (1985) [5, 6], translated into German by Schädler et al. (2011) [29]. The procedures followed for the VPIT were as described by Fluet et al. (2011) [17], with the following adaptations: 1) only 3 repetitions of the test were carried out as opposed to the 5 suggested by Fluet et al. (2011); 2) The force threshold to grasp and release pegs was set to 2 Newtons.
This force threshold was empirically tested in our previous work with neurological patients, where it was shown to be adequate for most participants with mild to moderate hand impairment to perform the task [18, 19]; and 3) participants who needed assistance for the affected arm to perform the test were allowed to do so. The third adaptation was not initially planned, but proved necessary to implement after we observed major difficulties in performing the VPIT during the test trial in some participants (e.g., cursor alignment or regulation of the grasping force while holding the handle not precise enough). The participants were requested to first perform the BBT and the NHPT as per protocol. In some we noticed that participants needed support for the VPIT while the 2 clinical tests were already completed. It was too much due to fatigue to repeat them. The validity of this adaptation was addressed by a later subgroup analysis (see results section).

To minimize the effect of the required adaptation to a novel tool (VPIT), we gave patients a test trial prior to the measurements (on both days), where they were given all the time they needed to explore the virtual environment and get familiar to the functioning of the test. According to our previous work, this proved to be sufficient for patients to understand the task and the use of the VPIT.

All participants were tested twice with a 3–7 day interval between assessments. This time interval (mean 6.4 ± 0.5 days) was defined to minimize any learning effect that may occur from repeating the tests within a short time frame. During the first assessment session, demographic data were collected and the MMSE, FLANDERS and LST were carried out. Then, motor tests were performed in the following order of (1) BBT, (2) NHPT and (3) VPIT. For each test, one test trial (not timed) was allowed, which was then followed by three repetitions of the test. If needed, participants could have a rest between the motor test performances. All tests were done with the affected arm only where possible, with participants being allowed to use both hands for the VPIT as necessary. Therefore, depending on the result of the VPIT test trial (support of the affected arm needed/not needed), the following three repetitions of the VPIT were all performed accordingly. The second assessment session was composed of the test trial followed by three repetitions of the VPIT.

Data analysis

SPSS version 22.0 (SPSS Inc, Chicago, Illinois) was used for data analysis. The study population and clinical characteristics were defined adopting descriptive statistics. The
average of the 3 trials of each test was calculated and used for data analysis, as test-retest reliability is highest in all tests when the mean of three trials is used; lower correlations are known to occur when one trial or the highest score of three trials are utilized [30]. Normality of data was evaluated using the Shapiro Wilk test [31]. The level of statistical significance was set to $p \leq 0.05$.

Validity
Concurrent validity was assessed by determining Spearman’s rank correlation coefficient ($\rho$) for the relationship between the VPIT/BBT and the VPIT/NHPT, respectively [31], namely for the parameters measuring a similar construct in both tests: (1) mean time in seconds ($\text{Tex}[s]\text{VPIT} / \text{Tex}[s]\text{NHPT}$), (2) mean number of dropped pegs ($\text{NdpVPIT} / \text{NdpNHPT}$), (3) mean number of dropped pegs/cubes ($\text{NdpVPIT} / \text{NdcBBT}$) and (4) mean number of transported cubes ($\text{Tex}[s]\text{VPIT} / \text{NtcBBT}$). The following correlation classification was used: no or very low: $\rho = 0–0.25$; low: $\rho = 0.26–0.40$; moderate: $\rho = 0.41–0.69$; high: $\rho = 0.70–0.89$; very high: $\rho = 0.90–1.0$ [32]. To measure the correlation between the mean number of dropped pegs/cubes during the VPIT/NHPT and VPIT/BBT, Cohen’s kappa was computed using GraphPad software (www.graphpad.com/quickcalcs) [33]. We hypothesized that (1) the correlations between $\text{Tex}[s]\text{VPIT} / \text{Tex}[s]\text{NHPT}$ and $\text{Tex}[s]\text{VPIT} / \text{NtcBBT}$ would be high for chronic patients with stroke ($0.70–0.89$). We further hypothesized that (2) the correlations between the number of pegs dropped during the VPIT and the number of pegs/cubes dropped during the conventional NHPT and BBT for chronic patients with stroke would be moderate ($0.41–0.69$).

Reliability
Relative reliability was determined by calculating intra-class correlation coefficients (ICCs) separately for average measures. In particular, we used the ICC 2 (A,k) formula (two-way mixed effects model where people effects are random and measures effects are fixed; k because the average of the three tests was used) [34, 35]. We selected the option “absolute agreement” in order to take into account the systematic error between raters, as there were two raters (BCT and JH) involved in data collection [36]. In this study, the same rater performed the test and the re-test procedure with a single participant. Furthermore, intensive pilot-testing was performed prior to data collection by both raters. The following classification was used: $0.90–0.99$, high reliability; $0.80–$
0.89, good reliability; 0.70–0.79, fair reliability; 0.69 or below, poor reliability [37, 38]. We anticipated that the relative reliability of the 9 parameters measured with the VPIT would be good (ICC ≥ 0.80) [37].

To calculate absolute reliability, the ICCs were complemented by the Bland-Altman analysis, which can be used to show variation (or the magnitude of difference) of repeated measurements [39, 40]. The plots show the difference between test sessions 2 and 1 against the mean of the two test sessions for each subject [41, 42]. A free sample of the MedCalc statistical software version 14.8.1 (www.medcalc.org) was used to draw the Bland-Altman plots. The degree of heteroscedasticity was measured by calculating Kendall’s tau correlation (τ) between the absolute differences and the corresponding means of each VPIT parameter. When a positive τ > 0.1 was found, the data were considered heteroscedastic. When τ < 0.1 or negative, the data were considered homoscedastic [43]. The data were logarithmically or square root transformed when heteroscedasticity was found [44, 45]. Thereafter, we calculated Kendall’s tau again; if Ttrans decreased - indicating a more homoscedastic distribution of the data - reliability was analyzed using the transformed parameters [43].

To quantify the precision of individual scores on a test, we calculated the Standard Error of Measurement (SEM), using the formula \( \text{SEM} = \sigma \sqrt{1 - \text{ICC}} \), with \( \sigma \) being the total variance of the scores from all subjects [34, 36]. We then calculated the Smallest Detectable Difference (SDD) based on the SEM, as follows: \( \text{SDD} = \text{SEM} \times 1.96 \times \sqrt{2} \).

As a last step we calculated: \( \text{SDD\%} = \frac{\text{SDD}}{\text{grand mean}} \times 100 \). The grand mean is the mean of the means of each VPIT parameter. As agreement parameters (SDDs) are expressed on the actual scale of the assessments, they allow clinical interpretation of the results [34, 36]. Furthermore, the SDD\% can be used to compare test-retest reliability among tests [25]. We hypothesized that the SDD\% is ≤ 54 % of the mean average values of the VPIT, as Chen et al. (2009) found an SDD\% of 54 % for the affected hand using the NHPT in patients with stroke [25].

**Results**

The demographic and clinical characteristics of the participants are summarized in Table 1 and the results of the achieved scores (BBT, NHPT, VPIT) in Table 2. From the 33 chronic patients with stroke initially recruited, two participants dropped out: one only completed 2 of the 3 required VPIT test trials due to poor physical health, whereas
the second could not perform the VPIT task with the affected hand. In 26 participants (83.9 %), stroke had occurred for the first time, while 5 (16.1 %, 4 men and 1 woman) had suffered at least two stroke events. Although all participants fulfilled the inclusion criteria, 11 patients were able to perform all motor tests solely with the affected arm, whereas 20 patients also required support from the non-affected arm to perform the VPIT, as they possibly were fatigued by the duration of the tests. This support was needed to increase stability of the affected arm and therefore motor control of the affected hand. To account for these differences in test performance (with and without support), we decided to additionally conduct a subgroup analysis, subgroup 1 (n = 11) being without support of the non-affected arm and subgroup 2 (n = 20): being those who required support. Shapiro Wilk testing indicated too much difference within the validity and test-retest reliability data to be normally distributed.

**Table 1** Demographics and clinical characteristics of participants (n=31)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, n (%)</td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>8 (25.8)</td>
</tr>
<tr>
<td>male</td>
<td>23 (74.2)</td>
</tr>
<tr>
<td>Age, in years</td>
<td></td>
</tr>
<tr>
<td>mean ± SD</td>
<td>62.7 ± 15.1</td>
</tr>
<tr>
<td>Time onset stroke, in months</td>
<td></td>
</tr>
<tr>
<td>mean ± SD</td>
<td>51.1 ± 82.0</td>
</tr>
<tr>
<td>Affected hand, n (%)</td>
<td></td>
</tr>
<tr>
<td>right</td>
<td>17 (54.8)</td>
</tr>
<tr>
<td>left</td>
<td>14 (45.2)</td>
</tr>
<tr>
<td>Learned left-handedness due to stroke, n (%)</td>
<td>7 (22.6)</td>
</tr>
<tr>
<td>Affected hand = dominant hand, n (%)</td>
<td>18 (58.1)</td>
</tr>
<tr>
<td>LST, recognized images, n (%)</td>
<td></td>
</tr>
<tr>
<td>3 (out of 3)</td>
<td>5 (16.1)</td>
</tr>
<tr>
<td>2 (out of 3)</td>
<td>12 (38.7)</td>
</tr>
<tr>
<td>1 (out of 3)</td>
<td>11 (35.5)</td>
</tr>
<tr>
<td>0 (out of 3)</td>
<td>3 (9.7)</td>
</tr>
<tr>
<td>MMSE</td>
<td></td>
</tr>
<tr>
<td>mean ± SD</td>
<td>27.5 ± 2.4</td>
</tr>
</tbody>
</table>

*LST* Lang Stereotest 1, *MMSE* Mini Mental Status Examination *n* number, *SD* Standard Deviation
Table 2 Achieved scores of the VPIT, NHPT and BBT

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Mean ± SD of achieved scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean of 3 trials</td>
<td>All patients (n=31)</td>
</tr>
<tr>
<td>$T_{ex}[s]$ VPIT</td>
<td>119.7 ± 72.1</td>
</tr>
<tr>
<td>$T_{ex}[s]$ NHPT</td>
<td>54.4 ± 28.9</td>
</tr>
<tr>
<td>$T_{ex}[s]$ NHPT(TP1)</td>
<td>39.4 ± 20.5</td>
</tr>
<tr>
<td>$N_t$ BBT</td>
<td>35.8 ± 17.4</td>
</tr>
<tr>
<td>$N_{dp}$ VPIT</td>
<td>0.5 ± 1.1</td>
</tr>
<tr>
<td>$N_{dp}$ NHPT</td>
<td>0.8 ± 0.9</td>
</tr>
<tr>
<td>$N_{dc}$ BBT</td>
<td>0.7 ± 0.9</td>
</tr>
</tbody>
</table>

$T_{ex}[s]$ execution time in seconds, VPIT Virtual Peg Insertion Test, NHPT Nine Hole Peg Test, TP1 time point 1 (time stopped at the point when all 9 pegs were put in the 9 holes (without replacing them in the dish)), BBT Box and Block Test, $N_t$ number of transported cubes, $N_{dp}$ number of dropped pegs, $N_{dc}$ number of dropped cubes, SD Standard Deviation

Concurrent validity

The results of the concurrent validity calculations (ranges) are presented in Table 3. The correlations between the VPIT and BBT/NHPT were low ($\rho = -0.23$–$0.31$) and non-significant ($p = 0.09$–$0.51$) in all parameters. The correlations between the VPIT and BBT/NHPT for subgroup 1 were moderate ($\rho = -0.41$–$0.61$) and non-significant ($p = 0.07$–$0.60$). The correlations between the VPIT and BBT/NHPT for subgroup 2 were low ($\rho = -0.21$–$0.35$) and non-significant ($p = 0.13$–$0.51$). The strength of agreement for $N_{dp}$VPIT / $N_{dp}$NHPT and no (zero) dropped pegs in the VPIT and NHPT was considered to be poor (Kappa = 0.041; SE of kappa = 0.176; 95 % CI = $-0.30$–$0.39$) with 16 (51.6 %) observed agreements. Accordingly, the kappa agreement for the $N_{dp}$VPIT / $N_{dc}$BBT and no (zero) dropped pegs/cubes during those tests was poor (Kappa = 0.189; SE of kappa = 0.157; 95 % CI = $-0.12$–$0.50$) with 18 (58.1 %) observed agreements.
Table 3 Concurrent validity of the VPIT with the NHPT and the BBT

<table>
<thead>
<tr>
<th>Concurrent validity parameters</th>
<th>Spearman’s rank correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All patients (n=31)</td>
</tr>
<tr>
<td></td>
<td>Subgroup 1: without support (n=11)</td>
</tr>
<tr>
<td></td>
<td>Subgroup 2: with support (n=20)</td>
</tr>
<tr>
<td>Mean of 3 trials</td>
<td></td>
</tr>
<tr>
<td>$T_{ex}[s]$ VPIT / $T_{ex}[s]$ NHPT</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>0.09*</td>
</tr>
<tr>
<td></td>
<td>0.07*</td>
</tr>
<tr>
<td></td>
<td>0.13*</td>
</tr>
<tr>
<td>$T_{ex}[s]$ VPIT / $T_{ex}[s]$ NHPT$_{(TP1)}$</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>0.10*</td>
</tr>
<tr>
<td></td>
<td>0.47*</td>
</tr>
<tr>
<td></td>
<td>0.16*</td>
</tr>
<tr>
<td>$T_{ex}[s]$ VPIT / $N_t$ BBT</td>
<td>-0.23</td>
</tr>
<tr>
<td></td>
<td>-0.41</td>
</tr>
<tr>
<td></td>
<td>-0.21</td>
</tr>
<tr>
<td></td>
<td>0.22*</td>
</tr>
<tr>
<td></td>
<td>0.22*</td>
</tr>
<tr>
<td></td>
<td>0.37*</td>
</tr>
<tr>
<td>$N_{dp}$ VPIT / $N_{dp}$ NHPT</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>-0.18</td>
</tr>
<tr>
<td></td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>0.26*</td>
</tr>
<tr>
<td></td>
<td>0.60*</td>
</tr>
<tr>
<td></td>
<td>0.49*</td>
</tr>
<tr>
<td>$N_{dp}$ VPIT / $N_{dc}$ BBT</td>
<td>-0.12</td>
</tr>
<tr>
<td></td>
<td>-0.18</td>
</tr>
<tr>
<td></td>
<td>-0.16</td>
</tr>
<tr>
<td></td>
<td>0.51*</td>
</tr>
<tr>
<td></td>
<td>0.60*</td>
</tr>
<tr>
<td></td>
<td>0.51*</td>
</tr>
</tbody>
</table>

*p-values: level of significance: $p \leq 0.05$

$T_{ex}[s]$ execution time in seconds, $VPIT$ Virtual Peg Insertion Test, $NHPT$ Nine Hole Peg Test, $TP1$ time point 1 (time stopped at the point when all 9 pegs were put in the 9 holes (without replacing them in the dish)), $BBT$ Box and Block Test, $N_t$ number of transported cubes, $N_{dp}$ number of dropped pegs, $N_{dc}$ number of dropped cubes, $SD$ Standard Deviation

**Test-retest reliability**

The 9 test-retest reliability parameters of the VPIT are presented in Table 4 for the whole study population and in Table 5 (subgroup 1) and 6 (subgroup 2) for the subgroups. All VPIT parameters are illustrated in Fig. 2 by Bland-Altman plots. For the whole stroke sample, the correlations for the 5 parameters $T_{ex}[s]$, $F_{ggo}[N]$, $N_{zc}[1/s]go/return$ and $F_{cmean}[N]$ were good to high (ICCs = 0.83–0.94, SEMs = 0.07–0.63, except for the $T_{ex}[s]$ with SEM = 25.34). Fair reliability was found for the parameters $F_{greturn}$ (ICC = 0.75, SEM = 0.43) and $E_{trajgo}[cm]$ (ICC = 0.70, SEM = 0.19). Poor reliability was measured for $E_{trajreturn}[cm]$ (ICC = 0.67, SEM = 0.29) and $N_{dp}$ (ICC = 0.58, SEM = 0.14). The SDD were $\leq 54$ % in all VPIT parameters (SDD% = 1.37–21.42) except for $T_{ex}[s]$ with SDD% = 434.5 %.

The subgroup analysis of subgroup 1 ($n = 11$) showed good ICCs for the 5 VPIT parameters $F_{ggo}[N]$, $N_{zc}[1/s]go/return$, $E_{trajgo}$ and $F_{cmean}$ (ICCs = 0.82–0.89, SEMs = 0.07–1.79). Fair reliability was found in the parameter $F_{greturn}[N]$ (ICC = 0.78, SEM
While Tex[s], Ndp and Etrajreturn[cm] showed poor reliability (ICCs = 0.18–0.35 with SEMs = 0.20–37.01). All VPIT parameters showed SDD of ≤ 54% (SDD% = 1.27–32.97) except for Tex[s] with SDD% = 682.87%.

Subgroup 2 (n = 20) showed high ICCs in 5 out of 9 VPIT parameters (Tex[s], Fggo[N], Nzc[1/s]go/return and Fcmean[N] with ICCs = 0.92–0.97, SEMs = 0.02–19.25). Fair reliability was found in Fgreturn[N] (ICC = 0.73, SEM = 0.40). The remaining 3 parameters (Ndp and Etrajgo/return[cm]) showed poor reliability (ICCs = 0.49–0.69, SEMs = 0.10–0.22). Within subgroup 2, all VPIT parameters showed SDD% values of ≤ 54% (SDD% = 0.26–15.08%) except for Tex[s] with SDD% = 317.76.
Table 4 Test-retest reliability parameters of the whole stroke sample (n=31)

<table>
<thead>
<tr>
<th>VPIT parameters</th>
<th>Test Mean ±SD</th>
<th>Retest Mean ±SD</th>
<th>Test-Retest Mean difference ±SD</th>
<th>Kendall's tau correlation with transformed data</th>
<th>ICC (95% CI)</th>
<th>SEM agreement</th>
<th>SDD</th>
<th>SDD%</th>
</tr>
</thead>
<tbody>
<tr>
<td>T_ex [s]</td>
<td>119.71 ± 72.05</td>
<td>98.71 ± 50.77</td>
<td>21.01 ± 45.25</td>
<td>0.01</td>
<td>0.83 (0.61 – 0.92)</td>
<td>25.34</td>
<td>70.18</td>
<td>434.50</td>
</tr>
<tr>
<td>N_dp</td>
<td>0.54 ± 1.06</td>
<td>0.13 ± 0.19</td>
<td>-0.29 ± 1.55</td>
<td>0.16a</td>
<td>0.51 (-0.01 – 0.76)</td>
<td>0.95</td>
<td>2.62</td>
<td>16.22</td>
</tr>
<tr>
<td>F_g go [N]</td>
<td>8.98 ± 5.15</td>
<td>8.54 ± 4.96</td>
<td>0.44 ± 2.43</td>
<td>0.10</td>
<td>0.94 (0.87 – 0.97)</td>
<td>0.14</td>
<td>3.46</td>
<td>21.42</td>
</tr>
<tr>
<td>F_g return [N]</td>
<td>1.66 ± 0.83</td>
<td>1.60 ± 0.91</td>
<td>0.06 ± 0.80</td>
<td>0.08</td>
<td>0.75 (0.47 – 0.88)</td>
<td>0.43</td>
<td>1.18</td>
<td>7.31</td>
</tr>
<tr>
<td>N_zc [1/s] go</td>
<td>11.09 ± 1.52</td>
<td>10.80 ± 1.70</td>
<td>0.28 ± 0.89</td>
<td>0.21a</td>
<td>0.91 (0.82 – 0.96)</td>
<td>0.48</td>
<td>1.33</td>
<td>8.23</td>
</tr>
<tr>
<td>N_zc [1/s] return</td>
<td>11.36 ± 1.69</td>
<td>10.88 ± 1.76</td>
<td>0.48 ± 1.09</td>
<td>0.14a</td>
<td>0.87 (0.71 – 0.94)</td>
<td>0.63</td>
<td>1.75</td>
<td>10.84</td>
</tr>
<tr>
<td>E_trial go [cm]</td>
<td>0.74 ± 0.38</td>
<td>0.67 ± 0.29</td>
<td>0.07 ± 0.32</td>
<td>0.03</td>
<td>0.70 (0.39 – 0.85)</td>
<td>0.19</td>
<td>0.53</td>
<td>3.28</td>
</tr>
<tr>
<td>E_trial return [cm]</td>
<td>0.99 ± 0.40</td>
<td>1.00 ± 0.56</td>
<td>0.01 ± 0.49</td>
<td>0.18a</td>
<td>0.67 (0.30 – 0.84)</td>
<td>0.29</td>
<td>0.80</td>
<td>4.95</td>
</tr>
<tr>
<td>F_cmean [N]</td>
<td>1.32 ± 1.03</td>
<td>1.29 ± 0.95</td>
<td>0.02 ± 0.51</td>
<td>0.06</td>
<td>0.93 (0.86 – 0.97)</td>
<td>0.26</td>
<td>0.73</td>
<td>4.52</td>
</tr>
</tbody>
</table>

*T > 0.1, Ttrans decreased, a, b, c = log10 transformed, d = square root transformed

T_ex [s] execution time in seconds, N_dp number of dropped pegs during transport, F_g (go/return) [N] mean grasping force of the three force sensors integrated into the handle in Newton, N_zc (go/return) number of zero-crossings of the acceleration, F_cmean [N] mean collision force, SD Standard Deviation, ICC Intraclass Correlation Coefficient, CI Confidence Interval, SEM Standard Error of Measurement, SDD Smallest Detectable Difference
Table 5 Test-retest reliability parameters for subgroup 1 \((n=11)\) without support of the non-affected arm

| Subgroup 1 \((n = 11)\) VPIT test performance without support |  |
|---|---|---|---|---|---|---|---|
| VPIT parameters | Test Mean ± SD | Retest Mean ± SD | Test-Retest Mean difference ± SD | Kendall’s tau correlation | Kendall’s tau correlation with transformed data | ICC (95% CI) | SEM agreement | SDD | SDD% |
| \(T_{ex} [s]\) | 113.14 ± 64.54 | 83.56 ± 27.26 | 29.57 ± 61.17 | .19a | 0.31c | 0.35 (~0.90–0.81) | 37.01 | 102.50 | 682.87 |
| \(N_{dp}\) | 0.12 ± 0.23 | 0.12 ± 0.23 | 0.0 ± 0.34 | 0.0 | - | 0.25 (~6.04–0.69) | 0.20 | 0.55 | 3.66 |
| \(F_g\) go [N] | 10.49 ± 4.84 | 10.13 ± 4.09 | 0.36 ± 3.4 | 0.20a | 0.24c | 0.84 (0.40–0.96) | 1.79 | 4.95 | 32.97 |
| \(F_g\) return [N] | 1.74 ± 0.92 | 1.55 ± 1.22 | 0.19 ± 0.93 | 0.13a | 0.13d | 0.78 (0.18–0.94) | 0.50 | 1.39 | 9.26 |
| \(N_{zc}\) [1/s] go | 11.18 ± 1.79 | 10.79 ± 1.97 | 0.39 ± 1.21 | 0.06 | - | 0.89 (0.60–0.97) | 0.62 | 1.73 | 11.53 |
| \(N_{zc}\) [1/s] return | 11.72 ± 1.91 | 10.80 ± 1.83 | 0.92 ± 1.29 | 0.09 | - | 0.82 (0.28–0.95) | 0.79 | 2.20 | 14.66 |
| \(E_{traj}\) go [cm] | 0.56 ± 0.27 | 0.47 ± 0.15 | 0.09 ± 0.20 | 0.16a | 0.10b,c | 0.70 (0.00–0.92) | 0.12 | 0.32 | 2.13 |
| \(E_{traj}\) return [cm] | 0.79 ± 0.28 | 0.75 ± 0.35 | 0.04 ± 0.42 | 0.06 | - | 0.18 (~2.80–0.80) | 0.29 | 0.79 | 5.26 |
| \(F_{cmean}\) [N] | 1.08 ± 0.72 | 1.18 ± 0.66 | −0.10 ± 0.52 | 0.02 | - | 0.84 (0.41–0.96) | 0.28 | 0.77 | 5.13 |

\(^aT > 0.1, T_{trans}\) decreased\(^b,c = \log_{10} \) transformed, \(^d = \) square root transformed

\(T_{ex} [s]\) execution time in seconds, \(N_{dp}\) number of dropped pegs during transport, \(F_g\) (go/return) [N] mean grasping force of the three force sensors integrated into the handle in Newton, \(N_{zc}\) (go/return) number of zero-crossings of the acceleration, \(E_{traj}\) [cm] trajectory error, \(F_{cmean}\) mean collision force, \(SD\) Standard Deviation, \(ICC\) Intraclass Correlation Coefficient, \(CI\) Confidence Interval, \(SEM\) Standard Error of Measurement, \(SDD\) Smallest Detectable Difference
Table 6 Test-retest reliability parameters for subgroup 2 (n=20) with support of the non-affected arm

<table>
<thead>
<tr>
<th>VPIT parameters</th>
<th>Test Mean ± SD</th>
<th>Retest Mean ± SD</th>
<th>Test-Reetest Mean difference ± SD</th>
<th>Kendall’s tau correlation</th>
<th>Kendall’s tau correlation with transformed data</th>
<th>ICC (95% CI)</th>
<th>SEM agreement</th>
<th>SDD</th>
<th>SDD%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tex [s]</td>
<td>123.33 ± 77.24</td>
<td>107.04 ± 58.93</td>
<td>16.30 ± 34.62</td>
<td>0.03</td>
<td>-</td>
<td>0.92 (0.79–0.97)</td>
<td>19.25</td>
<td>53.32</td>
<td>317.76</td>
</tr>
<tr>
<td>Ndρ</td>
<td>0.77 ± 1.27</td>
<td>1.22 ± 1.87</td>
<td>-0.45 ± 1.91</td>
<td>0.18</td>
<td>0.11bc</td>
<td>0.45 (-0.37–0.78)</td>
<td>1.16</td>
<td>3.23</td>
<td>19.25</td>
</tr>
<tr>
<td></td>
<td>0.18 ± 0.22</td>
<td>0.26 ± 0.26</td>
<td>-0.07 ± 0.28</td>
<td></td>
<td>0.11bc</td>
<td>0.49 (-0.26–0.80)</td>
<td>0.17</td>
<td>0.48</td>
<td>3.11c</td>
</tr>
<tr>
<td>Fg go [N]</td>
<td>8.15 ± 5.25</td>
<td>7.67 ± 5.28</td>
<td>0.49 ± 1.79</td>
<td>0.11a</td>
<td>0.12d</td>
<td>0.97 (0.93–0.99)</td>
<td>0.91</td>
<td>2.53</td>
<td>15.08</td>
</tr>
<tr>
<td>Fg return [N]</td>
<td>1.62 ± 0.80</td>
<td>1.63 ± 0.73</td>
<td>-0.02 ± 0.72</td>
<td>0.10a</td>
<td>-</td>
<td>0.73 (0.29–0.89)</td>
<td>0.40</td>
<td>1.10</td>
<td>6.56</td>
</tr>
<tr>
<td>Nzc [1/s] go</td>
<td>11.03 ± 1.39</td>
<td>10.81 ± 1.59</td>
<td>0.22 ± 0.68</td>
<td>0.34a</td>
<td>0.32bc</td>
<td>0.94 (0.85–0.98)</td>
<td>0.02</td>
<td>0.04</td>
<td>0.26c</td>
</tr>
<tr>
<td></td>
<td>1.04 ± 0.06bc</td>
<td>1.03 ± 0.07bc</td>
<td>0.01 ± 0.03c</td>
<td></td>
<td>0.32bc</td>
<td>0.94 (0.85–0.98)</td>
<td>0.02</td>
<td>0.04</td>
<td>0.26c</td>
</tr>
<tr>
<td>Nzc [1/s] return</td>
<td>11.17 ± 1.57</td>
<td>10.93 ± 1.77</td>
<td>0.25 ± 0.91</td>
<td>0.18a</td>
<td>0.18d</td>
<td>0.92 (0.80–0.97)</td>
<td>0.47</td>
<td>1.31</td>
<td>7.81</td>
</tr>
<tr>
<td>Etraj go [cm]</td>
<td>0.84 ± 0.39</td>
<td>0.78 ± 0.28</td>
<td>0.06 ± 0.37</td>
<td>0.01</td>
<td>-</td>
<td>0.58 (-0.06–0.84)</td>
<td>0.22</td>
<td>0.60</td>
<td>3.58</td>
</tr>
<tr>
<td>Etraj return [cm]</td>
<td>1.10 ± 0.42</td>
<td>1.15 ± 0.61</td>
<td>-0.04 ± 0.53</td>
<td>0.16a</td>
<td>0.15bc</td>
<td>0.66 (0.13–0.87)</td>
<td>0.30</td>
<td>0.83</td>
<td>4.95</td>
</tr>
<tr>
<td></td>
<td>0.02 ± 0.16bc</td>
<td>0.02 ± 0.19bc</td>
<td>0.00 ± 0.17c</td>
<td></td>
<td>0.15bc</td>
<td>0.69 (0.21–0.88)</td>
<td>0.10</td>
<td>0.27</td>
<td>1.75c</td>
</tr>
<tr>
<td>Fcmean [N]</td>
<td>1.45 ± 1.16</td>
<td>1.36 ± 1.08</td>
<td>0.09 ± 0.50</td>
<td>0.10a</td>
<td>0.01bd</td>
<td>0.95 (0.87–0.98)</td>
<td>0.25</td>
<td>0.69</td>
<td>4.11</td>
</tr>
<tr>
<td></td>
<td>1.11 ± 0.47d</td>
<td>1.08 ± 0.47d</td>
<td>0.04 ± 0.23d</td>
<td></td>
<td>0.01bd</td>
<td>0.94 (0.84–0.98)</td>
<td>0.12</td>
<td>0.32</td>
<td>2.07d</td>
</tr>
</tbody>
</table>

* T > 0.1, Ttrans decreased = log10 transformed, d = square root transformed

TeX [s] execution time in seconds, Ndρ number of dropped pegs during transport, Fg (go/return) [N] mean grasping force of the three force sensors integrated into the handle in Newton, Nzc (go/return) number of zero-crossings of the acceleration, Etraj [cm] trajectory error, Fcmean mean collision force, SD Standard Deviation, ICC Intraclass Correlation Coefficient, CI Confidence Interval, SEM Standard Error of Measurement, SDD Smallest Detectable Difference
Figure 2 Bland-Altman plots of the 9 VPIT parameters

- **Figure 2a**: Execution time in seconds (Tex[s])
  - Mean of Test 1 and Test 2
  - Subgroups: 1 and 2

- **Figure 2b**: Number of dropped pegs during transport (Ndp)
  - Mean
  - Subgroups: 1 and 2

- **Figure 2c**: Mean grasping force (Fg go [N])
  - Mean
  - Subgroups: 1 and 2

- **Figure 2d**: Mean grasping force (Fg return [N])
  - Mean
  - Subgroups: 1 and 2

- **Figure 2e**: Number of zero-crossings of the acceleration (Nzc go)
  - Mean
  - Subgroups: 1 and 2

- **Figure 2f**: Number of zero-crossings of the acceleration (Nzc return)
  - Mean
  - Subgroups: 1 and 2
Discussion

This is the first study evaluating the test-retest reliability of the novel VPIT and its concurrent validity with conventional upper limb function tests in chronic patients with stroke. In this stroke sample \((n = 31)\), the VPIT presented seven out of nine (78 %) reliable parameters that passed the accepted minimal standards for group comparisons \((\text{ICC} \geq 0.70)\) with ICCs = 0.70–0.94 [46]. The SDD% values were small in all VPIT parameters (1.37–21.42 %) except for Tex[s] (SDD% = 434.5 %). The correlations of the execution time and the number of dropped pegs/cubes of the VPIT with the NHPT and the BBT, respectively, were \((\rho = -0.23–0.31)\), although non-significant \((\rho = 0.09–0.51)\).
Concurrent validity

For the concurrent validity part of this study, both hypotheses were rejected, as the correlations of the VPIT with the conventional tests NHPT and BBT were low. The rejection of hypothesis 1 (correlations for the Tex[s]VPIT / Tex[s]NHPT and Tex[s]VPIT / NtcBBT) might be due to the following reasons: (1) the diverse and unstable upper limb skills of the study sample led to a large inter-subjects variation, which can be seen in the high Standard Deviation (SD) of Tex[s]VPIT with 72.1 s and SD = 28.9 s for the NHPT, respectively (Table 2). Moreover, support of the affected-arm by the non-affected arm might have caused further variability in the data, (2) Although the test performance of the NHPT was stable within the 3 test trials (mean SD (average from each subject’s SD) = 5.2 s), there was variation within the 3 test trials of the VPIT (mean SD = 27.2 s). The higher intra-subject variation in the Tex[s]VPIT parameter could be attributed to the nature of the VPIT, which is not a real physical object manipulation test and involves tools that patients are not familiar with (e.g., the robotic handle and the computer for some patients). This is supported by Bowler et al. (2011), who achieved a more accurate and consistent data set for an embedded NHPT than for a purely hapto-VR-NHPT version [47]. As the VPIT was the last assessment administered in session one, it might be that patients were already more tired and less able to concentrate than in the NHPT. This observation is supported by the smaller SD of the VPIT during the second assessment session (see Table 4). Furthermore, compared to the BBT, the Tex[s]VPIT has no time limit, thus the participants with poorer upper limb function had to perform the assessment even longer than the more skilled ones, also resulting in increased fatigue of the affected arm from test trial to test trial (although they were allowed to have a rest between tests). Conversely, the correlations of the VPIT with the NHPT - the two tests with no time limit - were better than those with the BBT (time limit: 60 s). A reason for this could be the similar test procedures of the VPIT / NHPT, whereas the BBT has different test procedures [5].

The rejection of hypothesis 2 (correlations for NdpVPIT with the NdpNHPT / NdcBBT) occurred due to very low correlations between those validity parameters. The scatter plots of those parameters showed no linear correlation, as many participants did not drop any pegs or cubes during test trials (no drops $n = 17$ in VPIT (55%); $n = 10$ in NHPT (32%) and $n = 14$ in BBT (45%)). This ceiling-effect - considered to be present if more than 15 % of all participants received the highest possible score (here: dropping no peg/cube) [48] - did not allow a distinction from the participants with the highest
achievable score, indicating limited validity. This finding is supported by the poor strength of agreement for the Ndp and no dropped pegs in the VPIIT and NHPT (Kappa = 0.04) or BBT (Kappa = 0.19), respectively. Furthermore, the high SDs of the Ndp/c (VPIIT/BBT/NHPT) parameters indicate the high variance of the study sample (Table 2). The high variance in the VPIIT could be due to the difficulty of the coordination of the PHANTOM Omni arm and the virtual pegboard in the virtual 3D space, especially for subgroup 2 (those with the higher Ndp parameter than subgroup 1; see Table 2). Most participants in this group had difficulty in aligning the handle of the haptic display to the dropped peg lying on the virtual board, which resulted in several grasp attempts and the associated higher drop rate of pegs. Furthermore, the VPIIT handle remains in the patients’ hand during the entire test, but inserting a peg in the hole requires control of grasping force and to decrease applied grasping force below the 2 N force threshold. Also, patients have to be below the 2 N threshold before being able to grasp a new peg (i.e., the patient cannot just tightly grasp the handle during the whole test and “only” align cursor to pegs/holes to achieve the task, see Fluet et al. (2011) [17] and the force traces presented in that paper may illustrate how force control (grasping AND releasing) is required for achieving the task. Nevertheless, this may have influenced the performance time of the VPIIT and also the low values for concurrent validity. The degree of stereopositivity (Table 1), which we assumed to be important to perform the VPIIT [28], didn’t seem to be indicative of the participants being able to perform the VPIIT or not, as all participants could perform it equally. However, further investigations into this observation would be necessary to allow us to draw a comprehensive conclusion.

**Test-retest reliability**

In the hypothesis concerning the relative test-retest reliability, we expected the ICCs to be ≥ 0.80. This assumption was met by 5 out of 9 VPIIT parameters in the whole stroke sample (Table 4) and both subgroups (Tables 5 and 6). Those 5 parameters were the same for the whole sample and for subgroup 2 (Tex[s], Fggo[N], Nzc[1/s]go/return and Fcmean[N]), with slightly higher correlations in subgroup 2 than for the whole sample. This is not surprising, as subgroup 1 with the highest mean differences between test and retest (Table 5) is not included, indicating that supporting the affected arm allowed a more stable test performance in subgroup 2. The lower ICCs in sub-group 1 might also be due to the lower number of participants (n = 11).
[49], as well as due to a learning effect in some parameters; e.g., the important decrease in \(\text{Tex[s]}\) of 29.6 s from test to retest (Table 5). This finding is illustrated by the Bland-Altman plot (Fig. 2a), where 4 subjects in subgroup 1 improved by > 50 s. The kinetic parameters \(F_{\text{ggo[N]}}\) and \(F_{\text{cmean[N]}}\) and the kinematic parameters \(N_{ZC[1/s]}\go\text{return}\) achieved high test-retest reliability in the whole sample and in both subgroups. This means that the stroke participants were able to transport the pegs with a constant grip strength and movement coordination while transporting the peg (go). However, there were fair correlations when approaching a new peg (return) \(F_{\text{greturn}}\) in the whole sample and both subgroups. This might be due to the fact that participants still had to hold an object (the handle) on the way back to approach a new peg without actually carrying a peg. Furthermore, if we look at the overall test-retest reliability of the whole study sample for \(N_{dp}\), it was poor \((\text{ICC} = 0.58)\). This might be due to the artificial force threshold to grasp and release the pegs, which is quite unintuitive, as there is no feedback provided on the force applied by the subject and on the force threshold, although measured by the handle \((F_{\text{ggo/return[N]}})\). This might lead many subjects to drop the pegs. However, to achieve a conclusive statement regarding the clinical use of those unique kinetic and kinematic VPIT parameters, further research is needed. This could be done by comparing the reliability of stroke results with healthy controls, or by evaluating their validity with other outcome measurements quantifying force control and movement coordination.

From the 5 VPIT parameters meeting the hypothesis for relative test-retest reliability, all VPIT parameters fulfilled the hypothesis for the absolute test-retest reliability with SDD\% of \(\leq 54 \%\) (Tables 4, 5 and 6) except for the parameter \(\text{Tex[s]}\). This is quite surprising, as more complex and demanding upper limb function tests – such as the VPIT – have in general higher SDDs than simpler tests – such as the BBT or NHPT. Therefore, our results are in contrast with those of Chen et al. (2009) [25], whose SDD\% values were high for the affected hand, especially in the NHPT, with an SDD\% of 52\% for the nonspastic and 88\% for the spastic group. In our sample however, the only VPIT parameter not being susceptible to change was \(\text{Tex[s]}\), as its SDD was high and varied greatly within the whole sample and both subgroups \((SDDs = 53.3 – 102.5 s)\). In other words, only a change between two consecutive measurements exceeding at least 53 s for \(\text{Tex[s]}\) can be interpreted as a true clinical improvement when chronic patients with stroke perform the test with the affected hand. In addition, the SDD can be used as a threshold to identify statistically significant individual change [50, 51].
Thus, if a change between 2 consecutive measurements for an individual patient exceeds the SDD (e.g., 53.3 s in subgroup 2) the individual patient may be exhibiting significant improvement. In our sample, this was the case for 4 participants for the Tex[s] parameter in subgroup 2, while one participant exceeded 102.5 s in subgroup 1. The data show that allowing support of the affected arm increased performance stability in the test and retest. This can be seen in the almost twice as high SDD of subgroup 1 (SDD = 102.5 s) compared to subgroup 2 with an SDD of 53.3 s.

Limitations and future research

As the VPIT extracts much more (i.e., kinetic and kinematic) parameters than conventional, non-VR upper limb function assessments (i.e., time, number of dropped/transported objects), it was not possible to validate all of them. Therefore, future studies should focus on the validity of those parameters measurable by VR devices by comparing them with other VR-based upper limb function assessments. Furthermore, it would be of interest to evaluate the discriminant validity of the VPIT by comparing stroke participants’ performances with those of healthy controls.

Future versions of the VPIT should provide visual feedback on the force applied by the subject and on the force threshold. Another limitation might be the fact that two raters collected data, although the test-retest measurements of one participant were performed by the same rater. However, the major limitation of our study can be seen in the given permission to support the weight of the affected arm with the non-affected arm (and not use the other hand to steer), visually monitored by the supervising rater (occupational therapist). We are aware that this can be judged as a bias, because the test performance was not identical between the two subgroups, nor can we be sure how much the non-affected arm really supported (steered) the affected arm. Nevertheless, allowing support of the affected arm opens the use of the VPIT for a motorically weaker stroke population. If available, adjustable armrests (e.g., the Armon Ele- mento [52]) could be used instead of the non-affected arm, which might improve body stability and therefore the test performance by providing repeatable conditions from trial to trial and patient to patient. Randomisation of the several tests (NHPT, BBT, VPIT) may also reduce tiredness of the patients and increase validity.

Last, the sample size was relatively small and may have affected the values of the reproducibility and measurement error. A sample size of at least 50 is generally seen as adequate for the assessment of the agreement parameter, based on a general
guideline by Altman [53]. The sample size we used of 33 patients with chronic stroke is, however, a realistic group size to find first estimates for the assumed relation between stroke and the hemiparetic arm / hand for the VPIT. Future studies may therefore strive for a bigger study sample and evaluate responsiveness to treatment of the VPIT.

**Implications for practice**

The VPIT provides more objective and comprehensive measurements of upper limb function than conventional, non-computerized hand assessments. Receiving feedback concerning for example the patient’s ability to control force or movement coordination is essential for rehabilitation. Furthermore, getting information about movement smoothness, as provided by the VPIT, offers clinically relevant information, as smoothness has been shown to be a good indicator of upper limb coordination and stroke recovery [23]. Therefore, clinicians can use these information to work on those motor skills crucial for the independent performance of daily activities [54]. Thanks to its compact and transportable shape, the VPIT is easy to administer even in the patient’s home, thus allowing a broad use amongst clinicians working in different settings. As all test results are stored in the computer and could be graphically displayed immediately after completion of the test, performance can be discussed with the patient and adjustments made to the therapy program of a patient depending on the progress. To increase the ease of use of the VPIT, time limitations in the test duration for motorically weaker patients should be considered, together with the allowance to use adjustable armrests if needed.

An important aim of developing this new assessment is to improve the assessment of arm function in a clinical setting where the results of the assessment can be generalized to a population reflective of the “real world”. Seen from this perspective it is rather a strong point of our study that we had a rather heterogeneous sample because this can be considered more realistic for clinical settings.

**Conclusions**

The VPIT is a promising upper limb function assessment which has proved to be feasible for use with this diverse group of chronic patients with stroke. The low concurrent validity showed that the VPIT was inherently different from the conventional tasks, indicating that performing this hapto-virtual reality assessment requires other
components of upper limb motor performance than the NHPT and BBT. The high test-retest reliability in 5 and the low SDD% in 8 out of 9 VPIT parameters showed that those parameters remain consistent when performed by patients with chronic stroke and are susceptible to change, allowing diagnostic and therapeutic use in clinical practice for this patient group. The other 4 parameters (Ndp, Fgreturn[N] and Etrajgo/return[cm]) showed poor to fair ICCs when performed with the affected hand and require further research for this population. The high intra-subject variation indicated that the VPIT is a demanding test for this stroke sample, which requires a thorough introduction to this assessment. Allowing testing trials before starting with the assessment is a prerequisite for a reliable test performance. When using the VPIT as outcome measurement, clinicians may want to use the SDDs reported in this article as reference points for clinically important changes, and the SDD% results to compare test-retest reliability among other tests.

Abbreviations
BBT: Box and Block Test; CI: Confidence Interval; Etraj[cm]: Trajectory error in centimeters; ICC: Intraclass Correlation Coefficient; Fcmean: Mean collision force; Fg[N]: Mean grasping force in Newton; Ndc: Number of dropped cubes; Ndp: Number of dropped pegs; NHPT: Nine Hole Peg Test; Ntc: Number of transported cubes; NzcNzc[1/s]: Number of zero-crossings; SD: Standard Deviation; SDD: Smallest Detectable Difference; SEM: Standard Error of Measurement; Tex[s]: Execution time in seconds; VPIT: Virtual Peg Insertion Test.

Competing interests
MCF is shareholder of ReHaptix GmbH, Zurich, Switzerland, a startup company commercializing assessment technologies. All other authors declare that they have no competing interests.

Authors' contributions
BCT conceived the methodology and carried out data collection, quality assessment, data analysis and manuscript writing. RK supervised progress, participating in methodology conception and manuscript writing & revision. MCF & OL had developed the VPIT and maintained technical support during data collection. MCF, OL & RDB
contributed in manuscript revision. EDB supervised progress, helped with methodology conception, manuscript writing & revision. All authors read and approved the final manuscript.

**Acknowledgements**

The authors warmly thank all OT outpatient clinics who were willing to help recruiting the participants for this study. Many thanks go to all participants who were willing to share their time for this project. Further thanks go to Judith Häberli (JH) who helped with data collection and to Dr Martin J Watson for proof reading the manuscript on English and structure. Final thanks go to the Rehabilitation Engineering Lab, Department of Health Sciences and Technology, ETH Zurich, Switzerland, who provided the VPIT for this research project.
References


Chapter 4

The Zürich Maxi Mental Status Inventory (ZüMAX): Test-retest reliability and discriminant validity in stroke survivors

Pre exergames intervention results of the figure copying task (visuoconstruction) of the ZüMAX of P6, female, 78 years old, 5 weeks onset right-hemispheric stroke and visuo-spatial neglect

This chapter has been published as: Tobler-Ammann BC, de Bruin ED, Brugger P, de Bie RA, Knols RH. The Zürich Maxi Mental Status Inventory (ZüMAX): Test-Retest Reliability and Discriminant Validity in Stroke Survivors. Cognitive and Behavioral Neurology 2016;29(2):78-90.
Abstract

Objective: To examine discriminant validity and test-retest reliability of the Zürich MAXI mental status inventory (ZüMAX) in patients with stroke.

Background: The ZüMAX is a novel domain-specific cognitive assessment tool to screen for disturbances in neuropsychological function. The test can be used in stroke rehabilitation to estimate severity of cognitive impairment. Because evidence for validity and reliability is lacking, the tool’s clinical use is limited.

Methods: We administered the ZüMAX in a test-retest design to 33 community-dwelling stroke survivors, and once to 35 healthy controls matched for age and sex.

Results: We found significant group differences in subscores for the cognitive domains of executive functions and language as well as total score (P = 0.001 to 0.004); we did not find group differences for the domains praxia (defined as the ability to perform purposeful actions), visual perception and construction, or learning and memory. Test-retest reliability of the total score was good (intraclass correlation coefficient = 0.81), with the individual domain subscores ranging from poor to fair (0.59 to 0.79). The ZüMAX could detect changes in patients with low smallest detectable differences in executive functions, language, and praxia (0.05 to 1.49) and total score (0.09).

Conclusion: The ZüMAX has moderate to good test-retest reliability. Furthermore, the tool might discriminate between healthy persons and chronic stroke survivors on three of five subscales. The ZüMAX shows promise in measuring neuropsychological disturbances in stroke survivors; however, further trials are required with larger samples.

Keywords: stroke, post-stroke cognitive impairment, domain-specific, test-retest reliability, discriminant validity
Introduction

Cognitive impairment is common after stroke. Improving cognition is the highest clinical and research priority after stroke for patients, their caregivers, and health professionals (Pollock et al, 2012, 2014).

Rates of post-stroke cognitive impairment differ widely in the literature, depending on the group studied (eg, hospital-based versus population-based cohorts), criteria for cognitive impairment (eg, with or without pre- or post-stroke dementia or degenerative, vascular, or mixed cognitive disease), and the time between onset of stroke and testing (Blackburn et al, 2013; Brainin et al, 2015; Ihle-Hansen et al, 2011; Jacquin et al, 2014; Makin et al, 2013; Middleton et al, 2014; Sun et al, 2014).

A recent systematic review of domain-specific cognitive impairment in stroke revealed that impaired cognition involved all major cognitive domains (Edwards et al, 2013). To date, however, research evidence is insufficient to support clear recommendations for clinical practice, as post-stroke cognitive impairment is difficult to treat and manage (Gillespie et al, 2015).

The crucial first step in managing cognitive problems is recognition and diagnosis, with baseline cognitive testing for screening or triage. Depending on the results, patients can be offered specialist assessment to define the cognitive problem more precisely (Lees et al, 2013).

Standardized assessment batteries and cognitive screening tools play an important role in identifying underlying cognitive strengths and deficits in survivors of stroke (Van Heugten et al, 2015). Previous studies have provided comprehensive comparisons of the utility and psychometric properties of such tools (Van Heugten et al, 2015). One of the most widely used is the Montreal Cognitive Assessment (MoCA) (Nasreddine et al, 2005). Burton and Tyson (2015) consider the MoCA “the most valid and clinically feasible screening tool to identify stroke survivors with a wide range of cognitive impairments.” The Mini-Mental State Examination (MMSE) (Folstein et al, 1975) is widely used for quick dementia screening (Burton and Tyson, 2015).

The MMSE and the majority of other screening tools used to assess post-stroke cognitive impairment were originally developed to assess age-related cognitive decline. Although they are now also used to screen stroke survivors, because of their original purpose, which focused on global impairments, they lack domain-specific items. This might explain why many such tools are unable to detect mild cognitive impairment, an intermediate clinical state that often progresses to dementia (Petersen
et al, 1999). Furthermore, there is no consensus on either a preferred tool to test for post-stroke cognitive impairment (Lees et al, 2013) or a preferred timing of assessment (Brainin et al, 2015).

A screening tool for cognition in survivors of one or more left or right hemispheric strokes must first be able to discriminate between healthy people and those who have had a stroke (Chen et al, 2013). Testing discriminant validity for a novel instrument will thus help clinicians interpret stroke survivors’ performance.

The growing health care burden, and specifically the social and economic impact of post-stroke cognitive impairment, necessitate further clinical studies that go beyond discriminating between stroke and non-stroke to evaluating assessments and delineating their abilities to discriminate between people with and without post-stroke cognitive impairment. Studies are needed in both acute and long-term settings, to help clinicians plan and provide appropriate treatment throughout a patient’s course (Brainin et al, 2015).

The validity of future trials will depend on the ability to describe and quantify cognitive changes in stroke survivors. This ability, in turn, will depend on robust critical evaluations of the cognitive tests given to stroke survivors (Brainin et al, 2015; Quinn et al, 2009), as some commonly used tools may not be suitable for them (Godefroy et al, 2011). Any appropriate test would have to be sensitive to common post-stroke impairments such as difficulties in executive functions; thus, the test should assess different cognitive domains.

A new domain-specific cognitive assessment tool is the Zürich maxi mental status inventory, nicknamed the ZüMAX. It was developed by author P.B. at the Department of Neurology, University Hospital Zurich, in 2012. His goal was to “maximize” the “Mini”-Mental State Examination, to create a neuropsychological screening instrument that would allow a fast (approximately 30 minutes) but highly sophisticated assessment of the cognitive domains of executive function, language, praxia, visuoperception and construction, and learning and memory. Although P.B. designed the ZüMAX to be a quick, convenient first test for patients with suspected neuropsychological deficits of any neurologic origin, the tool is also suitable for long-term follow-up. The actual instrument is described in the Methods section.

P.B. modeled the ZüMAX on principles described by Schnider (2004) in his textbook on comprehensive yet time-efficient assessments in behavioral neurology. P.B. also incorporated suggestions from Hachinski et al (2006) that any neuropsychological
screening instrument for vascular cognitive impairment include the domains of executive function, language, visuospatial perception and construction, and learning and memory. P.B. added the fifth domain, praxia, defined as the ability to perform purposeful actions, because this domain is typically affected by stroke (Buxbaum and Coslett, 2009).

The ZüMAX has several advantages over both the MMSE and the MoCA. First, the ZüMAX examines the domains of perception and language in greater detail. Second, the ZüMAX tests praxia and executive function; the MMSE does not cover either of these domains, and the MoCA covers only executive function.

The ZüMAX has not yet been published (see the “Limitations and Future Directions” section of the Discussion). Readers who want to learn more about the tool or request a research copy can write to Peter Brugger at peter.brugger@usz.ch.

The only existing evaluation of the ZüMAX is an unpublished 2014 master’s thesis by Andrea Rust at the University of Zürich. Rust assessed norms and the construct validity of the ZüMAX compared with the MoCA and MMSE in 227 healthy adults. She found high correlations (P < 0.001) between the total scores of all the assessments.

The current study is the first peer-reviewed publication evaluating the ZüMAX. The tool has not been altered, meaning that we used the same version with our stroke survivors that Rust used in 2014.

Our purpose in this study was to evaluate whether the ZüMAX was a valid and reliable assessment for measuring cognitive impairment in chronic stroke survivors. We aimed to evaluate not only test-retest reliability and smallest detectable difference (SDD), but also differences in scoring between chronic stroke survivors and healthy controls.

We hypothesized that the total score and the subscores for the five cognitive domains would have good relative reliability, with an intraclass correlation coefficient (ICC) ≥ 0.80. Furthermore, we expected the SDDs (absolute reliability) to be ≤ 10% of the mean average values of the total score and subscores for the five cognitive domains.

For discriminant validity, we expected to find significant group differences (P ≤ 0.05) in the total scores and cognitive domain subscores.
Methods

Participants
To recruit patients who had neurologic deficits after stroke, we contacted 40 outpatient occupational therapy practices in the canton of Bern, Switzerland. We asked the occupational therapists to refer patients currently or recently under their care, and we provided a list of admission criteria for screening the patients. Of the practices we contacted, 23 either had no access to stroke survivors or did not want to participate in the study. The remaining 17 practices, which regularly treated stroke survivors, each referred one to three patients to us between September 2013 and March 2014. (Unfortunately, no data are available about the patients who were approached but refused to participate.) Our study coordinator (author B.C.T.-A.) kept in touch with the participating occupational therapists to discuss the patients’ eligibility and guarantee their privacy. Together, the examiner and therapist scheduled the appointments for the examiner to test the patients.

Through these means, we recruited a consecutive community-dwelling sample of 33 stroke survivors who met these eligibility criteria:

- Age older than 18 years
- A stroke diagnosis, confirmed by a physician, at least 6 months before the study
- Ability to speak and understand the German language
- Ability to sit in a chair or wheelchair with a backrest for up to 60 minutes
- MMSE score ≥ 20 (indicating mild vascular dementia) for people younger than 80 years of age, and ≥ 16 for people aged 80 years or older (Folstein et al, 1975; Tombaugh et al, 1996)

Exclusion criteria were a diagnosis of a brain injury other than stroke, an earlier diagnosis of neglect or aphasia by a physician, and a noncontrolled medical condition such as chronic pain or drug abuse.

We used the Flinders Handedness survey (FLANDERS) (Nicholls et al, 2013) to assess the patients’ handedness.

We recruited a consecutive sample of healthy controls, whom we matched to the age and sex of the patients. To find the controls, we handed out participant information sheets to colleagues and relatives of the research team, with a request that they further disseminate the information to other people they knew (snowball sampling). We
needed 33 controls, but this method was so successful that 35 people volunteered to take part. Because they fulfilled the inclusion criteria, we included all 35 in the control group.

The eligibility criteria for the controls were:

- Age older than 18 years
- No medical condition (such as a stroke or dementia) that could influence their cognitive skills
- Ability to speak and understand German

We asked the controls about their handedness, but we did not test it.

The study was approved by the Ethics Committee (Kantonale Ethikkommission [KEK]) (KEK-Nr. 119/13) of the Canton of Bern, Switzerland. All participants gave their informed consent before entering the study.

**Zürich maxi mental status inventory (ZüMAX)**

The ZüMAX is given as an interview. The total administration time of 30 minutes includes 5 minutes of instructions from the examiner. The examiner needs this equipment: a ZüMAX testing form with one coversheet, 24 stimulus cards (chimerical faces), and several demonstration, work, and record sheets; a stopwatch; a metronome; and a plain pencil (no eraser, no ruler) for patients to use.

Before starting the test, the interviewer assesses patients' orientation and obtains their demographic data, including education level. The interviewer also asks patients to score their mood on a visual analogue scale. The ZüMAX includes this “mood scale” (Figure 1) (Regard et al, 1982) because mood has a considerable impact on cognition (Kimura et al, 2000; Narushima et al, 2007).
What is your current mood?
(Please tick the vertical line)

![Mood Scale](image)

**FIGURE 1.** Affect: Self-reporting mood scale of the ZÜMAX inventory, in English translation. Participants rate both their current mood and their best recollection of their affective state before they suffered their stroke. The examiner rates the value from -6 to +6.

The ZüMAX profiles the five major cognitive domains typically covered in standard neuropsychological evaluations (Schnider, 2004): language, praxia, learning and memory, visual perception and construction, and executive function. The five domains are assessed in 15 separate tasks, which are given in the following order and some of which we illustrate in this paper:

1. Visual perception: degraded figures, unfamiliar scenes, face
2. Figural fluency (see Figure 3)
3. Visuoverbal memory: immediate recall
4. Visuoconstruction: copy figure (see Figure 4)
5. Verbal fluency with letter criterion (see Figure 3)
6. Visual-spatial memory: draw figure from memory
7. Reading aloud and praxia (see Figure 2)
8. Writing, calculation in writing, and oral repetition
9. Category fluency (see Figure 3)
10 – 11. Interference control (two tasks) (see Figure 3)
12. Cognitive flexibility (see Figure 3)
13. Visuoverbal memory: late recall
14. Nonverbal visual-spatial memory: draw figure again from memory
15. Asymmetry in perception: chimerical faces

As noted, some of the tasks test several domains at once. For example, Figure 2 illustrates a task that assesses both language and praxia. Some tasks require spoken responses, and others require responses written or drawn with pencil and paper, using the dominant hand. Because many of the tasks are timed, the ZüMAX also addresses speed of information processing.

**What finger movements does use of the black object require?**
**(Demonstrate this by pantomime)**

**FIGURE 2.** Reading aloud and praxia: ZÜMAX demonstration sheet for a task in the domains of language and praxia, in English translation. Participants are asked to read the instruction aloud and then act on it. Then they are given four more actions to pantomime, and must act on them.

Participants can score a maximum of 6 points in each domain, with a total of 30 points representing optimal function for all five domains. Figure 3 shows the scoring for the domain of executive function and briefly explains the component tasks.
<table>
<thead>
<tr>
<th>Enter 1 point in the box for:</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fluency (nonverbal, letter, category)</strong></td>
<td></td>
</tr>
<tr>
<td>Figural fluency in 1 minute: ≥5 correct figures</td>
<td>1/1</td>
</tr>
<tr>
<td>P-words in 1 minute: ≥5 correct words</td>
<td>1/1</td>
</tr>
<tr>
<td>Supermarket items in 1 minute: ≥10 correct items</td>
<td>1/1</td>
</tr>
<tr>
<td><strong>Interference control and response stereotypy</strong></td>
<td></td>
</tr>
<tr>
<td>Digits time + stars time ≤3 AND digits time &lt;30 seconds</td>
<td>1/1</td>
</tr>
<tr>
<td>Digits: ≤2 errors</td>
<td>1/1</td>
</tr>
<tr>
<td>Mental Dice Task: &lt;15 ascended counting steps</td>
<td>1/1</td>
</tr>
<tr>
<td><strong>Subtotal &quot;executive function&quot;</strong></td>
<td>6/6</td>
</tr>
</tbody>
</table>

**Figural fluency** (Regard et al, 1982): 5-point figures: Participants are given a worksheet with 25 boxes, each containing a pattern of five dots, and are asked to draw as many different “figures” as they can in 1 minute. A figure is any straight line connecting 2 or more of the 5 dots. A figure does not need to look good; it just needs to be different from every other figure. After 1 minute, participants are asked to estimate the duration of the tasks. The examiner checks the worksheet and points out to participants the first repeated figure that they have drawn. Each correct figure counts for 1 point.

![5-point box](image)

**Verbal fluency with letter criterion**: P-words (Regard et al, 1982): Participants are asked to say as many word starting with the letter “P” as they can in 1 minute. After 1 minute, the examiner asks participants to estimate the duration of the task. The correct words are counted.

**Category fluency**: supermarket items (Regard et al, 1982): Participants are asked to say as many supermarket items as they can in 1 minute. The correct words are counted.

**Interference control**: digits and stars (Sedo, 2004): In the first of two tasks, participants are shown a worksheet with 25 boxes, each containing 1 to 5 stars, and are asked to say as quickly as possible how many stars are in each box. In the second task, the participants are shown a sheet with 25 more boxes, each containing 1 to 5 identical digits, and are asked to say as quickly as possible the number of digits in each box. Time is recorded for each trial. Self-corrections are considered as errors.

![stars]![digits](image)

**Response stereotypy and cognitive flexibility**: Mental Dice Task (Brugger et al, 1996): Participants are asked to say the most random possible sequence of numbers, using only the digits 1 to 6, as if they were rolling dice. They are to speak in time with a metronome, at a pace of 1 digit per second. They are stopped after they have said a total of 66 digits. The examiner counts correct digits and notes digits >6 and digits in ascending (eg 1, 2, 3) and descending (eg, 5, 4, 3) series.

**FIGURE 3.** Scoring for the ZüMAX executive function domain, in English translation, with task descriptions.
Immediately after completing the ZüMAX, the interviewer notes how long it took and characterizes both qualitatively and quantitatively the patient’s orientation, alertness, tempo, ability to cooperate, affect, and communicative behavior.

**Procedures**

We assessed the patients in the stroke group either at their occupational therapist’s clinic or in their home. We tested the patients twice. During the first session, we collected the demographic data and gave the MMSE and the FLANDERS. Then we administered the ZüMAX. As instructed, participants used their dominant hand to perform the writing and drawing tasks. During the second session, we repeated the ZüMAX. We spaced the two sessions about a week apart (mean 6.4 ± standard deviation 1.0 days) to minimize the effect of any learning or memory that might influence a patient’s performance on the repeat test.

We did not try to give the patients a thorough neuropsychological evaluation. The patients’ treating physician did not request this, and asking for such a service or providing it ourselves would have been beyond our scope.

After finishing the test-retest procedure with a patient, we recruited and tested an age- and sex-matched control participant. We tested the controls at a place of their choosing; most chose their home. At the controls’ one testing session, we gathered their demographic data and asked about their handedness. We gave them the ZüMAX only once, to evaluate its discriminant validity.

The examiners were author B.C.T.-A. and master student Judith Häberli (J.H.). Author P.B. had trained them to administer the ZüMAX. The same examiner did the test and re-test of an individual patient. The examiners divided up the testing of the controls depending on their availability.

**Statistical Design**

We used SPSS version 22.0 (SPSS Inc, Chicago, Illinois) for data management, and MedCalc statistical software version 14.8.1 (www.medcalc.org) to draw Bland-Altman plots. We set the level of statistical significance at $P \leq 0.05$.

We used descriptive statistics to define the study population and their clinical characteristics. We calculated normality of the data using the Shapiro-Wilk test for small sample sizes ($n < 50$) (Norman and Streiner, 2008). Our discriminant validity parameters were the subscores for the five cognitive domains and the total test score.
We determined testing differences between the stroke and control groups with the Mann-Whitney U test (Norman and Streiner, 2008). We calculated effect size using the formula for nonparametric tests:

\[ r = \frac{Z}{\sqrt{n}} \]

(Z being the converted U score) (Fritz et al, 2012). We then used the effect size estimate \( r \) to calculate Cohen \( d \) value:

\[ d = \frac{2r}{\sqrt{(1 - r^2)}} \]

We calculated the 95% confidence intervals for effect size (Ivarsson et al, 2013). Cohen’s (1988) guidelines for \( d \) and \( r \) are large effect (\( d = 0.80; r = 0.37 \)), medium (\( d = 0.50; r = 0.24 \)), and small (\( d = 0.20; r = 0.10 \)).

We analyzed relative reliability by calculating ICCs separately for single measures using the ICC\(_{2(A,1)}\) formula (Shrout and Fleiss, 1979; Weir, 2005). We selected the option “absolute agreement” to take into account the systematic error between raters B.C.T.-A. and J.H. (de Vet et al, 2006). We used the following criteria for reliability: high, 0.90 to 0.99; good, 0.80 to 0.89; fair, 0.70 to 0.79; and poor, ≤ 0.69 (Arnall et al, 2002; Denegar and Ball, 1993).

To calculate absolute reliability, we complemented the ICCs with Bland-Altman analysis, which can be used to show variation or the magnitude of difference of repeated measurements (Rankin and Stokes, 1998). We measured the degree of heteroscedasticity by calculating the Kendall tau correlation between the absolute differences and the corresponding means of the reliability parameter. When we found a positive \( t > 0.1 \), we considered the data to be heteroscedastic. When the Kendall tau was < 0.1 or negative, we considered the data homoscedastic (Brehm et al, 2012). When we found heteroscedasticity, we transformed the data logarithmically (Bland and Altman, 1996; Euser et al, 2008). Then we calculated the Kendall tau again; if \( t_{log} \) decreased, we analyzed reliability using the log-transformed parameters (Brehm et al, 2012).

To quantify the precision of individual scores on an assessment, we calculated the standard error of measurement (SEM) with the following formula:
SEM agreement = \( \sigma \sqrt{(1 - ICC)} \)

with \( \sigma \) being the total variance of the scores from all participants (de Vet et al, 2006; Weir, 2005). We then calculated the SDD based on the SEM:

\[
SDD = SEM \times 1.96 \times \sqrt{2}
\]

Here 1.96 defines the 95% confidence interval. The factor \( \sqrt{2} \) is included because it concerns the difference between the two measurements.

As a last step, we calculated the SDD %:

\[
SDD \% = \frac{SDD}{\text{grand mean}} \times 100
\]

The grand mean is the mean of the means of each ZüMAX parameter. Because agreement parameters (SDDs) are expressed on the actual scale of the assessments, they allow clinical interpretation of the results (de Vet et al, 2006; Weir, 2005). The SDD % can be used to compare test-retest reproducibility among tests (Chen et al, 2009).

**Results**

All patients and controls completed the study. Table 1 lists the participants’ demographic and clinical characteristics. Of our 33 stroke survivors, 29 (87.8%) had suffered one stroke and four (three men and one woman) (12.1%) had sustained at least two strokes.
<table>
<thead>
<tr>
<th></th>
<th>Stroke (n = 33)</th>
<th>Control (n = 35)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex (n [%])</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>9 (27.3)</td>
<td>9 (25.7)</td>
</tr>
<tr>
<td>Men</td>
<td>24 (72.7)</td>
<td>26 (74.3)</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td>63.2 ± 14.7</td>
<td>63.0 ± 14.6</td>
</tr>
<tr>
<td>Mean ± standard deviation</td>
<td>66 (33 to 89)</td>
<td>67 (33 to 85)</td>
</tr>
<tr>
<td>Median (range)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University education</td>
<td>6 (18.2)</td>
<td>20 (57.1)</td>
</tr>
<tr>
<td>Apprenticeship</td>
<td>24 (72.7)</td>
<td>15 (42.9)</td>
</tr>
<tr>
<td>No completed education</td>
<td>3 (9.1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Education (n [%])</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>22 (66.7)</td>
<td>33 (94.3)</td>
</tr>
<tr>
<td>Mixed</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Left</td>
<td>11 (33.3)</td>
<td>2 (5.7)</td>
</tr>
<tr>
<td><strong>Stated handedness (n [%])</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>22 (66.7)</td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>1 (3.0)</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>10 (30.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Handedness according to Flinders Handedness</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right (+5 to +10 points)</td>
<td>22 (66.7)</td>
<td></td>
</tr>
<tr>
<td><strong>Survey¹ (n [%])</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right (-4 to +4 points)</td>
<td>1 (3.0)</td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Left (-10 to -5 points)</td>
<td>10 (30.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Affected hand (n)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td><strong>Affected hand = dominant hand (n [%])</strong></td>
<td>20 (66%)</td>
<td></td>
</tr>
<tr>
<td><strong>Forced to use nondominant hand because of stroke (n)</strong></td>
<td>8</td>
<td></td>
</tr>
<tr>
<td><strong>Months since onset of stroke</strong></td>
<td>Mean ± standard deviation</td>
<td>49.4 ± 79.8</td>
</tr>
<tr>
<td>Mean ± standard deviation</td>
<td>49.4 ± 79.8</td>
<td>22 (6 to 387)</td>
</tr>
<tr>
<td>Median (range)</td>
<td>22 (6 to 387)</td>
<td></td>
</tr>
<tr>
<td><strong>Hemisphere affected by stroke (n [%])</strong></td>
<td>Left</td>
<td>19 (57.6)</td>
</tr>
<tr>
<td><strong>Mini-Mental State Examination² score</strong> (maximum, 30 points; 24 points = normal³)</td>
<td>Mean ± standard deviation</td>
<td>27.4 ± 2.3</td>
</tr>
<tr>
<td>Mean ± standard deviation</td>
<td>27.4 ± 2.3</td>
<td>28 (18 to 30)</td>
</tr>
<tr>
<td>Median (range)</td>
<td>28 (18 to 30)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 presents the two study groups’ ZüMAX scores, including the total testing time and scores for self-evaluation of mood. The results of the Shapiro-Wilk test (Norman and Streiner, 2008) showed no normal distribution.

<table>
<thead>
<tr>
<th>Domain (Maximum Points)</th>
<th>Unit</th>
<th>Stroke (n = 33)</th>
<th>Control (n = 35)</th>
<th>Z-score</th>
<th>Mann-Whitney U Test (P = 0.05)</th>
<th>Effect Size r and d (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Test</td>
<td>Retest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive function (6)</td>
<td>Mean ± SD</td>
<td>4.7 ± 1.4</td>
<td>5.0 ± 1.3</td>
<td>5.6 ± 0.7</td>
<td>0.29</td>
<td>0.04 (–0.20 to 0.28)</td>
</tr>
<tr>
<td></td>
<td>Median (range)</td>
<td>5 (1 to 6)</td>
<td>5 (2 to 6)</td>
<td>6 (3 to 6)</td>
<td>0.004</td>
<td>0.08 (0.04 to 0.12)</td>
</tr>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>5.1 ± 1.1</td>
<td>5.1 ± 1.2</td>
<td>5.8 ± 0.5</td>
<td>0.04 (–0.20 to 0.28)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median (range)</td>
<td>5 (1 to 6)</td>
<td>5 (1 to 6)</td>
<td>6 (4 to 6)</td>
<td>0.003</td>
<td>0.08 (0.04 to 0.12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language (6)</td>
<td>Mean ± SD</td>
<td>5.8 ± 0.5</td>
<td>5.8 ± 0.4</td>
<td>5.9 ± 0.3</td>
<td>0.11 (–0.13 to 0.35)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median (range)</td>
<td>6 (4 to 6)</td>
<td>6 (5 to 6)</td>
<td>6 (5 to 6)</td>
<td>0.87</td>
<td>0.386</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>3.3 ± 0.3</td>
<td>3.6 ± 0.2</td>
<td>3.9 ± 0.1</td>
<td>0.18 (–0.06 to 0.42)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median (range)</td>
<td>6 (2 to 6)</td>
<td>6 (1 to 6)</td>
<td>6 (5 to 6)</td>
<td>0.78</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>5.4 ± 1.2</td>
<td>5.6 ± 1.2</td>
<td>5.9 ± 0.4</td>
<td>0.22 (–0.02 to 0.46)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median (range)</td>
<td>6 (2 to 6)</td>
<td>6 (1 to 6)</td>
<td>6 (5 to 6)</td>
<td>1.78</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>25.7 ± 3.9</td>
<td>26.6 ± 3.4</td>
<td>28.4 ± 1.9</td>
<td>0.40 (0.16 to 0.64)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median (range)</td>
<td>26 (14 to 30)</td>
<td>28 (13 to 30)</td>
<td>29 (23 to 30)</td>
<td>3.33</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>33.8 ± 6.6</td>
<td>28.6 ± 6.2</td>
<td>28.7 ± 6.5</td>
<td>0.44 (0.40 to 0.48)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median (range)</td>
<td>35 (20 to 50)</td>
<td>30 (20 to 40)</td>
<td>30 (20 to 50)</td>
<td>3.33</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>3.9 ± 2.2</td>
<td>4.2 ± 2.3</td>
<td>4.4 ± 1.3</td>
<td>0.40 (0.16 to 0.64)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median (range)</td>
<td>4 (–1 to 6)</td>
<td>5 (–1 to 6)</td>
<td>5 (1 to 6)</td>
<td>0.44 (0.40 to 0.48)</td>
<td></td>
</tr>
</tbody>
</table>

NA = not applicable. SD = standard deviation.
**Discriminant Validity**

In addition to the ZüMAX results, Table 2 presents the group differences. Descriptive statistics showed that the controls scored higher in all five cognitive domains and in the total score. For example, for the domain executive function, the patients' median and mean rank scores were 5 and 28.3 versus the controls' 6 and 40.3. The Mann-Whitney U value was significant: \( U = 369.5 \) ($Z = 0.29$), $P = 0.004$. The difference between groups was small: effect size $r = 0.04$; $d = 0.08$.

Overall, we found a significant difference between the two groups in three of the six measures: executive function, language, and total test score ($P$-values ranging from 0.001 to 0.004. The other three domains (praxia, visual perception and construction, and learning and memory) had no significant differences between the groups ($P = 0.075$ to 0.386). We found a small to medium effect size in the cognitive domains ($r$ ranging from 0.04 to 0.22; $d$ from 0.08 to 0.44) and a large effect size for the total score ($r = 0.40$, $d = 0.95$).

**Test-Retest Reliability**

Table 3 shows the test-retest reliability of the ZüMAX for the stroke group, and Figure 4 illustrates the findings in Bland-Altman plots. (Because many participants had the same ZüMAX scores, their data points overlap, and fewer than 33 data points are visible in the plots.)
### TABLE 3. Test-Retest Reliability of the ZüMAX Subscores and Total Score in the Stroke Group (n = 33)

<table>
<thead>
<tr>
<th>Domain (Maximum Points)</th>
<th>Test Mean ± SD</th>
<th>Retest Mean ± SD</th>
<th>Test-Retest Mean Difference ± SD</th>
<th>Correlation</th>
<th>Correlation with Log-Transformed Data</th>
<th>Intraclass Correlation Coefficient (95% Confidence Interval)</th>
<th>Standard Error of Measurement Agreement</th>
<th>Smallest Detectable Difference No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive function (6)</td>
<td>4.7 ± 1.4</td>
<td>5.0 ± 1.3</td>
<td>−0.24 ± 0.97</td>
<td>0.30*</td>
<td>0.16†</td>
<td>0.73 (0.53 to 0.86)</td>
<td>0.69</td>
<td>1.92 (39.48)</td>
</tr>
<tr>
<td>Language (6)</td>
<td>5.1 ± 1.1</td>
<td>5.1 ± 1.2</td>
<td>−0.03 ± 0.77</td>
<td>0.96*</td>
<td>0.99</td>
<td>0.79 (0.62 to 0.89)</td>
<td>0.53</td>
<td>1.49 (28.99)</td>
</tr>
<tr>
<td>Praxia (6)</td>
<td>5.8 ± 0.5</td>
<td>5.8 ± 0.4</td>
<td>−0.03 ± 0.30</td>
<td>0.15*</td>
<td>0.14†</td>
<td>0.74 (0.53 to 0.86)</td>
<td>0.21</td>
<td>0.59 (10.08)</td>
</tr>
<tr>
<td>Visual perception and construction (6)</td>
<td>4.7 ± 1.4</td>
<td>5.1 ± 1.1</td>
<td>−0.36 ± 1.08</td>
<td>0.05</td>
<td>NA</td>
<td>0.59 (0.31 to 0.77)</td>
<td>0.78</td>
<td>2.16 (44.49)</td>
</tr>
<tr>
<td>Learning and memory (6)</td>
<td>5.4 ± 1.2</td>
<td>5.6 ± 1.1</td>
<td>−0.21 ± 1.00</td>
<td>0.05</td>
<td>NA</td>
<td>0.60 (0.34 to 0.78)</td>
<td>0.71</td>
<td>1.97 (35.89)</td>
</tr>
<tr>
<td>Total score (30)</td>
<td>25.7 ± 3.9</td>
<td>26.6 ± 3.4</td>
<td>−0.88 ± 2.25</td>
<td>0.26*</td>
<td>0.24†</td>
<td>0.79 (0.61 to 0.89)</td>
<td>1.66</td>
<td>4.61 (17.65)</td>
</tr>
</tbody>
</table>

**Bold type** indicates log10-transformed heteroscedastic parameters.

*<sup>t</sup> > 0.1 indicates heteroscedastic data.

†<sup>log</sup> decreased.

SD = standard deviation. NA = not applicable (because t < 0.1).
FIGURE 4. Bland-Altman plots of the five ZüMAX cognitive domain subscores and total score. The plots for executive function, praxia, and total score showed a heteroscedastic data distribution, and therefore needed to be transformed logarithmically for further analysis. Because many of the 33 stroke survivors shared the same score, fewer than 33 data points are visible in each plot; the overlap is most obvious in the praxia plot, which has only 5 data points, all falling between 4 and 6 (mean = 5.8 points) in the test and retest. The plots for executive function, visual perception and construction, learning and memory, and total score all showed a trend toward a slight improvement on the retest. Almost all domains had visible within-group differences, with outliers.
Our first hypothesis, about relative reliability (ICC ≥ 0.80), was confirmed by virtue of the ZüMAX total score having an ICC of 0.81 (95% confidence interval: 0.64 to 0.90) (Table 3). For subscores, the ICCs ranged from fair for executive function, language, and praxia (ICC = 0.72 to 0.79), to poor for visual perception and construction (0.59) and for learning and memory (0.60).

Our second hypothesis, concerning absolute reliability, was supported in four of the six ZüMAX measures: The SDDs for executive function, language, praxia, and total score (ranging from 0.05 to 1.49) were all ≤ 10% of the mean average values of the total score and subscores (Table 3). The SDDs of the two cognitive domains visual perception and construction (SDD = 2.16) and learning and memory (SDD = 1.97) exceeded the expected 10%, forcing us to reject the hypothesis. Calculating Kendall tau revealed that the three measures (executive function, praxia, and total test score) were heteroscedastic and therefore needed to be logarithmically transformed (Table 3).

**Discussion**

This study evaluated the discriminant validity and test-retest reliability of the novel cognitive assessment “ZüMAX” in survivors of stroke.

**Validity**

In the discriminant validity parameters with significant group differences, the stroke group scored on average 0.7 point lower than the controls for the domain language, 0.9 point lower for executive function, and 2.7 points lower for the total test score (Table 2). These results correspond in part to those of Kaya et al (2014), who compared MoCA scores for 114 patients with mild cognitive impairment and 246 healthy controls. The authors reported that the most useful domains in discriminating mild cognitive impairment from normal cognition were recall, visuospatial, and language.

In the ZüMAX, however, the differences between groups were nonsignificant: only 0.1 point for praxia and 0.5 point each for visual perception and construction and for learning and memory. It is possible that by the time we recruited our sample of stroke survivors, they had already recovered almost completely in these domains and thus reached near-perfect scores, creating a ceiling effect.

A test is considered to have floor or ceiling effects if >15% of the respondents score the minimum or maximum (Terwee et al, 2007). When we checked for such effects in
the ZüMAX, we found no floor effects but we saw ceiling effects in all three cognitive domains with nonsignificant group differences. For praxia, for example, 84.9% of our stroke group scored the full 6 points, as did 91.4% of the controls. For the total score, 34.3% of the controls scored the full 30 points (and 25.7% reached 29 points), while only 6.1% of the stroke sample scored the maximum (and 30.3% reached 29 points). These ceiling effects make it impossible to distinguish among participants who had perfect scores, indicating limited validity.

Still, it seemed important for the ZüMAX to have a maximum of 6 points for each of the five relevant neuropsychological domains. This would make the total scale score of 30 points comparable to other cognitive screening instruments, especially the MMSE and MoCA, both of which also have a maximum score of 30. Before this scoring is made final for the ZüMAX, however, the scale should be examined in larger samples of stroke survivors and healthy controls, correcting for possible confounding factors such as cognitive status, to find out if ceiling effects remain.

We found that our stroke group needed on average of 5.1 minutes longer than the controls to perform the ZüMAX (Table 2). Needing more time to complete an assessment requires a longer period of concentration, disadvantageous in a population most of whose members likely have cognitive deficits (Van Heugten et al, 2015). One reason for the prolonged test could be impairments in patients’ information processing speed, a domain often affected after stroke (Edwards et al, 2013; Middleton et al, 2014). Another reason is that 66% of our stroke sample had to perform the ZüMAX writing tasks using their affected hand as the dominant hand.

Furthermore, the stroke group self-rated their mood an average of 0.5 point lower than the controls (Table 2). There is evidence that mood can affect cognition and that depression is quite common after stroke (Kimura et al, 2000; Narushima et al, 2007). This might explain our stroke sample’s lower self-assessment on the mood scale and their slowed and less accurate test performance.

It is also well known that education has a main effect on cognitive assessment scores, while sex does not (Kaya et al, 2014). We suggest correcting the ZüMAX for education effects, eg, by adding one point to the total score of participants who have at most 12 years of education, as is done with the MoCA (Nasreddine et al, 2005; Pendlebury et al, 2010).

In our study, the control group was much more educated than the patients (Table 1). We were unable to correct for these education differences, however, for two main
reasons. First, our older participants in particular could not remember exactly how many years of education they had, and we were reluctant to include estimations in our study. Second, most participants, especially those with a university degree or apprenticeship, shared an attitude of life-long learning and had attended several continuing education programs, thus making it impossible for us to determine a true cutoff for years of education. To address this problem, we recommend considering education in the inclusion criteria of future studies.

**Reliability**

The ZüMAX showed fair to good reliability in four of the six measures, the exceptions being the domains visual perception and construction (ICC = 0.59) and learning and memory (ICC = 0.60). In the visual perception and construction domain, copying a 10-element figure (Figure 5) may have been subject to learning effects despite our effort to prevent them by spacing the test and re-test about a week apart. Likewise, in the learning and memory domain, the break between test sessions may not have been long enough to prevent a learning effect.

The other three cognitive domains (executive function, language, and praxia) and the total score, all of which had fulfilled our first hypothesis, had generally good test-retest reliability. Their low SDDs indicate that the ZüMAX is sensitive to individual change (Lexell and Downham, 2005).

Our ICCs were lower than those of other cognitive assessments, eg, the MoCA. The MoCA’s test-retest reliability for patients with cerebrovascular disease was good, ranging from 0.75 to 0.96 (Tu et al, 2013). However, it is unclear which ICC formula Tu and colleagues applied in their statistics. This might be of importance in interpreting the results, as we used single measures (ICC 2,1 formula) to calculate our ICCs, which might have given us lower ICCs than if we had used averages for calculation (eg, ICC 2,k formula) (Shrout and Fleiss, 1979; Weir, 2005).
Copy the geometric figure as well as you can and try to remember it!

![Geometric Figure]

**FIGURE 5.** Visuoconstruction and speed of information processing: ZüMAX worksheet for a task in the domain of visual perception and construction, in English translation. Participants are asked to copy the geometric figure as accurately as they can and to memorize it for later. They are told that “beauty doesn’t matter.” What does matter is that they draw the entire figure with all the parts in the right places and in approximately correct proportions. Participants earn one point for each of these elements: (1) the big rectangle, (2) a correctly placed small triangle, (3) the horizontal middle line, (4) the small rectangle on the far right side, (5) the “plus” sign in the upper left corner, (6) the three small diagonal lines in the lower left corner, (7) the oval, (8) the vertical line tangent to the left side of the oval, (9) the lower diagonal line, and (10) the upper diagonal line. One point is deducted for each markedly displaced or confabulated added element.

**Limitations and Future Directions**

Our study had some limitations. First, we should emphasize that conclusions about the sensitivity of the ZüMAX must await larger investigations that include both a population with cognitive deficits and a comparable population of matched healthy participants. Second, we could have used stricter inclusion and exclusion criteria to minimize heterogeneity between and within the stroke and control groups. For example, stricter criteria could distinguish between survivors of a first-ever or a recurrent stroke, as well as among patients with different subtypes of mild cognitive impairment (vascular, degenerative, mixed). With stricter criteria, we could also stratify the sample by age group, education level, or time since stroke onset, and we could consider pre-stroke cognitive decline (Brainin et al, 2015).
Third, accessing and assessing our patients’ exact stroke diagnosis and medical history was difficult because some of the patients had suffered their stroke many years earlier.

A possible final limitation was our small sample size, which may have affected the values of reliability and measurement error. While the guidelines by Altman (1990) would consider a sample size of 50 as adequate for our purpose, we think that our sample of 33 patients was of reasonable size in this first attempt to evaluate the general usefulness of the ZüMAX in patients with chronic stroke.

We see several directions for future research. One idea is to evaluate the value of the ZüMAX for patients during the acute phase of their stroke, or for an entirely different population, such as people with traumatic brain injury. Cutoff scores have not yet been established for this novel instrument, unlike other similar tools such as the MMSE (≥ 24 points = normal) (Burton and Tyson, 2015) or the MoCA (> 26 points = normal) (Dong et al, 2010; Nasreddine et al, 2005). More work is required here.

It would also be valuable to evaluate the ZüMAX’s ecological validity, defined as the degree to which results obtained under experimental conditions relate to those obtained in natural environments (Tupper and Cicerone, 1990). Chaytor and Schmitter-Edgecombe (2003) reviewed the research on ecological validity of neuropsychological tests, and concluded that many instruments offer moderate ability to predict patients’ everyday cognitive function; the limitation of many tests is that they do not target individual cognitive domains that reflect specific aspects of daily function. However, it is unclear from the literature how strong the relationship between neuropsychological tests and measures of everyday function should be for a test to be considered ecologically valid.

To extend its reach, the ZüMAX could be digitized and integrated into a virtual exercise program for stroke survivors, allowing them to follow their progress when exercising independently at home so that they could further improve their upper limb motor and cognitive skills.

Finally, to increase awareness of the ZüMAX among clinicians, the tool should be translated into other languages. As of May 2016, author P.B. had written an English-language version of the test manual for clinicians, and he was planning a German-language publication on normative data for the ZüMAX in healthy adults. Author B.C.T.-A. had been involved with a Spanish-language translation, which was being tested in...
stroke survivors in Seville, Spain. There was no plan to publish the original German-language version of the test and manual.

In summary, the ZüMAX is a brief, yet comprehensive, domain-specific cognitive assessment for measuring disturbances of neuropsychological function in patients with chronic stroke. The instrument shows moderate to good reliability. In this preliminary study, the total test score showed better results than the subscores for the five single cognitive domains measured. The domains of executive function, language, and praxia showed fair reliability; the domains of learning and memory and of visual perception and construction showed poor reliability. As for discriminative validity, the ZüMAX might discriminate chronic stroke survivors from healthy controls on the three subscales of executive function, language, and learning and memory, as well as on the total score. In conclusion, our study indicates that the ZüMAX can be used as a single direct assessment to provide a “snapshot” of the current state of cognition in survivors of chronic stroke, but further research is required with larger sample sizes.

**Acknowledgments**

The authors thank the occupational therapists who helped recruit the stroke group, the families and friends who helped recruit the healthy controls, and the patients and controls who participated in the study. Further thanks go to Judith Häberli (J.H.), former master’s student in the Department of Health Sciences and Technology, ETH Zurich, for her invaluable support in data collection and analysis, and to Professor Andreas Luft, MD, for his continuous encouragement to apply the ZüMAX in stroke survivors. Finally, thanks to Martin J. Watson, PhD, for proofreading the manuscript for English and structure.
References


Chapter 5
Exergames encouraging exploration of hemineglected space in stroke patients with visuospatial neglect: A feasibility study

Post exergames intervention results of the figure copying task (visuoconstruction) of the ZüMAX of P6, female, 78 years old, 7.5 weeks onset right-hemispheric stroke and visuo-spatial neglect

Abstract

Background: Use of exergames can complement conventional therapy and increase the amount and intensity of visuospatial neglect (VSN) training. A series of 9 exergames – games based on therapeutic principles – aimed at improving exploration of the neglected space for patients with VSN symptoms poststroke was developed and tested for its feasibility.

Objectives: The goal was to determine the feasibility of the exergames with minimal supervision in terms of (1) implementation of the intervention, including adherence, attrition and safety, and (2) limited efficacy testing, aiming to document possible effects on VSN symptoms in a case series of patients early poststroke.

Methods: A total of 7 patients attended the 3-week exergames training program on a daily basis. Adherence of the patients was documented in a training diary. For attrition, the number of participants lost during the intervention was registered. Any adverse events related to the exergames intervention were noted to document safety. Changes in cognitive and spatial exploration skills were measured with the Zürich Maxi Mental Status Inventory and the Neglect Test. Additionally, we developed an Eye Tracker Neglect Test (ETNT) using an infrared camera to detect and measure neglect symptoms pre- and postintervention.

Results: The median was 14 out of 15 (93%) attended sessions, indicating that the adherence to the exergames training sessions was high. There were no adverse events and no drop-outs during the exergame intervention. The individual cognitive and spatial exploration skills slightly improved postintervention ($P=.06$ to $P=.98$) and continued improving at follow-up ($P=.04$ to $P=.92$) in 5 out of 7 (71%) patients. Calibration of the ETNT was rather error prone. The ETNT showed a trend for a slight median group improvement from 15 to 16 total located targets (+6%).

Conclusions: The high adherence rate and absence of adverse events showed that these exergames were feasible and safe for the participants. The results of the amount of exergames use is promising for future applications and warrants further investigations - for example, in the home setting of patients to augment training frequency and intensity. The preliminary results indicate the potential of these exergames to cause improvements in cognitive and spatial exploration skills over the course of training for stroke patients with VSN symptoms. Thus, these exergames are proposed as a motivating training tool to complement usual care. The ETNT showed to be a promising assessment for quantifying spatial exploration skills. However,
further adaptations are needed, especially regarding calibration issues, before its use can be justified in a larger study sample.

**Keywords:** exergames; eye tracking; virtual reality; visuospatial neglect; feasibility

**Introduction**

Unilateral spatial neglect (USN) is characterized by the inability to detect, respond, or orient toward stimuli presented on the contralateral side of a brain lesion [1]. Being a neurological disorder of attention, USN can affect the auditory, visual, or motor system [2-4]. With 43% in the acute phase and 17% at 3 months poststroke, USN is the most common and persistent problem associated with lesions of the right temporoparietal cortex [5]. Furthermore, USN patients share an unawareness of their deficits to different extents. This anosognosia, combined with an associated reduction in the ability to cope with activities of daily living [6], typically results in longer rehabilitation periods [7-10]. Therefore, USN is a predictor of poor outcome in stroke patients and an added burden for the health care system [7,8,11] requiring efficient treatment modalities [2,12-14].

A variety of accepted and proven traditional methods exist to treat USN [2,15-19], such as pharmacological interventions [20], different physiological sensory stimulations [21-23], and cognitive behavioral training [24]. A combination of multiple approaches to develop a personalized rehabilitation process is recommended [25,26] together with use of a battery of tests to assess USN rather than a single sole assessment [14,27]. However, none of these traditional methods could completely rehabilitate the condition, and rehabilitation methods investigating new approaches are warranted.

Virtual reality (VR), defined as “an advanced form of human-computer interface that allows the user to ‘interact’ with and become ‘immersed’ in a computer-generated environment in a naturalistic fashion” [28] shows some preliminary evidence favoring its use, and further investigations in stroke rehabilitation may complement traditional USN treatment methods [29-32]. VR methods provide a safe copy of the real environment while allowing the creation of customized rehabilitation programs through progressive, repetitive training with immediate feedback [13,19,26,33,34]. Promising VR instruments exist both for the assessment [35-39] and rehabilitation [30-32,40-44] of neglect [26]. The VR assessments were not only able to accurately detect USN
patients but also made USN-related symptoms visible that were previously not identified with conventional assessments [26,33,34]. The VR systems tested for rehabilitation, however, were mostly complex to set up or used rather expensive tools (eg, head-mounted displays or cyber gloves), restricting their use to laboratory settings.

The European research consortium Rehabilitative Wayout in Responsive Home Environments (REWIRE) developed a nonimmersive VR system for stroke patients using portable devices with good performance and affordable equipment [45]. By creating such a VR system, REWIRE aimed to facilitate its use for stroke patients discharged from the hospital to allow continuation of the rehabilitation process within their own homes. A variety of home-based VR systems already exist for stroke patients mainly focusing on motor recovery [46-49], but none exist for USN. Therefore, the consortium designed, among others [50], exergames for the treatment of visuospatial neglect (VSN) (VSN being a subtype of USN [4]) [41,51]. Exergame is a portmanteau of the words exercise and game [52], allowing the patients to exercise their skills through gaming. In contrast to games that are designed for diversion for healthy persons, exergames should follow therapeutic principles - for example, the principles of exercise training, such as specificity and progression [53], or adopting the training method of shaping [54], including frequent feedback and the selection of tasks addressing the individual deficits of patients. The REWIRE consortium adopted the principles of exercise training as described by Hoffman [53] to design the neglect exergames. Therefore, the games include the principles of (1) specificity, implying that the required performance of each game corresponds to the goal of the game (to explore the neglected space) and (2) overload and progression, stating that the components being used must be exercised at a level the patient is not normally accustomed to and the patient should progress once accustomed to a level. In order to quantify training progression from simple to complex within each game, the REWIRE consortium used Gentile’s taxonomy of motor skills as a template to develop the exergames (see Borghese et al [55] and Wüest et al [56] for more detail). Due to the nature of neglect and the related unawareness of neurological deficits [4], it is important to test such a novel intervention with the target patient population in a surrounding where close monitoring is possible and feasibility of the approach can be tested. Feasibility may cover aspects such as adherence, safety, and attrition to the novel intervention or whether the intervention and assessments all run smoothly [57].
Mainetti et al [41] and Sedda et al [32] already tested a former version of the REWIRE VSN exergames in a single-case study design involving a neglect patient in the chronic stage. The results were promising in terms of a positive attitude of the patient toward the exergames and in showing a trend for improvement of the VSN-related deficits in daily life. Based on the experience with the exergames of this single user together with feedback on their usability, the exergames were adapted and improved and then tested in this study for the first time in a case series of patients. We aimed to test the exergames in early stroke patients shortly before their discharge to home. This time point was chosen to include as realistic a target population as possible while still guaranteeing safety and supervision of the patients playing the exergames in the supportive environment of the rehabilitation clinic. Specific aims were to determine the feasibility of the exergames with minimal supervision in terms of (1) implementation of the intervention, including adherence, attrition, and safety, and (2) limited efficacy testing, aiming to document possible effects on VSN symptoms in patients early after stroke.

Methods

Study Design

We adopted a quasi-experimental pretest-posttest design with a subsequent follow-up to test the feasibility of the exergames in a case series of stroke patients with VSN symptoms. As we aimed to assess implementation of the exergames, thus testing if our intervention can be fully implemented as planned and proposed, an uncontrolled pretest-posttest design is appropriate [57]. A broad variety of definitions exist for the concept of case series in literature [58]. For our study, we used the definition of a case series as being a “report on a series of patients with an outcome of interest” [59]. Recruiting a small convenience sample was ideal for the planned limited efficacy testing, as we aimed to gain intermediate rather than final outcomes in this feasibility project, which allowed us to plan a shorter follow-up period [57].

Patients

Identification of potential patients for this project was carried out by staff neuropsychologists and occupational therapists in 2 collaborating rehabilitation clinics (Klinik Bethesda Neurorehabilitation, Parkinson-Zentrum, Epileptologie, Tschugg,
Bern, and Zürcher RehaZentrum Wald, Faltigberg-Wald, Zurich). They screened all incoming stroke patients with a diagnosis of VSN for eligibility in this study. We aimed for at least 5 participants, as this amount is considered the minimum reasonable number of independent subjects in a group to combine their data [58]. Fewer than 5 patients are usually presented in a descriptive, narrative form of individual case reports. We strived for a maximum of 10 patients as recommended by Abu-Zidan et al [58]. We included patients with a right brain lesion (RBL) due to a first stroke 15 to 180 days after the cerebral event and a diagnosed VSN as measured by the Catherine Bergego Scale (CBS) [60]. Inclusion criteria were being able to sit in a chair or wheelchair with a backrest for 45 minutes, being at least 18 years old, and having a clear view (with or without vision aids) of a computer screen placed at a distance of 60 to 65 centimeters from patient’s face. VSN patients were excluded if their neglect was diagnosed as due to brain injury other than stroke, if severe apraxia was present—measured as less than 5 points on the Apraxia Screen of TULIA (test for upper limb apraxia) (AST) [61]—or if other noncontrolled medical conditions (eg, chronic pain, drug abuse) were present. Patients with a left brain lesion due to a first stroke were excluded because the exergame difficulty levels were designed to progress from the right (easy) to the left (difficult) side of the computer screen. An option to run the games vice-versa (from left to right) was not available.

All patients signed written informed consent before study entry. Ethical approval for the study was received from the local ethics committees (Zurich No. 2014-0543 and Bern No. 389/2014) as well as from the Swiss agency for the authorisation and supervision of therapeutic products (Swissmedic, 2015-MD-0003). The latter approval was required as the software was not yet certified with the European Community marking for medical devices. The study is registered at ClinicalTrials.gov [NCT02353962].

Setup

Patients were seated at a table in front of a 21-inch computer monitor at a distance of 60 to 65 cm in order to provide optimal eye tracking (Figure 1). We chose a seated position to allow more patients to participate (eg, wheelchair users) and avoid exhaustion through standing in an upright position. A height-adjustable chin rest (Novavision GmbH) was mounted on the table to avoid compensatory head movements while playing the exergames. Instead of a mouse to control games, a haptic Falcon Novint device (Novint Technologies) was used. This enabled individuals...
to experience a realistic sense of touch by providing simulated sensory feedback when reaching for and grasping virtual objects [62]. The Falcon Novint device can be handled with one hand only, allowing stroke patients to play the exergames with their nonaffected hand. The device was placed at the side of the computer monitor at a distance allowing ease of reach for the patients. The nonaffected upper extremity was positioned in approximately 45° shoulder abduction, 70° to 90° elbow flexion, and the forearm fully pronated. All participants were expected to independently complete 15 training sessions while being monitored by a supervising therapist. The supervision included observation of the patient during the intervention giving assistance where appropriate (eg, using the menu to start a new game). Observation was necessary for assistance if potential software difficulties occurred and for safety reasons for the patient, the latter being a regulation of the collaborating clinics.

Figure 1. Setup of the exergames training station.
Exergames

The intervention program consisted of a series of 9 exergames performed while seated. The games were designed to simulate real-world tasks, such as cooking from a recipe, going for a walk with a dog, or doing a puzzle (for detailed game information on 4 games, see Pirovano et al [51]). The Falcon Novint represented, for example, a dog leash by simulating a pull from the dog to the left or right side of a virtual walking path. The increase in difficulty of all games during the training course was accomplished according to Gentile’s taxonomy of motor skills [56]. Using this systematic classification to design the exergames allowed us to design a theory-based rehabilitation program that followed the principles of exercise training (see Hoffman [53] and Ammann et al [63] for detailed descriptions of these principles). Playing time per game was adjusted from 1 to 10 minutes per game depending on the patient’s ability to concentrate playing the VR game while maintaining a seated position. After initial training and instructions were given by the research team, all subsequent game adjustments during the intervention were performed by clinic staff (occupational therapists and neuropsychologists) in accordance with the patient’s wishes.

Intervention Protocol

The VR-based VSN training intervention took place in the 2 collaborating rehabilitation clinics serving as an additional therapy option to the standard program, which comprised daily occupational, physical, and neuropsychological therapy. Each patient was asked to attend 5 30- to 45-minute sessions per week for 3 weeks. The supervising therapist individually adjusted the intensity of playing the exergames by changing the difficulty level or game duration in the game menu and by deciding if short breaks between each game would be necessary or not. In accordance with the training principle of individuality [53], which states that people respond differently to the same training stimulus, the patient selected 3 to 4 REWIRE VSN exergames from the game menu to be played in each session. The choice was based on personal interest of the patient, which was assumed to enhance motivation while playing. During the 3-week intervention time, patients were allowed to change games if they wanted to test another one or felt bored with the previously played game. After a break of 4 weeks, a follow-up measure was performed aiming to test the training principle of reversibility, which states that the ability to maintain performance is reduced when the training stimulus (the exergames) is removed.
Assessments

Primary Outcome
In order to measure the likelihood and extent to which our intervention can be fully implemented as planned and proposed [57], we designed a training diary as a protocol to document attrition, adherence, and safety issues. This training diary was on hand in the collaborating clinics and completed after each training session by the clinic staff in presence of the participating patient. The type of games played including difficulty level according to the Gentile’s taxonomy, effective training time, and patient subjective statements regarding their perceived health condition after training (posed question: “How do you feel after training: fit or tired?”) were all noted in the training protocol. Additionally, any adverse events related to the exergames intervention were noted. Potential adverse events could have been a recurrent stroke or other medical emergencies due to the early stage of recovery or an epileptic seizure or cybersickness due to playing the exergames [64]. For attrition, the number of participants lost during the intervention was registered. For adherence, participant engagement with the intervention was noted. We expected a good adherence to the intervention, defined as an attendance of at least 50% of the maximum 15 possible training sessions. Adherence was then calculated as the number of completed training sessions as a percentage of the maximum 15 possible training sessions.

Secondary Outcomes

Overview
In order to test limited efficacy of our intervention [57], the Eye Tracker Neglect Test (ETNT), Zürich Maxi Mental Status Inventory (ZüMAX), and Neglect Test (NET) were administered by the research staff at baseline and after the intervention. After a 1-month follow-up, the ZüMAX and NET were repeated either in one of the collaborating clinics or at the patient’s (new) residence (home or retirement home), depending on the length of rehabilitation stay.
Spatial Exploration Skills

The ETNT is an adapted version of the cancellation test developed by Rabuffetti et al [34], aiming to assess symptoms of neglect. In contrast to the original test, in which targets are tagged by finger touch as measured by a touch screen, target detection is operated by eye fixation as measured by an eye tracker (Figure 2). The ETNT display consists of a uniform distribution of 60 stimuli including 20 targets (squares) and 40 distractors (other than square shapes) divided equally into 30 stimuli (10 targets, 20 distractors) on the right and left sides of the computer screen. The eye point of gaze was tracked with the Eye Tribe Tracker camera (Eye Tribe). Eye tracking is the process of using sensors to locate features of the eyes and estimate where someone is looking [65]. The technology relies on infrared illumination so that it does not interfere with the visual scenario. Since the system tracks eye movements relative to the sensor/screen, it is necessary to fix the head position, since head movements would be wrongly assumed as eye movements. Therefore, the patient’s head was fixed on a chin rest. In
our study, each participant was seated in front of the screen with the midsagittal plane of the trunk aligned with the center of the screen. An initial calibration of the Eye Tribe Tracker camera was then followed by 1 test trial with only 4 targets and 8 distractors and the ETNT with 60 stimuli. A stimulus was counted as being found (being circled) if the patient maintained his or her point of gaze for at least 0.4 seconds within an area surrounding the stimulus with a diameter of 7% of the total screen width, in keeping with Blignaut et al [66]. There was no time constraint; patients were instructed to inform the researcher when they had finished the test. This procedure was chosen to impose no stress on the patient while exploring the targets on the screen. However, if the patient got lost or became tired, the researcher present during the test asked the patient if he or she had the impression of having found all targets and then stopped the test depending on the patient’s response.

The Neglect Test (NET) consists of 7 paper-and-pencil subtests (letter and star cancellation, copying 3 figures, and line crossing and bisection with a total possible score of 70 points) and 10 behavioral subtests (representational drawing, scanning 3 pictures, menu and article reading, telling time, setting time on a digital and analog watch, and address copying with a total possible score of 100 points) designed to identify a wide variety of visual neglect behaviors [67]. It has been shown to be a robust predictor of VSN and is a predictor of functionality after stroke [68]. To assess the level of anosognosia for VSN after stroke, self-ratings of performance in 6 subtests of the NET (figure copying, star cancellation, line crossing and bisection, representational drawing, and article reading) were contrasted with external performance ratings of the examiner on a 5-point Likert scale (ranging from 1=severe difficulties to 5=no difficulties). The degree of unawareness for VSN was quantified as proposed by Vossel et al [69] (see Figure 3).

This anosognosia index (AI) will be smaller than 0 if the patient suffers from anosognosia, indicating an overestimated self-performance to what objectively has been performed. If the patient is able to rate his or her performance realistically, thus being below or matching the external rating, the index becomes equal to or greater than 0, indicating no signs of anosognosia.

**Figure 3.** Formula for anosognosia index.

\[
\text{Anosognosia index} = \frac{\sum \text{external rating} - \sum \text{self-rating}}{\sum \text{external rating}}
\]
Cognitive Skills

The ZüMAX is a domain-specific assessment tool measuring cognitive impairment by evaluating executive function, language, praxia, visual perception and construction, and learning and memory (see Tobler-Ammann et al [70] for a detailed description of the test). Each of the 5 domains allows a maximum score of 6 points, with a maximum possible test score of 30 points, representing optimal cognitive functioning. The ZüMAX has moderate to good test-retest reliability for the total test scores in patients 6 months or more poststroke and may discriminate between this patient group and healthy age and gender matched persons [70]. The ZüMAX visual perception and construction domain is the one indicating VSN symptoms. The task for visual perception is to recognize and name degraded figures, unfamiliar scenes, and a face. The task for visual construction consists of copying a figure. This assessment was chosen due to its advantage of providing both general information about poststroke cognitive impairment and neglect-specific information and because of its origin in Switzerland where the study took place and, therefore, matching the cultural background of the participants.

Data Analysis

Data analysis was carried out using SPSS for Windows version 23.0 (IBM Corp). A Shapiro-Wilk test was administered and quantile-quantile plots were drawn to test normality of the data. The results confirmed our assumption of nonnormally distributed data due to the small sample size ($P \leq .05$ for most parameters). We therefore used nonparametric tests for data analysis. Accordingly, a Wilcoxon signed rank test was adopted to compare post- with preintervention results and follow-up with postintervention results. The Friedman test was used to test for differences in the NET and ZüMAX between all 3 measurement time points [71]. We analyzed the data for each individual and for the whole group. The fact that the ZüMAX and NET comprise subtests and the NET additionally provides a conversion table to transform raw scores into standard scores allowed us to use the Wilcoxon signed rank test not only on a group level but also on an individual level. For calculations per patient, we used the achieved standard scores of each subtest as variables, resulting in 17 variables (corresponding to the 17 NET subtests) for the NET and 5 variables (corresponding to the 5 ZüMAX domains) for the ZüMAX. For the analysis on a group level, we compared the achieved total scores per measurement point of the 7 patients.
The ETNT data were provided by the software described in Rabuffetti and colleagues [34]. A subset of the relevant indexes - namely those that were related to visual perception - was used for data analysis as only these items were suitable for the adapted test version. As the ETNT software provided 1 value per index and patient, we used the Wilcoxon signed rank test to analyze post- to preintervention changes within the sample. Additionally, we graphically displayed the individual changes post-pre intervention by drawing the performed search path and fixation points and creating heat maps to visualize group changes post-pre intervention.

In order to perform an a priori power analysis to determine the minimum sample size for a future randomized controlled trial, we calculated the effect sizes for the secondary outcome measures. We applied the Cohen formula for nonparametric tests [72] (see Figure 4). Accordingly, small, medium, and large effect sizes were labeled as $r=0.1$, $0.3$, and $0.5$, respectively [73]. The level of significance was set at $P \leq 0.05$.

**Figure 4.** Cohen formula for nonparametric tests.

$$r = \frac{Z}{\sqrt{n}}$$

where $n$ is the number of pairs and $Z$ being the converted U score.
Results

Overview

From the 18 VSN patients consecutively screened for eligibility in both clinics from March 2015 to March 2016, 7 patients (39%) were eligible and consented to participate in this study, therefore taking part in the VR exergaming program including baseline, postintervention, and 3-month follow-up measures. Reasons preventing patients from participating were suffering from a right-sided VSN due to a left brain lesion, having a severe apraxia (fewer than 5 points on the TULIA (AST) screening instrument, and being in a poor health condition confining them to bed. Patient characteristics are presented in Table 1.

Table 1. Patient characteristics.

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age, years</th>
<th>Sex</th>
<th>Days post-stroke at study entry</th>
<th>RBL stroke type</th>
<th>Handedness/affected hand (function)</th>
<th>Education</th>
<th>Locomotion</th>
<th>CBS b,i</th>
<th>AST c,j</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>64</td>
<td>M</td>
<td>25</td>
<td>ischemia</td>
<td>R/L (none)</td>
<td>PE d</td>
<td>WC e</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>P2</td>
<td>74</td>
<td>M</td>
<td>29</td>
<td>ischemia</td>
<td>R/L (back)</td>
<td>PE</td>
<td>W f</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>P3</td>
<td>64</td>
<td>M</td>
<td>114</td>
<td>ischemia</td>
<td>R/L (none)</td>
<td>PE</td>
<td>WC</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>P4</td>
<td>70</td>
<td>M</td>
<td>32</td>
<td>ischemia</td>
<td>R/L (back)</td>
<td>SE g</td>
<td>WC</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>P5</td>
<td>53</td>
<td>M</td>
<td>42</td>
<td>ischemia</td>
<td>R/L (none)</td>
<td>PE</td>
<td>W</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>P6</td>
<td>78</td>
<td>F</td>
<td>35</td>
<td>hemorrhage</td>
<td>R/L (back)</td>
<td>SE</td>
<td>W</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>P7</td>
<td>77</td>
<td>F</td>
<td>47</td>
<td>hemorrhage</td>
<td>R/ L (back)</td>
<td>PE</td>
<td>WC</td>
<td>16</td>
<td>9</td>
</tr>
</tbody>
</table>

IQR h

| 25 | 64 | — | 29 | — | — | — | — | 5 | 12 |
| 50 | 70 | — | 35 | — | — | — | — | 7 | 12 |
| 75 | 77 | — | 47 | — | — | — | — | 16 | 12 |

aRBL: right brain lesion.  
bCBS: Catherine-Bergego Scale.  
cAST: Apraxia Screen of TULIA.  
dPE: primary education.  
eWC: wheelchair.  
fW: walker.  
gSE: secondary education.  
hIQR: interquartile range.  
iMaximum score = 30 (severe neglect); 0 points = no neglect.  
jMaximum score = 12 (no apraxia); threshold for apraxia: <9 points; severe apraxia: <5 points.
Primary Outcome
An overview of individual (P1-P7) and group (interquartile range [IQR], mean) results in the training protocol is shown in Multimedia Appendix 1. There were no adverse events and drop-outs during the intervention. A median attendance of 14 (IQR 12-15) training sessions (maximum 15 sessions) was achieved, which corresponds to a median adherence of 93% (IQR 80%-100%). Reasons for nonparticipation were of organizational or medical nature (eg, overlap with other therapy sessions or due to fatigue) rather than because of motivational factors. All patients played 2 to 4 games and repeated at least 1 game per training session. The supervising therapists adapted and individually progressed the patient training protocols on a weekly basis during the exergames intervention in accordance with patient progress. If, for example, the patient got bored with the current difficulty level of the played game or the therapist observed that the game was played without effort, the therapist modified the difficulty level within each game. However, if the patient had reached the most difficult level, the therapist replaced easy games with more complex ones (ie, games including more distractors or moving objects). An analysis of the progress as measured by the achieved game scores was therefore not feasible, as progression in difficulty resulted in a temporary decrease in game scores. Instead, the progress in the exergames training of the 7 individual patients was documented weekly according to the Gentile’s taxonomy of motor skills. These results are shown in Multimedia Appendix 1.

Secondary Outcomes
Spatial Exploration and Cognitive Skills
An overview of the individual ETNT scores and group changes post- to preintervention is shown in Multimedia Appendix 2. Figure 5 shows 2 examples of pre-post intervention search path strategies and fixation points as measured by the eye tracker camera (see Multimedia Appendices 3 and 4 for all graphs of individual post-pre ETNT search paths and fixation points). Figure 6 shows the heat maps of the pre- and postintervention and differences in post-pre detected targets of the ETNT. An overview of ZüMAX, NET, and AI scores and group changes pre-, postintervention, and at follow-up is summarized in Multimedia Appendix 5 and graphically displayed in Figure 7 (overview) and Multimedia Appendix 6 (individual results per outcome measurement). The individual changes in the ZüMAX and NET assessments are presented in Multimedia Appendix 7.
**Figure 5.** Examples of pre- and postintervention results of the Eye Tracker Neglect Test search paths and fixation points of P3 and P4.

<table>
<thead>
<tr>
<th>patient</th>
<th>preintervention</th>
<th>postintervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3 – search path</td>
<td><img src="image1" alt="P3 preintervention" /></td>
<td><img src="image2" alt="P3 postintervention" /></td>
</tr>
<tr>
<td>P3 – fixation points</td>
<td><img src="image3" alt="P3 fixation preintervention" /></td>
<td><img src="image4" alt="P3 fixation postintervention" /></td>
</tr>
<tr>
<td>P4 – search path</td>
<td><img src="image5" alt="P4 preintervention" /></td>
<td><img src="image6" alt="P4 postintervention" /></td>
</tr>
<tr>
<td>P4 – fixation points</td>
<td><img src="image7" alt="P4 fixation preintervention" /></td>
<td><img src="image8" alt="P4 fixation postintervention" /></td>
</tr>
</tbody>
</table>

The bright green square signifies the first located target of the ETNT, thus the first square the participant looked at for ≥ 0.4 sec in the target-area (defined as a 7% area of the computer screen).

A circle around a tagged item represents the latency associated with its hit. Latency is the time passed from the hit on the previous item. Therefore, the bigger the circle, the more time passed. A single item may have several circles because it can be hit more than once (perseverations for the same target). Except for the case of quite similar latencies on repeated hits – apparently only one circle shows up because they superimpose - the number of circles marks the number of repeated hits.

The bright blue line represents the middle line of the two computer screen sides, dividing the 20 total targets into 2x 10 items.

The dark blue dots are the visualization of the “points of gaze” (estimates where someone is looking, as measured by the sensors of the Eye Tribe Tracking infrared camera). A fixation point is identified as the average of a “bunch of points of gaze” which stay close to each other for at least 100 ms.
**Figure 6.** Heat maps of the preintervention, postintervention, and difference post-pre results of the Eye Tracker Neglect Test.

The a) pre- and b) post treatment heat maps refer to a percentage ratio of detection (0% = no detection to 100% = full detection) of targets. The blue squares represent the positions of the targets.

c) The values in the post-pre treatment heat map refer to the differences between percentage target detection rates as obtained post- vs pre treatment.

**Figure 7.** Graphical overview of preintervention, postintervention, and follow-up Zürich Maxi Mental Status Inventory, Neglect Test, and anosognosia index results.
Eye Tracker Neglect Test

**Group Level**

The pre-post assessment showed a median group trend of slight improvement in the total located targets from 15 to 16 (+6%) pre-post, which was concomitant with an increasing median total test duration (+33.9 seconds pre-post) (Multimedia Appendix 2). The postintervention performance showed a median decline of 2 (IQR 0-4) missed targets on the left side of the screen and a median increase of 1 (IQR 0-3) missed target on the right side of the screen. The heat maps (Figure 6) indicate that the ability to detect targets in the upper left portion of the computer screen increased postintervention and remained substantially unmodified otherwise. The results further showed an unchanged median group trend pre-post intervention regarding the neglect score, the median latency, and proximity. There were no statistically significant changes pre-post intervention for any ETNT parameters (Multimedia Appendix 2).

**Individual Level**

P4 improved from 5 to 15 total located targets pre-post intervention (Multimedia Appendix 2 and Figure 5). The only individual worsening pre-post regarding the left spatial exploration skill was P3, with no target missed preintervention and 4 targets missed postintervention (Figure 5). Most participants (6/7, 86%) started their search in the right sector and scanned leftwards vertically (Multimedia Appendices 3 and 4), while P3 started in the upper left sector and scanned top-down horizontally (Figure 5). Additionally, 4 out of 7 (57%) started their search on the right side of the screen, 2 out of 7 (29%) were able to change the starting point from right to left, and 1 participant (P3) started the search on the left side of the screen pre- and postintervention (Multimedia Appendices 3 and 4).

**Neglect Test**

**Group Level**

The NET showed statistically significant improvements pre- to postintervention ($P=.01$ to $P=.03$) in the total score and both subtests as well as in the Friedman test ($P=.01$ to $P=.02$) (Multimedia Appendix 5). The AI results showed a group trend toward zero (median 0.08 pre, 0.04 post, 0 follow-up), indicating a trend toward perfect awareness of oneself (Multimedia Appendix 5).
**Individual Level**

P1 and P2 declined at follow-up in the NET total scores, while P3 to P7 showed further improvement. P1, P5, and P7 significantly changed in their NET total scores, P1 post–follow-up ($P=.04$) and P5 and P7 pre-post–follow-up (Friedman test; $P=.01$ and $P<.001$, respectively) (Multimedia Appendix 7). The post–follow-up result of P1 also showed a large effect ($r=-.80$), indicating a large decline in the NET scores postintervention, especially in the paper-and-pencil subtasks (Multimedia Appendix 6). Analyzing NET scores by subtests, all patients improved pre- and postintervention in the paper-and-pencil subtests (Figure 7), while in the behavioral subtests, all patients improved pre- to postintervention except for P3 who remained unchanged (Multimedia Appendix 6). Two patients (P3 and P7) suffered from anosognosia with a trend of aggravation from pre- to follow-up assessments (Figure 7). P1, P2, and P4 showed no signs of anosognosia at follow-up (AI index=0), while P5 and P6 showed a trend toward an increasing AI index, underestimating their actual performance (Multimedia Appendix 6).

**Zürich Maxi Mental Status Inventory**

**Group Level**

The ZüMAX showed improvements in the total scores from pre- to postintervention to follow-up (Figure 7) that were not statistically significant ($P=.29$ to $P=.45$, median +2 points pre-post, +1 point post–follow-up) (Multimedia Appendix 5).

**Individual Level**

On an individual level, P1, 2, 5, and 6 improved from pre- to postintervention in the ZüMAX total scores, while P3 and 4 declined and P7 remained unchanged (Multimedia Appendix 6). The post–follow-up scores declined in P1 and P7, while P2 to P6 still improved their post–follow-up scores, with P3 and P4 declining pre-post intervention. In the ZüMAX visual perception subtask (recognizing and naming degraded figures), 3 of the 7 patients stayed unchanged from pre-post to follow-up (P3, P5, P6), while the largest progression was apparent in P7, with a clear decline of 3 points after termination of the exergames intervention (Multimedia Appendix 6). In the second ZüMAX subtask related to neglect (visual construction: figure copying), 3 patients performed the pre- and postintervention assessments with unchanged scores (P3, P6, P7), and 3 patients showed a slight decline post–follow-up (P1, P3, P5).
Discussion

Principal Findings

This study evaluated the feasibility of an exergames intervention aimed to affect VSN symptoms in patients early poststroke in terms of implementation (adherence, attrition, and safety) and limited efficacy testing by documenting changes in VSN symptoms. The exergames intervention was tolerated well by all participants and was mainly performed without major difficulties, showing that its implementation in the clinical setting was feasible. With 0 out of 7 (0%) attrition, no adverse events, and a median adherence rate of 14 out of 15 sessions (93%), the compliance of the patients to the exergames was excellent. Such a result was possibly due to the clinic staff’s commitment, as the VR intervention was smoothly integrated into the daily therapy schedule of the clinic. However, as we aimed to test the fit of our intervention in a real-world setting, we prioritized clinic constraints over optimal conditions and settings. As a consequence, this priority reduced potential omissions of training sessions as described in purely home-based VR interventions [46]. There, the level of use of the VR system was variable and fell far short of recommendations, despite the weekly or biweekly visits of a researcher to the patients’ homes to check progress and retrieve data.

Other studies testing novel VR systems for upper limb stroke rehabilitation have also shown high levels of adherence to the training intervention [74-76]. However, these patients were in the chronic stage of recovery and did not suffer from USN. The single participant with USN of the Duckneglect study (Mainetti et al [41]), who also was in the chronic stage of stroke recovery, showed an excellent adherence to the exergames, in keeping with our case series.

Regarding the exergames training, the median duration per session was 30 (IQR 23-30) minutes, which fell short of the planned 30 to 45 minutes of training time. For our study, we decided to set a timeframe rather than an exact exposure time, because little is known about the optimum duration and patterning of training exposure to virtual environments [77]. Possible reasons for the rather short training sessions in our sample were twofold: either patients (eg, P2 and P3) were quite fit and finished the planned exergames session early or, conversely, were too tired to play longer (eg, P1). In our sample, 5 out of 7 (71%) patients indicated being fit after most trainings (median 13 out of 15 [87%] possible sessions), excluding fatigue as being a main reason for the rather short training time. P7, however, being the only participant with a highly
distinctive VSN as measured with the NET, needed on average 34 minutes to accomplish the games. These results highlight the importance of adjustability of the difficulty level of the exergames to the functional ability of the stroke patient [76], thus avoiding a decline in enjoyment level while playing [77]. Given the uncertainty about optimal exposure together with our small and highly varied sample (Table 1, CBS scores), future trials should focus on investigating optimal thresholds of exposure time corresponding to the degree of VSN severity.

The fact that there were no adverse events during the training period in our case series was encouraging. Besides being lucky that no recurrent stroke or other medical emergency happened during the intervention, the design of the games might have contributed as well to the safety of our participants. For example, implementing both stationary and in-motion conditions of the virtual scenario together with the option of choosing between intertrial variability and no intertrial variability while gaming allowed the patient to choose the optimal virtual environment to be challenged on the one hand but not be overwhelmed on the other hand (see Multimedia Appendix 1). Allocating these options in difficulty level might have contributed to reducing the risk of cybersickness while playing despite the stationary condition of the patient [64]. When designing the setup of our study, we intentionally planned a seated position to play the games. This allowed patients to fully concentrate on the exploration of the neglected space without having to invest energy standing in an upright position. Furthermore, playing the games in a seated position contributed to the prevention of falls. Prahm et al [78], for example, also designed a game-based intervention in a seated position reporting no adverse events. However, their participants were able-bodied adults. Wiloth et al [79] reported no adverse events in their game-based assessment to measure motor-cognitive function in people with dementia while they were standing on a movable platform. Although falls are highly prevalent in people with cognitive impairment such as dementia [80], people poststroke additionally suffer from motor impairment. Despite their hemiparesis, our participants were all able to perform the games and handle the Falcon Novint with the nonaffected hand. The clinic staff reported that sometimes the more concentrated or – toward the end of the training session – the more tired the participants became while playing the exergames, the more they tilted to the left side with their upper body. As this is a well-known phenomenon in VSN patients poststroke [81], we think our approach of offering
Gameplay in a seated position guaranteed patient safety. The rather tilted position, however, did not prevent the participants from continuing to play the exergames. Our limited efficacy testing showed a group trend of improvements in cognitive and spatial exploration skills. However, these changes cannot be exclusively attributed to the exergames intervention. One reason is the ongoing VSN treatment in the rehabilitation clinic that might also explain some of the improvements. A further possible reason is spontaneous recovery of VSN symptoms not only during the acute phase after stroke but also during the following few weeks. Paolucci et al [82], for example, reported a decrease of VSN symptoms to 20% from 45% after 1 month poststroke, which may also have occurred in our sample. Additionally, the heterogeneity in our sample regarding neglect severity – 4 out of 7 (57%) were only mildly affected with CBS scores between 5 to 7 points (Table 1), while 2 out of 7 (29%) were severely affected (16-17 scores on the CBS) – might have influenced the rate of improvements, too. However, our sample did not show a ceiling effect – being present if 15% or more participants achieve the highest possible score [83] – as no participant achieved the highest score in any of the outcome measurements. Furthermore, most patients continued improving their scores in the NET and ZüMAX assessments postintervention, achieving the highest scores after a break of 4 weeks (follow-up). Only P1 in both tests and P2 (NET) and P7 (ZüMAX) showed a decline in scores from post to follow-up, which would correspond to the training principle of reversibility [53, 63] that states that once a training stimulus is removed, performance levels will eventually return to or below baseline. Comparing our efficacy testing results with literature was difficult, as studies with a similar setup, time point of measurement, and target group are scarce. There is evidence that training with VR methods improve spatial attention and show transfer of improved spatial attention in activities of daily living in chronic neglect [30, 32]. Kim et al [84] showed additional benefit for treating cognitive impairment in stroke patients without VSN in the subacute phase of recovery when adding VR training to classical cognitive rehabilitation. The evidence supports our findings that using VR systems to treat cognition in stroke patients is promising and feasible; however, further research is warranted and necessary to test its use in patients with VSN symptoms early poststroke. Future studies with a focus on treatment effects using controlled research designs should be used to assess causal relationships between the game-based interventions and important patient outcomes.
The median AI values in our sample ranged from 0.08 preintervention to 0.0 at follow-up. Comparing those results to bigger RBL stroke samples with USN – mean (SD) lower AI –0.28 (0.5) for n=34 and mean (SD) higher AI –0.47 (0.5) for n=22; Vossel et al [69] and mean (SD) AI –0.16 (0.38) for n=55; Vossel et al [6] – our 7 patients showed quite a high level of self-awareness for their visuospatial deficits, including P3 and P7 who scored below zero (indicating anosognosia) during all 3 measurements. Of the 7 patients, 3 continued increasing their level of self-awareness postintervention to zero (P1, P2, P4), suggesting positive effects of time poststroke on anosognosia rather than our training intervention. However, time poststroke might not be a viable indicator for those continued improvements in self-awareness of neglect, as Vossel and colleagues [6] found no differences in their AI values across their 3 patient subgroups differing in time since stroke onset. Therefore, continued rehabilitation might be a plausible reason for those further improvements. However, looking at the other 4 patients, AI values showed a tendency toward worsening from preintervention to follow-up, with P3 and P7 overestimating and P5 and P6 underestimating their NET performance, although they also had continued rehabilitation. There is evidence that anosognosia for spatial deficits is not predominant, with different tasks evoking different degrees of awareness about the neglect symptoms [85]. As the AI is calculated on the basis of 4 paper-and-pencil and 2 behavioral tasks of the NET, it might be that this mix of tasks also influenced the miscellaneous AI results. For example, Ronchi et al [85] found that anosognosia level improved after performance of complex visuomotor (eg, cancellation and drawing) and reading tests. By contrast, the self-rating in line bisection tasks was not related to actual task performance [85]. Furthermore, as the AI test is designed to be performed after task execution, it might be that the patients were able to correct their erroneous self-rating to some extent at least in the complex visuomotor tasks. Last but not least, the repetition of the NET test over a relatively short time span might have influenced the AI results, too, as the patients knew that a self-rating would follow after certain tasks.

The ETNT results are to be considered preliminary and should be interpreted with caution. Calibration difficulties with the Eye Tribe Tracker system (eg, most patients only reached 3 out of 5 points in calibration quality scores) may have influenced the reliability of the setup and the accuracy of the results. The calibration consisted of eye-tracking a circle that moved around the whole display. The difficulties experienced by patients in following the rapidly moving circle and the requirement to look at calibration
points at the very left of the computer screen were the main reasons for the rather poor calibration results. Such difficulties may produce visible effects (in Figure 5, for example, P4 evidences a dense cloud of gaze points on the bottom right corner during the postintervention assessment, where P4 unsuccessfully tried to tag the 3 targets he could point at with his hand). Baheux et al [86] also reported calibration problems with their 3-D haptic VR system coupled with an eye-tracking device. They assumed that the VSN patient spectacle wear or eye color might have been reasons for calibration difficulties. However, these calibration difficulties notwithstanding, our ETNT results showed trends toward slight improvements in both total located and missed targets on the left side of the screen. The heat maps display the increased ability to detect targets in the upper left portion of the computer screen postintervention but remained substantially unmodified otherwise. However, the preintervention performance was already fairly good in our sample. Furthermore, the heat maps show that the ETNT can identify neglected areas. This is in line with the Rehabilitation Gaming System by Maier et al [31] using the Kinect motion capture system being equally able to measure symptoms of neglect. In contrast to our test, the stroke patients in the chronic stage of recovery explored the neglected side with the paretic arm.

The individual search strategy (Multimedia Appendices 3 and 4) of most patients was comparable with those described by Müri et al [65] and Rabuffetti et al [34], namely to start in the (extreme) right sector and scan leftwards vertically. This contrasts with the search strategy demonstrated by the control subjects in these studies, which started in the upper left sector and scanned top-down horizontally (like reading). Only P3 showed a nonneglect-specific search strategy (Figure 5). The 4 missed targets in the left lower corner postintervention, which P3 was able to detect preintervention, were due to calibration difficulties, as P3 was able to point at those 4 targets with his hand. P1 and P6 were the only patients able to shift their search starting points from the right to the left side from pre- to postintervention, indicating improvements in exploring the neglected side [34]. Not surprisingly, the increased total test duration pre- to postintervention in our sample went along with slightly more detected targets. When the participant increases the number of detected targets, it can be the case that previously neglected targets are now detected albeit with a fairly large latency. We interpret this as positive since more visual space is actively explored but latency and concomitantly the total test duration may therefore also increase. A future study
including a larger sample and a control condition is, however, warranted to substantiate or refute these findings.

We performed an a priori power analysis to determine the minimum sample size for such a future trial [87]. Specifically, we assessed the requirements for a randomized controlled study with an experimental group (receiving exergame-based therapy and usual stroke rehabilitation) and a control group (receiving usual stroke rehabilitation only). Assuming an effect size of $r=0.9$ (based on our observed value for NET total scores, Multimedia Appendix 5), acceptable type I and II error probabilities (0.05 and 0.20, respectively) may be obtained with a minimum sample of 34 subjects per group for a 2-group pre-post-test design. To account for attrition, initial sample size should be increased to 45 subjects per group [88]. Given the fact that only 18 potential participants were available within 12 months of recruitment, we recommend collaborating with more than 2 clinics for such a trial.

**Limitations and Future Work**

The length of the training phase was rather short (ie, 3 weeks). We deliberately did not choose a longer training period, as we primarily wanted to test the exergames' feasibility and not their effect on VSN symptoms. On the other hand, the rather short training phase allowed us to keep the drop-out risk relatively low (eg, due to discharge home during the training phase). In a next step, it would be important to test the exergame system’s feasibility in patient homes to evaluate adherence, safety, and attrition to using the system in this setting, as the provision of novel home-based rehabilitation options was the main goal of REWIRE. In this setting, a longer training phase could be tested. Furthermore, a progression as measured by the game scores should be implemented together with an immediate graphical feedback after each training session to enhance motivation for playing the exergames. For this implementation, ideas from the rehabilitation method of shaping [54], where frequent feedback and encouragement during training are central, could be adopted. In order to maximize confidence that changes in outcomes can be attributed causally to the exergames intervention, a control group in a pilot randomized controlled trial design would be needed. The neglect exergames should further be designed to switch levels of difficulty (ie, progressing from the right to the left side of the screen or vice-versa). Designing this option would allow recruiting stroke patients with a left-sided brain lesion and VSN symptoms, too. By excluding them in our project we were aware that we
would probably miss some patients having ipsilesional neglect [89], which would have made a participation in the exergames intervention feasible. However, as left-sided neglect is quite rare compared to right-sided neglect [5], the risk of missing such an ipsilesional neglect patient was relatively low.

The ETNT could further be developed regarding the following:

- Calibration procedure of the Eye Tribe Tracker by reducing the speed of the circle to be followed, for example
- Software indexes, which were initially designed for the touchscreen (hand-eye coordination) test. Indexes important for eye-tracking would be, for example, the cumulative fixation duration, spatial distribution of fixations in the horizontal and vertical plane, or the number and amplitude of exploratory saccades as explored by Müri et al [65]
- Collection of the search strategy patterns of age-matched controls

Additionally, future work could correlate ETNT measures to scores in standardized clinical scales, such as the NET scores, in order to validate the derived ETNT measures of recovery after VSN.

**Conclusion**

This study showed that patients adhered well to the REWIRE neglect exergames intervention with no drop-outs, no adverse events, and an adherence rate of 14 out of 15 sessions (93%). We therefore judged this intervention to be safe and feasible for VSN patients early poststroke and appropriate for further testing. Cognitive and spatial exploration skills, as evaluated using ETNT, NET (spatial exploration), and ZüMAX (cognition) assessments, improved in most patients from pre- to postintervention. The results of the amount of exergames use is promising for future applications and warrants further investigations, for example, in the home setting of patients as a motivating training tool to complement usual care and support augmenting training frequency and intensity in RBL stroke patients with VSN.

**Abbreviations**

AI: anosognosia index  
AST: Apraxia Screen of TULIA  
CBS: Catherine Bergego Scale  
ETNT: Eye Tracker Neglect Test  
IQR: interquartile range
**NET:** Neglect Test
**RBL:** right brain lesion
**REWIRE:** Rehabilitative Wayout in Responsive Home Environments
**TULIA:** test of upper limb apraxia
**USN:** unilateral spatial neglect
**VR:** virtual reality
**VSN:** visuospatial neglect
**ZüMAX:** Zürich Maxi Mental Status Inventory

**Acknowledgments**

BCT-A designed the study; provided support in designing the exergames; conducted the acquisition, analysis, and interpretation of the data; and wrote the manuscript. EF and MR developed the software and exergames and contributed to the analysis and interpretation of data and writing of the manuscript. LW helped recruit the patients and edited the manuscript. EDdB, NAB, and RHK initiated the study and contributed to design, writing, and editing of the manuscript. All authors read and approved the final manuscript.

This work was partially supported by the REWIRE project (www.rewire-project.eu) and funded by the European Commission under the FP7 framework with contract 287713 and the Occupational Therapy Association of Switzerland. Author MR was supported by funding (Ricerca Corrente) from the Italian Ministry of Health.

Many thanks go to the collaborating clinics who made the patients recruitment possible, namely to Klaus Meyer, MD, Klinik Bethesda, Tschugg, Canton of Berne, Switzerland, and his team of neuropsychologists and to Andreas Sturzenegger, MD, Klinik Wald, Wald, Canton of Zurich, Switzerland, and his team of occupational therapists. Laura Wiederkehr, former master student of the ETH Zurich, is very much appreciated for her scientific and administrative support and for performing the measurements. Further thanks goes to Gabriel Baud Bovy (Istituto Italiano di Tecnologia, Genova) for his work on the Novint Falcon haptic device. Finally, many thanks to Martin J Watson, PhD, for proofreading the manuscript for English and structure.

**Conflicts of Interest**

None declared.
Multimedia Appendix 1
Overview of individual and group results in the training protocol.

<table>
<thead>
<tr>
<th>training protocol parameters</th>
<th>attended sessions</th>
<th>total training time (min)</th>
<th>mean ± SD duration/session (min)</th>
<th>health condition after training n (%)</th>
<th>progression over the 3 week intervention</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>patient</td>
<td>n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>12 (80)</td>
<td>276</td>
<td>23 ± 3.0</td>
<td>2 (17)</td>
<td>10 (83)</td>
<td>2x</td>
</tr>
<tr>
<td>P2</td>
<td>15 (100)</td>
<td>392</td>
<td>26 ± 4.0</td>
<td>13 (87)</td>
<td>2 (13)</td>
<td>3x L5 / 2A</td>
</tr>
<tr>
<td>P3</td>
<td>12 (80)</td>
<td>278</td>
<td>21 ± 2.8</td>
<td>10 (83)</td>
<td>2 (17)</td>
<td>4x L4 / 3A</td>
</tr>
<tr>
<td>P4</td>
<td>15 (100)</td>
<td>445</td>
<td>30 ± 1.3</td>
<td>14 (93)</td>
<td>1 (7)</td>
<td>3x L3 / 1A</td>
</tr>
<tr>
<td>P5</td>
<td>15 (100)</td>
<td>450</td>
<td>30 ± 0</td>
<td>10 (67)</td>
<td>5 (33)</td>
<td>3x L5 / 1A</td>
</tr>
<tr>
<td>P6</td>
<td>14 (93)</td>
<td>420</td>
<td>30 ± 0</td>
<td>14 (100)</td>
<td>0 (0)</td>
<td>3x L3 / 2A</td>
</tr>
<tr>
<td>P7</td>
<td>10 (67)</td>
<td>335</td>
<td>34 ± 10.3</td>
<td>0 (0)</td>
<td>10 (100)</td>
<td></td>
</tr>
<tr>
<td>IQR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>12 (80)</td>
<td>278</td>
<td>24.5</td>
<td>2 (17)</td>
<td>1 (7)</td>
<td>3x L1 / A1</td>
</tr>
<tr>
<td>50</td>
<td>14 (93)</td>
<td>392</td>
<td>30</td>
<td>13 (87)</td>
<td>2 (13)</td>
<td>3x L2 / A2</td>
</tr>
<tr>
<td>75</td>
<td>15 (100)</td>
<td>445</td>
<td>30</td>
<td>14 (93)</td>
<td>10 (83)</td>
<td>3x L3 / B1</td>
</tr>
<tr>
<td>mean ± SD</td>
<td>13.3 ± 2.0 (88.6 ± 13.2)</td>
<td>370.9 ± 74.8</td>
<td>27.7 ± 4.6</td>
<td>9.4 ± 5.9 (64.3 ± 39.8)</td>
<td>4.3 ± 4.2 (35.7 ± 39.8)</td>
<td></td>
</tr>
</tbody>
</table>

a maximum training sessions: n = 15

b calculated as the total training time divided by the total number of attended sessions

c to be read as:

2-4x = number of exergames.

L1-5: difficulty levels (the higher the more exploration needed to the neglected left side).
1A-5A or 1B-5B: Gentile’s Taxonomy progression (boxes A: no force feedback from Falcon; boxes B: force feedback from Falcon;
1A/1B: stationary conditions of virtual scenario, no intertrial variability during the game;
2A/2B: stationary conditions, intertrial variability;
3A/3B: in-motion conditions (i.e. moving virtual objects during the game, e.g. balloons), no intertrial variability;
4A/4B: in-motion conditions and intertrial variability)

IQR = Inter Quartile Range; min = minutes; n = number; NA = not applicable; % = percentage
Multimedia Appendix 2

Overview of individual Eye Tracker Neglect Test scores and group changes pre- to postintervention.

<table>
<thead>
<tr>
<th>ETNT indexes</th>
<th>total located targets(^a) (n)</th>
<th>missed targets left(^b) (n)</th>
<th>missed targets right(^b) (n)</th>
<th>neglect score(^c) (%)</th>
<th>median latency(^d) (sec) IQR 25/50/75</th>
<th>median proximity(^e) (cm) IQR 25/50/75</th>
<th>total test duration(^f) (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>patient</td>
<td>pre</td>
<td>post</td>
<td>pre</td>
<td>post</td>
<td>pre</td>
<td>post</td>
<td>pre</td>
</tr>
<tr>
<td>P1</td>
<td>20</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.3</td>
</tr>
<tr>
<td>P2</td>
<td>10</td>
<td>16</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>-10</td>
</tr>
<tr>
<td>P3</td>
<td>20</td>
<td>15</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>5.2</td>
</tr>
<tr>
<td>P4</td>
<td>5</td>
<td>15</td>
<td>9</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>34.0</td>
</tr>
<tr>
<td>P5</td>
<td>20</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td>P6</td>
<td>15</td>
<td>20</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>P7</td>
<td>4</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>IQR</td>
<td>5</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>WSR</td>
<td>P</td>
<td>.21</td>
<td>.27</td>
<td>.22</td>
<td>.83</td>
<td>.80</td>
<td>.31</td>
</tr>
<tr>
<td>Z</td>
<td>-0.26</td>
<td>-1.10</td>
<td>-1.23</td>
<td>-0.21</td>
<td>-0.25</td>
<td>-1.03</td>
<td>-1.35</td>
</tr>
<tr>
<td>r</td>
<td>-.10</td>
<td>-.42</td>
<td>-.47</td>
<td>-.08</td>
<td>-.09</td>
<td>-.39</td>
<td>-.51</td>
</tr>
</tbody>
</table>

\(a\) maximum detectable targets: 20 = 100% (a target is counted if the participant looked at it for longer than 0.4 sec)

\(b\) maximum detectable targets: 10 on the left and 10 on the right side of the screen

\(c\) defined as the difference between the number of targets cancelled on the left and on the right side of the screen, expressed as a percentage of the total number of targets. 0 = no laterality to search performance; negative numbers = focus more on the left side; positive numbers = focus more on the right side

\(d\) summary value of all time spans between the current and the previous item hit during the test

\(e\) number of not yet found targets that were closer to the previous one than the actual newly seen one

\(f\) from the beginning to the time point when the participant stated to have found all targets

cm = centimeters; ETNT = Eye Tracker Neglect Test; IQR = Inter Quartile Range; n = number; P = level of significance \((P \leq .05)\); r = effect size; sec = seconds; WSR = Wilcoxon signed ranks test; Z = Z-score (approximation of the observed difference in terms of the standard normal distribution)
Multimedia Appendix 3

Pre- and postintervention results of individual Eye Tracker Neglect Test search paths.

<table>
<thead>
<tr>
<th>ETNT</th>
<th>search path</th>
<th>patient</th>
<th>preintervention</th>
<th>postintervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The bright green square signifies the first located target of the ETNT, thus the first square the participant looked at for ≥ 0.4 sec in the target-area (defined as a 7% area of the computer screen).

A circle around a tagged item represents the latency associated with its hit. Latency is the time passed from the hit on the previous item. Therefore, the bigger the circle, the more time passed. A single item may have several circles because it can be hit more than once (perseverations for the same target). Except for the case of quite similar latencies on repeated hits – apparently only one circle shows up because they superimpose - the number of circles marks the number of repeated hits.

The bright blue line represents the middle line of the two computer screen sides, dividing the 20 total targets into 2x 10 items.

The dark blue dots are the visualization of the “points of gaze” (estimates where someone is looking, as measured by the sensors of the Eye Tribe Tracking infrared camera). A fixation point is identified as the average of a “bunch of points of gaze” which stay close to each other for at least 100 ms.
Multimedia Appendix 4

Pre- and postintervention results of individual Eye Tracker Neglect Test fixation points.

<table>
<thead>
<tr>
<th>ETNT</th>
<th>fixation points</th>
</tr>
</thead>
<tbody>
<tr>
<td>patient</td>
<td>preintervention</td>
</tr>
<tr>
<td>P1</td>
<td><img src="image1.png" alt="Image" /></td>
</tr>
<tr>
<td>P2</td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>P3</td>
<td><img src="image5.png" alt="Image" /></td>
</tr>
<tr>
<td>P4</td>
<td><img src="image7.png" alt="Image" /></td>
</tr>
<tr>
<td>P5</td>
<td><img src="image9.png" alt="Image" /></td>
</tr>
<tr>
<td>P6</td>
<td><img src="image11.png" alt="Image" /></td>
</tr>
<tr>
<td>P7</td>
<td><img src="image13.png" alt="Image" /></td>
</tr>
</tbody>
</table>

The bright green square signifies the first located target of the ETNT, thus the first square the participant looked at for ≥ 0.4 sec in the target-area (defined as a 7% area of the computer screen).

A circle around a tagged item represents the latency associated with its hit. Latency is the time passed from the hit on the previous item. Therefore, the bigger the circle, the more time passed. A single item may have several circles because it can be hit more than once (perseverations for the same target). Except for the case of quite similar latencies on repeated hits – apparently only one circle shows up because they superimpose - the number of circles marks the number of repeated hits.

The bright blue line represents the middle line of the two computer screen sides, dividing the 20 total targets into 2x 10 items.

The dark blue dots are the visualization of the “points of gaze” (estimates where someone is looking, as measured by the sensors of the Eye Tribe Tracking infrared camera). A fixation point is identified as the average of a “bunch of points of gaze” which stay close to each other for at least 100 ms.
Multimedia Appendix 5

Overview of Zürich Maxi Mental Status Inventory, Neglect Test, and anosognosia index scores and group changes preintervention, postintervention, and follow-up.

<table>
<thead>
<tr>
<th>Assessment</th>
<th>IQR (25/50/75)</th>
<th>WSR</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>intervention</td>
<td>pre</td>
<td>post</td>
<td>FU</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZüMAX</td>
<td>total scores(a)</td>
<td>16</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>visual perception</td>
<td>6</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>subtest(b)</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>visual construction</td>
<td>7</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>subtest(b)</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>NET</td>
<td>total scores(c)</td>
<td>127.5</td>
<td>144</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>paper-and-pencil subtests(d)</td>
<td>149.5</td>
<td>154.5</td>
<td>156.5</td>
</tr>
<tr>
<td></td>
<td>behavioral subtests(e)</td>
<td>152.5</td>
<td>156</td>
<td>161.5</td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>58</td>
<td>53.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>61.5</td>
<td>65.5</td>
<td>65.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>66</td>
<td>68.5</td>
<td>67.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>71.5</td>
<td>86</td>
<td>82.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>86.5</td>
<td>88</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>88.5</td>
<td>89.5</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>Anosognosia Index(f)</td>
<td></td>
<td>-0.04</td>
<td>-0.04</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.08</td>
<td>0.04</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.16</td>
<td>0.12</td>
<td>0.18</td>
</tr>
</tbody>
</table>

\(*\) = statistically significant; † large effect size (≥ 0.5)

AI = Anosognosia Index; FU = follow-up; IQR = Inter Quartile Range; NET = Neglect Test; \( P \) = level of significance (\( P \leq 0.05 \)); \( r \) = effect size; VSN = visuo-spatial neglect; \( X^2 \) = Chi-square test; \( Z \) = Z-score (approximation of the observed difference in terms of the standard normal distribution); ZüMAX = Zürich maxi mental status inventory

\(a\) maximum score = 30 points
\(b\) maximum raw score = 10 points
\(c\) maximum score = 170 points
\(d\) maximum score = 70
\(e\) maximum score = 100
\(f\) index < 0: patient suffers from anosognosia; index ≥ 0: no anosognosia
Multimedia Appendix 6
Graphical display of individual Zürich Maxi Mental Status Inventory, Neglect Test, and anosognosia index scores.

ZüMAX total scores

NET total scores

ZüMAX visual perception scores

ZüMAX visual construction scores

NET paper-and-pencil subscores

NET behavioral subscores

Anosognosia index scores

Legend

a maximum score = 30 points
b maximum score = 170 points (0-72: highly distinctive VSN; 73-135: distinctive VSN; 136-166: slightly distinctive or suspected VSN; >166 points: no VSN)
c maximum raw score = 10 points; task: to recognize and name degraded figures, unfamiliar scenes and a face
d maximum raw score = 10 points; task: to copy a geometric figure
e maximum score = 70 points
f maximum score = 100 points
g Anosognosia Index: < 0 = patient suffers from anosognosia; ≥ 0 = no anosognosia
### Multimedia Appendix 7

Individual changes in the Zürich Maxi Mental Status Inventory and Neglect Test assessments.

<table>
<thead>
<tr>
<th>Patient</th>
<th>test</th>
<th>ZüMAX</th>
<th>NET</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pre-post-FU</td>
<td>pre-post</td>
<td>post-FU</td>
</tr>
<tr>
<td>P1</td>
<td>.61</td>
<td>1.00</td>
<td>.41</td>
</tr>
<tr>
<td>P2</td>
<td>.10</td>
<td>4.67</td>
<td>.18</td>
</tr>
<tr>
<td>P3</td>
<td>.93</td>
<td>0.15</td>
<td>.66</td>
</tr>
<tr>
<td>P4</td>
<td>.11</td>
<td>4.43</td>
<td>.36</td>
</tr>
<tr>
<td>P5</td>
<td>.29</td>
<td>2.46</td>
<td>.28</td>
</tr>
<tr>
<td>P6</td>
<td>.10</td>
<td>4.67</td>
<td>.32</td>
</tr>
<tr>
<td>P7</td>
<td>.06</td>
<td>5.69</td>
<td>.98</td>
</tr>
</tbody>
</table>

* = statistically significant; † large effect size (≥ 0.5)

P = level of significance (P ≤ 0.05); r = effect size; $X^2$ = Chi-Square; Z = Z-score (approximation of the observed difference in terms of the standard normal distribution)
References


15. Hill D, Coats RO, Halstead A, Burke MR. A systematic research review assessing the effectiveness of pursuit interventions in spatial neglect following stroke. Transl Stroke Res 2015 Dec;6(6):410-420. [CrossRef] [Medline]


23. Rusconi ML, Carelli L. Long-term efficacy of prism adaptation on spatial neglect: preliminary results on different spatial components. ScientificWorldJournal 2012;2012 [FREE Full text] [CrossRef] [Medline]


43. van Kessel ME, Geurts ACH, Brouwer WH, Fasotti L. Visual scanning training for neglect after stroke with and without a computerized lane tracking dual task. Front Hum Neurosci 2013;7:358 [FREE Full text] [CrossRef] [Medline]


54. Taub E, Mark VW, Uswatte G. Implications of CI therapy for visual deficit training. Front Integr Neurosci 2014;8:78 [FREE Full text] [CrossRef] [Medline]

55. Borghese NA, Pirovano M, Lanzi PL, Wüest S, de Bruin ED. Computational intelligence and game design for effective at-home stroke rehabilitation. Games Health J 2013 Apr;2(2):81-88 [FREE Full text] [CrossRef] [Medline]


63. Ammann BC, Knols RH, Baschung P, de Bie RA, de Bruin ED. Application of principles of exercise training in sub-acute and chronic stroke survivors: a systematic review. BMC Neurol 2014 Aug 22;14:167 [FREE Full text] [CrossRef] [Medline]

64. Gavgani AM, Nesbitt KV, Blackmore KL, Nalivaiko E. Profiling subjective symptoms and autonomic changes associated with cybersickness. Auton Neurosci 2017 Mar;203:41-50. [CrossRef] [Medline]


70. Tobler-Ammann BC, de Bruin ED, Brugger P, de Bie RA, Knols RH. The Zürich Maxi Mental Status Inventory (ZüMAX): test-retest reliability and discriminant validity in stroke survivors. Cogn Behav Neurol 2016 Jun;29(2):78-90. [CrossRef] [Medline]


78. Prahm C, Vujaklija I, Kayali F, Purgathofer P, Aszmann OC. Game-based rehabilitation for myoelectric prosthesis control. JMIR Serious Games 2017 Feb 09;5(1):e3 [FREE Full text] [CrossRef] [Medline]

79. Wiloth S, Lemke N, Werner C, Hauer K. Validation of a computerized, game-based assessment strategy to measure training effects on motor-cognitive functions in people with dementia. JMIR Serious Games 2016 Jul 18;4(2):e12 [FREE Full text] [CrossRef] [Medline]


83. Terwee CB, Bot SDM, de Boer MR, van der Winot DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007 Jan;60(1):34-42. [CrossRef] [Medline]


Chapter 6

User perspectives on exergames designed to explore the hemineglected space for stroke patients with visuospatial neglect: Usability study

Follow-up exergames intervention results of the figure copying task (visuoconstruction) of the ZüMAX of P6, female, 78 years old, 11 weeks onset right-hemispheric stroke and visuo-spatial neglect

This chapter has been published as: Tobler-Ammann BC, Surer E, Knols RH, Borghese NA, de Bruin ED. User Perspectives on Exergames Designed to Explore the Hemineglected Space for Stroke Patients With Visuospatial Neglect: Usability Study. JMIR Serious Games. 2017;5(3):e18.
Abstract

Background: Visuospatial neglect due to stroke is characterized by the inability to perceive stimuli emerging in the area opposite to the side of brain damage. Besides adopting conventional rehabilitation methods to treat neglect symptoms, the use of virtual reality (VR) is becoming increasingly popular. We designed a series of 9 exergames aimed to improve exploration of the neglected side of space. When new VR interventions are designed, it is important to assess the usability aspects of such management strategies within the target population. To date, most studies used questionnaires to assess user satisfaction with the intervention or product being tested. However, only a combination of both quantitative and qualitative data allows a full picture of user perspective.

Objective: The purpose of this study was to quantitatively and qualitatively assess patient and therapist perspectives of a VR intervention based on the series of 9 exergames designed to explore hemineglected space. Specifically, we wanted to evaluate (1) perceived-user friendliness of the exergames, (2) attitude towards using the exergames, and (3) intention to use the exergames in the future.

Methods: A total of 19 participants (7 patients, 12 therapists) evaluated the exergames they had used 5 times a week during 3 weeks. The Technology Acceptance Model (TAM) questionnaire was filled out after the intervention. Based on those responses, we conducted focus group interviews (with therapists) and individual interviews (with patients). To analyze the TAM questionnaires, we used descriptive statistics. We adopted content and comparative analysis to analyze the interviews and drew illustration maps to analyze the focus group interviews.

Results: The therapists took a more critical stance with a mean TAM questionnaire total score of 48.6 (SD 4.5) compared to the patients who had a mean total score of 56.1 (SD 12.3). The perceived user-friendliness score was 5.6 (SD 1.4) for patients and 4.9 (SD 1.4) for therapists. The attitude towards using the exergames was rated 4.8 (SD 1.9) by patients and 3.6 (SD 1.4) by therapists, respectively. The intention to use the exergames in the future was rated 3.9 (SD 2.1) by patients and 3.7 (SD 1.8) by therapists. We gained information on how to improve the exergames in the interviews.

Conclusions: Patients and therapists perceived the exergames as user-friendly; however, using the games further with the actual test version was not perceived as conceivable. The therapists were generally more critical towards future use than the
patients. Therefore, involving both users to achieve acceptable and user-friendly versions of game-based rehabilitation for the future is deemed crucial and warranted.

**Keywords:** usability; user perspective; mixed-methods; exergames; visuo-spatial neglect; stroke

**Introduction**

Stroke-related visuospatial neglect (VSN) due to a right-sided brain lesion (RBL) is characterized by the inability to perceive stimuli emerging in the area opposite to the side of brain damage [1,2]. VSN patients usually have lower scores on disability tests and require longer rehabilitation periods compared to stroke patients without neglect [3,4]. Thus, VSN influences most activities of daily living such as eating, reading, and getting dressed [2,5].

Besides adopting conventional rehabilitation methods to treat stroke-related VSN symptoms, the use of virtual reality (VR) in their assessment and treatment is becoming increasingly popular [6-8]. VR is defined as “an advanced form of human-computer interface that allows the user to ‘interact’ with and become ‘immersed’ in a computer-generated environment in a naturalistic fashion” [9]. Reasons for this increasing popularity might be found in the many advantages attributed to VR, for example, the ability to provide a safe but engaging environment [10], immediate feedback on performance, and repetitive task training with quantifiable continuous progression of training [9]. For example, VR training in isolation or in combination with conventional therapy approaches proved to be superior for the improvement of lower extremity function in stroke patients [11]. However, despite this, evidence for VR therapies being superior to conventional intervention methods for treating VSN is so far somewhat limited [6-8]. Evidence shows that VR has the capacity both to enhance current methods for the assessment and rehabilitation of VSN and to provide new ones. Tsirlin et al [8] presented three major challenges for successful implementation of VR systems in VSN therapy: (1) ergonomic aspects in the sense that mobile, lightweight VR systems are required for rehabilitation, (2) the complexity of VR systems insofar as treating clinic staff do not necessarily have programming skills, and (3) the prohibitive costs of VR devices (eg, for immersive VR systems with head-mounted displays or cyber gloves) [7,8,12]. For these reasons, VR rehabilitation platforms have been
mainly restricted to laboratories and to prototypical systems [8] and have not been widely implemented in patients’ homes.

A European research group, Rehabilitative Wayout In Responsive Home Environments (REWIRE), developed a game-based VR rehabilitation intervention trying to account for those challenges [13]: (1) the exergame station was designed as a computer workplace, allowing the patient to practice the exergames in a seated position, (2) the complexity of the user interface was reduced to a minimum by designing a game menu with large and clear icons to select a game, difficulty level, and playing time (Figure 1), and (3) the costs of the VR systems are relatively low, as the exergames are played on a personal computer, using the Novint Falcon haptic device (Novint Technologies) to control the games [14] (Figure 2). The Novint Falcon enables people to experience a realistic sense of touch by providing force and haptic feedback when reaching for and grasping virtual objects [14]. Furthermore, it can be operated with one hand only, thus permitting VSN patients to play the exergames with their unaffected upper limb.

When new VR interventions are designed, it is important to follow a phased iterative approach, wherein the usability aspects of such a management strategy, within the target population, are first assessed [15]. Usability is defined by the International Organization for Standardization (ISO) as “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” [16]. “Specified users” do not only include patients but also therapists, as their requirements for the use of such games may differ from those of the patients [17]. Therapists may, for example, need an easy startup and configuration procedure, or stress that the games should be supportive not only for the patient during play but also for the therapist in tracking a patient’s performance [17]. It is therefore important to assess both the therapists’ and patients’ opinions, as they will use the exergames at least as often as a patient.

Assessing opinions from users can be done by using questionnaires or by means of interviews. The former has the advantage to assess many opinions of a representative sample but cannot tell us about the meaning behind a response. The latter is usually applied in a smaller sample but provides personal thoughts from an insider’s perspective [18]. To date, most studies used questionnaires to assess users’ satisfaction with the intervention or product being tested [19-22]. Currently, there is a lack of evidence from studies examining the users’ perspective via application of
qualitative methodologies [23]. King et al [24] and Lewis et al [25], for example, used focus group and semistructured interviews to assess patients’ satisfaction with their game intervention. All participants enjoyed playing the computer games. However, all stroke patients were in the chronic stage of recovery and none were diagnosed with VSN symptoms. Another qualitative report explored the perceptions and personal experiences of stroke survivors regarding a leisure-based VR program [26], reporting improved self-efficacy belief in leisure activities after the VR experience in in-depth interviews. However, the reported evidence was based on a single game session only.

Results of a recent systematic review state that in the posttest stage of usability evaluation, performing interviews to evaluate user perceptions of games is recommended, whereas the use of questionnaires is considered useful for evaluating user acceptance and satisfaction [27]. As a consequence, only a combination of both quantitative and qualitative data allows a full picture of user perspectives, as it is done in mixed methods research [18]. Therefore, the purpose of this study was to quantitatively and qualitatively assess the patients’ and therapists’ user perspective when using REWIRE exergames for rehabilitation of VSN symptoms due to a stroke. Specifically, we wanted to evaluate the (1) perceived-user friendliness of the exergames, (2) attitude towards using the exergames, and (3) intention to use the exergames in the future.

**Figure 1.** Game menu of the 9 neglect exergames.
Figure 2. REWIRE exergames training station.

Methods

Study Design

For this usability study, we used a mixed methods design adopting the “sequential explanatory” design strategy [18]. This design strategy is characterized by an initial collection of quantitative data followed by a collection and analysis of qualitative statements. The purpose of this strategy is to use the qualitative results to assist in explaining and interpreting the findings of the quantitative data.

Participants

There were 2 groups of users involved in this study: patients as end users and therapists as experts.

The patient group included 7 adults with an ischemic (n=5 men) or hemorrhagic (n=2 women) RBL due to a first stroke with accompanying VSN symptoms as measured with the Catherine Bergego Scale (CBS) [28]. The CBS includes direct observation of the patient’s functioning in 10 real-life situations. The functioning is rated from 0-30, where 0 indicates no neglect symptoms. These patients simultaneously participated in a feasibility study in which the exergames were evaluated, while taking part in this
usability evaluation [29]. Their mean age was 68.6 (SD 8.9) years. Their stroke incidence took place 46.3 (SD 30.8) days before study entry. All patients were right-handed. Three participants were able to walk, while the others used a wheelchair for locomotion. Their CBS mean score was 9.4 (SD 5.1) points. All but one of these 7 patients identified themselves as having a computer at home prior to participating in the study.

The expert group consisted of therapists responsible for the treatment of the stroke patients during their inpatient stay. The 12 therapists (6 occupational therapists from one rehabilitation clinic and 6 neuropsychologists from another clinic) supervised and trained the patients in the use of the REWIRE exergames during the 3-week intervention phase. Their mean age was 33.3 (SD 5.7) years (range 27-45) with a mean work experience of 6.8 (SD 5.8) years (range 0.5-20). All therapists stated being familiar with the use of computers, rating their computer knowledge as excellent (n=3), good (n=8), and poor (n=1). All 19 participants signed informed written consent before study entry. We obtained ethical approval for the study from the local Ethics Committees (Zurich No. 2014-0543 and Bern No. 389/2014) as well as from Swissmedic (2015-MD-0003). The study is registered with ClinicalTrials.gov.

Setup of the Exergames Training Stations

We installed 2 exergames training stations, one in each of the collaborating clinics (Figure 2). The games were played at a table in a seated position either in a chair or wheelchair depending on the patient's motor skills. We used a 21-inch computer monitor at a distance of 60-65 cm to display the games and the haptic Falcon Novint device to control the games. The haptic feedback enabled the patients to experience a realistic sense of touch, for example, by feeling some resistance simulating the weight of the currently held virtual object in the virtual hand displayed on the screen (force feedback of the Falcon [14]) or a vibration when dropping, for example, a virtual apple in a virtual basket. The Falcon Novint was placed at the side of the computer monitor at a distance allowing the patients ease of reach with their nonaffected hand. A height-adjustable chin rest (Novavision GmbH) was mounted on the table to avoid compensatory head movements while playing the exergames.
REWIRE Visuospatial Neglect Exergames

We designed a series of 9 exergames aimed to improve exploration of the neglected side of space. During the development of the exergames, we regularly tested them in healthy controls prior to implementing them in a clinical setting. Their feedback was constantly integrated in the development process until consensus was reached. In order to maintain principles of training, game progression was individually adjustable through the selection of appropriate different levels of difficulty (more demanding meant more exploration towards the hemineglected side was required) [30]. The exergames content aimed to imitate activities of daily living (ADL), such as cooking a meal, following a recipe, gathering apples, walking a dog, and doing a puzzle. An overview of the 9 games and corresponding short instructions supporting their independent use are shown in Multimedia Appendix 1. A detailed description of the exergames can be found elsewhere [13].

Intervention

Both therapists and patients had the opportunity to test the exergames before entering the study, followed by a training event organized by the research team to learn, for example, how to handle the game menu and Falcon Novint haptic device. During the whole intervention phase, the research staff provided telephone or personal support whenever needed, for example, to handle technical problems with the training station. The exergames intervention lasted 3 weeks and included 15 training sessions each of approximately 30 minutes duration. Patients exercised with the games under supervision of the therapists depending on their required level of support, for example, to start a new game. The neuropsychologists included exergames playing in their computer group, meaning that participating patients played the REWIRE exergames while other group members performed alternative computer tasks. The occupational therapists (OTs) supervised their patients in a one-to-one setting during individual therapy sessions. Additionally, the supervising therapist individually adjusted the intensity of playing the exergames. This was done, for example, by changing the difficulty level or game duration in the game menu or by implementing short breaks between each game if needed. Each patient selected up to four REWIRE VSN exergames from the game menu to be played in a gaming session. The choice was based on personal interest of the patient, which was assumed to enhance motivation while playing. Therefore, during the 3-week intervention time, the patient was also
allowed to change games if they wanted to test another game or felt bored with the previously played one. However, we suggested the patients test all of the 9 games at least once. Rehabilitation continued during the study intervention, our exergames serving as an additional therapy option to the standard program comprising daily occupational, physical, and neuropsychological therapy.

**Outcome Measurements**

Both patients and therapists completed a questionnaire at the end of the intervention. This included 12 questions with a 7-point Likert scale (1 point=strongly disagree; 7 points=strongly agree), evaluating (1) perceived user-friendliness of the exergames, (2) attitude towards using the exergames, and (3) intention to use the exergames in the future. The questionnaire design was based on an abridged version of the Technology Acceptance Model (TAM). TAM is an intention-based model developed specifically for explaining user acceptance of computer technology [31] that we considered useful for evaluating user acceptance and user satisfaction [27]. Patients received physical assistance from clinic staff to complete the questionnaire when incapable of writing or reading due to neglect. The therapists filled in their questionnaires independently. We analyzed the completed questionnaire responses and thereafter used them as a basis to prepare the individual interview [32] with the patients and the focus group interviews [33] with the therapists.

BC-T-A performed the audio-recorded individual interviews during the follow-up assessment planned for the feasibility study 4 weeks post-intervention. They focused on the patients’ everyday life experiences with right hemispheric stroke and VSN symptoms during active rehabilitation and served as an opportunity to deepen, clarify, or confirm answers that were given in the TAM questionnaire. BC-T-A, who is an occupational therapist, took an active role during the interviews, aiming to build a relationship with the participants based on confidence and co-creation. Thanks to the many opportunities to meet the patients in the past (eg, while introducing the exergames to the patients or during data generation for the simultaneously running feasibility study [29]), a good basis to achieve this aim was already established.

We conducted the two focus group interviews in the collaborating rehabilitation clinics after the last patient had finished the REWIRE exergames intervention. We audio-recorded and filmed them. In comparison to individual interviews, focus groups represented group opinions influenced by social interactions and team dynamics, as
therapists already knew each other well [34]. Therefore, they were an important complement to individually given answers via TAM questionnaires and served as a means for therapists to gain, share, or dispute experiences made with those exergames from their perspectives. During each interview, the moderator (BC-T-A) summarized the given answers on a flip chart, allowing therapists to add, complement, or change statements if needed. Additionally, the role of the moderator was (1) to balance the therapists’ statements in terms of allowing/encouraging everybody to speak, while intervening when someone would have claimed too much time to speak and (2) to take care that all therapists’ TAM answers were discussed that needed clarification while keeping the time set for the focus group interview.

We expected the acceptance of and satisfaction with the neglect exergame intervention to be good, which we defined as a total mean score of more than 4 points on the Likert scale, where 1-3 points meant no agreement, 4 points neutral, and 5-7 points agreement. The maximum achievable score was 84 points, indicating perfect agreement. For the individual and focus group interviews, we hypothesized that the users (therapists and patients) would have experienced the VR-based neglect training as supportive in treating VSN.

Data Analysis
To analyze the TAM questionnaires, we used descriptive statistics in SPSS software version 23. We calculated means, standard deviations, medians, and interquartile range values as appropriate.

We transcribed the individual interviews verbatim. Subsequently, we selected text passages from the entire conversation in which the interviewer (BC-T-A) and the patient discussed the use of the exergames and stored them separately. We used content and comparative analysis to analyze those passages [35,36], taking the following analysis steps: (1) reading the interview passages’ transcriptions, (2) highlighting significant statements that provide an understanding of how the patient experienced the use of the exergames, (3) comparing those statements with the TAM questionnaire answers and assigning them to the three subcategories (user-friendliness, attitude, and intention to use in the future), and (4) writing a composite description of the patients’ perspectives of using the exergames, while using quotes to underpin the interpretation.
We analyzed the two focus group interviews by drawing “Focus group Illustration Maps” (FIMs) [33]. The aim was to summarize the complex variety of statements and opinions without losing information or knowledge. Therefore, capturing the whole range of group knowledge is the essence of knowledge mapping, rather than highlighting the single statements of individuals. We took the following analysis steps: (1) listening to the audio recording while watching the video and taking notes, (2) comparing the notes and audio recordings together with the flipchart notes, (3) drawing the FIMs, re-watching the video to check the accuracy of the FIMs, (4) sending the FIMs to the participants for member checking, (5) incorporating feedback from participants into the FIMs if representative for the whole group, and (6) merging FIMs from both clinics into one FIM per subcategory from the TAM questionnaires, representing the opinions from all 12 participating therapists.

**Results**

We summarized the answers from the TAM questionnaires in Table 1 for patients and in Table 2 for therapists. Generally, the therapists took a more critical stance with a mean TAM questionnaire total score of 48.6 (SD 4.5) compared to the patients with a mean total score of 56.1 (SD 12.3). Their statements are presented according to the three subcategories of the TAM questionnaire. These are “perceived user-friendliness,” “attitude towards using the exergames,” and “intention to use the exergames in the future.”
Table 1. Postintervention patients’ TAM questionnaire responses.

<table>
<thead>
<tr>
<th>Statement</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
<th>Mean (SD)</th>
<th>Median (Q1/Q3)</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perceived user-friendliness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The exergames were easy to use.</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>6.1 (0.9)</td>
<td>6 (5/7)</td>
<td>5.6 (1.4)</td>
</tr>
<tr>
<td>The exergames manual was clear and understandable.</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6.1 (0.9)</td>
<td>6 (5/7)</td>
<td></td>
</tr>
<tr>
<td>Learning to use the exergames independently would be easy for me.</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>5</td>
<td>4.6 (1.7)</td>
<td>5 (3/6)</td>
<td></td>
</tr>
<tr>
<td><strong>Attitude towards using the exergames</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I generally have a positive attitude towards using the exergames.</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>6</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>4.9 (2.4)</td>
<td>6 (2/7)</td>
<td>4.8 (1.9)</td>
</tr>
<tr>
<td>I enjoyed exercising with the exergames.</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>4.7 (1.8)</td>
<td>5 (3/6)</td>
<td></td>
</tr>
<tr>
<td><strong>Exercising with the exergames...</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>was motivating.</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>5.0 (1.8)</td>
<td>5 (4/7)</td>
<td></td>
</tr>
<tr>
<td>was exhausting.</td>
<td>3</td>
<td>7</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>3.9 (2.3)</td>
<td>3 (2/7)</td>
<td></td>
</tr>
<tr>
<td>was a stupid idea.</td>
<td>5</td>
<td>7</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>5.4 (1.3)</td>
<td>5 (4/7)</td>
<td></td>
</tr>
<tr>
<td><strong>Intention to use the exergames in the future: If I had access to the...</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I would use them in the future.</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>3.0 (1.9)</td>
<td>4 (1/5)</td>
<td>3.9 (2.1)</td>
</tr>
<tr>
<td>I would use them regularly.</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2.7 (1.7)</td>
<td>3 (1/4)</td>
<td></td>
</tr>
<tr>
<td>I’m convinced that my family/friends would support me using the exergames.</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>6</td>
<td>4.7 (1.8)</td>
<td>5 (3/6)</td>
<td></td>
</tr>
<tr>
<td>I would recommend the exergames to other patients.</td>
<td>6</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td>2</td>
<td>6</td>
<td>5.0 (2.5)</td>
<td>6 (2/7)</td>
<td></td>
</tr>
<tr>
<td><strong>Total score</strong></td>
<td>54</td>
<td>70</td>
<td>36</td>
<td>64</td>
<td>69</td>
<td>50</td>
<td>50</td>
<td>56.1 (12.3)</td>
<td>54 (50/69)</td>
<td></td>
</tr>
<tr>
<td><strong>Mean (SD)</strong></td>
<td>4.5</td>
<td>5.8</td>
<td>3.5</td>
<td>3.3</td>
<td>5.8</td>
<td>5.3</td>
<td>4.2</td>
<td>2.0 (2.0)</td>
<td>2.0 (2.0)</td>
<td></td>
</tr>
<tr>
<td><strong>Q1</strong></td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>4.8</td>
<td>2.3</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Q3</strong></td>
<td>5</td>
<td>7</td>
<td>5.5</td>
<td>6</td>
<td>7</td>
<td>5.8</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(a\) Patient / \(b\) Quartile 1=Q_{0.25}; Quartile 3=Q_{0.75} / \(c\) Positive statements: 1=strongly disagree / 7=strongly agree; Negative statements: 1=strongly agree / 7=strongly disagree
Table 2. Postintervention therapists’ TAM questionnaire responses*.

<table>
<thead>
<tr>
<th>Statement</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>Mean (SD)</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perceived user-friendliness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a)</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>5.6 (0.9)</td>
<td>5</td>
<td>5.5</td>
<td>6</td>
<td>4.9 (1.4)</td>
</tr>
<tr>
<td>b)</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>5.2 (1.3)</td>
<td>4</td>
<td>5</td>
<td>6.8</td>
<td></td>
</tr>
<tr>
<td>c)</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4.3 (1.7)</td>
<td>2.5</td>
<td>4</td>
<td>5.8</td>
<td></td>
</tr>
<tr>
<td>d)</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4.7 (1.4)</td>
<td>3.3</td>
<td>4.5</td>
<td>5.8</td>
<td>3.6 (1.4)</td>
</tr>
<tr>
<td><strong>Attitude towards using the exergames</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e)</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>3.8 (1.0)</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>f)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3.3 (0.8)</td>
<td>3</td>
<td>3.5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>g)</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>4.7 (1.7)</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>h)</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2.1 (1.1)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>i)</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>3.8 (0.9)</td>
<td>3</td>
<td>4</td>
<td>4.8</td>
<td></td>
</tr>
<tr>
<td><strong>Intention to use the exergames in the future</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j)</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>2.4 (0.8)</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>k)</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6.0 (0.9)</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>3.7 (1.8)</td>
</tr>
<tr>
<td>l)</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>2.8 (0.9)</td>
<td>2</td>
<td>2.5</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td><strong>Total scores</strong></td>
<td>53</td>
<td>53</td>
<td>53</td>
<td>50</td>
<td>50</td>
<td>54</td>
<td>43</td>
<td>50</td>
<td>42</td>
<td>45</td>
<td>42</td>
<td>48</td>
<td>48.6 (4.5)</td>
<td>42.3</td>
<td>50</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>4.4</td>
<td>4.4</td>
<td>4.4</td>
<td>4.2</td>
<td>4.1</td>
<td>4.5</td>
<td>3.9</td>
<td>4.2</td>
<td>3.5</td>
<td>3.8</td>
<td>3.5</td>
<td>4.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q1</td>
<td>2.5</td>
<td>3.25</td>
<td>2</td>
<td>2.25</td>
<td>2</td>
<td>4</td>
<td>2.25</td>
<td>3.25</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>5</td>
<td>4.5</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3.5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td>5.75</td>
<td>5</td>
<td>7</td>
<td>6.5</td>
<td>7</td>
<td>5.75</td>
<td>4.75</td>
<td>5</td>
<td>4</td>
<td>4.75</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Positive statements: 1=strongly disagree / 7=strongly agree; Negative statements: 1=strongly agree / 7=strongly disagree.

b=T-Therapist
c=Quartile 1=Q0.25
d=Quartile 3=Q0.75

Statements:
- Perceived user-friendliness
  - a) The exergames manual was clear and understandable.
  - b) I was easily able to train my patients for using the exergames.
  - c) Learning to use the exergames independently was easy for my patients.
  - d) I experienced learning to use the exergames as easy.
- Attitude towards using the exergames
  - e) I generally have a positive attitude towards using the exergames.
  - f) The exergames were a gain for my patients.
  - g) The exergames were an unnecessary burden for my patients.
  - h) The exergames were a relief of responsibility for me.
  - i) The supervision of my patients was a pleasure for me.
- Intention to use the exergames in the future
  - j) I can imagine using the exergames regularly as a training for my patients.
  - k) I generally believe that my workplace supports the use of VR training possibilities for my patients.
  - l) I would recommend using the exergames to other colleagues.
Perceived User-Friendliness

This subcategory was the most positively judged among patients with a mean 5.6 (SD 1.4) points and therapists with a mean 4.9 (SD 1.4). All patients agreed on the clarity of the manual and the ease of use of the exergames, while learning to use the exergames independently would not have been easy for patients 1 and 6 (P1 and P6; see Table 1). A possible explanation was the poor general computer knowledge that both patients identified. Thus there were associated uncertainties of what to do when sudden difficulties arose while playing, as stated by P6: “Sometimes I had difficulties in discharging the apples, and when I wasn’t able to keep them steady over the basket, then the apples refused to drop into the basket” (P6). P2 had the same experience in another game with the “faulty pieces”:

Yes, they [the exergames] sometimes weren’t technically well set.
You touched it [the fruit], but it didn’t stick [to the virtual hand] [P2]

To solve those difficulties in game control, P6 and P2 indicated that they contacted the supervising therapist for help. Using the Novint Falcon as haptic device to play the games was perceived as good by all patients. Patient P3 described that the device “acted up” now and then (not running smoothly or skidding of the whole device while playing), suggesting use of the mouse instead of the Novint Falcon as input device in order to solve this problem.

The therapists all agreed on the game manual’s good understandability (Figure 3). They rated themselves as being capable of introducing the exergames to their patients (Table 2). However, they also stated that only the fitter patients were able to use the exergame station independently. For the more severely affected ones, guidance of the arm was necessary to play the games. This was experienced as no relief of responsibility for the therapists, which negatively influenced their attitude towards using the exergames (Figure 4). They further reported that some patients had difficulty understanding the purpose of the games. Therapists’ reasons for this were given as being either (1) due to patients’ poor self-awareness of their actual skills, making it difficult for them to explain to the patients the necessity of exercising their visuospatial exploration skills, or (2) due to the game design, which was experienced as being unappealing by both patients and therapists in combination with the abstract game control. The breakdown susceptibility of the software and the suboptimal posture, especially the use of the chin rest, were major critique points mentioned by most therapists (Figure 3). The former explained why therapists reported experiencing loss
of valuable therapy time, as they often had to re-boot the system, keeping patients waiting while the computer restarted. Thus the exergames were rated as being somewhat impractical, although therapists recognized that they were testing in a pilot phase.
Figure 3. Focus group illustration map: perceived user-friendliness.
Figure 4. Focus group illustration map: attitude toward using the exergames
Attitude Toward Using the Exergames

The general attitude of the patients toward the use of the exergames was 4.8 (SD 1.9) points (Table 1). The therapists rated it a mean total score of 3.6 (SD 1.4) points (Table 2). The patients experienced exercising with the games to be motivating and interesting and also as a “welcome change” to the conventional therapy methods provided in the Rehabilitation clinics: “Those games were a welcome diversion to normal neuropsychology, where you just sit face-to-face and have to do exhausting things all the time” (P5).

However, most patients preferred the conventional therapy methods over the exergames intervention, as they described it difficult to understand the purpose of the games:

*In the beginning, I didn’t really understand what all this meant; for what the games were good for. Then they [the therapists] explained it to me. After they had told me what aspects I had to pay attention to, then it was all good.* [P2]

*I couldn’t make sense out of it [the exergames]. I always had the feeling that the things there—those tests—were meant for ones with very severe brain damage, who weren’t back on their feet yet. […] But not for me—I don’t have such severe damage!* [P3]

*I don’t know what they [the exergames] would have been useful for. And no matter how much you have scored, you couldn’t see the progress you had made.* [P6]

The patients experienced conventional methods as more effective than the VR intervention:

*I didn’t have the impression that it [the exergames intervention] did yield much. I experienced it as being a bit silly. I had the impression that the other things simply helped me much, much more. […] The eyes were not equally challenged to move back and forth.* [P6]

The content of the games was judged as “not bad” (P4). P4 perceived the exergame “puzzle” as being difficult because the puzzle template was displayed only once at the beginning, requiring the player to piece together the puzzle out of memory. The speed of the games and the related short reaction time was another difficulty mentioned by most patients as being experienced during play. They described being initially very motivated to play the games, but over the course of the 3-week intervention, their
enthusiasm decreased, as they started to perceive playing the games as “boring” (P1, P4) and even “childish” (P1, P3):

You know, piling the ABC can be done by a first- or second-former!
And to burst balloons that pop up out of a hole isn’t very demanding either” (P3); and “Boring! In the beginning, it was good. But most recently… it was complicated to look through this thing [chin rest], you know. The other games they had were more interesting in a way. [P4]

Using the chin rest while playing was the main reason why most patients experienced the exergames as exhausting (mean 3.9 [SD 2.3]):

This [chin rest] wasn’t useful! I couldn’t sit in an upright position and look through [the chin rest] to scan the whole computer screen. This was exhausting. You also weren’t able to turn your head. [P4]

The suboptimal posture of the patients while using the exergames was also problematic for the therapists (Figures 3 and 4). It was the main reason why some therapists rated them as an unnecessary burden. They described observing their patients sitting tilted to the left side in their wheelchair due to their VSN symptoms, watching past the chin rest instead of looking through it while playing. Furthermore, the therapists expressed reservations regarding the therapeutic use of the exergames. Those reservations were based on their uncertainty of achieving a carryover effect of visuospatial room exploration skills trained in a virtual environment into the real world. Additionally, they had difficulties in perceiving the VR intervention as supportive to achieve the patient’s rehabilitation goal, namely to regain independence in daily life as well as possible. The therapists further described that their patients fully trusted them in the choice of therapy intervention to improve their skills (Figure 4). This blind trust gave the therapists a dilemma: on the one hand, they wished to use conventional therapy methods instead which they knew to be effective, but on the other hand, they recognized that the patients had agreed (and were eager) to participate in the study using this novel intervention. Despite this rather negative attitude towards exergames use, some therapists rated them as being a motivating alternative for fit patients to exercise independently, although they also rated them as being too easy for some patients.
Intention to Use the Exergames in the Future

Using the exergames regularly in the future was not viewed as conceivable yet, either among patients (mean 3.9 [SD 2.1] points) or therapists (mean 3.7 [SD 1.8] points). Most patients perceived the exergames as a good pastime and diversion that helped shorten the long days in the rehabilitation clinic (P4, P6, P7) but indicated that they would prefer doing activities other than gaming once back at home: “Up there [in the rehabilitation clinic], I thought that it is way better to do this [playing the exergames] than lying in bed or sitting on a chair while doing nothing” (P6). P5 and P7 described themselves as not being “a computer freak” (P7) or “a gamer” (P5) and therefore not wanting to use the games further at home. P3 missed the relevance to real life of the games, making the following suggestion for improvements:

Well, maybe tests that are more related to practice. You know, where you see: “Ah, this could be useful!” […] For example doing an exercise you will need in the future when you want to drive a car again. Reaction or such things… which will help me to go ahead. [P3]

For some patients, the games could have been more challenging and entertaining. P7 did not experience much pleasure while playing:

Not really…well, when I was successful, then I felt pleasure anyhow. Then I thought: ‘Indeed, I am not as dull as I thought!’ […] It simply worked out somehow, but not as good that I would have felt pleasure to play more. [P7]

Remarkably, most patients were nevertheless convinced that their family and friends would support them using the exergames at home. The patients also stated that they would recommend the exergames to other patients (Table 1). Reasons might be that the support of their relatives is taken for granted—no matter what they were doing to get better—and that they believe that trying everything to get better is the best rehabilitation strategy, including novel therapy methods like the exergames: “One should leave nothing undone, and try out everything!” (P7).

The therapists were not yet ready either to use the exergames in the future or to recommend their use to other colleagues—at least in the version used for this study—although all therapists were convinced that their workplace supports the use of VR training methods (Table 2). The neuropsychologists in particular were experienced in using the computer as a means of therapy and therefore accustomed to high-tech VR
methods. The OTs, however, were rather restrained towards VR methods, some even fearful of being replaced by computers in the future (Figure 5). Barriers to future use of exergames were diverse and numerous; for example, the benefits of virtual versus equivalent real-life tasks was mentioned by the OTs, who expressed preference for the latter therapy option. The nonadaptiveness of the software was another barrier highlighted by the neuropsychologists, as they were used to exergames with this feature.

The therapists proposed suggestions for improvements for all mentioned barriers (see Figure 6), which, given that those improvements are implemented in a new version of the exergames, indicates that a new version of those exergames would be used in the future.
Figure 5. Focus group illustration map: intention to use the exergames in the future.
Figure 6. Focus group illustration map: suggestions for improvements of the game-based virtual reality intervention.
Discussion

Principal Findings

This usability study aimed to quantitatively and qualitatively assess user perspectives (patients and therapists) of using REWIRE exergames as a novel rehabilitation intervention to treat VSN symptoms due to stroke. The findings showed that the patients as end users generally rated the use of the exergames more highly than did the therapists. Most patients experienced the games as motivating, interesting, and a welcome diversion in their daily routine during their inpatient stay in the rehabilitation clinic. The feeling of joy and motivation while playing was also described in other studies assessing user perspective in stroke patients testing a novel VR intervention [23,37,38]. Those studies tested games aiming to improve motor control in the affected arm due to hemiparesis following stroke without VSN symptoms. In our study, the patients controlled the games with a haptic device using their unaffected arm, in order to focus on improvement of cognitive skills. Most participants liked using the Novint Falcon instead of the mouse to control the games. However, some patients described having difficulties in grasping and releasing virtual objects. They confirmed that in one game, they had to touch an object with the index finger of the virtual hand to grasp it and in another game with the palm of the hand. This discrepancy was experienced as being misleading. In order to standardize the game control, Mainetti et al [39] suggest optimizing the degree of overlap between the virtual hand collision region and the target collision region.

Our sample, however, suffered from neglect and a certain related level of anosognosia [40]. They nevertheless experienced pleasure while playing. This is in line with other findings from a satisfaction questionnaire where stroke participants with neglect symptoms indicated enjoyment of the VR experience [19]. Despite having fun while playing, the presence of anosognosia in our sample negatively influenced their understanding of the purpose of the exergames. Inability to fully understand the purpose of the intervention was also a topic in a focus group interview with stroke patients without VSN symptoms [24]. It is important to make sure that patients understand the game purpose, so as to meet their expectations and to avoid frustration [23]. Although we paid weekly visits to participating rehabilitation clinics to discuss progress and progression of the exergaming with patients and therapists, it nevertheless seemed difficult for some of the former group to understand the purpose...
of the treatment strategy. This was particularly the case for the more severely affected patients.

Mainetti et al [39] tested exergames in a single patient with chronic stroke who had VSN symptoms. This patient liked the exergames and was not bored while playing them. Most of our patients, however, experienced a decreasing enthusiasm during the 3-week intervention and started to perceive the activity as boring, even though games regularly and individually progressed and were designed according to therapeutic principles [13]. It seems that basing the selection of games on personal interest of the patient could not enhance motivation while playing either. Paying attention to the diversity and progression of game complexity is no guarantee of constant use and engagement over time. Other studies testing different VR interventions with patients with cerebral palsy also described a reduction in engagement over time [41,42]. Therefore, reasons for this decreasing enthusiasm other than a suboptimal balance of providing a challenge while still enabling success might be the time point of the intervention and the lack of feedback in the achieved game scores. Compared to other stroke samples [24,25,37,39], our patients were still in the early stage of recovery and were still hospitalized, therefore in the situation of receiving daily therapy sessions with which they could compare the REWIRE exergames. In this context, it is perhaps understandable that the exergames - still a test version - fell behind other VR therapy options that are long-established in the market. Furthermore, testing a novel therapy option with stroke patients in their chronic stage, when regular therapy often might have stopped, evokes hope for further motor or cognitive improvements and therefore increases motivation [25]. Another reason for the decreasing interest might be seen in the fact that our games did not display the achieved results after each training session, unlike Lewis et al’s [25] submarine game, for example. Although this option was provided by the software, we had decided not to activate it, as the achieved scores after each game/session were not yet storable. This prevented the patients seeing progression over time. We were aware that being unable to see the achieved scores equaled a lack of feedback regarding the patient’s personal progress. However, positive feedback and measures of success are critical components to enhance engagement [23,24]. There was a rewarding system after each REWIRE game (Figure 7), but as this was random and not performance-based, patients did not care for it.
The lack of feedback experienced by fitter patients, combined with their perceptions of being insufficiently challenged while playing the exergames, might be reasons why they indicated preference for conventional therapy methods with a “real” therapist over this VR intervention. Their experiences of the games as “a good pastime and diversion” suggests that most patients did not see this as a rehabilitation intervention per se, supporting their preference for conventional therapy. This is in line with other findings, where stroke survivors experienced the novel games as supplementary to conventional therapy, the latter being viewed as providing beneficial rehabilitation [25]. Conversely, the majority of patients in another study reported experiencing VR interventions as useful as conventional therapy [43]. The nature of play that is inherent to games may be perceived differently among adult patients, as the therapeutic benefit may not be as obvious as during conventional therapy, also depending greatly on how the virtual environments were designed [23]. It might be that having prepared a predefined set of games to be played during several training sessions while only progressing difficulty levels within the same games - as it was suggested by some of the therapists - might have helped patients perceiving the exergames as a (repetitive) rehabilitation intervention. However, we preferred letting them choose and switch games according to their individual preferences to (1) keep motivation as high as possible and (2) give them the opportunity to test all exergames during the intervention.

The REWIRE game design was rated as having limited appeal by both therapists and patients. The therapists in particular wished to have games that would be self-adaptive to patient progress in order to experience ease of responsibility. For example, they missed the opportunity to prepare a series of games that would then automatically run
through during an intervention session. The lack of facility within software for patients to save and return to previously achieved difficulty levels between sessions was also noted. Those features would allow the patient to start directly at the right difficulty level and subsequently play the game independently without the therapist needing to adjust the settings before and during the training session. Although the goal of such VR interventions is to create a game menu that patients can run themselves with little input from others, it is nonetheless imperative that an expert (ie, therapist) guides progression of the games to maintain the therapeutic basis of the intervention. In one study, where game speed and progression advanced automatically, the users were overwhelmed, which negatively influenced motivation and engagement in the game intervention [44].

Some of our patients indicated that they missed the exergames’ relevance to real-life tasks - feedback also given by other stroke patients testing similar interventions [25]. This is despite the fact that we had tried to design them to be as much alike as possible. Male patients in particular perceived being able to drive a car as very important to them and therefore wished to be able to train those driving skills on the computer. Such conflicts with real-life expectations have also been described by Lewis and Rosie [23], suggesting a selection of environments that are deliberately unreal. Such simple environments have the advantage of avoiding unnecessary distractors by providing a restricted amount of stimulation, thus targeting the required rehabilitation effect. On the other hand, they comprise a risk of boredom and a related reduction in engagement for both patients and therapists.

When supporting patients to play the REWIRE exergames, some OTs expressed uncertainty in achieving carryover effects into real-life tasks. This uncertainty was one of the reasons why they would have preferred to use time for the training of real ADL rather than game play to achieve the rehabilitation goals set for their patients. Indeed, evidence for positive carryover effects of VR interventions into real life is limited [45,46]. For example, Gruskin et al [46] observed increased awareness of the involved extremity as well as greater carryover into ADL when using an auditory feedback device to alert a patient with left hemineglect when his flaccid upper extremity was in a dependent position. Gates et al [45] compared walking overground and on a treadmill surrounded by a virtual environment that applied optic flow in individuals with and without transtibial amputation. They found that both groups walked with similar overall kinematics (eg, knee flexion/extension) and kinematic variability (ankle, knee, or hip).
on the treadmill as they did overground. Their results suggest that treadmill training in a virtual environment should be sufficiently similar to overground walking in the real world that changes carry over.

Further reasons why the therapists would have preferred use of rehabilitation time for conventional therapy rather than for testing the novel VR intervention was the breakdown susceptibility of the software - giving them the feeling of wasting too much therapy time. Additionally, the use of the chin rest forced the patients to sit in a nonergonomic posture. This posture was also the reason why many patients got tired while playing, rather than because of cognitive challenge. When planning this study, we did not expect the chin rest to be a major problem when playing the exergames. Its use was precipitated by a need to avoid compensatory movements of the head. However, according to the feedback of all participants, the use of the chin rest for a whole therapy session of approximately 30 minutes was too exhausting. We therefore recommend the use of a chin rest for short assessments only rather than for a whole therapy session [47]. Such technology limitations have also been described in other studies testing VR interventions [42,48]. For example, Wille et al [48] found a correlation between software failures and reduced ratings of fun while playing. Li et al [42] have described difficulties in positioning patients with postural impairments so that they were able to operate the VR system. Not surprisingly, such technology limitations are associated with negative feedback from the users, as was the case in our sample. Those limitations were also determinative of participants’ ratings of limited intention to use exergames in the future. Other perceived barriers were not being a “gamer” (patients), as well as the fear of being replaced by computers in the future (OTs). The neuropsychologists did not share this fear, as they were more used to computer-based interventions than the OTs. As a consequence, the neuropsychologists as computer experts were the most critical users of our exergames.

**Limitations**

Some limitations of this study should be discussed. First, most of the stroke patients needed assistance in completion of the TAM questionnaire, either in retaining the paper-based questionnaire while it was on the table due to their hemiparesis, or in being helped to read the questions due to their VSN symptoms. Both of these issues may have influenced their responses. A touch-screen version on a tablet fixed on a table to avoid side slipping for those stroke patients who suffer from hemiparesis would
allow questionnaire completion with one hand only. A button placed on the right margin of the tablet could be designed to audio-display the questions, making reading of the questions unnecessary. Second, the fact that the therapists participating in the focus group were working colleagues from the same team might also have influenced their interactions and utterances during the interview. For example, in both interview groups, the team leader was also present, which might have inhibited some participants in expressing what they really thought about the exergames. We therefore chose focus group illustration maps for data analysis. Together with the flip chart notes taken directly during the interview, those FIMs allowed a precise summary of the group statements without exposing someone through using quotes, where they might recognize the person who had said that. Third, the fact that the main researcher (BC-T-A) knew all participants quite well at the time point of the interview influenced her way of conducting the individual and focus group interviews. Holding preunderstanding about the patients’ life from former meetings during data acquisition for the feasibility study might have influenced her way of formulating questions differently than when she would have met the patient for the first time. However, the interview quality probably had improved thanks to the already established relationship. Being an occupational therapist like half of the participating therapists further influenced the flow and conduct of the focus group interviews. However, speaking the same professional language might have facilitated formulating experiences made with the patients and exergames. Fourth, the recruitment of stroke patients with VSN symptoms in a clinical setting who were fit enough to test the game-based VR intervention was quite difficult. Testing such an intervention in a later, chronic stage where most patients are in a better health condition might have been easier. However, all patients were excited to take part in a research project during their inpatient stay and they cherished being asked for their personal opinion not only in a questionnaire, but also in a face-to-face interview.

**Future Work**

Lewis and Rosie [23] were entirely correct in their statement that “it may appear impossible to design a system that appeals to all users” (p. 1884). However, we should not overlook the fact that, despite all the critiques mentioned by users, most patients enjoyed playing the exergames. The criticisms identified are a motivator to improve the existing game design in order to achieve an optimal rehabilitation effect. Therefore, before thinking about testing the REWIRE exergames in a larger controlled trial of
stroke patients with VSN, for example, the game design should first be modified according to the suggested improvements. Decisions should be made regarding the degree of realism of the virtual environments: should we design environments as unreal as possible [23], or as real as possible by using a tool such as Google Street View [19], for example? Immediate feedback of the achieved game scores should be implemented together with a graphical overview of the changes over time to enhance engagement and motivation. The flexibility of the software should be increased, for example, by creating a function to save the chosen difficulty level for each game. Future work could examine if the frequency and time of game play - in our study on a daily basis over 3 weeks - or if providing a predefined set of exergames to be played instead of having free choice of game selection, influences user perspectives on the exergames. Results have shown that the use of a chin rest to control compensatory movements of the head is not recommended for a whole therapy session. Furthermore, evidence is needed to explore possible carryover effects of such VR interventions into real life in order to enhance acceptance of such interventions among therapists.

**Conclusion**

This study provided insight into user perspectives based on quantitative and qualitative statements of stroke patients suffering from VSN and therapists using novel exergames to explore the hemineglected left space in an inpatient setting. The results showed that all users perceived the REWIRE exergames as user-friendly, but that they would not necessarily entertain their use in their current format. The general attitude toward using the exergames was more positive among the patients than among the therapists. Recommendations for improvements of the exergames were mainly formulated by the therapists. Feedback suggests that once those recommendations could be realized, then the REWIRE exergames intervention could be explored using further trials. It is therefore of the utmost importance that end users (patients) and experts (therapists) are involved in order to achieve acceptable and user-friendly VR game-based rehabilitation methods.
Abbreviations

ADL: activities of daily living
CBS: Catherine Bergego Scale
FIM: Focus group Illustration Map
OT: occupational therapist
RBL: right-sided brain lesion
REWIRE: Rehabilitative Wayout in Responsive Home Environments
TAM: Technology Acceptance Model
VR: virtual reality
VSN: visuospatial neglect

Acknowledgments

This work was partially supported by the Rehabilitative Wayout in Responsive Home Environment (REWIRE) project (www.rewire-project.eu), funded by the European Commission under the FP7 framework with contract 287713, and by the Occupational Therapy Association of Switzerland [no grant number available].

We warmly thank the 19 participants who were willing to share their time and opinions with us to realize this study. Further thanks go to the two collaborating clinics where the study took place, namely the Zürcher Reha Zentrum Wald, Canton of Zurich, Switzerland, and the Klinik Bethesda, Tschugg, Canton of Berne, Switzerland. Further thanks go to Laura Wiederkehr, former masters student of the ETH Zurich, for her invaluable help in this research project, and to Martin J. Watson, PhD, for proofreading the manuscript for English and structure.

Authors' Contributions

BC-T-A designed the study; provided support in designing the exergames; conducted the acquisition, analysis, and interpretation of the data; and wrote the manuscript. E-S provided support in designing the study, developed the exergames, and contributed to editing the manuscript. ED-dB, NA-B, and RH-K initiated the study and contributed to editing the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.
Multimedia Appendix 1 REWIRE VSN exergames: Short instructions for the patient.

Bubbles
Stir all the bubbles with the stick. Only stir those matching the dragon’s color.

Balloon Popper
Pop the red and yellow balloons. Avoid the dark blue ones.

Mix Soup
Take all ingredients from the shelf on the left according to the recipe on the right and put them in the saucepan.

Gathering Apples
Pick all apples from the tree and put them inside the basket.

Faulty Pieces
Pick all fruits displayed on the computer screen from the assembly line and put them in the basket.

Shape Follower
Clean the dirty path around the house. Remember how many raccoons crossed your way. (dual task exercise)

Sequence Builder
Build the right sequence of hay bales: 1, 2, 3, ... or A, B, C, ....

Puzzle
Put together the puzzle.

Dog Leash
Take your dog for a walk. Let him collect the frisbees on the way, but avoid the trees, stones and the chicken crossing the path.
References


23. Lewis GN, Rosie JA. Virtual reality games for movement rehabilitation in neurological conditions: how do we meet the needs and expectations of the users? Disabil Rehabil 2012;34(22):1880-1886. [CrossRef] [Medline]


27. Simor FW, Brum MR, Schmidt JDE, Rieder R, De Marchi ACB. Usability Evaluation Methods for Gesture-Based Games: A Systematic Review. JMIR Serious Games 2016 Oct 04;4(2):e17 [FREE Full text] [CrossRef] [Medline]


30. Ammann BC, Knols RH, Baschung P, de Bie RA, de Bruin ED. Application of principles of exercise training in sub-acute and chronic stroke survivors: a systematic review. BMC Neurol 2014 Aug 22;14:167 [FREE Full text] [CrossRef] [Medline]


36. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [CrossRef] [Medline]


42. Li W, Lam-Damji S, Chau T, Fehlings D. The development of a home-based virtual reality therapy system to promote upper extremity movement for children with hemiplegic cerebral palsy. Technology and Disability 2009;21(3):107-113 [FREE Full text] [CrossRef]


47. Gamberini L, Seraglia B, Priftis K. Processing of peripersonal and extrapersonal space using tools: evidence from visual line bisection in real and virtual environments. Neuropsychologia 2008 Apr;46(5):1298-1304. [CrossRef] [Medline]

Chapter 7
From unawareness to emergent awareness: experiences of stroke patients with unilateral spatial neglect near the end of inpatient rehabilitation

Pre exergames intervention results of the Eye Tracker Neglect Test (ETNT) search paths of P4, male, 70 years old, 4.5 weeks onset right-hemispheric stroke and visuo-spatial neglect

This chapter has been submitted to: Disability and Rehabilitation / March 2018
Bernadette C. Tobler-Ammann, Andrea Weise, Ruud H. Knols, Martin J. Watson, Judith M. Sieben, Rob A. de Bie & Eling D. de Bruin
Abstract

Purpose: The purpose of this qualitative study was to explore everyday life experiences of persons with right hemispheric stroke and unilateral spatial neglect near the end of inpatient rehabilitation. Specifically, we aimed to investigate how they perceived their performance of daily activities, their body and self.

Materials and methods. We conducted individual semi-structured interviews with seven purposively sampled participants. All interviews were transcribed verbatim and analyzed using thematic analysis.

Findings. Participants’ experiences were captured in three over-arching themes: “unawareness of neglect”, “emergent awareness for neglect-related difficulties”, and “comparing the new life with the old one”. Findings showed that participants progressed from initial unawareness to emergent awareness for their neglect-related difficulties over the course of rehabilitation. Comparing their current life situation with the one before their stroke helped them to adapt to perceived changes.

Conclusions. This study informs health professionals regarding personal experiences of orientation in and reorganization of life of inpatients after a first right hemispheric stroke with unilateral spatial neglect symptoms. The findings highlight that being aware of neglect-related deficits is a prerequisite for using coping strategies and incorporating them in daily life. Possible therapeutic strategies that fit the current stage of recovery are discussed.

Keywords. Stroke, unilateral spatial neglect, awareness, rehabilitation, daily activities, thematic analysis
Introduction

Experiencing a stroke causes disruption and alteration in a person’s everyday life [1]. The post-stroke consequences are multifaceted, including motor [2] and cognitive [3] impairments, psychosocial consequences [4], and reduced quality of life [5, 6]. Additionally, certain stroke symptoms are more typical depending on the side of the brain lesion. Communication problems, for example, are more commonly related to left hemispheric stroke, while symptoms of spatial neglect or extinction are more common after a right-sided lesion [7, 8, 9].

Spatial neglect is a cognitive disorder having an attentional and perceptual basis [10], being a failure to respond to stimuli presented in the contralesional side of space [11]. Its prevalence varies widely from 13 to 81% of patients who had right hemispheric stroke [11]. Neglect can involve several modalities, including the visual, auditory, olfactory, tactile and gustatory senses [12]. As a consequence, patients may have difficulty in navigating their environment (e.g. they bump into objects) [13], or in carrying out activities of daily living (ADL) (e.g., shaving the left cheek) [14]. Between 9 to 22% of neglect patients additionally suffer from a reduced self-awareness (anosognosia), both for motor impairments [7] and visuospatial deficits [15]. Patients who are unaware of having a hemiparesis, for example, are convinced of being able to walk. Neglect and anosognosia are, therefore, negative predictors of functional outcome in stroke rehabilitation [6, 16, 17, 18, 19].

After the stroke event, the majority of patients are referred to a rehabilitation program to regain maximum independence. Therapies include the training of specific motor [20, 21], sensory [22], and cognitive [23, 24] skills, also focusing on improvements in social participation and occupation [25, 26, 27, 28]. Neglect-specific therapy methods include video feedback [29], visual-scanning training [30, 31, 32], right half-field eye-patching [33], and limb activation therapy [12, 34]. Promising new tools for the treatment of spatial neglect symptoms include the use of interactive virtual reality training [35], and strategies involving visuomotor imagery [36]. The combination of treatments is also recommended [37].

In addition to provision of an evidence base for treatment modalities, it is also crucial to have expertise about stroke survivors’ perceptions of their actual life situation, if optimal rehabilitation outcomes are to be achieved [38]. Knowledge of this insider’s perspective can help health professionals to gain a better understanding of their patients’ experiences, priorities and goals. Cultivating such an awareness of patient
experiences could promote effective strategies for care focused on what is most relevant to the individual during rehabilitation, which might enhance motivation for participation [38].

Previous research investigating stroke survivors’ experiences of rehabilitation highlights the importance of including these experiences in the rehabilitation process, due to its positive correlations with quality of life and treatment compliance [38, 39, 40, 41, 42, 43]. The focus has mostly been on patients' experiences of physical rehabilitation [41], the description of psychosocial processes that influence identity post-stroke [39], or on resilience of stroke patients [42]. Other research has focused on the impact of dysphagia during stroke recovery [44], or has included stroke patients with aphasia [45]. Few studies have examined the subjective experiences of persons with right hemispheric stroke and symptoms of unilateral spatial neglect during inpatient rehabilitation [46, 47, 48, 49]. These studies focused on the impact of the social environment on the personal experiences of stroke patients with unilateral spatial neglect [48], on the patient’s interactions with the physical environment in ADL [47], on neglect experiences adjacent to the first weeks following a stroke [49], and on the adjustment process of discovering their disabilities over the course of rehabilitation [46]. Less is known, however, about the personal experiences of stroke patients with unilateral spatial neglect, in relation to their performance of daily activities, their body and self near the end of inpatient rehabilitation. Knowledge of this area is important as stroke patients with unilateral spatial neglect often have different perceptions of their body and self to those of the external world, and are, therefore, oblivious to the contradictions in their own activity performances [50]. Furthermore, knowledge of patients’ views at this time-point of rehabilitation might allow health professionals to improve the discharge experience for service users [51]. Thanks to this knowledge, health professionals may gain access to patients’ concerns during inpatient therapy and, thus, can address these accordingly, aiming to reduce daily challenges encountered when returned to home to a minimum [51, 52]. Therefore, the purpose of the present study was to explore everyday life experiences of persons with right hemispheric stroke and unilateral spatial neglect symptoms, looking particularly at the period towards the end of inpatient rehabilitation. Specifically, we aimed to investigate how they experienced their performance of daily activities, their body and self.
Methods

Participants and setting

For this study, participants were recruited from an ongoing feasibility and usability project, where they tested and evaluated a novel game-based exercise intervention aiming to improve neglect symptoms [53, 54]. Occupational therapists and neuropsychologists of two stroke rehabilitation clinics in Switzerland had consecutively recruited them between March 2015 and March 2016. At the end of the project, participants were asked if they would be willing to take part in interviews towards the end of their inpatient rehabilitation. Criteria for participation were (a) diagnosed with a first right hemispheric hemorrhagic or ischemic stroke between 2 weeks to 6 months prior to the time point of study inclusion; and (b) unilateral spatial neglect as measured by the Catherine Bergego Scale (CBS) [55]. We additionally administered the individual level of anosognosia as described by Vossel et al. [14], where negative scores reflect higher self-ratings than external ratings, thereby reflecting anosognosia. All participants were independent and living at home prior to the stroke event. Clinicians usually referred stroke patients to the rehabilitation clinics after discharge from acute care. Both provided different individual and group treatments offered on a daily basis (excepting Sundays), including occupational, physical, speech and neuropsychological therapy. The first author of this study (BCT-A) conducted the interviews with the patients. The patients had already met BCT-A multiple times and had established a relationship prior to conducting the interviews, as they had also participated in the feasibility and usability project [53, 54]. BCT-A is an occupational therapist with a Master of Science degree and a PhD student, and the participants were aware of her credentials. They received written information regarding the background and purpose of this study. If needed, clinic staff (i.e. nurses) not involved in the study assisted in reading the material. Thereafter, patients decided whether to participate in this qualitative study. All seven patients from the ongoing feasibility and usability project agreed to participate. No one refused and there were no dropouts during the interviews. Before study entry, they signed written informed consent. There was no financial compensation for participating in the study. We received ethical approval from the local Ethics Committees (Zurich No. 2014-0543 / Berne No. 389/2014).
Data collection
For this study, we adopted a qualitative design, exploring data gathered from single, semi-structured, open-ended individual interviews [56]. Each interview lasted 18 to 47 minutes (mean 32 mins), duration being dependent on each patient’s health condition. Interviews were conducted between 9 and 22 weeks post-stroke (mean 13 weeks), timing being dependent on each participant’s length of rehabilitation. BCT-A conducted all of the interviews, which occurred towards or shortly after the end of inpatient rehabilitation. She performed a pilot interview with a healthy person prior to those of the main study, aiming to deepen her knowledge in conducting individual interviews [57]. The “Canadian Occupational Performance Measure” (COPM) served as a basis to design the interview guide [58]. We chose this assessment thanks to its client-centered design and predefined structure in three occupational performance areas (self-care, productivity and leisure). BCT-A initially asked each participant to describe his/her daily activities based on a daily routine post-stroke in the inpatient rehabilitation setting. Subsequently, the course of the interview touched topics such as perceived difficulties and changes in performance of daily activities, and in relation to the body and self.

With participants’ verbal approval, we audio-recorded all interviews for later verbatim transcription. BCT-A took notes during all interviews, and carried out an instant debriefing after each, aiming to record immediate observations and impressions [56]. The location of the interviews was set according to each participant’s personal preferences (see table 1 for the places of interviews). No other persons (e.g., clinic staff or relatives) were present during the interviews to allow for privacy. We made no plans for returning the interview transcripts to participants, as such member checking is unusual in persons with cognitive disorders [59].

Data analysis
We adopted a form of thematic analysis that was data-driven, therefore paying no attention to pre-existing coding frames [60]. Thematic analysis is a flexible and useful research method examining the ways in which life events and experiences are described within a social context [60]. A 6-phase guide to doing thematic analysis provided by Braun et al. [60] was followed in this study:
1. **Familiarizing yourself with your data**: transcribing the data, reading and re-reading the data, noting down initial ideas

2. **Generating initial codes**: in a systematic fashion across the entire data set, collating data relevant to each code

3. **Searching for themes**: collating codes into potential themes, gathering all data relevant to each potential theme

4. **Reviewing themes**: checking if the themes work in relation to the coded extracts and the entire data set, generating a thematic map of the analysis

5. **Defining and naming themes**: ongoing analysis to refine the specifics of each theme / generating clear definitions and names for each theme

6. **Producing the report**: selection of vivid, compelling data extract examples / final analysis.

BCT-A initially prepared and analyzed each phase of analysis before showing it to the second author (AW) for contributing to coding and analysis and subsequent discussion. AW is an occupational therapist with a Master of Science degree, being an expert in neurorehabilitation. The actual phase of data analysis was always discussed with due consideration of all previous ones. This allowed us to check if the condensed data units still represented the meaning of the whole data set. Additionally, we checked the interview field notes to complement the interview data, integrating those notes as memos to the corresponding data extract. Only after we had reached consensus between authors did data analysis move forward to its next phase. We searched for themes or patterns across all interviews, rather than within one single interview. The aim was to provide a more detailed and nuanced account of one particular theme, or group of themes, within the data [60]. We considered data saturation as reached when further coding relevant for the scope of our study was no longer feasible [61, 62]. We used Office Word to structure the data analysis phases 1-3 and 5-6, and Power Point to generate thematic maps (phase 4). An example of how we conducted our data analysis is given in figure 1. We exchanged all documents in an anonymized format via a secured email-address and stored them in a password-protected computer. We followed the “consolidated criteria for reporting qualitative studies (COREQ)” 32-items checklist to produce this report [63].
**Figure 1:** Example of thematic data analysis

<table>
<thead>
<tr>
<th>Phase 2</th>
<th>Phase 3</th>
<th>Phase 4</th>
<th>Phase 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Extract</strong></td>
<td><strong>Condensed Meaning Unit (Descriptive)</strong></td>
<td><strong>Coded for (Analytic)</strong></td>
<td><strong>Potential Theme: Unawareness of Hemiparesis</strong></td>
</tr>
</tbody>
</table>
| "I simply couldn’t walk – not a single step. This was very clear to me. But I didn’t realize I was paralyzed, and even less so my hand." (P6/78/f) | - Realizing that walking does not work  
- Being unaware of the reason why | - Experiencing physical limitations  
- Unawareness of hemiparesis | - Being confronted with limitations  
- Being unaware of the hemiparesis  
- ... |
| "They [the therapists] don’t let me do the transfer on a chair independently yet. They say that it is too dangerous. I however have the impression I can do so." (P4/70/m) | - Needing assistance for transfer on a chair  
- Having the impression of being able to do the transfer on a chair independently | - Experiencing mismatches in perception  
- Overestimating one’s skills | - Experiencing mismatches in perception  
- Overestimating one’s skills  
- Putting oneself in self-endangerment  
- ... |

**Potential Theme: Discrepancy between Self-Awareness and External Feedback**

- Being unaware of the paralyzed side of the body
- Trusting themselves to do more than they were able

**Unawareness of Neglect**
Findings
In total, seven participants (5 men and 2 women; mean age 69 years, range from 53-78 years) participated in this study. All were diagnosed with a right hemispheric stroke and had symptoms of unilateral spatial neglect of varying degrees of severity (CBS score range from 5-17 points). Five out of seven participants had gained emergent awareness for their neglect at the time point of the interview near the end of inpatient rehabilitation. The demographic data of the participants are presented in table 1.

Descriptions of daily rehabilitation routines
During the interviews, all participants described a comparable daily routine during inpatient rehabilitation, with meal times, therapy units and leisure time structuring their days. Therapies included computer-based therapeutic exercises, workouts in the weight room, cooking groups and exercises for the hemiparetic limbs. Neglect-specific therapy comprised visual scanning training, the training of dual tasks, and manual grasping exercises towards the left side of the body. P4 gave an example of using a red sticker as a landmark for reading during therapy: “They [the therapists] always stuck a red sticker on the left margin of the computer screen for me. They then said to pay attention to this sticker [while reading]. Because of the [neglected] left side.” The participants reported how the clinic staff treated them well in the clinic. P2 characterized his physical therapist as an “excellent and tough woman” and P7 felt that the clinic staff were “being very nice. They kept looking after me, helped me.” Furthermore, they all experienced therapy progress since they first came to the rehabilitation clinic. “It’s encouraging when you don’t have to call a nurse for every trifle.” (P4). During the interviews, most participants talked about other stroke patients in the clinic with which they compared themselves. P5, for example, realized that many patients were in wheelchairs: “They all had to firstly learn how to walk again. This was great luck for me that I was able to walk again so quickly.” Observing others in a similar situation helped them to deal with their own situation in a more optimistic way. The data analysis revealed three over-arching themes in the participants’ discourses about how they experienced everyday life post-stroke with unilateral spatial neglect, nearing the end of inpatient rehabilitation. Each theme incorporated specific sub-themes (figure 2).
# table 1: Demographic data of participants

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>64</td>
<td>74</td>
<td>64</td>
<td>70</td>
<td>53</td>
<td>78</td>
<td>77</td>
</tr>
<tr>
<td>sex</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>f</td>
<td>f</td>
</tr>
<tr>
<td>weeks of inpatient rehabilitation</td>
<td>6</td>
<td>8</td>
<td>14</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>CBS(^1)</td>
<td>7</td>
<td>17</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>neglect awareness(^2)</td>
<td>0</td>
<td>0</td>
<td>-0.07</td>
<td>0</td>
<td>0.27</td>
<td>0.18</td>
<td>-0.92</td>
</tr>
<tr>
<td>weeks post-stroke at time point of interview</td>
<td>9</td>
<td>12</td>
<td>22</td>
<td>12</td>
<td>14</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>interview place</td>
<td>clinic</td>
<td>home</td>
<td>clinic</td>
<td>clinic</td>
<td>clinic</td>
<td>home</td>
<td>nursing home</td>
</tr>
<tr>
<td>stroke type</td>
<td>ischemic</td>
<td>ischemic</td>
<td>ischemic</td>
<td>ischemic</td>
<td>ischemic</td>
<td>hemorrhagic</td>
<td>hemorrhagic</td>
</tr>
<tr>
<td>hand function of affected left hand</td>
<td>none</td>
<td>full</td>
<td>none</td>
<td>full</td>
<td>clenching fingers</td>
<td>full</td>
<td>full</td>
</tr>
<tr>
<td>locomotion</td>
<td>WC</td>
<td>W</td>
<td>WC</td>
<td>WC</td>
<td>W</td>
<td>W</td>
<td>WC</td>
</tr>
<tr>
<td>profession</td>
<td>farmer</td>
<td>divisional bank director</td>
<td>catering service holder</td>
<td>blue-collar worker</td>
<td>newspaper editor</td>
<td>secretary</td>
<td>bookseller</td>
</tr>
<tr>
<td>family status</td>
<td>married / 2 adult children</td>
<td>widowed / no children</td>
<td>married / 2 adult children</td>
<td>domestic partner / no children</td>
<td>married / 2 children</td>
<td>married / 2 adult children</td>
<td>widowed / 3 adult children</td>
</tr>
<tr>
<td>leisure time activities</td>
<td>dog caring/ riding a bike/ timbering</td>
<td>gardening/ visiting relatives</td>
<td>skiing/ biking/ sailing/ playing music</td>
<td>singing in a choir/ playing cards and bowls</td>
<td>gardening/ cooking/ hiking</td>
<td>visiting friends/ reading/ listening to music</td>
<td>reading/ knitting/ sewing/ travelling</td>
</tr>
</tbody>
</table>

Legend: CBS: Catherine-Bergego-Scale; f: female; m: male; P: participant; WC: wheelchair; W: walker

\(^1\)CBS: maximum score = 30 (severe neglect); 0 points = no neglect

\(^2\)Level of neglect awareness: anosognosia index ≥ 0: no anosognosia; < 0: patient suffers from anosognosia
Unawareness of neglect

During the early stage after stroke, all participants reported being unaware of having neglect. They described their unawareness by experiencing incomprehensible mismatches between personal capabilities and external demands. We identified three sub-themes where those mismatches were manifest: in (a) being unaware of the paralysed side of the body; (b) perceiving incomprehensible mismatches between personal capabilities and (i) results of actions, (ii) content of therapy; and (c) trusting themselves to do more than they were able.

Being unaware of the paralysed side of the body

Most participants expressed an initial unawareness of being paralysed on one side of the body. P4, for example, told his wife not to call the ambulance after the stroke event, because he perceived that his hands worked perfectly: “She nevertheless called [the ambulance]. I told them to go: Nothing happened, you see, I can walk again and the arm functions well, too. But it was the right arm that functioned well and that still works fine today.” During acute care, P6 recounted not realizing that he had a paralyzed leg:
“I simply couldn’t walk – not a single step. This was very clear to me. But I didn’t realize I was paralyzed, and even less so my hand.” P5 wanted to get out of bed in hospital without having the capacity of doing so: “I always kept the nurses busy. As soon as I woke up at night [after the stroke event], I wanted to get up. I somehow had the feeling that I was able to walk.”

Perceiving incomprehensible mismatches between personal capabilities and (i) results of actions, (ii) content of therapy
During the interviews, all participants described how they perceived discrepancies between their personal expectations and the actual achieved results. Some participants, for example, expressed how they got exasperated with their confusing arm, because it did not do what they wanted. While taking a shower, for example, P5 realized that “everything I tried to hold in my left hand fell on the floor, which extremely upset me.” He mentioned having the same feeling when he unsuccessfully tried to place his rucksack on his back after his stroke: “I was extremely upset because I could not thread in my left arm through the strap of my rucksack. This didn’t work out.” For most participants, the neglect-specific therapy made no sense in the early stage of recovery: “Initially I didn’t understand what this was all about [the computer-based visual scanning therapy]. For what purpose it was meant to be useful. I’ve always seen all objects to the left and to the right. But she [the therapist] kept telling me that I have a ‘left-handicap’.” (P2). Another example was provided by P6, who had to attend an “arm-group” without understanding why: “Because I can do everything with my arm, no problem. But they told me that it was paralysed, too, which I didn’t realize.” Despite experiencing such mismatches between the external feedback they received from clinic staff and their self-awareness, all participants were nonetheless compliant with therapy. “Well, I thought: As I’m here anyway [at the rehabilitation clinic], I might just as well improve “this thing” [neglect] with those therapists.” (P2).

Trusting themselves to do more than they were able
Most participants talked about trusting themselves to do more than they were able. This experience was described in terms of having to deal with constraints imposed by clinic staff for safety reasons which they did not understand, and with prognoses made by clinicians which they did not agree with. P2, for example, explained: “Throughout my life I’ve always read a lot. That’s why it initially bothered me not being allowed to
go to the cafeteria independently [to read the newspaper]. And well, on the second-to-last day, he [the neurologist] told me: ‘We simply had to thwart you a bit. I don’t like having people in casts in the clinic.’” Another example was provided by P4: “They [the therapists] don’t let me do the transfer on a chair independently yet. They say that it’s too dangerous. However, I’ve got the impression I can do so.”

Driving was one activity most participants wanted to resume in the near future, but the doctors kept telling them that they were not allowed to do so in the current situation. P3, for example, was convinced: “I’m able to drive, I know that. […] I see everything. I’m not restricted in my field of view. I don’t have all those limitations. My responsiveness is still good.” P1, who used to be a farmer, took the opportunity of climbing onto one of his agricultural machines when on a weekend at home during his inpatient rehabilitation. He had the impression of being able to drive, too. Furthermore, he used to cut his fruit trees alone in the past. “I then asked my therapist: ‘Will I be able to climb those trees, use the ladder next year to cut them?’ I should forget about that and hand it over to someone else, she replied. Well, I nevertheless thought I would be able to do it. I considered trying it.”

Emergent awareness for neglect-related difficulties

During the course of rehabilitation, most participants recounted experiencing an emergent awareness for their neglect symptoms. This process started by (a) searching for a reason behind difficulties experienced during daily activities. The participants further described how they (b) gained partial awareness for the neglect-related difficulties through doing, and (c) slowly started to implement coping strategies for mastering those difficulties during daily life.

Searching for a reason behind difficulties experienced during daily activities

All participants spoke about trying to make sense of the difficulties they experienced when doing familiar activities. P4, for example, explained why some cards unexpectedly fell out of his left hand while playing: “In those days, they [the playing cards] just fell once or twice to the floor. […] Not all of them, but two to three cards I had taken. I’d probably grasped them a bit too hard.” Some participants pictured having difficulties when reading the newspaper, or when working on a computer. They believed that their “poor eyesight” (P7), forcing them to “look two to three times” (P6) until they were able to read something, was the cause for those difficulties.
Gaining partial awareness for perceived difficulties

With time, some participants narrated that they started becoming aware of difficulties. P5, for example, became aware of having problems with calculating and reading the time: “I have a problem with reading the time. Often I don’t look properly. I’m then sometimes too early and wait for the therapists. [...] The same story with calculating. I do have problems there, too. Adding and subtracting [figures].” P6 described her experience when she tried for the first time to type in her name: “I felt as if struck by a hammer when I sat at a computer for the first time and realized that I wasn’t able to write my name. Awful! This was terribly awful. Really. ‘You’ve now become a basket case’ I thought.” She further recalled having been tasked with overseeing half of the flowers in the rehabilitation unit: “I had to water flowers, somewhere [in the rehabilitation clinic] for a day; and forgot half of them.” At this stage of rehabilitation, however, the participants were able to name their perceived problems, but it still seemed hard for them to find solutions to solving those difficulties.

Implementing coping strategies for neglect-related difficulties

At the time point of the interview, most participants had started implementing coping strategies for perceived neglect-related barriers to mastering daily activities. Strategies used ranged from delegating tasks to others, to doing them together with other people, to adopting an increased level of awareness towards their left side. P6, for example, explained how she now organized tasks with her husband, whereas she had previously completed these independently prior to the stroke incidence. “We now do quite a lot of things together.” Besides cooking, they now shared doing some computer work, too: “I cannot pay our bills [on the computer] at the moment, because I can’t properly read the figures. Well, this would result in very funny bills if I did pay (laughs). No, my husband simply has to type them now. Yes - teamwork.” P5 expressed experiencing a raised awareness towards his environment, which “happened unknowingly”: “I take more care of what is on the left. This was already a problem from the beginning in the hospital. Because of the neglect, you know. That I don’t always bump into the door frame. My awareness for that got raised. This doesn’t happen to me anymore. [...] You simply take much more care of that. Also when being in a crowd so as not to bump into people.”
Comparing the new life with the old one

During the interviews, the participants often compared their current life situation with the one before their stroke. For P2, his current life situation was “no better or worse than before”, as he was able to cook, dress and shower again independently, plus was able to work at the computer on his own once more. Most participants, however, perceived that changes had occurred, with associated decreases in pleasure in some instances.

Perceiving changes

The participants talked about perceived changes. For P7, “the whole of life has changed. It’s simply not the same as it used to be. I’m not as fit and adventurous as I used to be. The body isn’t the same anymore.” P6 perceived herself being slower than before the stroke: “Well, you just became slower in everything, you know.” Another example was given by P3, who used to be very active before his stroke: “I always did something from morning till evening. I’ve been working 12 to 14 hours a day and additionally did lots of things during my leisure time. I always used to be active – and now all of a sudden you’re just a potato bag.” Feeling tired was another issue most participants talked about. P6, for example, explained: “I was tuckered out in the evenings. From 6 p.m. onwards, I would have been able to fall asleep standing up. […] So many sensations bombarded me each day - they frazzled me out. It is getting better now with every week.”

Experiencing a decrease in pleasure for certain activities

Some participants expressed a decrease in pleasure for beloved activities, due to perceived neglect-related difficulties. Both women, for example, talked about reading, one of their leisure time activities before the stroke: “Then you try to read the newspaper. I don’t like reading that much anymore, because I have much more difficulty reading. […] I then think: ‘is it worth reading it?’ (laughs). This effort, you know. […] No - reading isn’t as nice as it used to be.” (P6). This change was perceived as “difficult for someone who likes reading” (P7). Another activity P6 did not like doing that much anymore was her computer work, “because I have so much bother now doing my computer work”.


Discussion
This study identified three themes that described everyday life experiences of seven persons with right hemispheric stroke and symptoms of unilateral spatial neglect in the period approaching the end of inpatient rehabilitation. The findings showed that all participants reported being unaware of having neglect during the early stage post-stroke. They expressed this unawareness in perceived mismatches between personal capacities and external demands, causing feelings of incomprehension and confusion. During the course of rehabilitation, most participants experienced an emerging awareness for their neglect symptoms. They described this experience as a stepwise process from searching for a reason behind experienced difficulties, to implementing coping strategies for mastering everyday life. The participants further narrated comparing their current life situation with the one before their stroke, with perceived changes and a consequent decrease in pleasure for some activities. The study findings add new insights to the body of knowledge within neglect research concerning perceptions of rehabilitation from the patient’s perspective [46, 47, 48, 49]. Several of the findings of this qualitative study have previously been reported in the neglect literature. In particular, the described process from initial unawareness to emergent awareness for neglect-related difficulties, over the course of rehabilitation, is well known [17, 64, 65, 66, 67]. Findings therefore confirm existing clinical knowledge of unawareness of neglect as being a transitory phenomenon [67]. The study findings, however, supplement existing literature by providing evidence of rehabilitation experiences from an insider’s perspective, with respect to stroke patients with an emergent level of awareness who are nearing the end of their inpatient rehabilitation. This consolidation of existing knowledge might help to reduce barriers to adoption of best practice in the management of unilateral spatial neglect [68]. Furthermore, it might enhance the therapy process by, for example, planning and implementing the therapy in accordance with the patient’s stage of recovery, needs, and preferences [34, 69]. Some of the findings in the present study are of particular interest for health professionals working with stroke patients and unilateral spatial neglect in an inpatient rehabilitation setting, which we would like to discuss below.

Unawareness of neglect
The initial unawareness of hemiparesis described in this paper triggered perceptions of mismatch between personal capabilities and results of actions, causing feelings of
confusion and exasperation in some participants. Other patient groups, such as the traumatically brain injured, have also described similar feelings of loss of control over one side of the body [70, 71]. Like our participants, they also experienced their body as ‘something that can’t be trusted’, thus hampering their emotional response towards their suddenly disabled selves.

According to the ‘pyramid model of awareness’, restoration of this faculty occurs on three levels [72, 73, 74]: (1) intellectual awareness, where patients are aware of the types of deficits following stroke; (2) emergent awareness, where patients have regained partial awareness of deficits during task performance; and (3) anticipatory awareness, where patients are able to predict their task performance. Adopting this model in clinical practice might assist therapists in choosing the right intervention that fits the current stage of awareness recovery. Therefore, at this early stage, the study findings suggest implementing an educational approach for the patients to learn about the types of impairments following stroke to gain intellectual awareness [73].

The findings revealed that lack of understanding for neglect-specific therapy by patients at an early stage of recovery might further aggravate their confusion. In order to diminish confusion, experts recommend choosing specific treatments by recovery phase and, thus, waiting with neglect-specific therapy interventions until the relevant level of awareness has been established [34]. For example, some participants described how they were convinced of being able to do more than the clinic staff allowed, thus potentially increasing their risk of injury. There, structuring the environment and providing greater levels of supervision to enhance neglect patients’ safety during the acute stage of recovery are effective interventions [74], as was the case for P2’s visits to the cafeteria.

**Emergent awareness for neglect-related difficulties**

During the course of rehabilitation, the participants described an emergent awareness for their neglect-related difficulties experienced during performance of everyday tasks such as reading and watering flowers. The more they engaged in doing daily activities, the more aware of their disabilities and the effects of these on their lives they became. This stage of recovery has been variously described as “fostering autonomy” [41], “adjustment” [42] and “constant strife for mastery” [47] over the problematic physical world and body. It is described as a process of adjustment often marked by setbacks [42], which triggers new modifications and simplifications of daily activities whenever
possible. During this process, provision of an appropriately timed and measured therapeutic input can be challenging for health professionals. Allowing patients to practice daily activities on their own and – thus – experience setbacks, and providing sufficient time and support for them to gain familiarity with new therapeutic appliances, might both be examples of strategies designed to meet their specific needs during this adjustment process [40, 43].

Only after having recovered to a level of emergent awareness is the patient ready for a multicontextual approach to neglect-specific interventions [74]. The combination of active learning of new skills to compensate for impairments with receipt of passive stimulation training seems to be most effective [64, 74, 75]. Examples for passive stimulation training include sensory cueing [76, 77], prism adaptation [78, 79], and virtual reality [53, 80]. Evidence also supports the benefits of providing video and observer feedback [81]. Confronting anosognostic patients with audio-visual recordings where they can observe themselves when performing daily activities, for example, might help to reinstate motor awareness. For feedback to be successful, it is, however, important to have established a good relationship with the patient and to know his/her actual level of awareness. This knowledge might allow health professionals to adjust their feedback and support to the patient’s current level of awareness [48]. Furthermore, adopting a feedback culture that is concordant among different health professionals (e.g., nurses, therapists, clinicians), might reduce the described perceptions of mismatches and, thus, enhance the patient’s comfort.

Towards the end of inpatient rehabilitation, some participants were able to predict their task performance and implement appropriate coping strategies to accomplish those activities successfully. These participants had regained a certain level of anticipatory awareness [72, 73, 74]. For example, doing tasks together that have been done independently prior to the stroke event, such as paying bills on the computer for P6, was a coping strategy that worked well for her and her husband. This creation of a ‘shared life-world’ has previously been described [48], where family members have learned how best to assist their relative during task performance. For health professionals, it is therefore important to integrate families within the rehabilitation process to increase an understanding of what it means to live with a stroke and unilateral neglect [82]. Embracing an increased awareness for the neglected environment is another example of a coping strategy, as for example adopted by P5 in order not to “bump into objects”. This finding is in line with previous neglect studies [46, 230]
Comparing the new life with the old one

During the interviews, participants often compared their current life situation with the one before their stroke. This experience of change, transition and transformation has previously been described, challenging the participants’ sense of self by being confronted with a new post-stroke self [38, 39]. Participant P3, for example, perceived his body as a ‘sack of potatoes’ early after stroke, with inherent feelings both of anger and frustration. Salter et al. [38] and Hole et al. [39] described this experience as an ongoing process of re-interpretation of the post-stroke self. They highlight professional and family support as crucial at this time to assist patients to rebuild their lives through this experience. For health professionals, it is important to acknowledge that the present self of stroke survivors is not a desired self, which can cause tensions between the past and present self [71]. Engagement in meaningful occupations in familiar environments can facilitate this self-adaptation process [48]. In our study, the opportunity to go home over the weekend during inpatient rehabilitation supported this process.

Needing more time to perform everyday activities and being more tired thereafter were experiences reported by all participants. For example, P6 described being “tuckered out in the evenings”. The stroke event creating effort where there had previously been none is a well-known phenomenon [41]. There is evidence that stroke-related fatigue can overwhelm participants’ ambitions to be more active or independent [41]. Therefore, it is important for health professionals to recognize that recreational activities, such as reading, crafts, sleep, and electronic games, also form part of rehabilitation in order to reduce fatigue and, thus, improve the patient’s capability to participate in therapy.

The study findings also emphasised how some patients can experience a decrease in pleasure for certain activities due to their neglect-related difficulties, as for example P6’s comments about reading the newspaper. Experiencing difficulties in reading were
also described by Lampinen et al.[47] and Klinke et al. [52], who recount resultant feelings of powerlessness. Interestingly, our participants talked mainly about leisure time activities, thus, about tasks that were important and meaningful for them, as they associated joy and pleasure during performance. Therefore, addressing and integrating such meaningful activities during rehabilitation is deemed crucial to meet the patients’ priorities, expectations and goals.

**Limitations and future research**

Inviting participants from another study to participate in our qualitative study implies the risk of limited specificity of knowledge and experiences within the sample. However, the concept of ‘information power’ developed by Malterud et al. [83] states that the less information a sample holds, the higher the number of participants which is needed. We were fortunate to have a group of participants with a diversity of experiences thanks to their differences in age, gender and social status. Nevertheless, one might argue that a sample of seven participants and conducting a single interview is not enough to reach data saturation. This might be true in terms of reaching “meaning saturation” to “understand it all”, where 16 to 24 interviews are recommended [62]. However, potential participants, who were both well enough and willing to participate in this study, were hard to find. We were, therefore, happy having reached at least “code saturation”, indicating that researchers have “heard it all” after seven interviews [62].

During thematic analysis, distinguishing neglect-specific experiences from other depicted impairments - for example those due to the hemiparesis - was quite challenging, as those experiences are always intertwined. Therefore, a constant dialogue between the co-authors of this study was a prerequisite for the selected chunks of data and related codes to match. A mix of data collection methods (triangulation), such as for example conducting interviews together with participation observation, might facilitate this process in future studies and increase validity of the findings. This approach has been taken in other studies [47, 48, 49, 50]. In our sample, however, we considered adopting more than one data collection method as expecting too much of our patients, because of their participation in the feasibility and usability study.

Future studies with this target group should consider applying a longitudinal design by conducting at least two interviews per participant. Follow-up interviews might give
insight into how participants’ lives have developed over time, if anosognosia still persists, and how they reason about their inpatient rehabilitation from a more distant point of view. If adopting a longitudinal design, a narrative approach might be a suitable method of data analysis, focusing on change over time [84].

Conclusion
This study presents everyday life experiences of seven persons with right hemispheric stroke and unilateral spatial neglect nearing the end of inpatient rehabilitation. The findings showed that all participants initially suffered from unawareness of having unilateral spatial neglect during the early stage post-stroke. During the course of rehabilitation, most participants experienced an emerging awareness for their neglect symptoms, while comparing their current life situation with the one before their stroke. Only after having recovered to this level of awareness is the patient ready for a multicontextual approach to neglect-specific interventions. The study findings further highlight that being aware of neglect-related deficits is a prerequisite for using coping strategies and incorporating them in daily life. They inform health professionals regarding personal experiences of orientation in and reorganization of life in patients with right hemispheric stroke and unilateral spatial neglect near the end of rehabilitation. Health professionals should provide a multimodal, client-centered treatment approach fitting the current stage of recovery, while maintaining a collaborative relationship and open communication style with service users.

Acknowledgments
We would like to thank the two collaborating clinics who made recruitment of potential participants possible, namely the Klinik Bethesda Tschugg and the Zürcher Höhenklinik Wald in Switzerland. Many thanks go to the seven participants for sharing their time and experiences with the first author. Final thanks go to Anina Röthlisberger for helping in transcribing the interviews.
Author contributions

BCT-A designed the study, conducted the interviews, performed interview transcribing, data analysis, interpretation of the data and wrote the manuscript. AW provided support in analyzing and interpreting the data and contributed to editing the manuscript. EDdB, RHK, MJW and JS contributed to critically reviewing and editing the manuscript. All authors read and approved the final manuscript.
References


235


Post exergames intervention results of the Eye Tracker Neglect Test (ETNT) search paths of P4, male, 70 years old, 8 weeks onset right-hemispheric stroke and visuo-spatial neglect
Introduction

The primary objective of this thesis was to evaluate a newly developed series of rehabilitation exergames, examining their feasibility and usability in stroke patients with visuo-spatial neglect (VSN) symptoms. Secondary objectives were to gain knowledge regarding reporting of exercise training principles for sub-acute and chronic stroke patients and their adherence to that training, plus the testing of psychometric properties of two novel assessments: the Virtual Peg Insertion Test (VPIT) and the Zürich maximal mental status inventory (ZüMAX). A final study explored everyday life experiences of persons with right hemispheric stroke and unilateral spatial neglect symptoms near the end of inpatient rehabilitation.

This chapter discusses the main findings of the six studies presented in this thesis. Subsequently, limitations and methodological considerations in relation to these studies are examined. Finally, implications for clinical practice and recommendations for future research are provided, followed by a general conclusion to the thesis.

Main findings

Reporting of and adherence to stroke interventions

When reporting the results of an intervention, it is important that the precise principles of the exercise training are consistently and accurately reported [1]. Application of such principles in the design of an exercise intervention ensures that the dose and type of exercise is planned such that benefits for the recipient are maximized [2, 3]. Equally important is the reporting of the intervention in sufficient detail to permit intervention replication and results interpretation, including the prescription of the components of the training program and participants’ adherence to that exercise prescription. Detailed reporting of the dose of exercise prescribed (and received) allows for a more adequate interpretation of results, as well as providing information about the tolerability and safety of the intervention.

The systematic review in CHAPTER 2 provides an overview of the currently available evidence on the application of principles of exercise training in sub-acute and chronic stroke survivors [4]. The results showed that randomized controlled trials (RCTs) on aerobic and/or resistance rehabilitation training in this target group incompletely and inconsistently reported the prescription of, and especially the adherence to, exercise. Detailed reporting of the training program is crucial to determining the optimal volume,
frequency and intensity required in an exercise intervention aiming to achieve durable long-term benefits for stroke patients [5]. Furthermore, there was discrepancy between the reporting of the Frequency, Intensity, Time, Type (FITT) components in the exercise intervention and adherence to those components. This general failure of reporting adherence with respect to the FITT components all but obscures this crucial aspect of an intervention [6], and hence prohibits important considerations that must be made before replicating an intervention in a clinical setting. There are several reasons why exercise programs might not be performed as prescribed, including but not limited to patient-related factors [7, 8], environmental factors, and health concerns [9]. Another reason might be that the dose and type of exercise do not meet the patient’s needs and skills, resulting in excessive demands or boredom. Without detailed information on both the type and dose of exercise that is prescribed and actually received, developing optimally designed and dosed exercise prescriptions for a desired level of benefit (i.e., response) remains difficult, thus ultimately limiting the implementation of evidence-based training programs.

**Basing clinical decisions on reliable and valid (stroke) assessments**

In order to base clinical decisions on results of clinical instruments, the latter should be both psychometrically sound and clinically useful. Before using newly developed instruments with patients, empirical evidence is, therefore, needed to demonstrate that they assess the desired outcome in a reproducible fashion (reliability), and measure what was intended to be measured (validity) [10].

**CHAPTER 3** introduces a novel virtual reality (VR) assessment tool, the VPIT, for measuring upper extremity (UE) function post-stroke [11]. The concurrent validity of this tool, together with its test-retest reliability, was evaluated with 31 chronic stroke patients utilizing the Nine Hole Peg Test (NHPT) and the Box and Block Test (BBT). A subgroup of 11 patients was able to perform the VPIT solely via use of their affected arm, while 20 patients also utilized support from their unaffected arm. The results showed that the VPIT is a promising upper limb function assessment, with demonstrable feasibility for use with this diverse group of chronic stroke patients. With high test-retest reliability in five and low SDD% in eight, many of the VPIT’s nine parameters demonstrated consistency when performed by patients with chronic stroke. Furthermore, these parameters were susceptible to change. This allows diagnostic and therapeutic use of the VPIT in clinical practice for this patient group. The high intra-
subject variation indicated that the VPIT is a demanding test for this stroke sample, who thus require a thorough introduction to the assessment. The VPIT’s low concurrent validity suggested that it is inherently different to more conventional tasks, indicating that performing this hapto-virtual reality assessment perhaps requires other components of upper limb motor performance to those evaluated by the NHPT and BBT.

CHAPTER 4 examines the discriminant validity and test-retest reliability of the novel cognitive assessment tool “ZüMAX” in patients with stroke [12]. It was administered in a test-retest design to 33 community-dwelling stroke survivors, and once to 35 healthy controls matched for age and sex. Significant group differences were found in the subscores for the cognitive domains of executive functions and language as well as total score. No group differences, however, for the domains praxia, visual perception and construction, or learning and memory could be detected. It is possible that by the time of recruitment, most participants had already recovered almost completely in these domains and thus achieved near-perfect scores. This fact created a ceiling effect, being present if >15% of the respondents scored the maximum [13]. This made a distinction among participants who had perfect scores impossible, indicating limited validity. Furthermore, results showed that the stroke group needed more time to perform the ZüMAX than the controls. Needing more time to complete an assessment requires a longer period of concentration, disadvantageous in a population containing some members likely to have cognitive deficits [14]. One reason for the prolonged test could be impairments in patients’ information processing speed, a domain often affected after stroke [15, 16]. Another reason was that 66% of our stroke sample were performing the ZüMAX writing tasks using their affected hand. The ZüMAX showed fair to good reliability in four of the six measures, the exceptions being the domains visual perception and construction and learning and memory. These domains may have been subject to learning effects despite efforts to prevent them by spacing the test and re-test assessments about a week apart. The other three cognitive domains (executive function, language, and praxia), as well as the total score, all had good test-retest reliability. Their low SDDs indicated that the ZüMAX is sensitive to change on an individual level [17]. To conclude, the main findings of this study showed that the ZüMAX is a promising instrument in measuring neuropsychological disturbances in stroke survivors.
Testing the fit of a novel game-based stroke intervention in a real-world setting

When selecting and adapting novel therapy interventions for stroke patients, evaluation is an obvious prerequisite before their broad implementation in the clinical setting is justified. Therefore, using feasibility research as a determinant for accepting or discarding an intervention approach is a key way to advance only those interventions that are worth testing [18].

After the REWIRE game-engineers had finished designing nine exergames, feasibility of these exergames was evaluated in CHAPTER 5 [19]. Those exergames were aimed at improving exploration of the neglected space in patients with VSN symptoms post stroke. A total of seven patients attended the 3-week exergames training program on a daily basis during their inpatient rehabilitation. All patients tolerated the exergames intervention well, which they mainly performed without major difficulties. This finding suggested that the implementation of those rehabilitation exergames in the clinical setting was feasible in this target group. Furthermore, the results showed that patient compliance for the exergames was excellent. Such a result was possibly due in part to clinic staff’s commitment, as the VR intervention was smoothly integrated into the daily therapy schedule. However, as we aimed to test the fit of our intervention in a real-world setting, we prioritized clinic constraints over optimal conditions and settings. The fact that there were no adverse events during the training period in our case series was encouraging. The design of the games may have contributed to the safety of our participants, including options that allowed them to choose a virtual environment fitting their current skills and needs. Furthermore, we think that our approach of offering gameplay in a seated position further augmented patient safety. Our efficacy testing showed a group trend towards improvements in cognitive and spatial exploration skills. However, these changes cannot be exclusively attributed to the exergames intervention, due to ongoing stroke rehabilitation as well as spontaneous recovery of stroke symptoms [20]. This is further supported by the fact that most patients continued improving their scores in the NET and ZüMAX assessments post intervention, achieving the highest scores four weeks later during the follow-up measurement.
Not to be neglected: the user’s perspective

When new VR interventions are designed, it is important to assess the target users’ perspectives of such management strategies, in order to determine their expectations and perceptions of the system being tested. For our series of VSN exergames, the “target population” not only included patients but also therapists, as the latter’s preferences and preconditions warranting the use of such games may differ from those of the patients [21]. An evaluation of users’ perspectives is presented in CHAPTER 6 [22].

The findings indicated that the patients as end users generally rated the use of the exergames more highly than did the therapists. Most patients experienced the games as motivating, interesting, and a welcome diversion in their daily routine during their inpatient stay in the rehabilitation clinic. This feeling of joy and motivation while playing is in line with other studies assessing user perspectives in stroke patients testing novel VR interventions [23-25]. Conversely, the therapists stated experiencing the games as being no relief of responsibility for them, as most patients needed support while playing. They further reported that some patients had difficulty understanding the purpose of the games, most likely due to the presence of anosognosia. It is, nevertheless, important to make sure that patients understand game purpose, so as to meet their expectations and to avoid frustration [23].

Most of our patients did however experience a decreasing enthusiasm during the 3-week intervention and started to perceive the activity as boring, even though games regularly and individually progressed and were designed according to therapeutic principles [26]. Reasons for this decreasing enthusiasm other than suboptimal balance of challenge versus success might be the time point of the intervention and the lack of feedback in the achieved game scores. It is perhaps understandable that the exergames - still a test version - fell behind other established VR therapy options available during active rehabilitation, such as for example the Amadeo Robotic System for improving hand motor function (Tyromotion GmbH, Austria) [27].

The lack of feedback experienced by fitter patients, combined with their perceptions of being insufficiently challenged while playing the exergames, might be reasons why they indicated preference for conventional therapy methods over this VR intervention. Furthermore, the therapeutic benefit may not be as obvious as during conventional therapy due to the nature of play that is inherent to games [23]. The REWIRE game design was rated as having limited appeal by both therapists and patients. The
therapists in particular wished to have games that would be self-adaptive to patient progress so as to reduce time for support. Although some gait and balance REWIRE exergames had successfully demonstrated real-time adaptation to modify the gameplay in healthy subjects [28], this option was not yet available for the neglect games at the time-point of our study. Game modifications were tested according both to the patient’s current performance and progress and to the exercise plan specified by the therapist [28].

When supporting patients to play the REWIRE exergames, some therapists expressed uncertainty in achieving carryover effects into real-life tasks. The therapists indicated measuring such effects using the standardized Canadian Occupational Performance Measure (COPM), a client-centered instrument examining self-perceived changes in occupational performance [29]. This uncertainty was one of the reasons why they would have preferred to use time for the training of real ADL rather than game play to achieve the rehabilitation goals set for their patients. Another reason why therapists preferred using rehabilitation time for conventional therapy rather than testing of the novel VR intervention was the software’s susceptibility to breakdown. Unsurprisingly, such technology limitations are associated with negative user feedback, as was the case in our sample. Those limitations were also determinative of participants’ low ratings for limited future intentions to use exergames.

Providing a more comprehensive picture of the consequences of stroke

Qualitative research has an important role in health care, representing the human dimension by exploring the lived experiences of individuals [30]. Therefore, the use of qualitative – as well as quantitative – research provides a more comprehensive picture of the consequences of stroke, grounded in the experiences of affected persons [31]. This expanded understanding of stroke is essential in the development of interventions that are relevant to the individual [32].

In this respect, CHAPTER 7 used a qualitative study design to explore the everyday life experiences of persons with right hemispheric stroke and unilateral spatial neglect symptoms who were in the later stages of inpatient rehabilitation [33]. Semi-structured individual interviews were conducted, then analyzed using thematic analysis [34]. Participants’ experiences were captured in three over-arching themes: unawareness of neglect, emergent awareness for neglect-related difficulties, and comparing the new life with the old one. The findings indicated that all participants reported being unaware
of their neglect during the early stage post-stroke. They expressed this unawareness in perceived mismatches between personal capacities and external demands, causing feelings of incomprehension and confusion. During the course of rehabilitation, most participants experienced an emerging awareness for their neglect symptoms. They described this experience as a stepwise process from searching for a reason behind experienced difficulties, to implementing coping strategies for mastering everyday life. Furthermore, participants narrated constantly comparing their current life situation with that experienced before their stroke.

The study findings supplement existing literature by providing insiders’ perspectives on the rehabilitation experiences of stroke patients with emerging levels of awareness in the later stages of inpatient rehabilitation. This knowledge might reduce perceived barriers to adoption of best practice in the management of unilateral spatial neglect [35]. Furthermore, it might enhance the therapy process by, for example, facilitating the planning and implementation of therapy in accordance with the patient’s stage of recovery, needs, and preferences [36, 37].
Limitations and methodological considerations

Study sample

Sample size

Setting an appropriate sample size for quantitative studies is important. If a study is too large, it may be judged to be unethical, as participants may be unnecessarily exposed to risks, or unnecessary costs may be incurred. If a study aims to detect clinically important effects, however, it will be unable to do so if the sample size is too small [38].

In this thesis, the two methodological studies testing the VPIT [11] and ZüMAX [12] assessments had rather small sample sizes. The guidelines by Kottner et al. [39] would consider a sample size of 50 as adequate for our purpose. We were, however, unable to recruit more than 31 patients for the VPIT study [11], and more than 33 stroke survivors together with 35 healthy controls for the ZüMAX study [12]. The main reason was due to difficulties in accessing more patients within a reasonable time in the occupational therapy practices that recruited patients for us. We were, therefore, aware that our relatively small samples might have affected the values of reproducibility, reliability and measurement error of the study results, influencing statistical analysis and interpretation. Therefore, in order to assess the magnitude of such errors, we recommend calculating a reliability coefficient for future methodological studies [40]. If possible, trying to increase the amount of recruitment places, together with the time span dedicated for recruitment, would be advantageous. Our study results can nonetheless be used to calculate appropriate sample sizes for future studies. Furthermore, we think that our samples were of reasonable size in this first attempt to evaluate the general usefulness of both assessments in chronic stroke patients.

Heterogeneity

Methodological heterogeneity impedes comparison among studies [41]. This was one of the reasons why we had to forego performing a meta-analysis for our systematic review, as the methodological diversity among the found RCTs was large [4]. Heterogeneity within one study sample, however, complicates straightforward testing and conclusion making for this group. Lowering heterogeneity can, thus, remove potential confounders and increase sensitivity for the measured outcome. When conducting research with stroke survivors, however, it is very difficult to recruit a homogenous sample due both to diverse pathophysiology (e.g., cerebral ischemia or
hemorrhage), and diversity in patients’ pre-stroke health condition. Therefore, high heterogeneity is often more realistic for modeling the real world, and may identify (unexpected) subpopulations, like it was the case in the VPIT stroke sample, where we had to split the study group in two subgroups [11].

To minimize heterogeneity between experimental and control groups (e.g., in the ZüMAX study) or within one group (e.g. in the feasibility study regarding neglect severity), stricter inclusion/exclusion criteria could have been used. Stricter criteria could distinguish between survivors of a first-ever and recurrent stroke, or among patients with different subtypes and severity of symptoms. With stricter criteria, our study samples could have been stratified by age group, education level, or time since stroke onset, whilst pre-stroke cognitive decline could also have been considered [42]. However, accessing and assessing the ZüMAX study patients’ medical history was difficult, because some of them had suffered their stroke many years earlier. Most patient charts were either inaccessibly stored in the GP surgeries, or were generally untraceable. Furthermore, applying stricter inclusion criteria would have further aggravated difficulties in finding suitable participants and in sufficient numbers for the studies presented in this thesis. The recruitment of stroke patients for our feasibility study in a clinical setting, for example, which required people fit enough to test the game-based VR intervention, was already quite difficult with the existing in- and exclusion criteria [19, 22]. Conversely, testing such an intervention in a later, chronic stage, where most patients are in a better health condition, might have been easier, but would have increased the risk of a ceiling effect (see chapter “ceiling effect”).

In qualitative research, however, heterogeneity of a sample is eminently preferable, as it facilitates exploration of the diversity of experiences with the phenomenon of interest. Therefore, having too homogenous a study sample implies the risk of limited specificity of knowledge and experiences within the group [43]. In our qualitative study [33], we had a small but diverse group of participants thanks to their differences in age, gender and social status.

Selection bias

In the ZüMAX study [12], we identified a selection bias between the stroke and control groups, the latter being much more educated than the former. It is well known that education has a major effect on cognitive assessment scores [44]. We were however unable to correct for these education differences. This was for two main reasons.
Firstly, our older participants in particular could not remember exactly how many years of education they had received, and we were reluctant to include estimations within our work. Secondly, most participants, especially those with a university degree or apprenticeship, shared an attitude of life-long learning and had attended several continuing education programs. This made it impossible for us to determine a true cutoff for years of education. To address this problem, we recommend considering education in the inclusion criteria of future studies, and to correct the ZüMAX for education effects. This could be achieved, for example, by adding one point to the total score of participants who have at most 12 years of education, as is done with the Montreal Cognitive Assessment (MoCA) \[45, 46\].

**Ceiling effect(s)**

A test is considered to have floor or ceiling effects if >15% of the respondents achieve the minimum or maximum score [13]. When we checked for such effects in the ZüMAX study [12], we found ceiling effects in all three cognitive domains with non-significant group differences (praxia, visual perception and construction, learning and memory). Such ceiling effects make it impossible to distinguish among participants who had perfect scores, indicating limited validity. It is possible that by the time we recruited our sample of chronic stroke survivors, they had already recovered almost completely in those domains and thus reached near-perfect scores, creating this ceiling effect. In contrast, no ceiling effects were to be found in the feasibility study, where we recruited patients in the sub-acute stage post-stroke [19]. Nonetheless, it seemed important for the ZüMAX to have a maximum of 6 points for each of the five relevant neuropsychological domains, resulting in a total score of 30 points. This would allow comparison with other cognitive screening instruments, especially the Mini Mental State Examination (MMSE) and MoCA, both of which also have maximum scores of 30. Before this scoring is made final for the ZüMAX, however, the scale should be examined in larger samples of stroke survivors and healthy controls, correcting for possible confounding factors such as cognitive status, to find out if ceiling effects remain.

**Spontaneous recovery**

Most spontaneous recovery from motor impairment and perceptual neglect tends to occur in the first 3 months post-stroke [47]. Spontaneous cognitive recovery, however,
is more likely to take place beyond the first 3 months after stroke onset [20]. Mildly affected stroke survivors usually recover quicker than severely affected ones, while patients with anosognosia have a divergent rate of recovery, some continuing to improve for up to a year after stroke [48]. Because of spontaneous recovery, the group trend for improvement, which we found in cognitive and spatial exploration skills following exergames intervention, cannot be exclusively attributed to the intervention [19]. In order to maximize confidence that changes in outcomes can be attributed to the exergames intervention, a control group in a pilot randomized controlled trial design would be needed.

**Methodological considerations**

**Selection of outcome measurements**

While conducting the VPIT and usability study for this thesis, participants were observed to have difficulties in performing some of the used outcome measurements. In the VPIT study, for example, we realized that most (20/31) of the stroke patients needed assistance from their non-affected arm when using the control handle. We had to decide whether to exclude all of them from the study, or allow adaptation in test performance. As we decided on the latter (aiming to open the use of the VPIT for a motor-impairment-related weaker stroke population), we introduced a performance bias in our study, as the test performance was no longer identical within the sample. We accounted for this bias by conducting a subgroup analysis (see **CHAPTER 3** of this thesis for more detail).

In the usability study of this thesis (**CHAPTER 6**), most of the stroke patients needed assistance for completion of the Technology Acceptance Model (TAM) questionnaire, either in retention of the paper-based questionnaire upon the table (due to hemiparesis), or in reading the questions (due to VSN symptoms). Both of these issues may have influenced participant responses. Ideas that would allow independent questionnaire completion include (i) creating a touch-screen version on a tablet fixed on a table to avoid side slipping for those stroke patients who suffer from hemiparesis; and (ii) providing a button on the right margin of the tablet display for audio-play of the questions, making reading of the questions unnecessary.
**Composition of focus groups**

Focus groups usually consist of a number of participants who have been selected because they are in some respect homogenous [49]. In our focus groups, the participants were therapists from the same team who already knew each other well. This fact allowed them to interact as members of a so-called “aggregated group”; i.e. a group of individuals sharing some general social features (their profession) and experiences (the use of exergames and clinical experiences) [49]. Therefore, a common ground was already established, allowing a permissive, nonthreatening environment for group discussion. Conversely, the fact that they knew each other might have negatively influenced some of their interactions and utterances during the interview. For example, in both interview groups, the team leader was also present, which might have inhibited some participants in expressing what they really thought about the exergames. To address this, we chose focus group illustration maps for data analysis [50], allowing a precise summary of the group statements without exposing someone through use of direct quotes.

**Feasibility design**

During the publication review process for our feasibility study, a recurring criticism was the lack of a control group. This criticism might have resulted from the diversity of use of the term “feasibility” and from a mixture of the two terms “feasibility” and “pilot”. While feasibility studies are pieces of research done before a main study, pilot studies are usually a miniature version of the main study [51]. The former is used to estimate important parameters that are needed to design the main study, for example, the number of eligible participants or adherence rate. The latter focuses on the processes of the main study, for example, randomization procedures. As we aimed to evaluate - among other parameters - the adherence to our exergames intervention, the term “feasibility” was appropriate for our purpose. To complicate things further, feasibility studies can have several areas of focus, including but not limited to acceptability, implementation, integration and limited-efficacy testing of the intervention of interest [18]. Thus, feasibility studies can look at how individuals react to the intervention, or they can assess the level of system change needed to integrate a new intervention into an existing infrastructure. Consequently, community constraints in feasibility studies are prioritized over optimal conditions and settings, as is required in highly controlled efficacy trials [18]. The advantage is that feasibility studies usually test the
fit of novel interventions in real-world settings, allowing an evidence-based
determination of whether interventions truly work in the intended environment and
whether they are accepted by their intended users. Depending on the results, a
subsequent controlled (pilot) study is justified to find out if the observed changes in
outcome measures are solely due to the tested intervention.

**Interviewing participants with spatial neglect**

Interviewing participants with cognitive disorders is a challenge. Stroke patients, for
example, may have difficulties describing their life situation because of cognitive
problems or aphasia [52]. In our sample, participants who had sustained a stroke
additionally had symptoms of spatial neglect, a cognitive disorder having an attentional
and perceptual basis [53]. Furthermore, the presence of anosognosia, a “lack of
awareness (insight) or underestimation of a specific loss in sensory, perceptual, motor,
affective or cognitive functions caused by a brain lesion” [54], added an additional
challenge to the interview situation. Unsurprisingly, during thematic analysis within our
qualitative study, it was somewhat challenging to distinguish between participants’
neglect-specific experiences and other recounted impairments, for example those due
to the hemiparesis. Although most participants had good levels of self-awareness, with
little or no signs of anosognosia, getting them to reflect on the specific consequences
of spatial neglect was somewhat difficult [33]. Using a combination of data collection
methods, for example, audio- and video-recording the interview, or combining these
with participation observation [55], would have facilitated data analysis and further
strengthened credibility of the findings.

**Role of the researcher**

A research interview is a social event between two or more people [56]. Therefore, the
role of the researcher is crucial when gathering qualitative data. It is the researcher’s
responsibility to establish an unthreatening environment and good dialogue climate. If
the informant feels neither comfortable nor fully understood, he/she will not open
his/her life-world for the researcher [56]. The fact that the thesis author was already
known to participants, due to the prior usability and feasibility studies, further facilitated
individual and focus group interviews. Conversely, it might be that the therapists
participating in the focus groups did not contribute completely honest opinions about
the usability of the exergames in order not to alienate the researcher. However,
speaking the same professional language – the thesis author and most participating therapists were all occupational therapists – might have facilitated the formulation of useful responses regarding patient engagement with exergames. Nevertheless, the relationship between the researcher and informant is one of inequality, as the researcher decides on the topic of conversation and guides through the dialogue, requiring the informant to talk about his/her situation [52]. In structured focus groups, for example, the moderator takes an active role in controlling not only the topic but also the group’s dynamics [49]. One might even call an interview situation “entrusting oneself to someone else without receiving the same trust in return” [52], highlighting the informant’s vulnerability. Overall therefore, understanding of the participants’ disease, plus sensitivity, plus an ability for the researcher to be open for what the participant is going to tell, are all prerequisites for a good quality of both the interviews and subsequent data analysis.

Technology

REWIRE IGER system

The Intelligent Game Engine for Rehabilitation (IGER) system was designed by engineers of the REWIRE research group. It was developed together with clinicians to specifically support rehabilitation at home through gaming [57]. IGER comprises the game engine, supplies all the typical gaming functionalities including a virtual therapist, has a game control mechanism, and consists of several haptic devices for the training of UE or lower extremity (LE) functions, respectively [26]. While testing the feasibility and usability of the nine exergames designed to treat neglect symptoms, the breakdown susceptibility of software was a major critique point raised by most therapists. The latter reported experiencing loss of valuable therapy time, as they often had to re-boot the system, keeping patients waiting while the computer restarted. Consequently, the exergames were rated as somewhat impractical, although therapists recognized that they were testing a preproduction model. Furthermore, the game design was perceived as unappealing by both patients and therapists, in comparison with long-established commercial market leaders (e.g., Japanese Nintendo Wii technology [58]). Implementing this user critique is imperative for future use of these exergames, such as the use at the patient’s home.
Devices

In this thesis, we tested three devices, namely (i) the PHANTOM Omni as performance control of the VPIT assessment [59], (ii) the Novint Falcon as game control to play the exergames [60], and (iii) an infrared eye tracking camera to perform the Eye Tracker Neglect Test (ETNT) [19]. All three devices are commercial, low-cost and transportable, which we adapted and used for research and rehabilitation purposes in this thesis.

The PHANTOM Omni (SensAble Technologies, Inc., USA) device includes a stylus that has two switches which can be pressed to perform actions in the virtual environment, such as selecting virtual objects [59] (Figure 1). In order to measure grasping force, a new handle was designed at the EPFL in Switzerland and placed on the stylus. Despite its spherical shape, which is easier to grasp and manipulate, many stroke survivors – including ones in the later stages of recovery – were unable to control the handle with their affected hand. Specifically, release of the on-screen virtual pegs with the affected hand, whilst still maintaining a grasp of the handle, was difficult for our participants. This unexpected observation forced us to introduce a performance bias in our test procedures (see chapter “selection of outcome measurements”). Future adjustments in the force threshold required from 2 to 3 or 4 Newtons might facilitate the distinction between grasp and release of the virtual pegs.

Figure 1. The PHANTOM Omni device

The Falcon Novint (Novint Technologies, Inc., USA) is originally built for gaming purposes (Figure 2). It is a haptic 3-D touch device intended to replace the mouse in
video games. It has removable handles the user holds onto to control the Falcon. The grips are movable in three dimensions (right-left, forwards-backwards, up-down). The Falcon’s software creates forces users can feel while playing [60].

To play the neglect exergames, a spherical handle was used without application of any force feedback in order to allow full concentration on the cognitive aspects of the exergames. Force feedback, however, was applied in some of the easy exergames levels. The aim was to support users in finding virtual objects they were unable to find within a certain time by pulling the handle towards these. The haptic feedback was implemented in such a way as to simulate grasp and release of objects by vibrating when touching, for example, a virtual apple.

Some of the stroke patients experienced difficulties using this device instead of the mouse; for example, it sometimes tended to slip on the table while playing, and the handle sometimes jammed. A major criticism, however, was the discrepancies in collision detection which were experienced by users [61]. The patients reported that the system did not always recognize the overlap of the virtual hand and objects. A possible future approach to this limitation might be to ensure that the circular collision area is always at the same place of the virtual hand for all games (e.g., within the palm) and does not vary from game to game (e.g. once at the index finger and then within the palm).

Figure 2. Falcon Novint device and game “Sequence Builder”
The Eye Tribe Tracker camera (Eye Tribe, Copenhagen, Denmark) technology relies on infrared illumination so that it does not interfere with the virtual scenario (Figure 3). Prior to performing the ETNT, an initial calibration of the camera is necessary, as each person has different eye characteristics. For this, the user has to track a circular target (red circle) that is displayed briefly at different locations on the screen on a blank background (Figure 4). A minimum of nine calibration locations covering most of the screen is recommended. For those patients with VSN symptoms, however, discovering the three locations on the left margin of the screen was difficult, limiting the calibration procedure’s accuracy and, thus, the subsequent ETNT performance. In future, the calibration procedure of the camera might be improved by, for example, reducing the speed of the circular target to be followed, or by including visual feedback while following the red circle.

**Figure 3.** Eye Tribe Tracker camera and ETNT.

**Figure 4.** Calibration procedures of the ETNT.
Implications for future research

Study sample
Future studies may strive for bigger and more homogeneous study samples for further evaluation of the VPIT and ZüMAX assessments and the neglect exergames. The ZüMAX and exergames could further be tested in other phases of stroke recovery, for example, in the acute phase of stroke. Stratifying their clinical utility according to the acute, sub-acute and chronic stage may help clinicians to conduct (technological) interventions and assessments relative to stroke recovery, as is recommended [62]. Use of the ZüMAX could also be tested in an entirely different population, such as people with traumatic brain injury. The ETNT could be evaluated for stroke patients who have hemianopia but who are without VSN symptoms. The VPIT should be re-evaluated in a more homogenous study sample regarding stroke severity, including a screening of motor capacity of the affected arm prior to study inclusion. Future studies could investigate the optimum duration and patterning of training exposure to virtual environments. Given the uncertainty about optimal exposure [63], together with our small and heterogeneous feasibility study sample, future trials should focus on investigating optimal thresholds of exposure time corresponding to the degree of VSN severity.

Methodological considerations
Further research is necessary to evaluate discriminant validity of the VPIT by comparing the performance of stroke participants with those of healthy controls. A randomisation of the administration of tests (NHPT, BBT, VPIT) is recommended, to reduce patient tiredness and increase validity. For further development of the ZüMAX, cutoff scores should be established to increase its clinical utility. It would also be valuable to evaluate the ZüMAX’s ecological validity to investigate the tool’s ability to predict patients’ everyday cognitive function. To extend its reach, the ZüMAX should be digitized and integrated into a virtual exercise program for stroke survivors, as was planned in the REWIRE consortium. Designing the ZüMAX as a home-based digital assessment would allow the clinician to monitor the patient’s progress. To increase awareness of the ZüMAX among clinicians, the tool should be translated and tested in other languages.
After demonstration of feasibility of the exergames in a clinical environment [19], the next step is to test them in patient homes, to evaluate adherence, safety, and attrition when using the system in a domestic setting. Future studies with a focus on treatment effects using controlled research designs could be used to assess causal relationships between the game-based intervention and pertinent patient outcomes. For the ETNT, collecting search strategy patterns of healthy persons of different age could be of interest to create norms. Additionally, future work could explore the correlation between ETNT measures and scores in standardized clinical scales, such as the Neglect Test (NET) scores, in order to validate the derived ETNT measures of recovery after VSN. Future work could examine if the frequency and time of game play, and the provision of a predefined set of exergames rather than a free choice of game selection, influences user perspectives. Furthermore, evidence is needed to explore possible carryover effects of such VR interventions into real life in order to enhance acceptance among service users.

Future qualitative studies with stroke patients with VSN symptoms could consider applying a longitudinal design. Conducting at least two interviews per participant, for example during different points in time of inpatient rehabilitation, might give a deeper understanding of the phenomenon under study (e.g. rehabilitation experiences). Follow-up interviews might give insight into how stroke patients’ lives have developed over time, and generate knowledge regarding anosognosia’s persisting effects. If adopting a longitudinal design, a narrative approach might be a suitable method of data analysis, focusing on change over time [64].

**Technology**

To increase VPIT ease of use, setting time limits for the test duration for motor impaired weaker stroke patients could be considered, together perhaps with use of an adjustable armrest for those needing additional arm support (e.g., the Armon Elemento [65]). Such an armrest might improve body stability and therefore test performance, by providing repeatable conditions from trial to trial and patient to patient. Before testing the neglect exergames in future trials, the game design could further be modified according to the degree of realism of the virtual environments, together with an enhanced flexibility of the software, for example, by creating a function to save the chosen difficulty level for each game. A progression as measured by the game scores could be implemented together with immediate graphical feedback after each training.
session to enhance game-playing motivation. Furthermore, implementing an option for game progression from the right to the left side of the screen would allow recruiting stroke patients with a left-sided brain lesion and VSN symptoms. Additionally, the calibration procedure of the Eye Tribe Tracker camera should be adapted to better meet the needs of stroke patients with VSN symptoms by, for example, reducing the speed of the circular target that has to be followed during this process.
Implications for clinical practice

Evidence for the use of VR technologies in the rehabilitation of stroke patients is rapidly growing with promising results [66-69]. This development, however, could potentially intimidate clinicians, creating doubts about the use of conventional therapy methods as being “out-of-fashion”, or generating fear of clinical replacement by automated therapy systems. Before presenting possible implications for clinical practice, therefore, it is important for the author of this thesis to highlight that such VR technologies should be valued as an augmentation of rather than a substitute for existing patient-centered rehabilitation programs in the clinic and the home.

Using VR technology to evaluate stroke impairment, such as the VPIT’s use in assessing UE function, provides objective and comprehensive measurements of cognitive and motor function. Receiving feedback concerning, for example, the patient’s ability to control force or movement coordination is essential for rehabilitation. Furthermore, getting information about movement smoothness offers clinically relevant information, as it has been shown to be a good indicator of UE coordination and stroke recovery [70]. Consequently, clinicians can implement this information in their treatment, and then reassess progress by, for example, repeating the VPIT.

All VR devices presented in this thesis are of compact and transportable form, permitting easy administration even in the patient’s home. Together with their relatively low cost and limited space requirements, these devices could potentially be used by clinicians working across different settings. Wider use of the neglect exergames in the clinical setting could support more active patients by providing augmented and – to a certain degree – independent rehabilitation. Furthermore, as all results are stored in a computer, clinicians can discuss performance with their patients, and individual adjustments can be made to the therapy program as required.

Irrespective of the clinician’s decision to use either VR or conventional methods in the rehabilitation of stroke, the application of the principles of exercise training is crucial for ensuring that the dose and type of exercise fits each patient’s current fitness level. Therefore, knowledge about these principles is important for providing rehabilitation according to each patient’s specific stage of recovery post-stroke. Using the FITT components may additionally facilitate therapy planning according to the patient’s needs and bodily capacities, and allow a regular check-up of the accuracy of the intervention.
Improvements in cognition as well as motor performance are crucial to post-stroke recovery. Using valid and reliable assessments, such as the ZüMAX, are therefore potentially helpful for enabling clinicians to gain an overview of the cognitive state of the patient. Subsequently, therapy interventions to train cognitive skills can be planned accordingly and re-evaluated on a regular basis.

Setting client-centered goals for stroke rehabilitation facilitates patient participation and motivation. Having knowledge regarding patients’ personal experiences during inpatient rehabilitation is crucial to meeting their personal needs and for facilitating formulation of realistic and achievable therapy goals. For clinicians, therefore, it is worth taking the time to talk in-depth to patients to build a collaborative professional relationship, and to empower service users to take an active part in their recovery process.
General conclusion

This thesis provides knowledge for the field of stroke rehabilitation. It presents study findings for the application of exercise training principles, for the suitability of novel (VR) assessments for this population, for the feasibility and usability of newly designed exergames to rehabilitate VSN symptoms, together with personal views of stroke survivors regarding their experiences of everyday life as they near the end of inpatient rehabilitation.

The incomplete and inconsistent reporting of prescription of, and especially adherence to, exercise prescriptions, seen when investigating training programs within existing RCTs, underlined the importance for researchers to plan and document exercise interventions in as detailed a manner as possible. As a result we therefore paid special attention to the detailed description of our planned intervention in the subsequent studies. The high adherence rate and absence of adverse events showed the feasibility and safety of our novel game-based exergames intervention for VSN patients early post-stroke. Additionally, the ETNT has shown to be a promising assessment for quantifying spatial exploration skills. The results of the usability study showed that all users perceived the exergames as user-friendly, though they would not necessarily entertain their use in their current format. We are nevertheless enthusiastic about the future and the role of VR applications in stroke rehabilitation. The results of this thesis showed that the VR intervention must be sufficiently challenging, progressive and should be adaptable, so that patients will continue to show improvements in their stroke-related deficits. Most important, the VR intervention must be sufficiently interesting and meaningful to engage the user and to keep motivation for training high. Both of the stroke assessments evaluated in this thesis showed promising results: the VPIT for measuring UE function, and the ZüMAX for measuring cognition. The fair test-retest reliability results showed that most parameters of both assessments remained consistent when performed by stroke patients and were sensitive to change, allowing diagnostic and therapeutic use in clinical practice for this target group. The low concurrent validity of the VPIT indicated that performing this hapto-virtual reality assessment required components of UE motor performance other than those in conventional tests. We showed that the ZüMAX might discriminate chronic stroke survivors from healthy controls. Our qualitative study provided insight into the stepwise process from initial unawareness to emergent awareness for neglect-related difficulties over the course of rehabilitation, seen from patients' perspectives.
Comparing their current life situation with the one before their stroke helped them to adapt to perceived changes. The study findings highlight the importance of providing a client-centered treatment to meet the patients’ needs during inpatient rehabilitation.
References


Summary

This thesis examined two virtual reality (VR) tools, which had been developed to assess and treat motor and cognitive impairment in sub-acute and chronic stroke patients with and without visuo-spatial neglect (VSN) symptoms. The two tools are (i) a series of nine exergames – computer games based on therapeutic principles – aimed at improving exploration of the neglected space, and (ii) the Virtual Peg Insertion Test (VPIT). It further tested a novel cognitive tool, the Zürich maxi mental status inventory (ZüMAX), and assessed the users’ perspectives on these exergames, together with patients’ experiences on how it is to live with a stroke during inpatient rehabilitation.

The primary objectives of the thesis were

1) to determine the feasibility of usage of exergames, with minimal therapist supervision, in terms of implementation of the intervention, including adherence, attrition, and safety, and limited efficacy testing, also aiming to evaluate possible effects on VSN symptoms in patient users (feasibility study).

2) to assess users’ perspectives (patients and therapists) when using the exergames for rehabilitation of VSN symptoms. Specifically, we wanted to evaluate perceived user-friendliness of the exergames, attitudes towards using them, plus intention to continue to use them in the future (usability study).

Secondary objectives of this thesis were

3) to gain knowledge about the quality and level of reporting of exercise training principles used in the rehabilitation of the patient group, plus patient adherence to the prescribed intervention (systematic review).

4) to test psychometric properties of a novel VR-based assessment to quantify upper extremity (UE) function in patients with chronic stroke (the VPIT) and of a novel cognitive paper-and-pencil tool to screen for disturbances in neuropsychological function (the ZüMAX) aiming to evaluate their clinical utility (methodological studies).

5) to explore everyday life experiences of persons with right hemispheric stroke and spatial neglect symptoms near the end of inpatient rehabilitation (qualitative study).

CHAPTER 1 presents the prevalence, pathophysiology and clinical features of stroke. It describes different consequences of a stroke event, including motor, cognitive and
psychosocial impairments. It provides an overview of the current state of the art of stroke rehabilitation and adherence to treatment. Finally, it introduces VR as a novel treatment and assessment tool for stroke rehabilitation.

Stroke rehabilitation is a dynamic process with the overall aim of reducing stroke-related disability. The current evidence base for the treatment of UE sensory-motor impairments is moderate. Interventions targeting improvements of gait- and mobility-related functions and activities have proven effectiveness. There is currently insufficient quality and quantity of evidence for the effectiveness of rehabilitation for cognitive impairments post-stroke to support clear recommendations for clinical practice. Furthermore, the effectiveness of rehabilitation interventions for reducing neglect symptoms remains largely unknown. VR training seems to be among the most effective and promising treatment methods to rehabilitate neglect.

Overall, further research is needed to advance this promising field, specifically, to provide further psychometric properties analysis of VR-based stroke assessments and to develop and test novel VR-interventions for feasibility of use with this target group. Furthermore, assessing users’ perspectives with this novel therapy method is important for maximizing acceptability of such rehabilitation methods. Gaining knowledge from personal experiences of users who have had a stroke is deemed crucial for designing custom-made VR-interventions. Therefore, the following thesis chapters (CHAPTER 2 to 7) aim to help filling these gaps of knowledge and thus contribute to an enhanced stroke rehabilitation.

**CHAPTER 2** comprises a systematic review regarding the application of exercise training in sub-acute and chronic stroke survivors. Their application enables the planning and implementation of custom-made interventions at an adequate dose, thus facilitates reaching the desired training effects for stroke survivors.

The results indicated incomplete and inconsistent reporting of the intervention components, underlying exercise training principles and patient adherence. The broad variation in the methodological quality of the included randomized controlled trials (RCTs) restricted further synthesis of the results. Furthermore, the results impact on the clinical reproducibility of trials, as health professionals cannot be confident whether non-significant findings are due to lack of efficacy or occur through limitations in treatment prescription. Therefore, this review comprises a call for transparency in the prescription of and adherence to treatment, to facilitate evidence-based practice.
CHAPTER 3 investigates for the first time the concurrent validity and test-retest reliability of the novel VPIT in patients with chronic stroke. The VPIT is designed to quantify UE motor function. The use of VR assessments are useful for precise and objective assessment of different functions affected by stroke, such as motor function and cognition. VR assessments present an effective alternative to conventional measurements, such as the Box and Block Test (BBT) or Nine Hole Peg Test (NHPT), which respectively measure gross and fine manual dexterity. The low concurrent validity demonstrated that the VPIT is different to conventional assessments and, thus, cannot be used as a substitute. Rather, the VPIT measures other components of UE motor performance to the NHPT and BBT, such as for example movement coordination or expended force control. The good test-retest reliability results are promising, thus suggesting diagnostic and therapeutic usefulness for the VPIT in clinical practice.

CHAPTER 4 introduces a novel, domain-specific cognitive assessment tool to screen for disturbances in neuropsychological function, the ZüMAX. Because evidence for this tool’s validity and reliability was lacking, thus limiting its clinical utility, first time examination of its discriminant validity and test-retest reliability in stroke patients was undertaken. The ZüMAX demonstrated reasonable discriminatory abilities between healthy persons and chronic stroke survivors on three of five subscales and showed moderate-to-good test-retest reliability. Therefore, the ZüMAX showed promise in measuring neuropsychological disturbances in stroke survivors. Accordingly, we used this assessment tool for our feasibility study (CHAPTER 5) to measure change in cognition post-stroke.

CHAPTER 5 determines the feasibility of a novel series of nine exergames in terms of implementation of the intervention, including adherence, attrition and safety, and limited efficacy testing in a case series of patients with VSN symptoms early post-stroke. The games aimed to improve exploration of the neglected space, were based on the Gentile’s Taxonomy, and designed according to therapeutic principles. Seven patients attended the 3-week exergames training program on a daily basis during their inpatient rehabilitation. Additionally, we developed an Eye Tracker Neglect Test
(ETNT) using an infrared camera to detect and measure neglect symptoms pre- and post-intervention.

The results showed that these exergames were safe and feasible for stroke patients with VSN symptoms. The preliminary results indicate the potential for these exergames to effect improvements in cognitive and spatial exploration skills over the course of training. Further adaptations are needed for the ETNT, especially regarding calibration issues, before its use can be justified in a larger study sample.

**CHAPTER 6** describes the user perspectives on this series of exergames tested for their feasibility in **CHAPTER 5**. We adopted a mixed methods design characterized by an initial gathering of quantitative data followed by a collection and analysis of qualitative statements. The purpose of this strategy was to use the qualitative results to assist in explaining and interpreting the findings of the quantitative data. Therefore, after initial evaluation of an abridged version of the Technology Acceptance Model (TAM) questionnaire, we conducted individual interviews with the patients and performed focus groups with the therapists. The latter had assisted the patients during the testing phase of the exergames.

The results showed that all users perceived the exergames as user-friendly. Nevertheless, they would not necessarily entertain their use in the current format. The general attitude toward using the exergames was more positive among patients than among therapists. These findings are encouraging on the one hand, as they show that even patients with VSN symptoms can enjoy a game-based VR-intervention. On the other hand, findings also highlighted the importance of therapists’ involvement in order to achieve acceptable game-based rehabilitation methods.

**CHAPTER 7** explores how persons with right hemispheric stroke and spatial neglect symptoms experience their everyday life near the end of inpatient rehabilitation. For this qualitative study, the author of this thesis conducted in-depth semi-structured interviews with the same patients as in **CHAPTER 5** and 6. Interviews addressed patients’ experiences of performance of daily activities, of body and of self since stroke onset. Results were analysed using thematic analysis.

The participants’ described their experiences of everyday life post-stroke in terms of three main themes: “unawareness of neglect”, “emergent awareness for neglect-related difficulties”, and “comparing the new life with the old one”. The findings showed
that all participants suffered from unawareness of having a neglect during the early stage of recovery. During the course of rehabilitation, most participants experienced an emerging awareness for their neglect symptoms, and constantly compared their current life situation with the one before their stroke.

The study findings highlight that being aware of neglect-related deficits is a prerequisite for using coping strategies and incorporating them in daily life. They further inform health professionals regarding personal experiences of orientation in and reorganization of life in patients with right hemispheric stroke and spatial neglect. Health professionals should provide a multimodal, client-centered treatment, integrate the stroke survivors’ families in the rehabilitation process, and maintain a collaborative relationship and open communication style with patients in order to best meet their needs and expectations.

CHAPTER 8 discusses the main findings of CHAPTER 2 to 7, examines limitations and methodological considerations of these studies, and recommends implications for clinical practice and future research. It closes with a general conclusion for the thesis. The findings highlight the importance of complete and replicable reporting of study interventions, together with the adherence of the patients to the prescribed training. The application of exercise training principles can facilitate reporting and enhance accuracy of dose and type for the planned intervention. Furthermore, basing clinical decisions on reliable and valid assessments is of utmost importance for the planning and implementation of an evidence-based stroke rehabilitation. With the VPIT and ZüMAX, this thesis provides two promising novel assessment tools for this target group.

The findings further emphasize that novel game-based stroke interventions should initially be tested for their feasibility and usability in a real-world setting, before large scale controlled trials are justified. In this thesis, we successfully tested a novel series of nine exergames for exploration of the neglected space. Finally, only qualitative research has the potential to provide a more comprehensive picture of the individual consequences of stroke. This knowledge is valuable for health professionals to gain a better understanding of their stroke patients, allowing a client-centered rehabilitation. Main limitations of this thesis were the experienced difficulties with some of the outcome measurements when used for stroke patients with VSN symptoms, and facing technical flaws in the game software and haptic devices. The “implications for future
In conclusion, we have presented two potential new assessment tools for stroke patients without VSN symptoms to assess UE motor function (VPIT) and cognition (ZüMAX), the former being a VR-based assessment tool. We have shown that the exergames are feasible and safe in a case series of stroke patients with VSN symptoms. In this way, we were able to receive valuable users’ feedback from the patients and therapists, which enables further development of the games. We have provided personal insights in the stepwise recovery and adaptation process of stroke survivors during their inpatient rehabilitation. Finally, the presented VR technologies in this thesis have shown to be a viable complement to existing conventional evidence-based stroke rehabilitation programs, warranting further research.
Zusammenfassung

Diese These hat zwei auf virtueller Realität (VR) basierende Instrumente untersucht, die zur Erfassung und Behandlung motorischer und kognitiver Einschränkungen von Schlaganfallpatienten mit oder ohne räumlich-visuellem Neglekt (VSN) während der subakuten und chronischen Phase entwickelt wurden. Die beiden Instrumente sind (i) eine Serie von „Exergames“ - Computerspiele, die auf therapeutischen Prinzipien aufgebaut sind und zum Ziel haben, den vernachlässigten visuellen Raum zu erkunden - und (ii) der „virtuelle Peg Insertions-Test“ (VPIT). Zudem testete die These ein neues kognitives Assessment, den „Zürich maximental Status Test“ (ZüMAX), wertete die Meinungen der Benutzer zu diesen Exergames aus und evaluierte Patientenerfahrungen zum Leben mit einem Schlaganfall während des Aufenthalts in einer stationären Rehabilitationsklinik.

Die Hauptziele der These waren:


2) die Meinungen der Benutzer (Patienten und Therapeuten) zu diesen Exergames zu erfragen. Insbesondere wollten wir wissen, ob die Benutzer die Exergames als benutzerfreundlich empfanden, welches die Grundhaltung zum Gebrauch dieser Exergames ist, und ob sie sich vorstellen könnten, diese in Zukunft weiter einzusetzen (Brauchbarkeitsstudie).

Untergeordnete Ziele der These waren:

3) zusammenzufassen, wie detailliert Trainingsprinzipien für die Rehabilitation dieser Patientengruppe sowie die Einhaltung der geplanten Therapieinterventionen in randomisierten kontrollierten Studien (RCTs) beschrieben sind (systematische Review).

4) die psychometrischen Gütekriterien von zwei Assessments zu untersuchen, um deren klinischen Nutzen zu steigern. Es handelt sich dabei um ein neues VR-basiertes Instrument, das die motorischen Funktionen der oberen Extremität bei
Schlaganfallpatienten in der chronischen Phase misst (VPIT) und um ein neues kognitives Papier-und-Bleistift Assessment, das neuropsychologische Hirnfunktionen überprüft (ZüMAX) (methodologische Studien).

5) alltägliche Erfahrungen gegen Ende des stationären Rehabilitationsaufenthalts aus der Sicht von Schlaganfallpatienten mit linksseitigen Neglekt Symptomen zu dokumentieren (qualitative Studie).

**KAPITEL 1** präsentiert Prävalenz, Pathophysiologie und klinische Eigenschaften eines Schlaganfalls. Es beschreibt verschiedene Folgen eines Schlaganfalls wie motorische, kognitive und psychosoziale Beeinträchtigungen. Es bietet eine Übersicht über den neuesten Stand der Forschung im Gebiet der Schlaganfall-Rehabilitation mit passenden Therapieinterventionen. Abschliessend führt es die VR als neue Behandlungsmethode und Erfassungsinstrument für die Schlaganfall Rehabilitation ein.


KAPITEL 2 beinhaltet eine systematische Review zur Anwendung der Trainingsprinzipien bei Schlaganfallpatienten während der subakuten und chronischen Phase. Deren Anwendung ermöglicht die Planung und Durchführung von massgefertigten Interventionen mit adäquater Dosierung, und vereinfacht somit das Erreichen des gewünschten Trainingseffekts für Schlaganfallpatienten.

Die Resultate zeigen, dass über die einzelnen Interventionskomponenten sowie über die zugrundeliegenden Trainingsprinzipien und die Therapietreue unvollständig und uneinheitlich berichtet wurde. Die grossen Unterschiede in der methodologischen Qualität der eingeschlossenen RCTs limitierten eine weitere Synthese der Resultate. Des Weiteren beeinträchtigen die Resultate die klinische Reproduzierbarkeit der Studien, da Fachkräfte im Gesundheitswesen nicht sicher sein können, ob die nicht-signifikanten Resultate auf einen Mangel an Wirksamkeitsnachweisen zurückzuführen, oder ob sie aufgrund einer ungenügenden Beschreibung der Therapieinterventionen entstanden sind. Daher beinhaltet diese Review einen Aufruf zu mehr Transparenz in der Beschreibung der Therapieinterventionen und der Therapietreue, um eine evidenzbasierte Praxis zu erleichtern.

KAPITEL 3 erforscht erstmals die Übereinstimmungsvalidität und die Test-Retest-Reliabilität des neuen VPIT in einer Gruppe von Schlaganfallpatienten während der chronischen Phase. Der VPIT wurde für die Messung von motorischen Funktionen der oberen Extremität entwickelt. Der Gebrauch von VR-basierten Assessments ist sinnvoll für die präzise und objektive Erfassung von verschiedenen betroffenen Funktionen nach einem Schlaganfall, wie zum Beispiel motorische und kognitive Funktionen. VR-Assessments stellen eine wirksame Alternative zu herkömmlichen Messinstrumenten dar, wie zum Beispiel der Box und Block Test (BBT) oder der Nine Hole Peg Test (NHPT), welche die Grob- und Feinmotorik messen.


KAPITEL 8 diskutiert die Hauptergebnisse aus den KAPITELN 2 bis 7, beleuchtet Grenzen und methodische Überlegungen dieser Studien und gibt Empfehlungen für die klinische Praxis und für weiterführende Forschung. Es endet mit einer allgemeinen Schlussfolgerung zu dieser These.


Die Studienresultate betonen weiter, dass neue spielbasierte Schlaganfallinterventionen zuerst in einer der echten Welt entsprechenden Umgebung auf ihre Durchführbarkeit und Brauchbarkeit getestet werden sollten, bevor gross angelegte, kontrollierte Studien gerechtfertigt sind. In dieser These untersuchten wir erfolgreich eine neue Serie von neun Exergames zur Erforschung des vernachlässigten Raums. Zudem hat nur die qualitative Forschung das Potential, ein umfassendes Bild der individuellen Schlaganfallkonsequenzen aufzuzeigen. Dieses Wissen ist für Gesundheitsfachleute wertvoll für ein besseres Verständnis ihrer Schlaganfallpatienten, was eine klientenzentrierte Rehabilitation ermöglicht.
Hauptlimitationen dieser These waren die erlebten Schwierigkeiten mit einigen der angewandten Messinstrumente bei Schlaganfallpatienten mit VSN-Symptomen, sowie technische Mängel in der Spielsoftware und den haptischen Geräten. Der Abschnitt „Empfehlungen für weiterführende Forschung“ zeigt Ideen auf, wie diese Limitationen in zukünftigen Studien angegangen werden könnten, zusammen mit Ideen für weiterführende Forschung, welche auf den Studienresultaten dieser These basieren.

Valorisation Addendum

This chapter discusses the social and economic relevance of the work presented in this thesis. Potential target groups for whom the novel assessment and treatment methods might be of interest are identified, together with planned activities and products derived from the study results. The innovation of the results is shown, complemented with a schedule for their implementation.

Relevance

This thesis examined different virtual reality (VR) tools to assess and treat motor and cognitive impairment in sub-acute and chronic stroke patients with and without symptoms of visuo-spatial neglect (VSN). Two research projects within this thesis showed the potential of newly developed stroke assessment tools in patients without VSN: the Virtual Peg Insertion Test (VPIT) investigating upper extremity motor function, and the Zürich maxi mental status inventory (ZüMAX) assessing cognition. Subsequent projects have demonstrated that neglect exergames are feasible and safe for stroke patients with VSN symptoms, also providing potentially valuable user feedback from patients and therapists for further future development. Therefore, application of such VR technologies in stroke rehabilitation is promising. Finally, this thesis explored everyday life experiences of persons with right hemispheric stroke and unilateral spatial neglect near the end of inpatient rehabilitation.

Annual health and social care costs are still increasing in countries with advanced economies such as Switzerland. Not surprisingly, the purchasers, such as health care insurances and cantons, are trying to find solutions for addressing this trend. Unfortunately, some of these efforts include staff reductions, which has substantial consequences not only for the health professionals involved, but also for their patients. Scientists and clinicians are, therefore, obligated to find solutions to nevertheless provide the best treatment options for patients. The use of VR technologies as a complement to existing conventional rehabilitation is a viable solution. The neglect exergames of the REWIRE consortium, for example, could provide aspects of continued domestic daily rehabilitation without the presence of a therapist. This provides the patient with individualized and progressive training to improve motor and cognitive skills post-stroke, resulting in greater independence in activities of daily living and, thus, improved quality of life. Furthermore, such enhanced training might possibly
prolong independent living at home, thus reducing or even avoiding further costs of extra care at home and transfer to nursing home care. From therapists’ perspectives, the use of these VR-based exergames provides a novel evidence-based treatment option for those stroke patients who are capable and willing to do more than conventional therapy. Its use allows the therapist to supervise the patient either during regular face-to-face meetings, or even via internet (e.g., Skype). Supervision via internet might avoid extra costs of traveling to the patient’s home, or vice-versa to the therapist’s practice for the patient. Additionally, the use of evidence-based assessment tools, such as the Zürich maxi mental status inventory (ZüMAX) or the Virtual Peg Insertion Test (VPIT) presented in this thesis, provide an overview of the patient’s progress. This is not only more motivating for the patient, but can further be of use for the therapists to prove the purchasers that their therapy is effective.

**Target groups**

The results presented in this thesis are of interest to multiple target groups. They are informative for health care purchasers, as they provide feasible, low-cost therapy options as an adjunct to conventional face-to-face therapy. The results from the users’ perspectives study about newly developed VR-based games for rehabilitation can be of interest for game designers. Such feedback is crucial for the development of effective, user-friendly and client-centered video games. The work of this thesis is also of interest for health care professionals, such as physical or occupational therapists, as it offers them novel and evidence-based ways of treating and assessing stroke patients. Furthermore, the findings inform health care professionals regarding personal experiences of orientation in and reorganization of life of inpatients after a first right hemispheric stroke with unilateral spatial neglect symptoms. Cultivating such an awareness of patient experiences could promote effective strategies for care focused on what is most relevant to the individual during rehabilitation. As the results of this thesis aim to further developing the treatment of stroke symptoms, they are not only of interest for stroke survivors, but also for their relatives. This further development might increase the families’ confidence that efforts are made both to improve rehabilitation, and to enhance understanding of the lived experience of stroke. Finally, other similar patient groups, such as people with a traumatic or non-traumatic brain injury (e.g., brain tumor, hypoxia) could profit from the assessment and treatment options presented in this thesis.
Activities, products and innovations

The newly developed REWIRE neglect exergames for stroke patients can be called innovative, as they were field-tested for the first time in this target group. They have shown to be a potential extension to conventional stroke therapy. Further activities are planned to enhance the use of those exergames; for example, the activation of the competition and networking option among stroke users. The idea is that patients can win notional rewards (‘medals’) for their achievements while playing the games, and can share scores with other users if wanted (competition). The option to create an online social network platform among patient users is another future planned activity, aiming both for enhanced training motivation and for avoidance of the social isolation, which occurs in persons restricted in mobility. To address some of the more specific requirements of the primary target group studied in this thesis – for example, having difficulties with reading – the development of an audio user manual for the exergames would facilitate independent use. Another project is the provision of further education for therapists and other health professionals who are interested in using the REWIRE platform and exergames with their patients. However, several steps need to be taken before the platform and its games would be ready for commercial and broad use within rehabilitation clinics and practices. Until then, the results of this thesis can be disseminated at national and international conferences and published in scientific journals.

Schedule for implementation of the results

A possible schedule for the implementation of the thesis results is presented in Figure 1. This schedule has potential for success, albeit with possibilities of failure also. For example, a possible success might be that knowledge from proposed future research may be useful for commercial activities with spin-off ventures. Such cooperation might be guided and supervised by experts in the field of stroke rehabilitation. Moreover, depending on the future research findings, the experts could create new spin-off ventures interested in dissemination and further development of findings. Additionally, the schedule affords the opportunity for future students to gain experience in a promising and evolving new research area. These students can spread the findings among other students, contributing to the growing community of researchers in the area, both within Switzerland and beyond. Potential risks for successful implementation of results might be difficulties in finding third-party funding, together
with the race against time. It might be that the further development and testing of the REWIRE neglect exergames might not keep pace with the fast growing VR-therapy commercial market. A further risk might be the high development costs of a disposable VR-therapy product.

**Figure 1.** Schedule for implementation of the results.

- **Generating third-party funding**
  - Possible stakeholders: Swiss National Fond; stroke associations; rehabilitation clinics
- **Further development of the REWIRE platform and neglect exergames**
  - Collaborators: Game designers of the REWIRE consortium; patients; therapists
- **Doing more research (RCT) dissemination of the findings at conferences**
  - Research assistants: PhD- and master students
  - Collaborators: Rehabilitation clinics
- **Marketing of the REWIRE platform and exergames**
  - Possible customers: Rehabilitation clinics; therapy practices; hospitals
Acknowledgment

This PhD thesis would never have been realized without support and guidance from many different people. I would like to take this opportunity to express my thanks and appreciation to them.

Prof. dr. Rob A. de Bie, my main supervisor, is professor of Physiotherapy Research at the Department of Epidemiology, Maastricht University, the Netherlands. I would like to thank him for his input, feedback and discussions, both during personal meetings at the ZHAW in Winterthur and via email correspondence. His input was always inspiring for me, whether strengthening me in going ahead in my intended PhD trajectory, or when thinking about other solutions and different ways of proceeding.

Both of my local co-supervisors in Switzerland, dr. Ruud H. W. Knols and Prof. dr. Eling D. de Bruin. Ruud is head of the Physiotherapy Occupational Therapy Research Center, Directorate of Research and Education at the University Hospital Zurich, my workplace. Eling is professor of Physiotherapy Research at the Karolinska Institute (Stockholm), Department of Neurobiology, Institute of Physiotherapy and senior lecturer-scientist at the Institute of Human Movement Sciences and Sport, Department of Health Sciences and Technology, ETH Zurich. I would like to warmly thank them both for their scientific and personal advice on my research path. They helped in my transformation from a pure clinician to a professional researcher, which at times was not an easy process. I would also like to thank them both for their patience and support during difficult times, including helping me to think in a broader context.

The European REWIRE project team, specifically Prof. dr. Nunzio A. Borghese, dr. Elif Surer, Renato Mainetti and dr. Michele Pirovano. Without their willingness and openness to share and discuss ideas from different professional points of view, the REWIRE exergames would have never been realized. I would like to thank them for their invaluable support during the game design process, as well as their support in solving technical problems with software during field-testing of the exergames. I also would like to thank them for their feedback during scientific paper preparation.

Dr. eng. Marco Rabuffetti, Polo Tecnologico, Istituto di Ricovero e Cura a Carattere Scientifico Fondazione Don Gnocchi Onlus, Italy. Thanks to him for sharing his computerised test adopting touch-screen technology with me. Without his knowledge, the adaptation to the Eye Tracker camera would not have been feasible. I also would
like to thank him for his invaluable help in data analysis, including his explanations of this to me in an easy and understandable way.

Prof. dr. Peter Brugger, professor in neuropsychology at the University Hospital Zurich.

I would like to thank him for sharing his ZüMAX assessment with me and for providing it for further research purposes. Thank to him for explaining the test to me, in teaching me how to best administer it to patients and for his scientific feedback during the article writing process.

Dr. Marie-Christine Fluet and dr. Oliver Lambercy, from the Rehabilitation Engineering Laboratory, Department of Health Sciences and Technology, ETH Zurich. I want to thank them both for explaining and introducing me in the VPIT assessment, as well as for providing it to me for further research purposes. I also would like to thank them for their invaluable help in data analysis and for their constant feedback during the VPIT project.

Dr. Martin J. Watson, retired UK-based physiotherapist and academic. Thanks to him for proof reading all manuscripts for English grammar and structure, including this thesis. I really appreciated our email correspondence, where I learned a lot not only about the English language, but also about cultural differences between the UK and Switzerland.

The collaborators of two external clinics, namely Klaus Meyer, MD, Klinik Bethesda, Tschugg, Canton of Berne, Switzerland, and his team of neuropsychologists, and Andreas Sturzenegger, MD, Klinik Wald, Wald, Canton of Zurich, Switzerland, and his team of occupational therapists. Without their willingness to field-test the neglect exergames with their patients in their clinic, this research project would not have been feasible. I would therefore like to warmly thank them all for their support in accessing patients, explaining the project to them, and in supervising them while testing the intervention. Special thanks go to the therapy team leaders, Mirjam Wattinger and (later) Lia Wittmer from Tschugg and Ingmar Anner from Wald, for their pleasant and open communications with me.

The occupational therapy outpatient clinics which were willing to help recruiting participants for the VPIT and ZüMAX studies. Thank you very much for time and interest in my work.

To all participants who were involved in the research projects for this thesis, including the patients, controls and therapists. Many thanks go to them for their time, motivation, trust and interest in my research.
Special thanks go to Judith Häberli and Laura Wiederkehr, former master's degree students in the Department of Health Sciences and Technology, ETH Zurich. I would like to thank them both for their time, enthusiasm and for the inspirational discussions we have had on research – and other – topics. They both were of invaluable help to me in realizing the research projects with little time resources.

I am very grateful for the scientific and personal discussions, feedback and input I received from many other persons: my fellow researchers, dr. Jaap Swanenburg and Pierrette Baschung, who is also a PhD candidate at Maastricht University; dr. Judith Sieben and dr. Carolien Bastiaenen, lecturers at Maastricht University; Andrea Weise, occupational therapist. They kept me going and helped me clear my mind.

Warm thanks go to my paranymphs Anke Langenfeld and Corinne Ammann-Reiffer, PTs and PhD candidates at Maastricht University. I wish to thank them both for their personal and scientific support during my PhD path. We have all been sailors on the same boat. Thank you so much for helping me to steer the PhD boat through windy weather and high waves, but also for celebrating and sharing the sunny moments and little successes along the way.

I would like to thank all of my friends and work colleagues for their continual support and for believing in me. Namely, Simone Blaser, Franziska Gisler, Jela Skerlak, Franziska Megert, Linn Fleury, Hanna Horwarth, Katrin Hartmann, Anina Röthlisberger, Rebecca Lang, Barbara Aegler, Angelika Beck, Beate Städler, Simone Rapp, Sonja Schneider, Sabine and Arian Nevzati, the teams of occupational and physical therapists from the hand unit of the University Hospital in Zurich, and Inselspital in Bern, and Astrid Schmid, owner of the “Ergotherapie Hand” practice in Thun.

Finally, special thanks go to my family, my parents, sisters and my husband for their open ears and friendliness during difficult phases, and to my little sunshine Amélie Lou who was born when I was a PhD student.
Curriculum vitae

Bernadette Tobler-Ammann was born on May 1st 1981 in Münsterlingen, Switzerland. After receiving her general qualification for university entry in 2001, she initially earned a degree in Occupational Therapy from occupational therapy college “Schule für Ergotherapie Biel”, graduating in 2005. In 2009, Bernadette was awarded a University of Applied Sciences (UAS) degree (BSc Berner Fachhochschule in Occupational Therapy). This required completion of a minimum of two years of subsequent professional experience, plus a tertiary-level course in the field of health. Her tertiary-level education was completed in 2010, when she was awarded the degree of Master of Science (MSc) in Occupational Therapy by the Amsterdam University of Applied Sciences (Amsterdam, the Netherlands), in cooperation with the University of Brighton (Eastbourne, UK), the University College Sjaelland (Naestved, Denmark), the Zurich University of Applied Sciences (ZHAW, Winterthur, Switzerland), and the Karolinska Institutet (Stockholm, Sweden).

Bernadette worked at a Red Cross institution (2005-2009), as a team leader at a private hospital (2009-2012), and at two different University hospitals (2012-2013 // 2013-2016). Since 2016, she has worked in private practice. Her main work focus is hand rehabilitation after trauma and disease. In 2012, she became a research fellow at the Physiotherapy and Occupational Therapy Research Center, Directorate of Research and Education, at University Hospital Zurich (Switzerland). In this role, she was admitted as a PhD candidate at the Care and Public Health Research Institute (CAPHRI), Department of Epidemiology, Maastricht University (the Netherlands). As a PhD student, she had the opportunity to become a member of the European research consortium “Rehabilitative Wayout in Responsive Home Environment” (REwire), where she developed her interest in stroke and virtual reality research. She supervised three Masters students between 2013 and 2015.

Since 2013, Bernadette has been a guest lecturer at the ZHAW, being involved in teaching occupational therapists to obtain their Certificates of Advanced Studies (CAS) in “best practice”, “occupational therapy today and tomorrow” and “occupation”. As an external examiner, she evaluates Master of Advanced Studies (MAS) theses in the field of hand rehabilitation, doing this since the first diploma was awarded in 2016. Between 2015 and 2016, Bernadette was a guest presenter (via internet) for the OTD OT 915 course, Occupational Science, School for Health Professions, Concordia University (Wisconsin, USA).
Since 2005, Bernadette has been a member of the Swiss Occupational Therapy Association (EVS), serving as co-president of the executive committee of the “Bern-Solothurn” department from 2007-2013. In 2010, she became a member of both the Swiss Hand Therapy Association (SGHR) and the European Federation of Societies for Hand Therapy (EFSHT). In 2014, she assisted the SGHR to organize the Swiss Hand Therapy Conference, and holds the mandate for the research fund of the SGHR since 2016. Bernadette contributed as co-worker and chair for the EFSHT scientific committee from 2011-2017, where she assisted (Paris, 2014) and headed (Budapest, 2017) the abstract reviewing process for the triennial European hand therapy conference. Bernadette has experience in reviewing scientific articles for peer-reviewed international journals, including the American Journal of Occupational Therapy, the British Journal of Hand Therapy and the Journal of Medical Internet Research. She served as an external consultant for a Swiss occupational therapy research and evaluation project from 2014-2015 (“Bieler Modell – Quo Vadis?”). In 2017, she organized a further education event in hand therapy (“occupation-based hand therapy: hype or hope?”).

Bernadette lives with her husband and daughter in Steffisburg, Switzerland.

Publications

International peer reviewed publications


Tschanz K, Zehnder T, Tobler-Ammann B. Zusammenhang des „Montreal Cognitive Assessment“ (MoCA) mit alltagsrelevanten Fertigkeiten – eine Pilotstudie mit Klienten nach Schlaganfall. (Relationship between performing the „Montreal Cognitive
Assessment” (MoCA) and everyday life tasks – a pilot study in people early post-stroke). ergoscience 2017, 12(3): 114-121.


National scientific publications


Conference proceedings


Tobler-Ammann BC. Occupation-based hand therapy - hype or hope? 18th Conference of the Swiss Society for Hand Therapy, November 24-25, 2016, Geneva, Switzerland. Oral presentation


Ammann BC, Knols RH, Baschung P, de Bie RA, de Bruin ED. Application of principles of exercise training in sub-acute and chronic stroke survivors: a systematic review. 13th Day of Clinical Research, June 12, 2014, Zürich, Switzerland. Poster presentation.
